Proceedings of the 4th Central European PhD Workshop on Technological Change and Development

Proceedings of the 4th Central European PhD Workshop on Technological Change and Development

Proceedings of the 4th Central European PhD Workshop on Technological Change and Development organized by the University of Szeged Faculty of Economics and Business Administration Doctoral School in Economics

SZEGEDI TUDOMÁNYEGYETEM
GAZDASÁGTUDOMÁNYI KAR

Proceedings of the 4th Central European PhD Workshop on Technological Change and Development

Edited by:
Beáta Udvari

Szeged, 2020

Editor:

Beáta Udvari

Reviewers:

Beáta Farkas
Péter Halmosi
Balázs Hámori
Mónika Imreh-Tóth
Beáta Kincsesné Vajda
Gábor Dávid Kiss
Andreász Kosztopulosz
Anita Pelle
Szabolcs Prónay
Sarolta Somosi
Izabella Szakálné Kanó
Miklós Szanyi
Beáta Udvari
Éva Voszka

ISBN: 978-963-306-741-3

Contents

Contributors

László Dinya professor, University of Szeged
Anikó Klausmann-Dinya PhD student, Eszterházy Károly University of Eger
Balázs Ács PhD student, University of Szeged, head of office, KTI Institute for Transport Sciences
Róbert Marciniak senior lecturer, Corvinus University of Budapest
Péter Móricz associate professor, Corvinus University of Budapest
Máté Baksa PhD student, Corvinus University of Budapest
Eszter Megyeri assistant professor, University of Szeged
Gabriella Tabajdi PhD student, University of Szeged
Margit Tarjányi PhD student, University of Szeged
Márton Vilmányi associate professor, University of Szeged
Ilkin Salmanli PhD student, Szent István University
Sevinj Omarli PhD student, Corvinus University of Budapest
Ilona Ida Balog PhD student, University of Szeged
Sarolta Somosi assistant professor, University of Szeged
Gabrialla Számfira PhD student, University of Szeged
Timothy Yaw Acheampong PhD student, University of Szeged
Éva Málovics associate professor, University of Szeged
Beáta Kincsesné Vajda assistant professor, University of Szeged
Elena Ştiubea PhD student, Doctoral School of Economic Sciences, Oradea
Péter Füzes PhD student, Corvinus University of Budapest
Roland Zs. Szabó associate professor, Corvinus University of Budapest
Zoltán Gódor PhD student, Corvinus University of Budapest
Marcell Zoltán Végh PhD student, University of Szeged
Ábel Czékus PhD student, University of Szeged
Senanu Kwasi Klutse PhD student, University of Szeged
Balázs Tóth PhD student, University of Szeged
Dániel Szládek PhD student, University of Szeged
John Kibara Manyeki PhD student, University of Szeged
Balázs Kotosz associate professor, University of Szeged
Ibrahim Niftiyev PhD student, University of Szeged
Sisay Demissew Beyene PhD student, University of Szeged
Isaac Kwesi Ampah PhD student, University of Szeged

Preface

The fourth international PhD workshop entitled Technological Change and Development, organized by the Doctoral School in Economics at the University of Szeged, took place in April 2019. Just like in the previous years, the workshop was held in parallel with the international conference of the Faculty of Economics and Business Administration. This offered PhD students a special opportunity for personal interactions with senior and young researchers from several countries. From the presentations of the PhD session and the conference, we selected and peer-reviewed 22 papers with authors coming from various doctoral schools and higher education institutions.

This time we invited again papers, addressing a better understanding of the impacts of technological change, industry 4.0 on the economy and development. The broadly defined subject offered the possibility to the participants to concentrate on the topic, most relevant from the point of view of their theses. This is why the topics and approaches of the volume are rich and varied.

The first chapter of the book deals with business development including the analysis of business and management environment and sectoral analyses. All seven papers are built on sophisticated methodological background. The geographical coverage includes Hungary, Central Europe and Azerbaijan. The next chapter analyses the impacts of innovation covering a wide range of topics: national accounts, agriculture, doctor-patient communication, middle-income trap. The third part is dedicated to financial issues like taxation, macroeconomic imbalance, fiscal distress at local municipalities. The fourth part focuses on the changing international relations covering topics of inequality, Dutch disease, misinvoicing and the practical test of Ricardian equivalence hypothesis.

We owe our thanks to the reviewers and to Bettina Ambrus and Judit Dús for formatting the document with high precision.

Szeged, 2020

The Editor

Chapter I

Business development

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 10–22.

Managing complexity in the era of Industry 4.0

László Dinya – Anikó Klausmann-Dinya

Complexity is one of the biggest barriers to success in organizations, whether in the business or nonbusiness sectors. Despite this fact there is very little research into the causes and consequences of this rapidly growing problem in the era of Industry 4.0. Similarly, there is very little practical information that provides actionable advice on how management in organizations can attack this problem.

Internal complexity challenges like economic turbulence, understanding changes in customer needs, coping with economic crises, successfully launching innovative new products or services, dealing with regulatory changes, and finding and keeping talent are all major issues of management. In coordinating internal complexity with a complex competitive external environment, management of organizations needs to continually respond in order to succeed.

We define complexity as the number of components in a system plus the variety of relationships among these components plus the pace of change in both the components and the relationships.

Larger systems are often more complex – but they may just be more complicated if their behavior is unpredictable. Based on the database of the Global Entrepreneurship Index (GEI) we compared the EU-member countries (especially Hungary) and how prepared they are for management of growing complexity. Simplicity in business exists when we have exactly the right number of essential components and connections to achieve a successful result – no more, no less. That means everybody has to find an optimal level of complexity, which is called simplicity, or good complexity, so we can talk about good or bad complexity, and their respective levels are changing continuously.

Investigating the countries, we have identified three clusters of displaying different management challenges: balanced, flexible, and vulnerable countries with regard to their capacity to face and manage growing complexity. Hungary is among the latter group.

Keywords: complexity challenges, management 4.0

1. Introduction and literature

Complexity will be one of the biggest barriers to success of organizations in the foreseeable future. There are many growing challenges like economic turbulence, understanding changes in customer needs, coping with emerging economies, successfully launching innovative new products or services, dealing with regulatory change, and finding and keeping talent, which are all major issues that preoccupying company leaders. In combination these amount to a complex competitive environment, which firms need to continually respond to in order to succeed (Bockelbrink et al. 2018).

We know that the competitive environment is becoming more complex, turbulent and unpredictable, and managers have little or no control over the underlying trends, from globalization, to increasingly segmented markets, to technological change. The most damaging kinds of complexity, however, come from within. More products and services, more strategic initiatives, more layers of management, more processes, procedures, disrupting innovations – until managers are overwhelmed (Dinya 2012).

We can see a natural evolution of the business models, driven by technological opportunities, new management practices, and the growing diversity of customer needs. The external complexity, which is called Industry 4.0 drives internal complexity.

We define the term complexity based on the literature as follows. **Complexity** is the **number of components** in a system, plus the **variety of relationships** among these components, plus the **pace of change** of both the components and the relationships. Complex systems are characterized by diversity, ambiguity, and unpredictability of outcomes relative to inputs, or changes in conditions. The interaction of three dimensions – number of components, variety of relationships and pace of change in both – means we cannot easily tell what a complex system is going to do. It also means it is more difficult to control. As a general rule, the more a system is made up of people, the more complex it is. The **simplicity in business** exists when you have exactly the **right number of essential components** and **connections** to achieve a successful result. No more, no less. This is the good complexity (Heywood et al. 2010).

These definitions are useful as they can be applied to most business systems, at any level: from firms in a supply chain, functional departments in a firm, machines in a production line, or people in an organization. In fact, you can look at any complex system and identify whether the overall complexity is being driven by the number of components, the variety of different components, the number of connections, the pace of change, or a combination of these factors. Once you know what type of complexity you are dealing with, the solution for the complexity problem becomes much clearer.

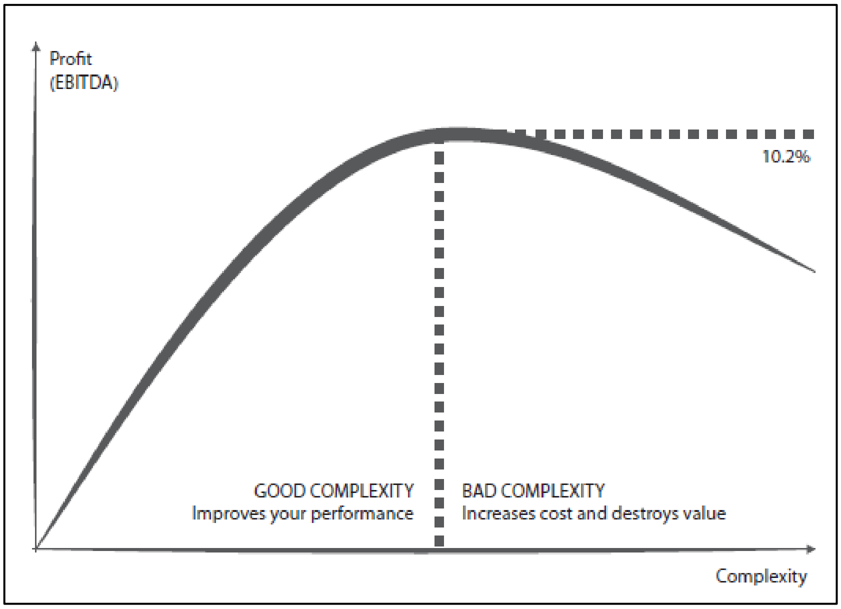
We have to talk about external and internal complexity specifically in relation to firms. Companies succeed, fail, or simply survive in complex competitive environments full of opportunities and threats, which they have to continually respond to. The scale of complexity facing organizations alone represents a significant challenge. It drives greater uncertainty, and unpredictability and makes decision-making more difficult. Firms have a broader range of options to choose from, but a more confused information picture on which to base decisions. Effective allocation of scarce resources becomes more challenging.

Two main forms of complexity relevant to business organizations are commonly discussed: **strategic complexity** and **organizational complexity**. Strategic complexity is about the positioning of the firm in a changing external competitive environment, and the management decision-making processes that try to navigate the best path through this environment. Focusing on dynamic capabilities managers have to improve agility and responsiveness in the face of chaotic or turbulent environments address this kind of complexity and the firm’s ability to survive (Gottfredson 2012).

Organizational complexity refers to internal sources of complexity stemming from the evolution of business divisions, processes, procedures and rules, and changing structural characteristics. Both forms are associated with positive (good) and negative (bad) performance effects of complexity. Managerial decisions are about finding the right balance – between 'good' complexity and 'bad' complexity. As successful firms grow, they add new products and services to their portfolios, enter new markets, engage in joint-ventures and acquisitions, and add new business units and lines of management; these strategic initiatives adding value and profits also growing. This is good complexity. At some point (we predictably called it the 'tipping point') added complexity – a new line of products, one more acquisition, an extra layer of management – does not add proportionate value. The firm does more things and the number of components and/or interrelationships grows, but the added value is outweighed by the added cost of the complexity. Bad complexity is costly complexity – and if it becomes too overwhelming, it can kill not just profits but the entire business (Kerr 2012).

A group of experts developed a complex indicator (called Global Simplicity Index, or GCI) some years ago and investigated hundreds of companies among the Top 500 (Collinson and Jay 2012) In total, they applied 18 proxy measures: nine for performance and nine for complexity – these being combined into the GCI. They found that a profit loss of companies with higher complexity (bad complexity) than the optimum was 10.2% (approx. 1.2 Billion USD) (Figure 1).

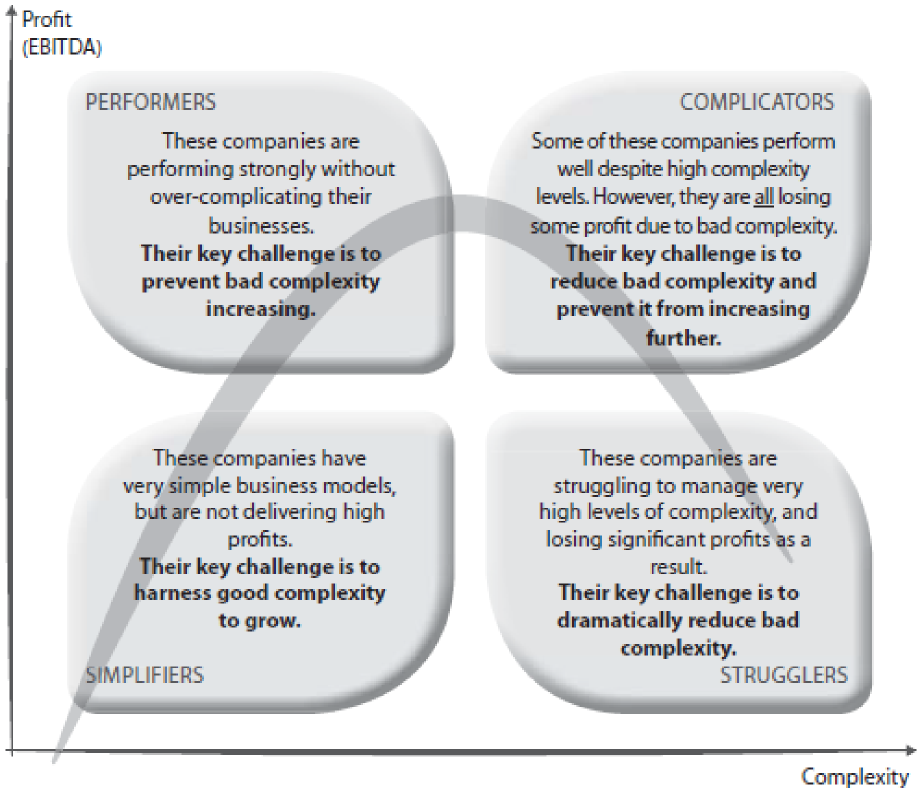
Figure 1 The relationship between performance and complexity



Source: Collinson and Jay (2012). Remark: EBITDA is the earnings before interest, taxes, depreciation, and amortization.

Investigating the top 200 companies globally they could identify four groups displaying different management challenges: simplifiers, complicators, struggles, and performers concerning their focus on complexity (Figure 2). They defined that complexity of organizations has six dimensions (internal: people – strategy – process – product – structure, plus external environment) and their priorities must be redefined in the era of Industry 4.0.

Figure 2 The performance–simplicity matrix – characteristic types of companies



Source: Collins and Jay (2012)

This led us to the idea that studying the challenges of complexity management should be started at the macro-level in order to take into account the differences of environmental (external) complexity. The potential proposed solutions and principles must be in harmony with the local (national) environment – there can be no best practice, just different good practices depending on the local situation. Familiar with the results of the Global Entrepreneurship Index (GEI), we thought that their database of 137 countries could be a good basis for our research for several reasons (Ács et al. 2017, Lafuente et al. 2019):

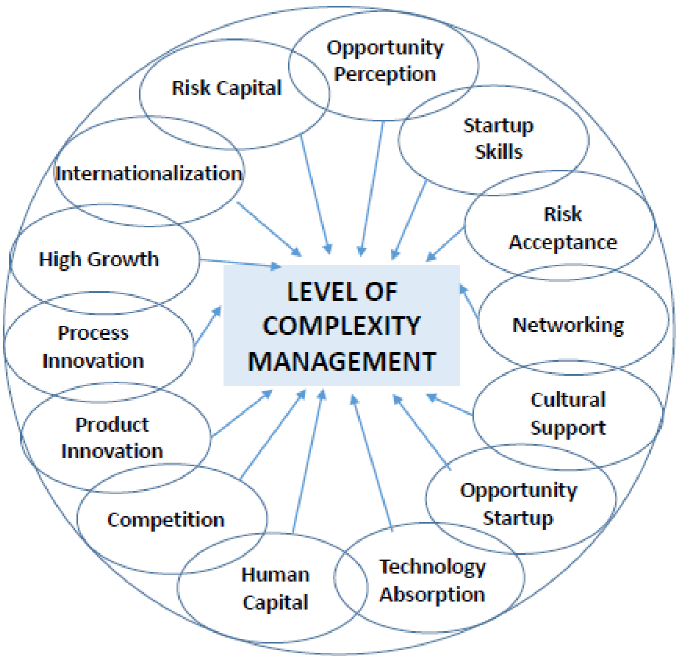
The Global Entrepreneurship Index (GEI) is a composite indicator of the health of the entrepreneurship ecosystem in a given country measuring the quality (level) of entrepreneurship and the extent and depth of the supporting entrepreneurial ecosystem.

- It gives excellent and detailed information on 137 countries concerning their level of entrepreneurship, which is a relevant pointer to the future competitiveness of their economy.
- But the indicators applied in the GEI have another dimension too: they are characterizing their relationship with future changes in the world of business. Namely: the lower the level of entrepreneurship in a country is, the more threatened they are by changes of growing complexity, in accordance with the literature.
- Each indicator of the GEI (see later) has some pertinence to the environmental and/or internal complexity. Companies and their management must be ready for continuous reconfiguration of their business model, for more and more disrupting innovations in every field of operation.

2. Investigation and results

We have elaborated a model to investigate the readiness of countries to manage the growing complexity of global business environment (Figure 3). It based on the combination of the international experiences and literature of complexity management over the last few years, and the database of the Global Entrepreneurship Index (GEI). As a first approach we have collected a database of 27 EU-member countries (only Malta has not been included) and the 14 indicators describing them as follows. The values of 14 indicators of GEI-database are normalized (their average is 0, standard deviation is +/–1,0). It would be possible to widen the focus of investigation globally in future, but firstly we tried to investigate the EU and especially the position of Hungary. We conducted a factor analysis of the 14 indicators to check their interdependence (if existent) and to form complex indicators (factors) from them for later use. Every group of interdependent indicators (variables) forms a certain factor (complex indicator), also with normalized values. After defining the professional meaning (content) of the factors (as complex indicators) we conducted cluster analysis based on the factor-weight matrix of the 27 countries. Investigating the different clusters (classes) of the countries it was possible to characterize the distribution of the EU-countries by their position regarding growing complexity.

Figure 3 The model of the investigation



Source: Own construction

We used these 14 variables ("pillars") for measuring the position of countries at 14 areas of internal conditions and external environment, and we characterized them (additionally to the original definition in the GEI) from the point of view of complexity management, as follows:

1 - Entrepreneurial Attitudes: This gives a picture of how a country thinks about entrepreneurship, or talking about quickly growing complexity, how familiar the culture is (value range) with the higher complexity of challenges.

1.1: *Opportunity Perception*. This pillar captures the potential "opportunity perception" of a population by considering the state of property rights and the regulatory burden that could limit the real exploitation of the recognized business opportunity.

1.2: *Startup Skills*. Launching a successful venture requires the potential entrepreneur to have the necessary startup skills, including how to deal with the complex environment?

1.3: *Risk Acceptance*. Of the personal entrepreneurial traits, fear of failure is one of the most important obstacles to a startup. Aversion to high-risk enterprises can retard nascent entrepreneurship.

1.4: *Networking*. Networking combines an entrepreneur's personal knowledge with their ability to connect to others in a country and the whole world.

1.5: *Cultural Support*. This pillar is a combined measure of how a country's inhabitants view entrepreneurs in terms of status and career choice, and how the level of corruption in that country affects this view.

2 - Entrepreneurial Abilities: It measures the level of how people are prepared for future uncertainties, do they have the necessary skills or competencies?

2.1: *Opportunity Startup.* This is a measure of startups by people who are motivated by opportunity but face red tape and tax payment. An entrepreneur's motivation for starting a business is an important signal of quality.

2.2: *Technology Absorption*. In the era of Industry 4.0, information and communication technologies (ICT) play a crucial role in adapting to complexity.

2.3: *Human Capital*. The prevalence of high-quality human capital is vitally important for ventures that are highly innovative and require an educated, experienced, and healthy workforce to continue to grow.

2.4: *Competition*. Competition is a measure of a business's product or market uniqueness, combined with the market power of existing businesses and business groups and the effectiveness of anti-monopoly regulation.

3 - Entrepreneurial Aspirations: This is for measuring the readiness and ambition of people to survive in global competition, in other words, are they clear with what is meant by continuous and disruptive innovation.

3.1: *Product Innovation*. New products play a crucial role in the economies of all countries. While countries were once the source of most new products, today developing countries are producing products that are dramatically cheaper than their Western equivalents. The high level of this indicator is a typical sign of affinity to rapid exchange of company portfolio, which is fundamentally important in a time of unpredictable (complex) environmental changes.

3.2: *Process Innovation*. Applying and/or creating new technology is another important feature of businesses with high-growth potential. It is another important dimension of readiness for disrupting renewal of the business model, or the operation.

3.3: *High Growth*. High Growth is a combined measure of the percentage of high-growth businesses that intend to employ at least 10 people and plan to grow more than 50 percent in five years with business strategy sophistication and the possibility of venture capital financing. It is a characteristic indicator of competition focus, not just the survival struggles.

3.4: *Internationalization*. Internationalization is believed to be a major determinant of growth. A widely applied proxy for internationalization is exporting. Exporting demands capabilities beyond those needed by businesses that produce only for domestic markets. We should not forget, that companies in more open (globalized) economies – like Hungary – are more vulnerable to the threats of complexity.

3.5: *Risk Capital*. The availability of risk finance, particularly equity rather than debt, is an essential precondition for fulfilling entrepreneurial aspirations that are beyond an individual entrepreneur's personal financial resources, even in the time of unpredictable challenges.

As a first step of factor analysis we wanted to ensure the homogeneity of the dataset to provide the most characteristic representation of the assumed interrelationships. Taking into account the values of MSA> 0.5 and KMO> 0.8), all of the 14 variables seemed to be important. So, it was advisable to involve these indicators into the investigation. The results of the factor analysis are summarized in Table 1. The names of indicators are abbreviated. The findings are:

- Based on the intrinsic values greater than 1, we found of the four factors that they had a high level of explanation: compressing 81.4% of the information content of the 14 indicators. Based on the Kaiser-test we found the first three factors to be the best approach, because they contain all 14 variables. Communality of the original indicators was also appropriate: each of the values ranging from 0.734 to 0.903 is well above the empirical rule of min. 0.25, and the Bartlett test was significant (0.00). Thus, the results of the factor analysis based on the database are methodologically correct.
- The maximum factor weights at the 14 indicators are higher than the expected min. 0.3 (between 0.614 and 0.926). According to this, the professional interpretation of the factors (based on the respective indicators) is as follows:
- *F1: Complex (socio-economic) readiness level.* With the exception of three of the 14 original indicators, 11 indicators are mutually synchronous and intertwined in this factor. It suggests that these indicators should be considered and treated as a complex, common changing system, and for example must be dealt with together in macro-level decisions aiming to increase readiness to manage the threats of growing complexity.
- *F2: this is related to one indicator that changes independently of all others, the level of internationalization*. This suggests that the degree of internationalization is independent from how companies (and the business ecosystem) are prepared for managing increasing complexity. Otherwise: companies could be prepared for it, or not, in an opened or a closed economy.
- *F3*: we find two variables here – *the level of start-up skills (0.656) and the risk capital (-0.614)* changing independently from the rest. Because their sign is opposite, this suggests that in countries where the start-up skills are higher there is less demand (or supply) of risk capital (and vice versa). And another message is: readiness for managing complexity is independent of start-up background. The *F4* is just a residual complex indicator without professional content.

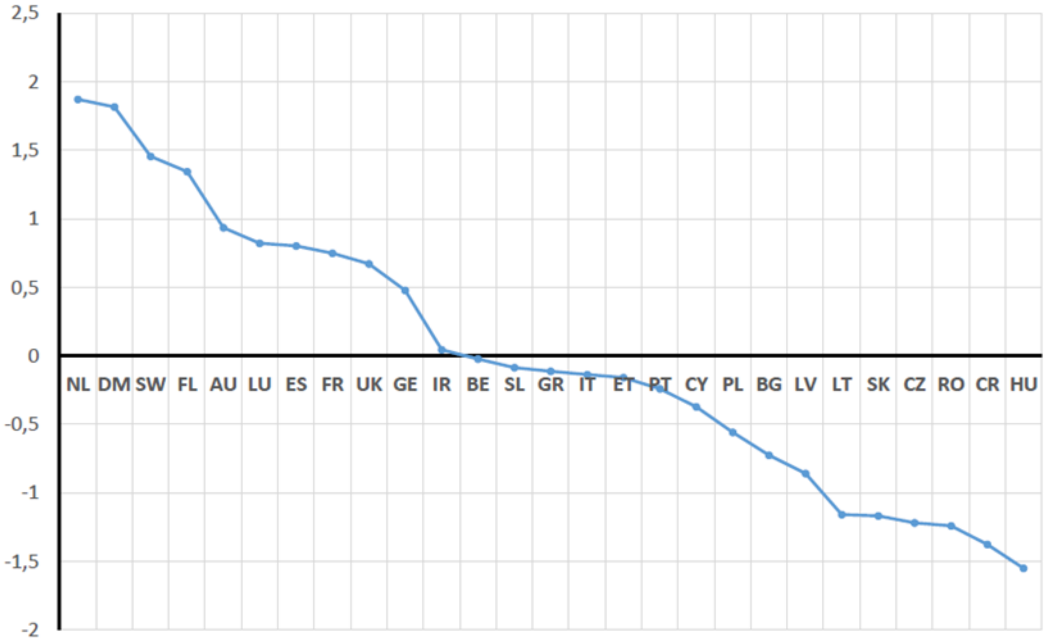
Table 1 Results of factor analysis

INDICATORS	Components			
	F1	**F2**	**F3**	**F4**
OPPORTPERCEPT	.918	–.158	–.007	.121
STARTUPSKILLS	.101	–.481	.656	.250
RISKACCEPT	.798	–.091	.326	–.176
NETWORKING	.688	–.507	–.138	–.130
CULTSUPPORT	.884	–.246	–.048	.094
OPPORTSTARTUP	.889	–.201	–.039	.137
TECHABSORP	.926	.056	.146	–.143
HUMANCAPITAL	.673	.143	–.251	.457
COMPETITION	.922	.065	–.052	–.061
PRODUCTINNOV	.724	.337	–.207	.103
PROCESSINNOV	.669	.086	.333	–.538
HIGHGROWTH	.624	.458	.192	.460
INTERNATIONAL	.327	.787	.245	–.211
RISKCAPITAL	.552	–.073	–.614	–.295
LOADINGS (%) ∑ = 81.4%	**53.4**	**11.5**	**9.1**	**7.4**

Source: Own calculation

If we take the F1 as the *complex readiness level* of a country (business ecosystem) to face the *threats of increasing complexity it seems to be useful for ranking* (Figure 4). *Decision makers* at every (macro- and micro-) level have to think through what the weakest (last) place of Hungary in the ranking means from the perspective of solving complexity management tasks! The vertical axis shows the normalized values of F1 factor in each country, where the average performance is 0, the positive values represent above average performance, while the negative ones are under average.

Figure 4 Ranking of EU-members based on their level of complexity management

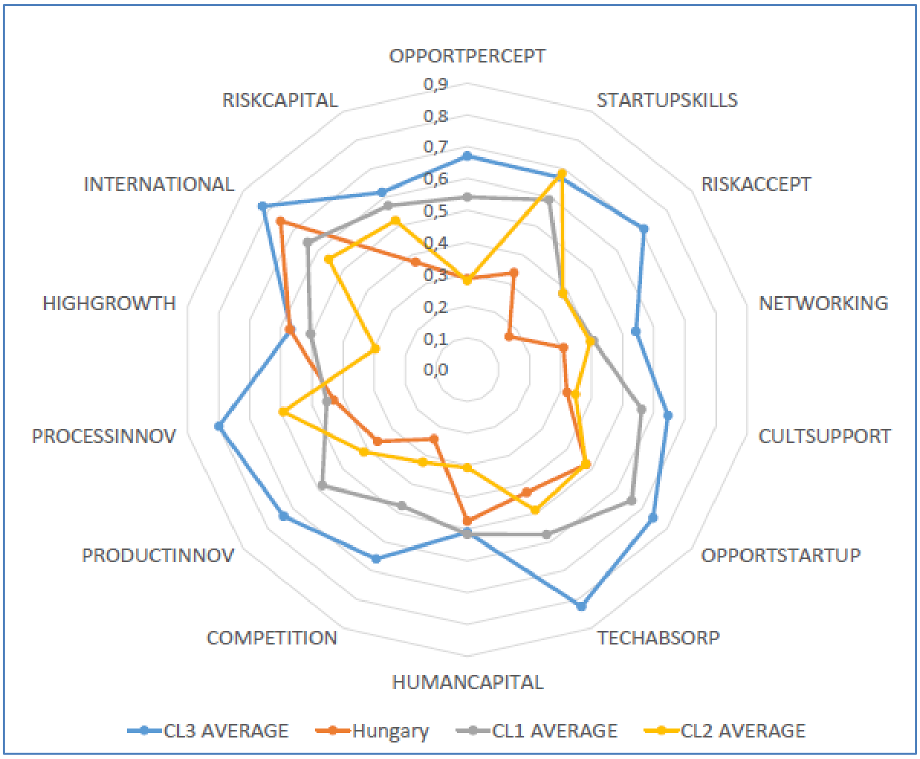


Source: Own construction

The cluster analysis with the four factors (as complex indicators) was performed on the 27-member database, and we determined the number and characteristics of the countries belonging to a distinct type (CL1, CL2, CL3 clusters) (Figure 5).

By trying different numbers of clusters, finally, three clusters gave the most definite results. Cluster member countries are referred to by their abbreviated names. The CL3 cluster of so-called "Prepared countries" (SW, FR, NL, FL, AU, GE, BE, LU, SL, SK, CZ) has the highest value of F1 (0.468) – they seem to be the most „robust" countries in the face of complexity challenges. The CL2 cluster (UK, DM, IR, ET, LV, PL, CY, LT, RO, including Hungary too, but not in Figure 5, because we wanted to show its profile separately) is the group of "Strongly threatened countries" based on their lowest value (–0.336) of F1. The CL1 cluster of "Vulnerable countries" (PT, ES, IT, GR, CR, BG, where F1 = –0.298) is practically the group of moderately threatened economies is characterized by some readiness (F1) and the least developed financial market (F2) among the three clusters.

Figure 5 Comparing the cluster profiles and Hungary



Source: Own construction

Hungary is a special case – its F1 value is the weakest in the EU (–1.549). Its weak points are:

– Concerning entrepreneurial attitudes: opportunity perception, risk acceptance, networking, cultural support
– Concerning entrepreneurial abilities: technology absorption, competition
– Concerning entrepreneurial aspirations: product innovation, process innovation
– Besides the above-mentioned, but independent of the F1 complex indicator, there are also weak points in start-up skills and risk capital.

The value (level) of these indicators are weaker than the average level of the CL3 cluster ("Prepared countries") or the CL1 cluster ("Vulnerable countries"), even the average level of our CL2 cluster ("Strongly threatened countries"). Only the level of high growth and the human capital could be termed acceptable – if we could maintain these levels into the future, but certain dangerous can be ascertained in the growing crises of the education and healthcare systems, the continuous migration of qualified workforce into Western Europe, and the possible future decrease in EU-subsidies.

3. Conclusion

We often meet with good-news reports in the media about the growing performance of the Hungarian economy. With the benefit of careful selective immersion, these reports can be supported by some indicators. But the statements can only be taken seriously if they stand in international comparison and in a broader context. Above all, the overall picture must be balanced, and provide sufficient guarantees for the country to respond effectively to the challenges of the future. This time we have studied one of these challenges, namely the threats of the accelerating global complexity in the economy as a consequence of the industrial revolution 4.0, and how countries are prepared for facing it.

As a result of our research, we have found that the pool of the EU member states shows a very mixed picture in this respect. A significant group of the 27 investigated countries (those with the most advanced business ecosystem, cluster CL3) are at a relatively acceptable level for all of the indicators examined. Some of the indicators in another group (cluster CL1) are weaker and that's why these are more vulnerable, but there is a chance for them to find the right answer to more complex challenges. However, there is a very vulnerable group of member states (cluster CL2) – which, if they do not try to build (rebuild) a viable, flexible, and entrepreneurial business ecosystem as quickly as possible, will not be able to successfully negotiate the maze of the fourth industrial revolution.

Hungary's situation is very specific: it is a member of the CL2 cluster, so in many respects it is highly threatened. Unfortunately, however, even in this cluster it differs in a negative sense, because most of those indicators are weaker (in some cases significantly) compared to the cluster average - indicators which are essential for the future competitiveness of the business ecosystem.

Finally, all EU member states (and all players in the globalized economy) face big and rapidly growing complex challenges. Countries whose business ecosystems are weaker than average (for various reasons) seem particularly vulnerable. All this underlines the responsibility of macro-level decision-makers and organizational-level managers to form business ecosystems by working together in the right direction and in a timely manner, as soon as possible.

The results of our research are a kind of diagnosis that draws attention to the points of business ecosystems where significant improvements and transformations are needed to survive in an era of growing complexity. The results also point to the fact that, because of the complex context of the ecosystem's characteristic features, it is pointless to find a single best rescue measure - the viable solution and the path to it may vary from country to country. And this is especially true in Hungary, because its position is very specific as our results show.

References

Ács, Z. J. – Szerb, L. – Lloyd, A. (2017): The Global Entrepreneurship Index – 2018. The Global Entrepreneurship and Development Institute, Washington, D.C., USA, 1–89.

Bockelbrink, B. – Priest, J. – David, L. (2018): Sociocracy 3.0 - A Practical Guide, 1–153. https://s3canvas.sociocracy30.org/s3-organization-canvas.html#organization-canvas-guide Accessed: October 27, 2019.

Collinson, S. – Jay, M. (2012): *From Complexity to Simplicity – Unleash Your Organization's Potential*. Palgrave Macmillan, London, 1–48.

Dinya, L. (2012): The green local economy. *International Journal of Business Insights and Transformations*, 5, 3, 84–91.

Gottfredson, M. (2012): The focused company. *Bain & Company*, 1–12.

Heywood, S. – Hillar, R. – Turnbull, D. (2010): *How do I manage the complexity in my organization?* Organization Practice - McKinsey & Company, 1–7.

Kerr, F. (2012): *Creating and leading adaptive organisations*. PhD-thesis, The University of Adelaide, 1–275.

Lafuente, E. – Leiva, J. C. – Moreno-Gomez, J. – Szerb, L. (2019): A *non-parametric analysis of competitiveness efficiency: The relevance of firm size and the configuration of competitive pillars.* BRQ Bus. Res. Q. 2019. https://doi.org/10.1016/j.brq.2019.02.002 Accessed: October 27, 2019.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 23–44.

New business models on the deregulated coach market in Central Europe

Balázs Ács

In many Central European countries, passenger transport markets have gone through a liberalization process, new business models and new brands have emerged, such as Flixbus, RegioJet, and Leo Express. Others have been less successful and have disappeared, such as Polskibus, Berlin Linien Bus and Megabus. Their story and the lessons learnt show that liberalized coach markets tend to return to oligopoly or monopoly, as in the case of Flixbus, which works on the deregulated coach market just like competent authorities do on regulated markets, however, without governmental control policy and use of public funds. Hungary is on the verge of deciding on liberalization as the public service contracts of incumbent bus operators expire at the end of 2019. This paper provides a general overview of the Hungarian coach system, (as there is hardly any reliable English language source on this topic), and tries to answer the question: could a commercial coach service – as a new business model – be viable in Hungary?

Keywords: long distance coach market, Flixbus CEE, liberalization, competition policy, public service obligation

1. Introduction

Liberalization of the coach market has happened in most European countries, however, in Hungary it is yet to come. The public service contracts of the incumbent Hungarian coach operators will expire by the end of 2019, which could be a perfect moment for a certain degree of deregulation. Considering the characteristics of the Hungarian bus (and railway) market, is liberalization a real option?

This paper first briefly discusses the main definitions, then recalls the evolution of liberalization through the experience of a selection of countries. Special attention is paid to the Visegrád (V4) countries, and the recently opened coach markets of Germany and France. The last section introduces the Hungarian passenger transport system, and tries to answer the research question: Could a commercial coach service – as a new business model – be economically viable in Hungary?

Some of those who fear coach liberalization worry about its projected negative effects on railways. However, the European experience has confirmed that coach liberalization does not harm the railways, as the coach boom did not continue indefinitely, but rather stabilized after a while at a certain sustainable level and went through a kind of consolidation.

2. Legal framework

The philosophy behind regulation is to provide free and fair competition in the market of public transportation services, which also includes the separation of commercial and public services, especially in terms of financing. In Hungary commercial passenger services are almost non-existent, and therefore unknown, due to the present structure of a regularized market system based only on public service obligation (PSO). This peculiarity has necessitated the development of an understanding between PSO and commercial services.

Regulation (EC) No 1370/2007 explains the division between PSO and commercial (non-PSO) services as follows: "many inland passenger transport services which are required in the general economic interest cannot be operated on a commercial basis. The competent authorities of the Member States must be able to act to ensure that such services are provided." Among the relevant mechanisms to ensure that public passenger transport services are provided, the most important is "the grant of financial compensation to public service operators".

The term 'Public service obligation' means "a requirement defined or determined by a competent authority in order to ensure public passenger transport services in the general interest that an operator, if it were considering its own commercial interests, would not assume or would not assume to the same extent or under the same conditions without reward" (Regulation (EC) No 1370/2007, Article 2, Definitions, e) paragraph).

This definition contains another important legal term, 'competent authority', which means "any public authority or group of public authorities of a Member State or Member States which has the power to intervene in public passenger transport in a given geographical area or any body vested with such authority" (Directive, Article 2, Definitions, b) paragraph). In the case of PSO, the competent authority has the right to plan timetables, choose operators (on competitive tendering, public procurement), and has the right to distribute the financial resources among the operators to cover their costs which are not covered by their income. Therefore, under PSO, entry to the market is regulated, while competition among operators is not on the market, but for the market and for the financial compensation. Conversely, in the case of commercial services, competition is on the market, in particular for ticket revenue from passengers.

In this paper, the term PSO is used for services that are subsidized and regulated by a competent authority, while the term 'commercial' is used for services on the deregulated, liberalized markets. This paper contains a description of the evolution of the commercial coach services through the examples of some relevant European countries. These services are generally deregulated, the market is liberalized, and there is no intervention of any competent authority. However, during the research a second research question arose: whether these liberalized markets tend to return to a status where one or several operators or brands start to behave just like competent authorities do in the PSO context.

Liberalization is considered complementary with deregulation. "The concept of 'deregulating' transport industries is examined to identify the elements and outcomes of the process by drawing examples of the principal modes (bus and coach, rail, air) in various countries. Experience in the British case is used as a starting point. A distinction is drawn between 'deregulation' (which may apply both to publicly and privately-owned operations), and 'privatization' (the transfer of assets and/or operations to the private sector. It is important to distinguish impacts of deregulation from those of other factors which will also affect the performance (such as trends in ridership) of the industries concerned, in drawing conclusions about its role. (White–Sturt 2009).

The philosophy behind regulation is to provide free and fair competition in the market of public transportation services, which also includes the separation of commercial and public services, especially in terms of financing. In Hungary commercial passenger services are almost non-existent, and therefore unknown, due to the present structure of a regularized market system based only on public service obligation (PSO). This peculiarity has necessitated the development of an understanding between PSO and commercial services.

Regulation (EC) No 1370/2007 explains the division between PSO and commercial (non-PSO) services as follows: "many inland passenger transport services which are required in the general economic interest cannot be operated on a commercial basis. The competent authorities of the Member States must be able to act to ensure that such services are provided." Among the relevant mechanisms to ensure that public passenger transport services are provided, the most important is "the grant of financial compensation to public service operators".

The term 'Public service obligation' means "a requirement defined or determined by a competent authority in order to ensure public passenger transport services in the general interest that an operator, if it were considering its own commercial interests, would not assume or would not assume to the same extent or under the same conditions without reward" (Regulation (EC) No 1370/2007, Article 2, Definitions, e) paragraph).

This definition contains another important legal term, 'competent authority', which means "any public authority or group of public authorities of a Member State or Member States which has the power to intervene in public passenger transport in a given geographical area or any body vested with such authority" (Directive, Article 2, Definitions, b) paragraph). In the case of PSO, the competent authority has the right to plan timetables, choose operators (on competitive tendering, public procurement), and has the right to distribute the financial resources among the operators to cover their costs which are not covered by their income. Therefore, under PSO, entry to the market is regulated, while competition among operators is not on the market, but for the market and for the financial compensation. Conversely, in the case of commercial services, competition is on the market, in particular for ticket revenue from passengers.

In this paper, the term PSO is used for services that are subsidized and regulated by a competent authority, while the term 'commercial' is used for services on the deregulated, liberalized markets. This paper contains a description of the evolution of the commercial coach services through the examples of some relevant European countries. These services are generally deregulated, the market is liberalized, and there is no intervention of any competent authority. However, during the research a second research question arose: whether these liberalized markets tend to return to a status where one or several operators or brands start to behave just like competent authorities do in the PSO context.

Liberalization is considered complementary with deregulation. "The concept of 'deregulating' transport industries is examined to identify the elements and outcomes of the process by drawing examples of the principal modes (bus and coach, rail, air) in various countries. Experience in the British case is used as a starting point. A distinction is drawn between 'deregulation' (which may apply both to publicly and privately-owned operations), and 'privatization' (the transfer of assets and/or operations to the private sector. It is important to distinguish impacts of deregulation from those of other factors which will also affect the performance (such as trends in ridership) of the industries concerned, in drawing conclusions about its role. (White–Sturt 2009).

3. Why long-distance services tend to be commercial?

Provision of passenger transport services used to be a good business until private car use became dominant and ridership began to shrink. Railway companies had started to suffer from losses at the beginning of the second half of the 20th century, while bus operators reached that level a few decades later. Nowadays most regional services - especially in rural areas where ridership is very low - require subsidy, or financial compensation. Suburban services in Europe have high, or even growing ridership, but when these are operated in fare alliances, the allocation of income is not transparent, and if the buses are stuck in traffic congestion, the cost of operation is high. Long-distance services, however, are generally capable of covering operating costs from revenue.

There are some reasons for this. Unlike regional buses, which run 40–60 thousand kilometers a year, a long-distance coach runs 200–400 thousand kilometers, and therefore depreciation and all other costs divided by this distance is significantly lower. Similarly, personnel costs are also lower, because a coach driver may drive 2–3 times further during his or her shift than regional bus drivers.

On the income side, regional and suburban tariffs are "a few" Euros for a ride, as there is not much chance to attract passengers with pricing techniques. However, in case of long-distance trips, passengers are easier to push into off-peak times by demand-responsive pricing.

4. Milestones of coach liberalization in Europe

Many European countries have undergone the deregulation process, each of them has unique solutions. The focus of this paper is on Central-Eastern European (CEE) countries; however, it is worth scanning through the main Western-European milestones leading to the almost Europe-wide liberalization process. The chronology of liberalization can be easily summarized as there was a well-planned first wave in the 80s-90s, followed by various solutions in CEE countries, and a second big Western wave originating from Germany in the 2010s, and finally the remaining states. This paper follows this structure.

4.1. United Kingdom

The pioneering role of bus liberalization was played by the Thatcher government in Great Britain in the 1980s. In the coach market, the incumbent National Express (NE) had to face some emerging operators, especially on the London-Oxford route, but NE remained the main player, and most challengers threw in the towel (van de Velde 2009). There were serious "bus wars" on regional routes, but after a while the situation has consolidated.

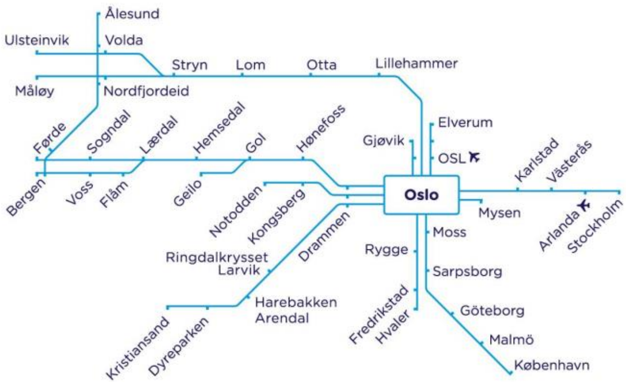
Among the challengers of NE there was one exception, Megabus, which was launched in 2003, and still has a significant market share (White–Robbins 2012). Megabus as a brand belongs to the Stagecoach group, one of the Big 5 transportation operators (the other 4 being National Express, Go Ahead, Arriva, and First). These companies are active not only at the bus and coach business, but also in the railway sector, and not only in the field of commercial services, but they operate services under PSO. Their business models contain both bidding for franchises and also real on-the-market-competition.

4.2. Scandinavia

Following Britain, the two main Scandinavian countries (Sweden and Norway) decided to liberalize their coach market. In Sweden the then incumbent Swebus had been a subsidiary of SJ (the Swedish state railways), serving routes where rail service was inadequate. It was privatized in 1996 and sold to Stagecoach. After the 1999 deregulation of the Swedish coach market, the new owner of Swebus became Concordia Bus, and in May 2018 the company was sold to Flixbus and the brand Swebus started to disappear. Swebus had around 50% market share in 2014. There are other brands and operators (e.g. Bus4You), their network is also Stockholm-centered, operating more or less in parallel with rail lines and each other. The competition forced both coach and railway operators to improve their services, and it resulted in growing patronage for public transport. Over the last two decades, the passenger kilometer figure for coach and rail has grown more than that of the private cars, and it represents the success of liberalization (Alexandersson 2014).

The evolution of the Norwegian coach market has provided some points of similarity to CEE countries. The process is described in detail by Aarhaug–Fearnley (2016): “Apart from its major cities, Norway is a country that is hard to serve by high capacity public transport, like rail. However, scheduled coach services were for a long time strictly regulated in order to protect the railways. Before 1998, it was the responsibility of coach companies to prove that they were not in competition with the railways, in order to obtain licenses to operate– similar to the requirements in Germany (Walter et al. 2011). In the 1980s, most express coach routes were extended local routes, and local bus companies operated most of them. Operators with area licenses in neighboring counties cooperated and joined their licenses in order to operate through services. Such cooperation was a de-facto requirement for establishing express coach routes (Leiren–Fearnley 2008). This regulation was first lifted in limited areas in 1999, when consideration of passenger benefit became important. Then, in 2003, entry regulation was abolished all together for county border crossing services, as is the official term (Leiren et al. 2007)”.

Figures 1 and 2 The route map of Nor-Way Bussekspress and Nettbus Express



Source: nor-way.no, visitnorway.com. (2019)

The express coach market carried 5.3 million passengers offering 34 million bus kilometers in 2010. For comparison, the total Norwegian scheduled bus market (including express coach services) has 314 million passengers and 249 million bus kilometers. (Aarhaug–Fearnley 2016).

The former Norwegian transport policy considered the express coaches to be a competitor for subsidized rail (which was seen as a natural monopoly), and there was a fear that deregulating the coach market would reduce rail ridership. This fear proved false, in spite of a rapid growth in the coach industry, “railways experienced steady growth after the express coach deregulation” (Aarhaug–Fearnley 2016).

The main lesson learnt in Norway is that coach liberalization did not harm railway patronage, rather it motivated railways to improve their service level and productivity.

The main bus brands in Norway are Nettbus Ekspress and Nor-way Bussekspress (see Figure 1 and 2). The former is a marketing umbrella for various operators, owned by the Norwegian State Railways (NSB). The brand Bus4You, which is known in Sweden, too, belongs to Nettbus. Nor-way is another umbrella company, partly owned by Nettbus. Small and medium enterprises (SMEs) in Norway may enter the coach market under these umbrella brand names as operators.

5. Experience of the Visegrád countries

After the changes in 1990, many former Eastern Block countries had to face difficulties in organizing and financing public transport services. Most of them found unique solutions, but generally the long-distance coach market became more or less liberalized, and kept their ridership, while the regional services suffered from decreasing ridership and growing costs.

In the Visegrád countries the initial situation was very similar in 1990. Each country had their bus and coach company groups, PKS in Poland, CSAD in Czech Republic, and VOLÁN in Hungary. The group of companies were divided according to administrative boundaries, by voivodships in Poland and by counties in the other two, however, some mergers and break-ups occurred. In Poland there were some 200 hundred PKS companies, while in Hungary 19 to 29 VOLÁN companies.

5.1. Poland

Poland is a pioneer in trying various business models for operating bus and coach services. Steps toward liberalization were taken in the early 90s. The PKS companies had gone through various reform processes, including simplification of company profiles, mergers, and privatization, and by the end some of them had gone bankrupt, while some are still suffering and only a few may be considered safe and sound. Poland was active in trying various business models for railway operation. “It was as late as in 2007 that a first private carrier appeared on the market, with today's Arriva RP providing services on behalf of Kujawsko-Pomorskie voivodship. Simultaneously, local and regional governments became, not only organizers of rail transport, but also owners of carriers” (Taylor–Ciechański 2018)

Figures 3, 4 and 5 A route map Polskibus, LUX Express, and PKS Express

Source: polskibus.pl (2015) and luxexpress.eu (2019) and autokar24.pl (2019)

The first challenge PKS had to face right after deregulation was the emergence of small private enterprises operating minibuses on regional and suburban lines, just a few minutes ahead of the PKS service. "A great threat here is the unfair competition engaged in as certain private carriers pick up passengers just prior to the due time of arrival of a PKS bus" (Taylor–Ciechański 2018)

On the coach market, first the National Express group launched a local brand called Polski Express, but this company eventually gave up competing with PKS, especially when PKS launched their own coach brand (PKS Express, see Figure 5). Many years later, in 2011, the Stagecoach group discovered Poland as an appropriate field to enter the market and launched the Polskibus brand, which survived for 7 years, right until the strong expansion of Flixbus.

"Despite the observed increase in competition, the bulk of scheduled passenger carriage continues to be done by PKS companies. Unfortunately, the future of these PKS companies is not very bright, partly because strategy accepted is concerned with survival, rather than development" (Taylor–Ciechański 2018)

Besides (the former) Polskibus (Figure 3), there are other commercial bus operators in Poland, such as LUX Express, Simple Express (Figure 4), etc., these companies being based in the Baltic states.

Poland has a wide range of experience in regulating liberalized markets, and there are many lessons learnt here. "Unfair competition appeared between private carriers and the PKS companies, ensuring bankruptcy of some of the latter, and hence also the transport-related exclusion of many areas, including even areas of relevance to tourism. In many cases, it is also possible to observe a sort of aversion on the part of local or regional authorities to support for public transport in their areas, or even for its limitation. This is an attitude quite the opposite of the one observed in the neighboring Czech Republic, for example" (Taylor–Ciechański 2018)

5.2. The Czech Republic and Slovakia

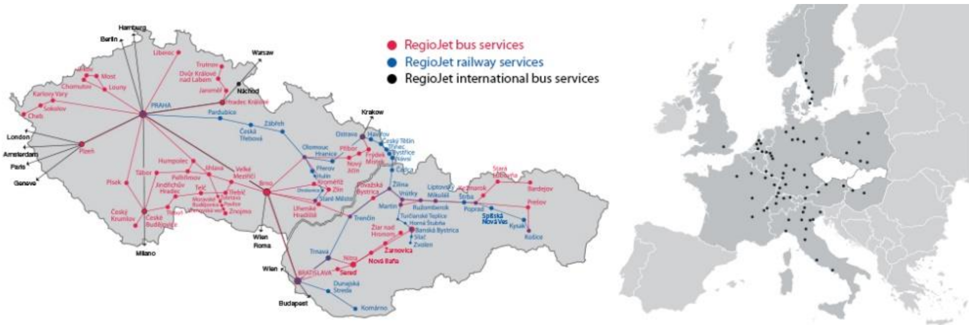
After the division of the country, CSAD remained the main bus and coach operator for the Czech part, and SAD was established for Slovakia. Some SAD companies were privatized, and now belong to Arriva. There are more than 200 Public Service Contracts in the Czech Republic for bus and coach services under PSO.

There is a private company which dominates the commercial coach market in both countries. It was initially called Student Agency[1], but since 2015 its name has been RegioJet (Figure 6). The big yellow buses run mainly on domestic routes in the Czech Republic, and as the liberalization process went through, also in Slovakia, but they also operate many international lines. The buses provide luxury onboard services and free drinks offered by an attendant.

[1] The company was founded and is still mainly owned by Radim Jančura. Their first money-maker business activity transporting young Czech citizens (mainly students, hence the name) to Great Britain for babysitting and au pair jobs. Later they opened other scheduled international routes, for example from Prague to Budapest via Brno and Bratislava.

The company is involved not only in the commercial coach operation, but also in commercial (open access) rail services within the Czech Republic and to Slovakia and Austria. Besides this, RegioJet operates rail services under PSO. The company won the contract for the Bratislava-Komarno line in 2013, and further rural lines in the Czech Republic from 2019. RegioJet is not just a brand-name or an umbrella, but a real vehicle operator, owning many buses and rail cars. RegioJet organizes its coach and railway timetable so that it feeds its own rail services by coaches and decreases parallel coach services on those routes it starts to serve by rail. Basically, RegioJet acts like a competent authority on the commercial long-distance market, therefore it has a very special business model.

Figure 6 The route map of RegioJet coach (red) and rail (blue) services

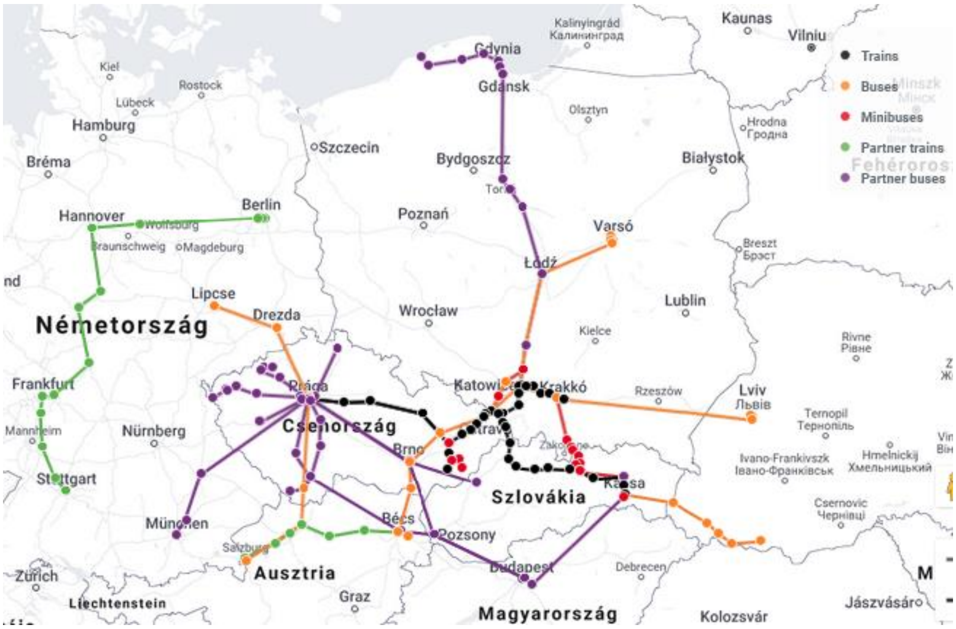


Source: regiojet.cz (2016)

LEO Express is another private Czech company which operates based on a very similar business model that of RegioJet. The main difference is historical: LEO Express was initiated as a commercial (open access) passenger railway operator, between Prague and Ostrava, right after the opening of the Czech domestic railway market in 2011. Later its network was extended by connecting feeder coach and bus services. Moreover, the company operates feeder minibuses and airport shuttle services, partly to feed its trains (Figure 7). Therefore, LEO Express also functions as a long-distance transport organizer competent authority (should) do.

LEO Express is the train operating company which took over Locomore's German domestic open access train services. Locomore's business model was supposed to be a special kind of community funding, but it failed, and now it operates under the brand name Flixtrain, in strong cooperation with Flixbus (green line on the western side of Figure 7). LEO Express also cooperates with the Austrian open access railway operator, Westbahn, (just like RegioJet does, it is the green line in Austria). LEO Express operated coach services under its own brand name between Krakow and Budapest for a few months in winter 2017, but currently LEO Express is represented in Hungary only as a partner of Eurolines on the Budapest–Prague route. LEO Express is a perfect example of a pioneering company which tries various business models especially in the field of cooperation.

Figure 7 The route map of LEO Express coach, bus and rail services



Source: leoexpress.com (2018)

6. The latest wave of liberalization

The recent changes in the Western-European domestic and international coach markets significantly redrew the transportation map of Europe. The deregulation process started in Germany in 2013, followed by France, Italy, Austria, and Switzerland. These countries used to be known for protecting their railways from competition, but recently the situation has completely changed.

6.1. Germany

Long-distance coach services in Germany used to be very limited. From 1934 to 2012 the law on public transport (Personenbeförderungsgesezt) basically prohibited the operation of passenger coach services, with the intention to protect the railways from competition. There were two main exceptions: some services from West-Berlin to West Germany (operated by Berlin Linien Bus), and the Romantische Strasse line connecting major tourist attractions.

From 2013, Germany raised the barriers and liberalized coach services, with very few remaining constraints (no trips are allowed for short distances (less than 50 km), or for less than one hour if there is proper rail service.)

Many new companies launched their services, and by 2016 the coach market has reached its zenith at some 20 million carried passengers a year. Even 2 of the big 5 British transport companies entered the market: National Express under the brand City to City, and Stagecoach as Megabus. Other competitors worth mentioning are ADAC Postbus, ALDI Reisen, DeinBus, and last, but not least, MeinFernbus and Flixbus.

The then incumbent domestic operator, Berlin Linien Bus, and international operators under the Eurolines umbrella had found themselves among many challengers. In January 2015 MeinFernbus and Flixbus merged, keeping the name of Flixbus and the green-and-orange color scheme of MeinFernbus, and they together gained a dominant role in the German coach market. By the end of 2016 almost all other competitors either ceased operation, or merged into Flixbus, which reached a market share higher than 90%.

Flixbus started a very fast and aggressive international expansion, and these days their green coaches serve almost all European countries. In 2017 the company opened routes in the USA, too. Flixbus is not just opening new services, but also integrates existing ones. Many former EurolinesUROLINES partners have joined Flixbus, which offers a different business model.

The main point of the Flixbus business model is sharing risks between operators (or as Flixbus calls them: bus partners) and the brand owner. Flixbus does not operates buses or coaches itself. Flixbus behaves like a transport organizer competent authority, selecting bus partners for services it plans, and does all the marketing, sales, customer service activities, the role of bus partners being only to operate the buses (Dunmore 2016). This is actually very similar to the business model of National Express, or the Norwegian Nettbus and Nor-Way.

Flixbus has eaten up not only the main German domestic operators, but also Polskibus, Swebus, and the whole Eurolines Romania operator group. There are rumors that RegioJet is also on their menu.

6.2. France

Having observed developments in Germany, from August 2015 the French government also decided to liberalize the domestic coach market. Flixbus is in fierce competition in France with Ouibus (Figure 8), which was founded and owned (under the name iDBus) by the French state railways, SNCF. In July 2016 the company bought Starshipper (which was the alliance of French bus operating SMEs), hence the competition on French domestic routes remained only between Flixbus and Ouibus, (and some other, smaller brands, under the Eurolines umbrella). However, in November 2018 Ouibus was purchased by BlaBlaCar, which is the main carsharing platform in Western Europe with a French background. In exchange for the deal, SNCF has some shares in BlaBlaCar. In March 2019 BlaBlaCar announced that it will attack Flixbus on the German domestic market, too. Time will show whether Flixbus may be beaten in its home playground.

Figure 8 The destinations served by Ouibus

Source: ouibus.com (2019)

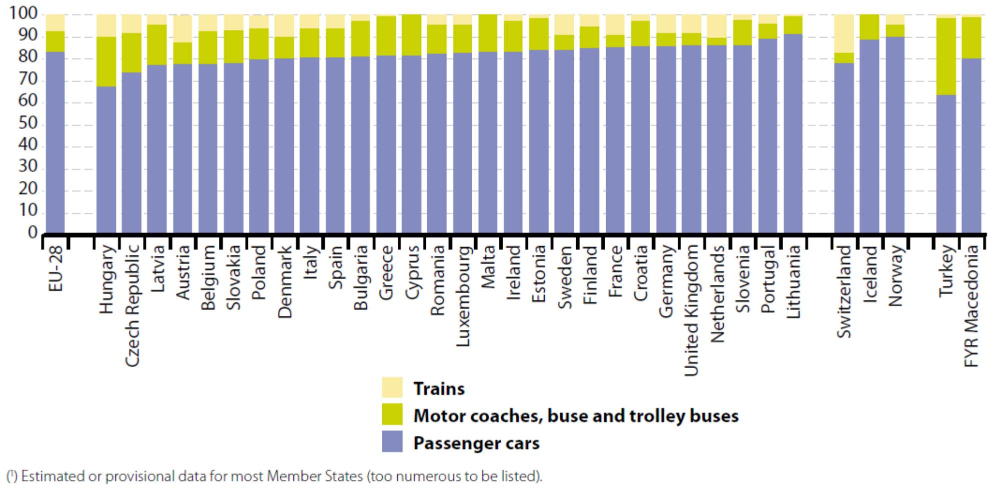
The entry of BlaBlaCar to the coach market represents a very new business model. For a carsharing company to enter the coach market, means that there has been a shift in the classic competition theory. The battle is not between rail and coach (and air), and not even between public transport and private car use, but rather between the classic transport services and the new transport modes based on the new buzzwords of digitalization, sharing economy, and the 4th industrial revolution.

7. Hungary – waiting for liberalization

Describing the Hungarian liberalized coach sector does not require much effort and space, as this is basically non-existent. All scheduled coach services are under PSO, there not being any commercial coach service in Hungary. However, in the near future, significant changes may occur in this field, this is why the case of Hungary is discussed in detail. One of the aims of this paper is to provide a general but detailed enough overview of the Hungarian coach system, because there is hardly any reliable English language source on this topic. To understand the special characteristics of Hungarian transport policy, this chapter also takes a look also towards the Hungarian railway sector.

According to Eurostat (2015), Hungary has the best modal split in the European Union (Figure 9). Based on the data of passenger-kilometer, the share of cars is around 80 per cent in the EU, while in Hungary it is around 67 per cent. In Hungary people tend to use public transport much more than the EU average, and also higher than the average of the 10 countries that joined the EU in 2004.

Figure 9 Modal split in EU countries



Source: Eurostat (2019) https://ec.europa.eu/eurostat/statistics-explained/images/5/57/Modal_split_of_passenger_transport_by_country_2013.jpg

However, if one takes a look at the state of infrastructure, the vehicles and the ticketing, the public transport system does not seem like one of the best in Europe. Buses are 14 years old on average, infrastructure is worn out, except for the recently built and renovated highways and railways. Trains are often late and not considered tidy (EU 2018).

7.1. Legal framework

The Hungarian law on passenger services was enacted in 2012 and categorizes the domestic services as local, suburban, regional and national. Local (urban) transport services (within the city limits) are authorized and financed by each local government, however, all others are run by the Ministry for Innovation and Technology (ITM). This strict division between the urban and interurban authorities might disappear in the near future.

There are hardly any private or commercial, open access operators in Hungary for bus and rail. Almost all transport services are provided under PSO, i.e. operators receive cost-compensation for those expenses which are not covered by income from passengers.

In the railway sector the public service contracts will expire in 2023, however, in the case of bus and coach operators the contracts expire at the end of 2019, currently the tendering process is ongoing. There are around 110 local (urban) PSO contracts, which have various expiry dates.

7.2. The significance of railways in transport policy

The current Hungarian government continuously emphasizes the significance of public - and especially railway - transport. The new law on passenger transport services (which came to effect in 2012) provides priority for railway operations, “as much as possible”. This unclear statement makes a perfect basis for justifying any railway project. The share of railways in the current infrastructure development programs is also very high. Railways are considered not as an economical issue but as a social one. In 2014, less than 20% of the cost of passenger railway operation was covered by passengers, Hungary is last on this list in the EU (European Commission 2016).

Railways are important for policy makers not only for domestic reasons, the development of international services is also an issue. There are millions of ethnic Hungarians living on the other side of the country borders, and the provision of adequate international public transport services for them is essential. The international railway services are integrated into the domestic timetable and fare system, and until the border station these operate under PSO., i.e. financed by taxpayer money. This policy is somehow contradictory to current EU regulations.

7.3. Market players

There are 3 main train operating companies. The biggest is MÁV-START, the passenger transport subsidiary of Hungarian State Railways (MÁV), carrying 137 million passengers a year, receiving an annual cost-compensation of HUF 144 billion (EUR 500 million, 0.4% of GDP, 3.2 EUR/passenger, 6 EUR/passenger-kilometer). There is another state-owned railway company, the „Győr-Sopron-Ebenfurthi Vasút” (GYSEV, or Raaberbahn), which is a Hungarian-Austrian joint venture, operating 8 lines in western Hungary, carrying 10 million passengers. Another subsidiary of MÁV, MÁV-HÉV operates 4 independent suburban railway lines in the Budapest area. These lines were taken over from the Budapest Transport Company (BKV) in November 2016. One of these lines is within the city limits of Budapest. There are some 20 narrow gauge forest railway lines with passenger services; however, these mainly serve as tourist attractions, so these are not operating as a PSO. There is currently no private passenger railway operator.

Regional buses and coaches had been operated for decades by 24 state owned „VOLÁN” companies, generally one (or two or three) company in each county. By January 2015 these companies were merged into 7 regional bus operators (called KKs), in preparation for the public tendering of services. The plan to further merging the companies into one big bus operator was announced in March 2019.

National, regional and suburban buses carry around 450 million passengers per annum and receive a cost-compensation of around HUF 65 billion (EUR 208 million, 0.4 EUR/passenger). This amount is increasing year by year; however, until 2006 these companies were able to operate profitably. VOLÁN companies also operate the urban buses in some 70 towns, carrying another few hundred million passengers. (Another 40 towns order local services from either small private operators or from their own bus company.)

The biggest VOLÁN company is the Budapest based VOLÁNBUSZ, which operates suburban and regional buses around Budapest, and also many nationwide highway coaches. VOLÁNBUSZ is the main international coach operator and had been the Hungarian partner for Eurolines for decades, however, by 2019 most of its international lines were operating under the brand Flixbus. There are also 4 small private bus operators with public service contracts, 2 of them operating highway coaches (once a day from Budapest to Nyíregyháza and to Zalaegerszeg), the other two operating 1-2 regional lines under PSO.

Some of the VOLÁN companies have tendered out the operation of some services to private firms (subcontractors). The maximum share of subcontracted services is limited to 49% in their contracts; VOLÁNBUSZ had been the closest to this threshold. By 2019 the share of subcontractors has been decreased to a just few per cent after a change in the government policy.

Between June 2014 and November 2016 VOLÁNBUSZ operated 150 (blue) buses in the Budapest suburban area as a subcontractor of the Centre for Budapest Transport (BKK, the competent transport authority for Budapest). In 2016 this contract was transferred into a PSO as ordered by the Ministry (then NFM, now ITM).

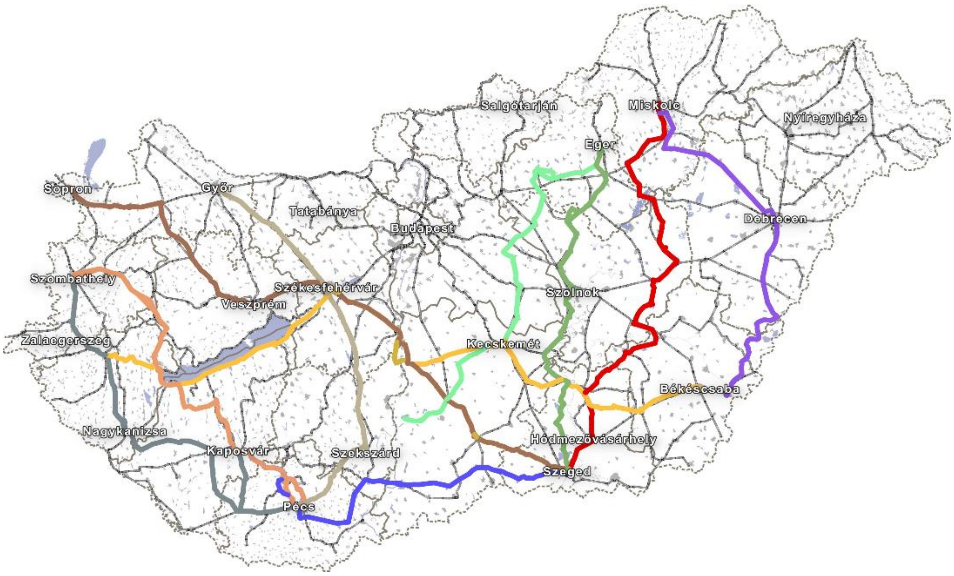
Entry into the bus and coach market is very limited. Liberalization of the market is yet to come, but it has low priority, mainly for the purpose of protecting railways from competition, just like in Norway in the late 90s. However, a certain number of lines, especially those considered to be parallel with subsidized railway services may be liberalized, or, at least, financial compensation would be withdrawn. Recently small and medium sized enterprises (SMEs) may enter the market as subcontractors. From 2020, they may have a chance to win national or regional tenders, if any are announced.

7.4. Coverage, accessibility

The Hungarian public transport network is extended to all but 4 villages, 3145 settlements are served by regular bus services (while 5 others only have rail). The timetable provides enough services for the vast majority of village residents to commute to the nearby town to school, to work in typical work shifts or for shopping.

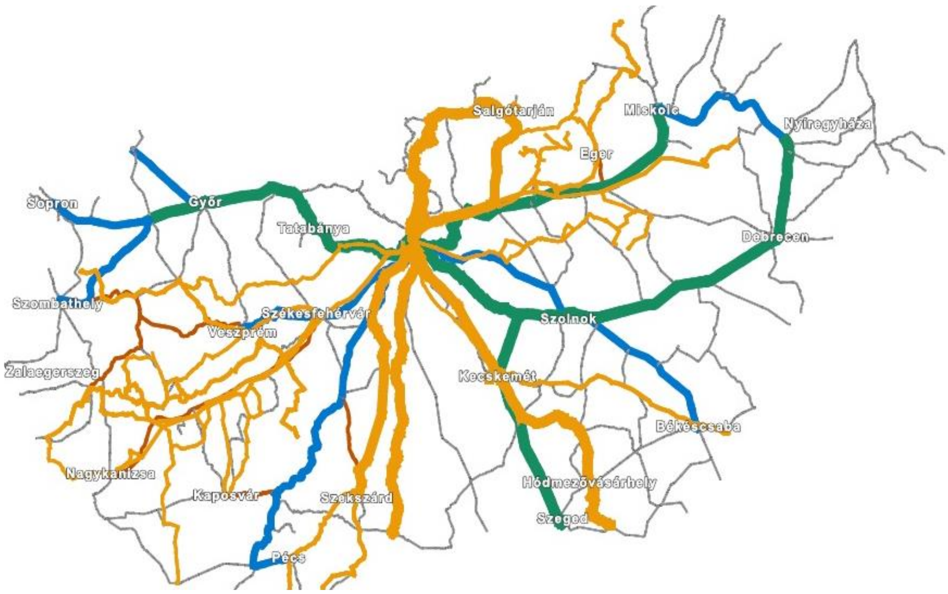
The frequency of the timetable on the main lines is usually attractive, and on some suburban and interurban lines buses operate frequently and periodically, based on the periodic, clock-face schedule, or as it is known, the “taktfahrplan-concept”. However, in remote rural areas villages are served only few times a day, while in some counties there are plenty of small villages where there is no bus service on Sundays, or even on Saturdays.

Figure 10 Diagonal long-distance (over 200 km) bus routes in Hungary. Not parallel with railways.



Source: KTI Institute for Transport Sciences

Figure 11 The long-distance Budapest-originated rail (green, blue, red) and coach (yellow) routes in Hungary. Hardly any seems to be parallel. The wideness of the lines represent significance



Source: KTI Institute for Transport Sciences

The coach network is also very extensive. Major towns and cities are directly connected with each other, even if they are some hundred kilometers away from each other. These diagonal coach routes are 150-300 kilometers long, and generally connect communities which are not served well by rail. (Figure 10). Generally, very few passengers travel all the way on these routes, as travel times are long, and coaches stop in each village on route. That is why most long-distance services carry significant number of short-distance passengers. These services are integrated into regional service structures, and serve needs which cannot be met by rail services. These routes cannot be considered as real long-distance routes rather "interconnected regional" ("láncolt helyközi") services, and therefore could hardly be deregulated.

The capital Budapest may be accessed from all bigger towns, either by rail or coach, or both (in very few cases, with reason). One of the main aims of the current Hungarian public transport policy is to dedicate only one mode of transport for a certain route as a public service, and in most cases it should be rail, as it is believed to be the most efficient solution.

7.5. Urban and other public transportation services

There are around 110 towns and cities with local urban bus networks. Most of them are operated by VOLÁN companies, and – especially in small towns – by private firms, under various circumstances. However, in 8 cities (Budapest, Miskolc, Debrecen, Szeged, Pécs, Kaposvár, from 2018 in Tatabánya, and from 2019 in Veszprém) the local buses (or even trams and trolleybuses) are operated by the internal operator company of the local government. Each city has different financial and operating models. Tariff alliances are non-existent; however, there are some tickets or passes, which combine the regional and local trips, or are valid for both operators in a city (e.g. in Szeged, urban passes are valid on buses operated by DAKK (formerly TISZA VOLÁN), and trams and trolleybuses operated by SZKT). In the Budapest area the suburban commuter passes may also be combined with the Budapest Pass (which covers the whole area of Budapest) since 2007, and hence passengers enjoy some discounts comparing to those travelling around smaller towns.

Some ferry services also operate under PSO on the rivers Danube and Tisza, but the share of waterborne services is very low. On Lake Balaton there are ferry and ship services, the latter only during the summer season, for tourism. There is no domestic air service.

7.6. Fare system

Fares in Hungary are distance based and slightly regressive. Fares used to have a small annual hike; however, there has not been any fare hike since 2010. In 2007 the rail and bus tariff level were unified in 2 steps, which meant a more than 30% increase in the rail tariff, and some 5% for buses and coaches. The relatively small surcharge which is applicable on InterCity trains (since 1991) was introduced on most highway buses in 2012. A minor surcharge for rapid trains – which was previously abolished in 1991 – was reintroduced in 2013. The same year a 15, then in 2014 a further 10 (altogether 25) percent discount was introduced on 30 rural lines with low traffic volume.

The railway company MÁV-START used to apply discounted tickets on lines parallel with highway buses, but it was almost completely abolished in 2013, and fares were set to be equal on trains and buses. Pre-purchased seat reservation on trains costs HUF 180 (EUR 0.6), however, for peak days (Friday and Sunday afternoons) and on the day of purchaseit costs HUF 300 (EUR 0.9).

Passengers pay only a very low percentage of the cost of operating the railway. The rest is paid either by the employer (in the case of commuter passes), or by the fare adjustment (SZMT) by the government (in case of the socio-political discounts), and the remaining part is covered by public service compensation. Commuter monthly pass holders are entitled to 86% compensation from their employer by law; however, some employers are reluctant to do so. Students (and many other groups like the handicapped or those with lower income) enjoy a 50% discount on the basic fare. The remaining 50% used to be covered as a social expenditure, however, nowadays it appears partly in the form of public service compensation. Pensioners may travel with 90% discounts (i.e. pay only 10% of the ticket price), and passengers over 65 (and under 6) years old may travel free, including EU citizens, however the surcharges must be paid.

7.7. Integrated Periodic Timetable (ITF)

The integrated and supply-oriented periodic timetable, also known as the ‘taktfahrplan concept’ (or Integrierter Taktfahrplan, ITF) is a special method of transportation network and timetable planning, which is based on the belief that increasing frequency and providing good connections (repeated every hour) is the key to attracting passengers. The Swiss transport system is fully based on this concept, and many other countries utilize this method. The philosophy behind the theory is that the income from new passengers would exceed the cost increase arising from the extended supply of trains (and buses).

ITF-based timetable was first introduced in 2004 on the Budapest-Vác-Szob suburban railway line and its sidings. In 2006 the taktfahrplan concept was extended onto most main railway lines in Eastern Hungary, then in the South East, and finally by 2009 on the Transdanubian (western) main lines. In 2009 the whole bus system in the western suburb of Budapest was adjusted to the new railway system, also based on the taktfahrplan concept. On the suburban railway lines around Budapest the increased frequency and the introduction of rapid train services attracted many new passengers, however, on non-Budapest lines the taktfahrplan has not been such a success story.

When promoting the idea of ITF, the main purpose was increasing ridership and income. In the last 30 years the number of train services from Budapest to the bigger towns have been doubled or tripled (with very few exceptions), based on the taktfahrplan concept. These extra services helped to improve the accessibility of rail services and contributed to stabilizing passenger numbers; however, the average distance of a rail trip has been decreasing sharply. Hence, implementing ITF could not turn back the process of losing long-distance rail passengers, however, it could attract passengers mainly on suburban routes.

7.8. Railway closures, reopening, and the 2012 passenger service supply reform

Hungary has a vast railway network; however, many lines were closed in the 70s. In 2007 and 2009, passenger railway operations were suspended on another 14 and 25 lines, 662 and 826 kilometers, respectively (before 2007 the route length was 7700 km). The new government elected in 2010 decided to reopen 11 lines (356 km), but later in 2012 the number of train services on most of these lines was cut to 2 per day per direction. Also, in April 2012, most or all highway bus services from Budapest to the following cities were cancelled: Győr, Sopron, Szombathely, Szeged, Miskolc and Nyíregyháza. These lines were considered to be parallel with rail services. On 15, April 2012, altogether 450 train and coach services were deleted from the timetable. The future of PSO highway coaches is unclear. There is a wide spread opinion that long-distance coach services which may be considered more or less parallel with good quality railway services should be cancelled. The less drastic view is that these should be transformed from PSO to commercial services.

Here is the point where the international and the Hungarian parts of this paper come together, and the following question arises:

7.9. Could a commercial coach service - as a new business model - survive in Hungary?

Taking into consideration the characteristics of the Hungarian passenger service market, the outlook is not rosy. First of all, there is some over-supply in the market of long-distance trains, the typical ridership of an InterCity train is around 50-60 passengers. This is a very low number for a train but would perfectly fill a coach. Would these passengers choose a coach instead of a train? Maybe if an hourly PSO train were replaced partly by a commercial coach, let us say, every second hour. By this solution, PSO compensation could also be dropped, however, it would contradict railway-oriented government policy.

Second, operators of PSO services are entitled to receive fare adjustment[2] for each passenger travelling with a socio-political discount. How would this be handled in the case of a commercial operator? In the Czech Republic commercial coach and open access rail operators receive a fare subsidy. If in Hungary a commercial operator had to cover its costs only from ticket income, it could hardly survive, especially because students travelling on 90% discounted monthly passes would not use the services. However, if a commercial operator were entitled to receive fare adjustment, there would have to be very strict control and a transparent ticketing system.

Third, if there were competition between PSO and commercial services, the PSO operators would have the chance to respond to the challenge, e.g. by decreasing their fares, or improving frequency, quality of cars, etc. As the financial losses of PSO operators have to be compensated by law, there is not much risk for them. For

[2] “szociálpolitikai menetdíjtámogatás”, or SZMT is a 50, 90 or 100% fare adjustment, which is paid by the Ministry of Finance based on the volume of discounted tickets sold to students, pensioners, the disabled, etc. It amounts to around HUF 90 billion per year.

commercial operators, entering a price war is riskier, and in the end, both of them may lose income, however, passengers could benefit.

Entering such a market is rather risky, especially if there is competition with state owned and state financed companies. However, there are good opportunities to test the market. Allowing cabotage, i.e. domestic travel on international coach services, may help understand market forces and the behavior of the players and passengers without risking state funds.

In conclusion, one may say that the attractiveness of the Hungarian public transport system is based on its good geographic and time-coverage and its relative cheapness for users. The implementation of the periodic timetable (taktfahrplan) offers a frequency which is greater than necessary. The wide range of discounts allows very cheap (or even free) rides for a significant number of residents. Commuter passes may be considered an extra perk from the employer, and with frequent suburban services – even on weekends – these passes offer great mobility to their holders. Such a service level makes living without cars possible for many citizens, and even families do not feel the need for a second car. Entering this fully PSO market on a commercial basis is therefore at least risky.

8. Conclusion

In most European countries long distance coach services are now liberalized and commercial, and hardly any PSO remains. After deregulation, these markets started to grow for a while, but soon reached a new equilibrium. Various kinds of new business models emerged, which use modern technology for sales, customer services, and data analysis, and there is a wide range of cooperation between the new entrants to the market.

Coach liberalization did not have a negative effect on railway ridership, rather motivated train operating companies to improve their services.

Operators and brand names must be distinguished. Some brands on the commercial markets do not even operate vehicles, they behave like a competent transport authority, selecting operators for the planned services. Other brands also behave like mobility managers, when they harmonize their timetable so that their buses and trains connect with each other.

In Hungary coach liberalization is yet to happen. There are fears that a liberalized coach market would harm the income of railways, but this has not been the case in countries where liberalization went through. Hungarian railways have the worst cost coverage ratio in the EU, meaning there is not much income to lose. On the other hand, there is probably not enough passenger income to launch a commercial coach service either, especially because many passengers are entitled to travel on subsidized fares on services under PSO. The rising car sharing solutions seem to be a more dangerous challenge for railways than coach liberalization. In France and in Germany, a brand-new business model is emerging: a car sharing company that operates, or at least, organizes coaches. Time will show whether this model is the future of commercial public transport.

Acknowledgments

I express my thanks to my guiding lights in the forest of competition, liberalization, and (de)regulation: Éva Voszka and István Szeri. I also express my gratefulness to my colleagues at KTI, Institute for Transport Sciences for their proposals, corrections, and moral support. Last but not least, I have to thank my wife for taking care of our children while I was preparing this paper.

References

Aarhaug, J. – Fearnley, N. (2016): Deregulation of the Norwegian long distance express coach market. *Transport Policy*, 46, 1–6.

Alexandersson, G. (2014): Reforms in the Swedish Bus and Railway Industries. *UITP Workshop*: New intercity bus lines: market for the bus, competition for the rail? Brussels

Dunmore, D. (2016): Comprehensive Study on Passenger Transport by Coach in Europe. Steer Davies Gleave. Prepared for DG MOVE, European Commission. https://ec.europa.eu/transport/sites/transport/files/modes/road/studies/doc/2016-04-passenger-transport-by-coach-in-europe.pdf Accessed: 11/12/2018.

Leiren, M. D. – Fearnley, N. (2008): Express coaches – the story behind a public transport success. Paper presented to the 2008 *European Transport Conference* in Leeuwenhorst Conference Centre, The Netherlands.

Leiren, M.D. – Samstad, H. – Fearnley, N. – Minken, H. (2007): Express Coach Routes – A Complex Market. *Institute of Transport Economics*, Norway, TØI report 904.

Taylor, Z. – Ciechanski, A. (2018): Systemic transformation and changes in surface transport companies in Poland: A synthesis after twenty-five years. *Journal of Transport Geography*, 70, 114–122.

Van de Velde, D. (2009).: Long-Distance Bus Services in Europe: Concessions or Free Market? *OECD International Transport Forum* Discussion Paper No. 2009–21,

Walter, M. – Haunerland, F. – Moll, R. (2011): Heavily regulated, but promising prospects: entry in the German express coach market. *Transportation Policy*, 18, 373–381.

White, P. – Sturt, A. (2009): *Transport and deregulation*. Elsevier.

White, P. – Robbins, D. (2012): Long-term development of express coach services in Britain. *Research in Transportation Economics*, 36, 30–38.

European Commission (2018): Europeans' satisfaction with passenger rail services http://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/survey/getsurveydetail/instruments/flash/surveyky/2172 Accessed: 11/03/2019.

EUROSTAT (2015): European Statistics. http://ec.europa.eu/eurostat/ Accessed: 15/03/2019

European Commission (2016): Fifth report on monitoring development of the rail market. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52016DC0780&from=EN Accessed: 11/03/2019.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 45–60.

Promotion of an Industry: Trends and Expectations of Digital Transformation in the Hungarian Business Services Sector

Róbert Marciniak – Péter Móricz – Máté Baksa

In this paper, we explore the current trends of and future expectations for digital transformation projects in Business Service Centres (BSCs) in Hungary. We carried out fifteen interviews with senior technology experts and executives in three Hungarian based BSCs of multinational parent companies to examine individual transformation projects. We also used quantitative data from large-scale surveys on the sector to get an overview of general practices. We reviewed the use of advanced technologies like robotic process automation, predictive analytics, chatbots, and artificial intelligence. We found that BSCs had mostly The consequences automated massively repeated processes and that this automation had liberated employees for more creative tasks. of this transition are threefold: (1) BSCs can reinforce their position as business partners of their global parents, (2) creative assignments are more attractive for prospective and current employees in a labour market characterized by a shortage of suitable personnel, (3) employees usually do not fear the possibility of job loss due to automation and digital transformation.

Keywords: service sector, business service centers, digital transformation, robotic process automation

1. Introduction

The expected effects of the 4th Industrial Revolution have been attracting increasing attention from researchers in economics, sociology, and organization studies. A critical combination of information technology, biotechnology, and robotics seems to be the most important driving force for the revolutionary changes around and ahead of us. Although not implied as such by the umbrella term ‘Industry 4.0’, digital transformation and automation are also substantial sources of competitive advantage for companies in the service sector. As the business service industry employs more than 50,000 people in Hungary and provides 1.2% of national GDP and 1.6% of exports alone, it is considered significant for the national economy (Drótos et al. 2018). Therefore, digital transformation in the industry and its effects on competitiveness and labour market trends require further examination.

With our research, we set out to explore current tendencies and future expectations for digital transformation in Business Service Centres (BSCs) in Hungary. To accomplish that, we investigated the following questions. Which advanced technologies are the most prevalent in BSCs? What are the motivations for and expected gains from automation? How do digitalization and automation affect their strategic positions? What tasks and processes get automated most frequently? What is the general attitude of employees towards automation? How do managers of BSCs attempt to shape and influence employee attitudes?

To answer our research questions and give a comprehensive understanding of the subject, we pursue the following line of thought. First, we propose a theoretical framework for digital transformation in BSCs, 'Business Services 4.0'. In this framework, we encapsulate the challenges and solutions offered by digital technologies in BSCs as well as indications of how these affect corporate strategies and development projects. We then present the overall trends of automation and digitalization in the business service industry in Hungary. Analysed data was collected in cooperation with the Hungarian Service and Outsourcing Association (HOA) and the Hungarian Investment Promotion Agency (HIPA). Next, we present the results of our interviews carried out in three different BSCs based in Hungary. Finally, we summarize our conclusions and propose directions for further research.

2. Theoretical framework

Due to their significance in the global supply chains of multinational companies, an increasing number of researchers are investigating the operation and strategy of BSCs. Digitalization and automation also have their stream of literature, with numerous management, information systems, and information technology journals dedicated solely to these subjects. Therefore, to understand the context of a Business Services 4.0 framework, we first give an overview of essential concepts related to the business services industry and digital transformation.

2.1. Concepts of business services and digital transformation

Business services are services that are primarily consumed by organizations. Based on input, outcome, and added value, we can establish two major categories for them: knowledge-intensive and operational services (McKinsey Global Institute 2017a). The former category includes functions such as accounting, corporate finance, research and development, and other professional services, while the latter comprises activities like facility management, temporary employment services, and contract staffing.

Corporations that provide business services are called Business Service Centres. BSCs are either independent of their client companies concerning ownership and management (i.e., outsourcing), or they are affiliates of a corporate group. In the latter case, they provide various activities and processes for other companies in the group. Their performance is usually assessed based on their output and efficiency, and measured by comparison to market prices.

The business services sector in Hungary has been growing steadily over the last decade, both in times of crisis and prosperity. It employs an increasing number of people, most of whom hold a college degree (Drótos et al. 2018). As 48% of BSCs have been present in Hungary for more than 11 years, and another 20% of them present for 7–10 years, the sector is considered to be mature. Due to automation (Lacity–Willcocks 2015b) and the appearance of advanced technologies, the whole industry seems to be moving towards higher added value creation and a more sophisticated business model.

Although neither automation nor digitalization is a new concept or phenomenon in business organizations, their volume and effect have become so significant that we can call it a new wave of digital transformation (Demeter et al. 2019, Losonci et al. 2019, Nagy 2019). In most manufacturing companies, automation preceded digitalization (e.g., assembly lines in Ford's factories). In BSCs, however, we see a reversed order: companies usually digitalize their processes first (by using document management and workflow systems) and automatize them later.

Digital transformation broadly refers to a change of organizational strategies, structures, processes, and business models in which digital contents and advanced technologies play a crucial role (Füzes et al. 2018, Hortoványi–Vilmányi 2018). Through this, companies strive to advance their adaptability and agility to keep or acquire a competitive advantage. Digital technologies now appear in every business function, changing the ways of operation and the means of value creation (Horváth–Szabó 2017).

According to Sebastian et al. (2017), digital strategy, the carrier of digital transformation may be understood as "*a business strategy, inspired by the capabilities of powerful, readily accessible technologies (like SMACIT), intent on delivering unique, integrated business capabilities in ways that are responsive to constantly changing market conditions*". They use a strategic management perspective and focus on organizational capabilities to emphasize the urgent need to adapt to an ever-changing competitive environment.

Digital transformation does not mean solving old problems with new technology: it may instead be understood as re-thinking old problems while considering novel possibilities (Andriole 2017). Even though the tools of technology are essential, it is the capability of their innovative use and combination that may result in lasting competitive advantage. In this way, digital transformation is the rethinking and renewal of the organization itself, and thus might not be implemented without the transformation of the human workforce (Eden et al. 2019, El-Khoury 2017).

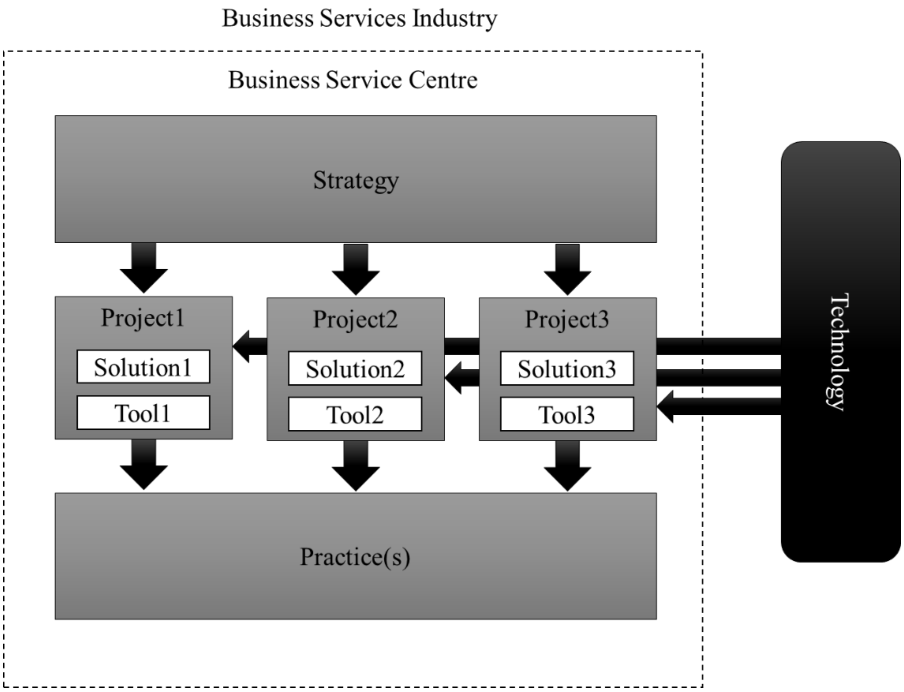
Digital transformation in the business services industry can be understood as a series of different stages and steps: (1) *digitization*, (2) *digitalization*, (3) *automation*, and (4) *robotization*. Digitization means a transformation from physically extant (i.e., hand-written or printed) documents to digital contents (Bhimani and Willcocks 2014). This step leads to digitalization, which is the application of various technologies (e.g., workflow systems) to exploit possible gains from electronic management and operation of processes (Drótos et al. 2018). Once processes are modelled and operated digitally, they can be automated. Automation permits decreasing the level of human interaction required for process completion (A.T. Kearney Inc. 2017, Everest Global Inc. 2014, Winroth et al. 2007). Robotization explicitly refers to the application of technologies that can substitute for the human workforce (Kukreja, 2016, Lacity–Willcocks 2015a, Unger et al. 2018). These technologies are primarily either physical, or software robots that specialize in specific tasks and functions otherwise carried out by humans. Robotic technology is generally based on machine learning, cognitive systems, and different kinds of artificial intelligence.

2.2. Business Services 4.0 framework

After explaining the key concepts of business services and digital transformation, we present the Business Services 4.0 framework. Based on the idea of Industry 4.0 (Nagy 2019, Szabó et al. 2019), Business Services 4.0 refers to the complex phenomena of how advanced technologies appear in the service sector, change and reshape business models, processes, and overall corporate practice (Keller 2017). We have summarized the concept of Business Services 4.0 and its main elements in Figure 1.

Our units of analysis are Business Service Centres that are affected by both the competitive environment (other players in the industry) and the most recent technological trends. These inputs might change strategic directions otherwise suggested by their parent companies. Development projects are initiated based on strategic goals. These projects use various pieces of advanced technology to create organizational solutions that will eventually alter everyday practice.

Figure 1 Business Services 4.0 framework



Source: authors' compilation

Digitalization and automation are not entirely new phenomena in the service sector (Chui et al. 2015, Héder 2014). Scripts and macros were used in different software environments to replace frequent routine tasks. These solutions, however, are limited in use. First, they cannot or can only partially cross software boundaries (e.g., MS Excel). Second, they can only use structured databases for computing. Third, they do not support end-to-end work processes as workflow systems do. Nowadays, robotic process automation (RPA) is the most widespread automation technology (Unger et al., 2018). It has already surpassed its predecessors as it can traverse its software environment, and use semi-structured databases for computing (Kukreja 2016). Robotic process automation technology encompasses software that is ready to support whole working processes and minimize human activity at each step.

Currently, cognitive automation (Davenport and Kirby 2015) represents the highest level of process automation. Armed with learning algorithms, it carries some aspects of artificial intelligence (AI). Cognitive automation can also process data from unstructured databases (McKinsey Global Institute 2017b). Chatbots, accessible and quite prevalent pieces of automation technology, make use of cognitive automation, for instance. Thanks to natural language processing, chatbots can translate spontaneously sentenced questions to queries of data. These questions can then be answered based on a database that is expanding continuously based on the questions asked.

Other technologies that we investigated in the Business Services 4.0 framework include enterprise resource planning (ERP) systems, business intelligence (BI) solutions, enterprise social media software, Big Data, and data mining. We also asked for data on the use of predictive analytics (PA), data privacy and security, simulations, service-oriented architecture, cloud computing, virtual or augmented reality, and machine to machine (M2M) systems. We found that amongst these technologies, ERP systems were the most widespread in the business services industry, although, many current and future development projects had aimed to implement workflow systems, RPA, and cognitive computing.

3. Methods

To explore digital transformation in the business services industry, we carried out a large-scale survey and several interviews with senior automation professionals and executives of BSCs. By doing so, we attempted to get an overall understanding of the industry, as well as to gain insight into motives, risks, and gains related to actual development projects. We present the process of data collection in both research phases below.

Data for our large-scale survey were collected in cooperation with the Hungarian Service and Outsourcing Association (HOA) and the Hungarian Investment Promotion Agency (HIPA). For the survey, all 110 companies operating in the business services industry in Hungary were asked to fill out a questionnaire focusing on general information, strategies and processes, employees, and technology. Seventy-one companies sent back their responses, which means a 64.5% coverage based on their absolute number and an 82.5% coverage based on total employee number (representing approximately 41,200 employees). Data for the survey was collected in autumn 2018. In some cases, responses can be compared with data from 2017 as a somewhat different version of the survey had been carried out in the previous year as well, however, with a lower level of coverage (60% based on the number of employees).

As most BSCs are located in Budapest this proportion was also apparent in our sample: approximately 80% of employees worked in the capital, while the others were located in Tier 2 cities (primarily in Debrecen, Székesfehérvár, Szeged, and Pécs). The parent companies of our respondents operate in various industries ranging from manufacturing (24%), through business services (15%) and telecommunications (10%), to energy (8%), IT (7%), and others. Further, most of our information providers perform other activities besides business services: other service activities (43%) or other manufacturing activities (21%).

Our survey sample appropriately covers the industry and adequately represents the population regarding the dimensions mentioned above. Therefore, our data can be used to describe general trends in the industry as well as cast light on similarities and commonalities among BSCs. Responses for some questions can be compared to those in the previous years as similar surveys had been carried out before with a slightly different focus and fewer respondents.

We carried out our interviews in three BSCs located in Budapest and Székesfehérvár. The companies were intentionally chosen according to the following criteria: variance in location, variance in the industry of parent company (manufacturing, IT), variance in the number of employees, and variance in technology adaptation strategy. In each case, we applied to managers of the companies with a request to support our research. Either they or their selected colleagues (senior professionals responsible for digitalization and automation projects) were then interviewed. The interviews usually took 1–1.5 hours and were semi-structured. We prepared a line of questions in advance, but as our main objective was to explore the corporate practice and learn about development projects, we sometimes deviated from the original items.

Table 1 Organizational function and affiliation of interviewees

Code	Organizational function	Company code
1	Managing Director	A
2	HR Services (external)	A
3	Procurement	A
4	Q2C (Sales Support)	A
5	Chief Information Officer	A
6	Indirect Tax	A
7	Accounts Payable	A
8	HR Transformation (internal)	A
9	Automation Team Lead	B
10	Global Compliance Lead	B
12	Security Lead	C
13	Expert Architect	C
14	IoT Portfolio Unit Lead	C
15	Managing Director	C

Source: authors' compilation

Table 1 shows a list of interviewees together with their organizational function and affiliation. Our fifteen interviewees are from three different companies (A, B, and C). Company A is in the information technology and services industry. It has its headquarters in Budapest and employs approximately 2,300 people. Company A has a global parent company with hundreds of thousands of workers worldwide. Company B operates in the manufacturing industry. Located in Székesfehérvár, it has about 500 employees. Its parent company is a global one, with a strong presence in Europe. Company C is also invested in information technology and services. It employs over 4,500 people at its Budapest headquarters. Company C is a subsidiary of a European multinational company with strong market interests in Hungary.

Each interview took between 1–1.5 hours. On some occasions, multiple interviewees were questioned at a time. All interviews were carried out by two or all three of us present. Notes and audio recordings were taken with the previously given consent of the interviewees. Recordings were transcribed by a research assistant and then reviewed and verified by senior researchers. Interview transcriptions were re-read several times during the research process to identify common themes and patterns.

4. Results

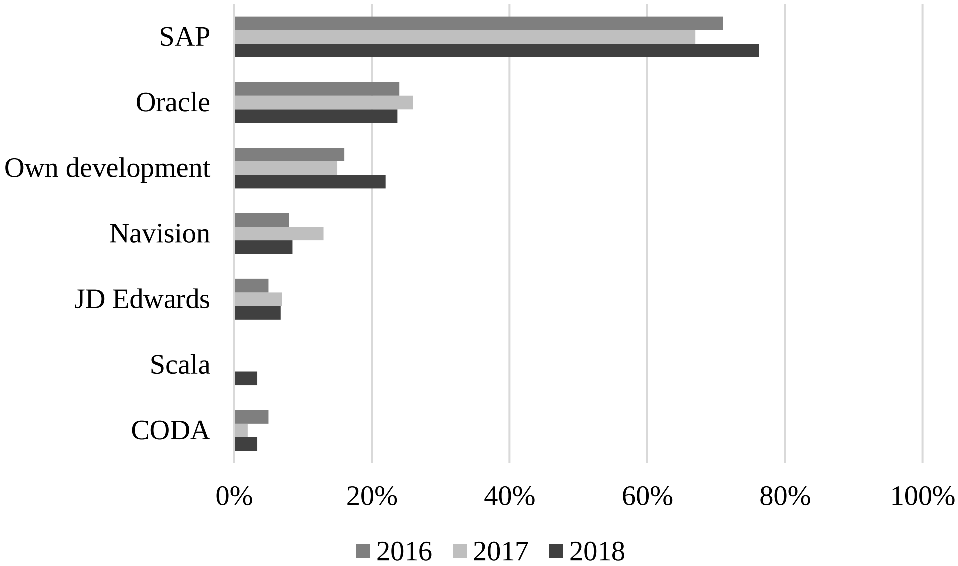
In this section of our paper, we first present the results of our large-scale survey; then we summarize the main messages and characteristic messages of the interviews. In the latter part, we give examples of specific development and digital transformation projects as well as meaningful citations from our interviewees.

4.1. Survey results

Technology has always played a significant role in the value creation of business services companies, but in the past few years, it became a key driver of growth and business development. An increasing number of BSCs have already implemented automation technologies, and even more of them are planning to do so. The operation of BSCs is becoming more streamlined than ever before: routine tasks are accomplished on a large scale with extreme efficiency. Automation of everyday tasks and use of integrated enterprise resource planning systems liberate the workforce for positions with higher added value.

Perhaps the most prevalent digitalization technology in the business services industry is enterprise resource planning (ERP). In 2018, 56% of our respondents claimed that they used a company-wide integrated ERP system, which is a 10-percentage point increase compared to the previous year. The most preferred ERP system providers are SAP and Oracle. However, more than one-fifth of BSCs, who use ERP, implemented their own developments (*Figure 2*).

Figure 2 Prevalent providers according to ERP system users – multiple choice



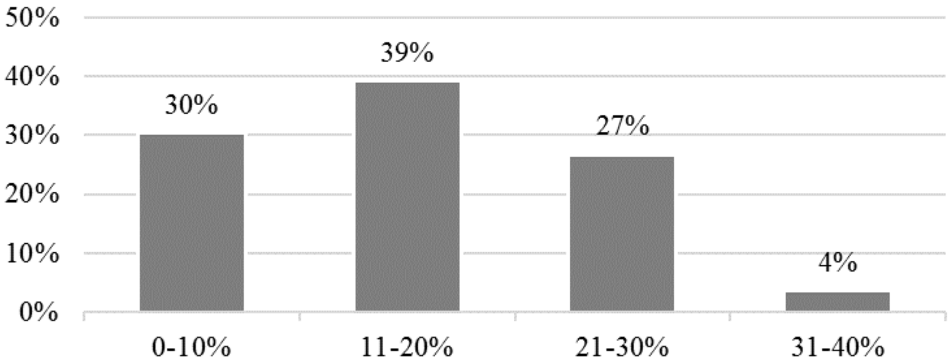
Source: authors' compilation based on survey responses

As BSCs deal with an extreme number of transactions, process automation seems a substantial means of reducing costs and lead time, while simultaneously increasing accuracy. Before automating specific processes or process steps, these should first be optimized or even re-designed to avoid any possible loss due to inadequate organization. Even so, it is usually difficult to fully automate processes that require data from different systems or that include non-routine elements. Thus, typically, human agents are kept to oversee, control, and connect separate process steps.

In the case of robotic process automation (RPA), however, software robots can completely substitute human agents in the entire process. Consequently, RPA proves to be even more beneficial in increasing efficiency and preventing errors. The main barriers against placing software robots in most administrative procedures remain their relatively high price as well as their need for thorough training and testing.

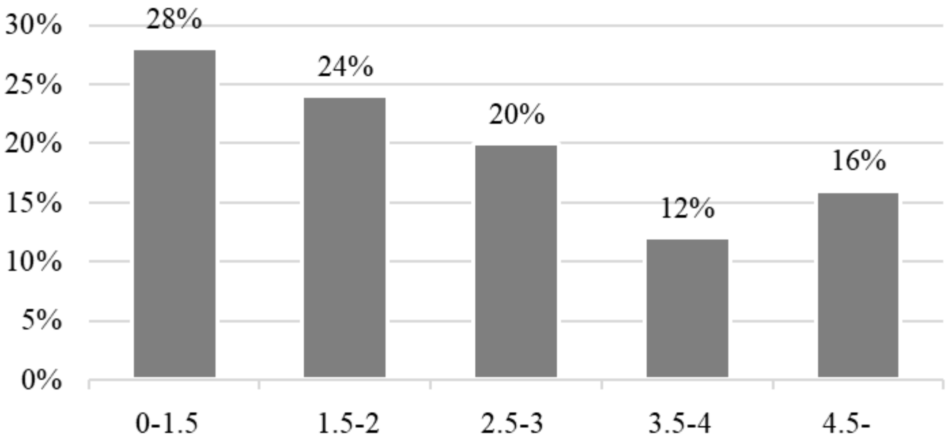
From the survey results, it becomes apparent that the motivations for and expected gains thanks to process automation and robotic process automation outweigh most objections. Considering possible cost savings on average due to automation technologies, most respondents (39%) expect 11–20% in savings (see *Figure 3)*. Approximately the same proportion (30 and 27%) of respondents are somewhat more or somewhat less optimistic, expecting 0–10 and 21–30% of average cost savings.

Figure 3 Expected cost savings on average in the next two years due to the use of automation technologies



Source: authors' compilation based on survey responses

Figure 4 Expected FTE replacement by one robotic unit on average

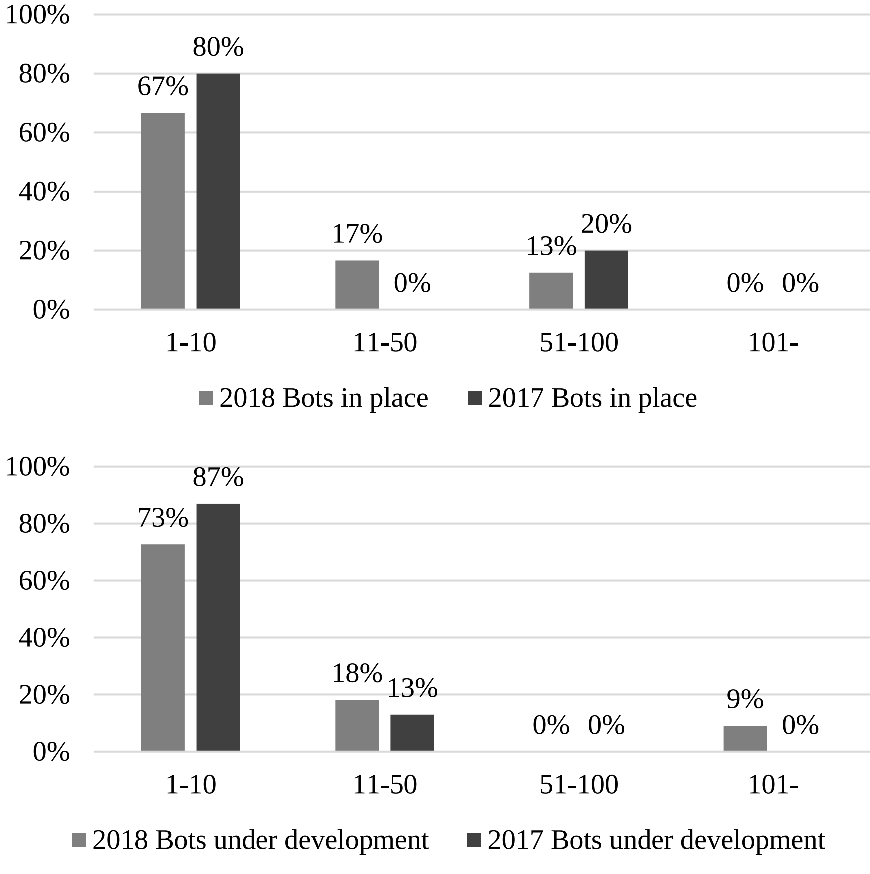


Source: authors' compilation based on survey responses

Regarding savings as full-time equivalents (FTEs) our respondents seem to be equally hopeful (see *Figure 4*). Although approximately half of them think that software robots may replace 0–2 FTEs per unit, half of them believe that 2.5 or more FTEs might be replaced. According to the most optimistic BSCs (16% of respondents), even 4.5 or more FTEs will be taken over by RPA technology. Differences in expectations might be due to divergent experiences during the testing phase and previous implementations, or as to dissimilarities in the properties of affected processes.

In 2018, 24 out of 71 BSCs responding to our survey claimed that they had already implemented RPA technology and had software robots in place or under development. This meant an advance both in the number of companies using RPA and in the number of robots in place compared to 2017. In *Figure 5*, it can be seen that those players who formerly had had 1–10 software robots assigned to various processes, now moved for new implementations and are currently developing an increased number of new robots. Two BSCs (9% of respondents using RPA technology) plans to apply more than 100 new software robots in automated processes soon.

Figure 5 Number of software robots in place and under development in 2017–2018



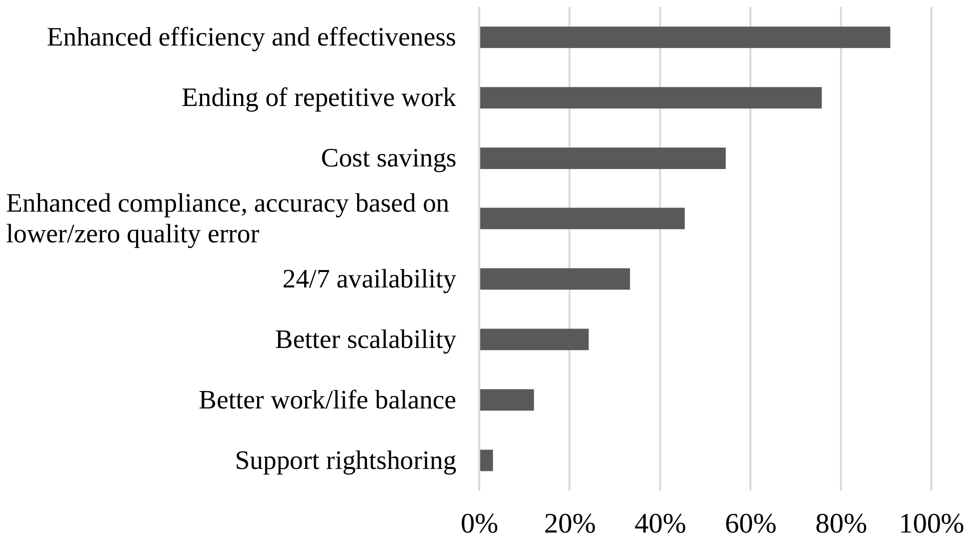
Source: authors' compilation based on survey responses

According to our respondents, the four most frequently mentioned criteria for selecting an automation technology vendor were (a) security and reliability (73%), (b) possibility of integration into extant IT environment (62%), (c) overall ease of implementation (62%), and (d) availability of global support. Concerning these criteria, among others, BSCs predominantly chose BluePrism, UIPath, and Automation Anywhere as automation technology vendors.

In most cases (38%) RPA is run in the cloud on own servers or premise-servers (28%). If using RPA solutions, most BSCs have 1–5 (35%) or 6–10 (35%) processes in scope. However, there is a committed segment (15%) who have over 51 processes currently automated by RPA.

Enhancing effectivity and efficiency, as well as cost savings, are the most important benefits of using RPA technology in BSCs. Apart from these, however, other expected advantages emerge: the amelioration of service levels (like zero error, better compliance, and all-time availability) as well as the relieving end to repetitive work, and because of this, a better work-life balance (see *Figure 6*).

Figure 6 Most important effects and benefits of using RPA in BSCs – multiple choice



Source: authors' compilation based on survey responses

Reducing the proportion of repetitive tasks in employee workload might be the main argument why most employees have an overall positive attitude towards automation. According to corporate respondents, 87% of them do not experience human resistance against developing and implementing automation tools within their organizations. Individual employees were also asked about their viewpoint: according to 566 employees from various BSCs, only 2% of them claimed that software robots would completely take over their jobs. A further 47% believed that parts of their jobs would be carried out by robots, while 42% of them thought that their current jobs would not be affected by RPA (the remaining 9% did not know or did not answer).

While this degree of optimism concerning the indispensability of the human workforce in various jobs seems to be precipitant, it certainly signals that managers of BSCs have been able to convince their employees. Reasons for this and the main arguments of official communication on automation will be presented in the next section of this paper, based on interview results.

4.2. Interview results

In this section, we present specific automation projects implemented in the three organizations in which we interviewed managers and automation experts. We also introduce common patterns regarding motivations of automation, its effects on corporate strategy, and official communication.

During the interviews, our general observation was that managers certainly embrace the idea of automation. As Interviewee 10 from Company B pointed out: "Being a service centre, our performance is measured by cost levels and the volume of tasks accomplished. We are expected to take over further activities from the core business areas while keeping the same cost levels. That would be impossible without automation and robotization." Other interviewees agreed, claiming that global management either articulates specific development requirements or declares savings targets. Company A, for instance, globally sets a so-called "annual challenge" every year, which means on average a 10% savings target to accomplish. As Company A is a market leader in multiple product lines, these challenges prevent it from getting too comfortable. Continuous improvements in efficiency are needed to match these requirements. Thus, in many cases, managers of BSCs face a dual challenge of successfully reducing process costs while taking over more and more tasks, and simultaneously maintaining quality standards.

Interviewee 9 from Company B said that when agreeing on taking over new tasks from business lines of the corporate group, they attempt to design possible automatization of them. "What you cannot automate; you should not take over." Interviewee 9 added that business needs to lead the automation processes, while IT merely attempts to catch up with. "IT systems in the company are very fragmented. The IT department is understaffed and struggles to integrate new applications in legacy systems as well as to lead new development projects. So, it is up to us to choose vendors and implement changes. However, when we are done, they want to take control."

As Company A and C operate in the information technology and services industry, IT departments not only provide functional support but co-create their main products. Thus, automation and robotization projects get more support and can even rely on their own IT solutions if available.

Company A, perhaps the most mature in the digital transformation of the three BSCs examined, is involved in dozens of more significant and hundreds of smaller automation projects. Project size is based on investment and workforce need: a project that requires 200–1,000 workhours is considered small, while a project that needs 1,000–5,000 are medium, and 5,000–20,000 are large. A smaller project had been the implementation of a chatbot that could answer questions related to the company's new travel system. After opening the new systems, employees got video and written learning material on how to use it. Much of the time, however, they did not put effort into processing these materials and preferred to ask IT-colleagues instead. Thus, they put a chatbot in use with natural language recognition that can understand human speech as well as written questions, translate them to data queries and answer accordingly. As the database behind the language interface of the chatbot was expanding with every new question, at some point, it grew capable of giving very accurate answers – saving time for IT personnel.

Another exciting project was their implementation of RPA technology in procurement requests to procurement order in Company A. Two types of procurement orders are classified as "hands-on" or "hands-free". The first one refers to orders that are unprecedented or for some other reasons should be negotiated with the provider, while the latter one refers to orders with a previously accepted general contract with the supplier. Hands-free procurements had already been automated for some time, but hands-on procurements required much human workforce to complete. Thus, RPA technology was introduced in case of hands-on purchases. Two software robots were needed for the process from end to end. The first one filled out a procurement order creation form after gathering data on same or similar orders for the same product, country, and supplier. The second one scanned the contract database, categorized suitable partners, and pulled data to PO form to complete it. Eventually, human interaction is still needed to approve orders. Robots would likely be able to do that as well, but as Interviewee 3 explained, special certificates would be needed for that, as decision-making and financial responsibility of robots is still a complicated legal issue.

According to our interviewees, other technologies like blockchains, cognitive systems, and artificial intelligence are also on the horizon but are still too expensive, unreliable, and under-studied for them to implement in the short term. As Interviewee 15 from Company C pointed out: "We have dreams to accomplish, but very few industries have the required capital in Hungary or in the countries our parent companies operate in. However, what is *dreamable* will soon be doable."

Until then, it seems that BSCs will continue to re-organize, automatize, and then robotize routine (and eventually non-routine) processes. Managers of BSCs, however, manage to frame these changes as a relief for employees who will be less loaded with monotonous tasks and will gain time for more creative and challenging activities. They have a good reason to think so: the number of tasks that business lines are willing to hand over to them seems to be almost infinite. So, they are confident that robots will not completely replace their employees.

Official communication on automation projects and digital transformation is also capitalizing on the argument that the human workforce will be liberated from routine tasks that will let them do more meaningful and more gratifying jobs. As Interviewee 3 mentioned, it might be the case that a particular employee is not suitable for tasks requiring higher levels of creativity and problem-solving. He underlined, however, that it is their very conscious decision not to dismiss anybody because of automation – as it would cast shadows on further development. If a department acquires a workforce surplus, they order a hiring freeze and "natural fluctuation generally solves the problems in a few months".

According to Interviewee 1, 10, and 15, senior managers of their companies, the automation of routine tasks will not only put employees in better positions but will promote the industry itself. As BSCs are carrying out tasks with increasing added value and a growing need for creativity, they can position themselves as *business partners* rather than mere suppliers in the global supply chain of their corporate groups. This may result in an increasing number of BSC managers receiving seats on the board of parent companies, better-negotiating positions, and greater prospects of future growth. As the business services sector is already a significant and steadily growing industry in Hungary, increased value creation of BSCs might have an overall beneficial effect on the economy: growing exports and tax revenues, as well as higher salaries, seem to be achievable targets.

5. Conclusions

In our paper, we analyzed the current trends of digital transformation projects in BSCs operating in Hungary. We introduced our theoretical framework for Business Services 4.0, a concept encapsulating the interrelation of technology, strategy, and organization in Business Service Centers (BSCs). We presented the overall picture of the industry based on quantitative data from large-scale surveys. We also presented the results of fifteen interviews with senior technology experts and executives carried out at three Hungarian based BSCs of multinational parent companies to examine individual transformation projects. We reviewed the use of advanced technologies like robotic process automation, predictive analytics, chatbots, and artificial intelligence.

We found that BSCs had mostly automated massively repetitive processes and that this automation had liberated employees for more creative tasks. Based on general trends and expert opinions, we argue that the consequences of this transition are threefold. (1) BSCs can reinforce their position as business partners of their global parents. This not only puts employees in a better position but opens the way to promote the whole industry. As business services is already a significant sector in the Hungarian economy, we argue that further economic benefits might arise from the digital transformation of the industry.

(2) Creative tasks are more attractive for prospective and current employees in a labor market characterized by a shortage of suitable personnel. As the business services sector is steadily growing, its need for more employees means a constant challenge for hiring. BSCs have already absorbed their prime targets from the labor market (young college graduates with excellent language skills); they will probably compete with other sectors for the workforce. Creative and meaningful jobs will likely be a key message when re-thinking their employer brands.

(3) Employees usually do not fear the possibility of job loss due to automation and digital transformation. Based on our research, this self-confidence is often well-grounded, although, in certain individual cases, employees will probably need to be re-educated. Considering all the possible gains and losses, we argue that the digital transformation of the business services industry is expected to have an overall positive effect on both organizational and national levels.

Based on the findings of our survey and interviews, we found that multiple arguments of previous research are also correct in the case of Hungarian BSCs. We observed that technological tools and solutions that are usually regarded as the leading force of Industry 4.0 (Bhimani and Willcocks 2014, Kukreja 2016), are also present in many BSCs. Similar to other research cited in our study, we found that digital transformation happens through different stages that are built upon each other. With our study, we also reinforced arguments (Eden et al. 2019, El-Khoury 2017) that the digital transformation of companies requires the transformation of the workforce. More creative and more knowledge-intense jobs can only be done by competent and educated employees who can learn and adapt. As BSCs traditionally employ college graduates, this is easier for them – as confirmed by our survey.

Acknowledgments

This research was supported by project No. EFOP-3.6.2-16-2017-00007, titled *Aspects on the development of intelligent, sustainable, and inclusive society: social, technological, innovation networks in employment and the digital economy*. The project has been supported by the European Union, co-financed by the European Social Fund and the government of Hungary.

References

A.T. Kearney Inc. (2017): *The Widening Impact of Automation*.

Andriole, S. J. (2017): Five Myths About Digital Transformation. *MIT Sloan Management Review*. Retrieved from https://sloanreview.mit.edu/article/five-myths-about-digital-transformation/

Bhimani, A. – Willcocks, L. (2014): Digitisation, "Big Data" and the transformation of accounting information. *Accounting and Business Research*, 44, 4, 469–490. https://doi.org/10.1080/00014788.2014.910051

Chui, M. – Manyika, J. – Miremadi, M. (2015): Four fundamentals of workplace automation. *McKinsey Quarterly*, 29, 3, 1–9.

Davenport, T. H. – Kirby, J. (2015): Beyond Automation. *Harvard Business Review*, (June), 2016.

Demeter, K. – Losonci, D. – Nagy, J. – Horváth, B. (2019): Tapasztalatok az ipar 4.0-val – egy esetalapú elemzés. *Vezetéstudomány*, 50, 4, 11–23. https://doi.org/10.14267/VEZTUD.2019.04.02

Drótos, G. – Marciniak, R. – Ránki-Kovács, R. – Jakab, D. – Willbrandt, N. – Baksa, M. … Mátrai, E. (2018): *Business Services Hungary: 360° view about the Hungarian Business Services Market 2018*. Budapest.

Eden, R. – Burton-Jones, A. – Casey, V. – Draheim, M. (2019): Digital Transformation Requires Workforce Transformation. *MIS Quarterly Executive*, 18, 1, 1–14. https://doi.org/10.17705/2msqe.00005

El-Khoury, D. D. (2017): Digital transformation and the world-class HR difference. *Strategic HR Review*, *16*(2), 86–88. https://doi.org/10.1108/SHR-01-2017-0001

Everest Global Inc. (2014): Automating Business Process : Using Robotics to Drive the Next Stage of Transformation.

Füzes, P. – Szabó, Z. R. – Gódor, Z. (2018): Szabadulás a kiaknázási csapdából a digitális jövő alakításával: Hogyan irányítható az iparági változás egy felhőalapú szolgáltatásra épülő kiaknázási és felderítési tevékenységgel? *Vezetéstudomány*, 49, 1, 54–64.

Héder, M. (2014): The machine's role in human's service automation and knowledge sharing. *AI & Society*, 29, 185–192. https://doi.org/10.1007/s00146-013-0474-y

Hortoványi, L. – Vilmányi, M. (2018): Üzletági stratégiák átalakulása a digitalizáció forradalmának forgatagában. In D. Horváth (Ed.), *A stratégiai menedzsment legújabb kihívása: a 4. ipari forradalom c. konferencia kiadvány* (1–5). Budapest.

Horváth, D. – Szabó, Z. R. (2017): A negyedik ipari forradalom vezetési aspektusai. In M. Veresné Somosi & K. Lipták (Eds.), *„Mérleg és Kihívások" X. Nemzetközi Tudományos Konferencia* (700–714). Miskolc: Miskolci Egyetem Gazdaságtudományi Kar.

Keller, A. (2017): Challenges and Directions in Service Management Automation. *Journal of Network and Systems Management*, 25, 4, 884–901. https://doi.org/10.1007/s10922-017-9437-9

Kukreja, M. (2016): Study of Robotic Process Automation (RPA). *International Journal on Recent and Innovation Trends in Computing and Communication*, (June), 434–437.

Lacity, M. C. – Willcocks, L. (2015a): *Robotic Process Automation : The Next Transformation Lever for Shared Services Research on Business Services Automation*.

Lacity, M. C. – Willcocks, L. (2015b): Stand to Gain from Automation. *Harvard Business Review*, (June), 14–17.

Losonci, D. – Takács, O. – Demeter, K. (2019): Az ipar 4.0 hatásainak nyomában – a magyarországi járműipar elemzése. *Közgazdasági Szemle*, 66, 2, 185–218. https://doi.org/10.18414/KSZ.2019.2.185

McKinsey Global Institute. (2017a): *Technology, Jobs, and the Future of Work*.

McKinsey Global Institute. (2017b). *What's Now And Next in Analytics, AI, and Automation*.

Nagy, J. (2019): Az ipar 4.0 fogalma és kritikus kérdései – vállalati interjúk alapján. *Vezetéstudomány*, 50, 1, 14–26. https://doi.org/10.14267/VEZTUD.2019.01.02

Sebastian, I. M. – Ross, J. W. – Beath, C. – Mocker, M. – Moloney, K. G. – Fonstad, N. O. (2017): How Big Old Companies Navigate Digital Transformation. *MIS Quarterly Executive*, 16, 3, 197–214.

Szabó, Z. R. – Horváth, D. – Hortoványi, L. (2019): Hálózati tanulás az ipar 4.0 korában. *Közgazdasági Szemle*, 66, 1, 72–94.

Unger, T. S. – Badillo, R. – Zeiss, S. – Gabriel, J. (2018): Expert Interviews on RPA. In *3rd AI & Robotic Process Automation World Summit 2018* (1–10).

Winroth, M. – Säfsten, K., – Stahre, J. (2007): Automation strategies : existing theory or ad hoc decisions ? *International Journal of Manufacturing Technology and Management*, 11, 1, 98–114.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 61–78.

Testing the importance of 4th Industrial Revolution characteristics in the Hungarian environment

Comparison of SMEs' and Business students' perception

Eszter Megyeri – Gabriella Tabajdi

The 4th Industrial Revolution, or IR 4.0 for short, is being embraced by academics and practitioners alike. However, business readiness to meet the changes of the 21st century shows us a very diverse landscape. The most responsive players are progressive companies who must deliver a new way of thinking, creating, and operating to stay competitive. In our research we chose to analyse the perception of these developments from the point of view of Hungarian small and medium-sized companies, and students, who are likely to generate the next new waves of innovation, technology orientation, and value creation. We approached Hungarian SMEs who are involved in and impacted by IR 4.0 challenges to analyse their understanding and ranking of the literature-driven characteristics of the 4th Industrial Revolution. In our study, we want to understand how Hungarian business students – soon to be employees and future entrepreneurs – perceive the pressure of IR 4.0 developments. Our aim was to identify links, dependencies, and gaps between Hungarian SMEs and business students who together will conjointly create new industrial opportunities.

Keywords: industrial revolution 4.0 characteristics, SMEs, entrepreneurship

1. Introduction

In today's turbulent environment, there is an increasing expectation on higher education to contribute to the development of competitive and high-quality human resource development. One component of the demanding external environment can be linked to the 4th Industrial Revolution. In this study we review the key features of this revolution in general, and detail its relevance for the Hungarian ecosystem with special emphasis on small and medium-sized enterprises. SMEs are potential future employers of graduates. As future employees, students bring significant value-added knowledge and competencies to SMEs which are meant to be developed during higher education training. We aim to articulate and deliver messages for lecturers in higher education to sensitize students for future technology demands in the context of local Hungarian business settings. In order to deliver that, we compared student and SME perception on the following three main themes (1) the perceived importance of technology challenges; (2) the relevance of 14 IR 4.0 characteristics from the respondents' point of view; (3) respondent group alignment on the perceived importance of IR 4.0 characteristics. The 4th Industrial Revolution, often also called Industry 4.0 or IR 4.0, is a new phenomenon which calls for clarification to assure that the responses in the primary analysis relate to the same business observations.

2. The 4th industrial revolution

The development of industry is of historical relevance regarding world growth. At the present time there are numerous discussions regarding the 4th Industrial Revolution both among academics and practitioners. The occurrence of this revolution can be at least partly ascribed to the expansion of globalization and to the technological changes that affect all spheres of life. Meanwhile researchers are trying to grab its elements in an exact way, the actors of economic life tackle with the opportunities and pressures of the operationalization of new factors in order to maintain competitiveness. We do not know how long this will last and what milestones we have already passed, and what others are in front of us, but it is sure that due to this new industrial revolution the global economy and society is going to change intensively. Therefore, it is inevitable to organize these changes and processes and to identify them in the various aspects of practical economic life. The aim of this section is to organize the relevant literature in a way that the elements of the 4th Industrial Revolution can be apprehended and understandable specifically regarding start-ups and small and medium-sized enterprises (SMEs).

In general, we can speak about an industrial revolution when due to new technology solutions the effectiveness of production systems increases considerably. Prior to IR 4.0 there were 3 other industrial revolutions; however, the length of time elapsing between two revolutions has decreased. While between the 1st and 2nd about one century elapsed but between the 3rd and 4th just about one decade did. Innovations and technologies affecting the economy, to a great extent, are evolving at an ever accelerating pace.

We can define IR 4.0 in a broad and in a narrower sense. In a broad sense it is a bundle of technologies adopted in manufacturing and in its related supporting activities over recent years, but in a narrow sense it refers to the adoption of cyber-physical systems (CPS) that result in the digitalisation of production (Szalavetz 2017, Kagermann et al. 2013, Monostori 2015). Angela Merkel, the German Chancellor defined Industry 4.0 as the following: "*the comprehensive transformation of the whole sphere of industrial production through the merging of digital technology and the internet with conventional industry*" (Davies 2015, p. 2.).

However, there are several other definitions for the ongoing 4th Industrial Revolution. Some concentrate on production, especially on the ever tighter connectedness of information technology and automated facilities, thus on the faster and more autonomous features of production[1]. Others include the processes of supply and logistics because due to Industry 4.0 through cyber physical systems we can connect real objects with data processors via virtual processes and information systems (Abonyi–Miszlivetz 2016). Another definition is: Industry 4.0 is such a phenomenon that raises the transparency of processes, integrates firms' value chains and the supply chain and increases customer value creation to a new level through

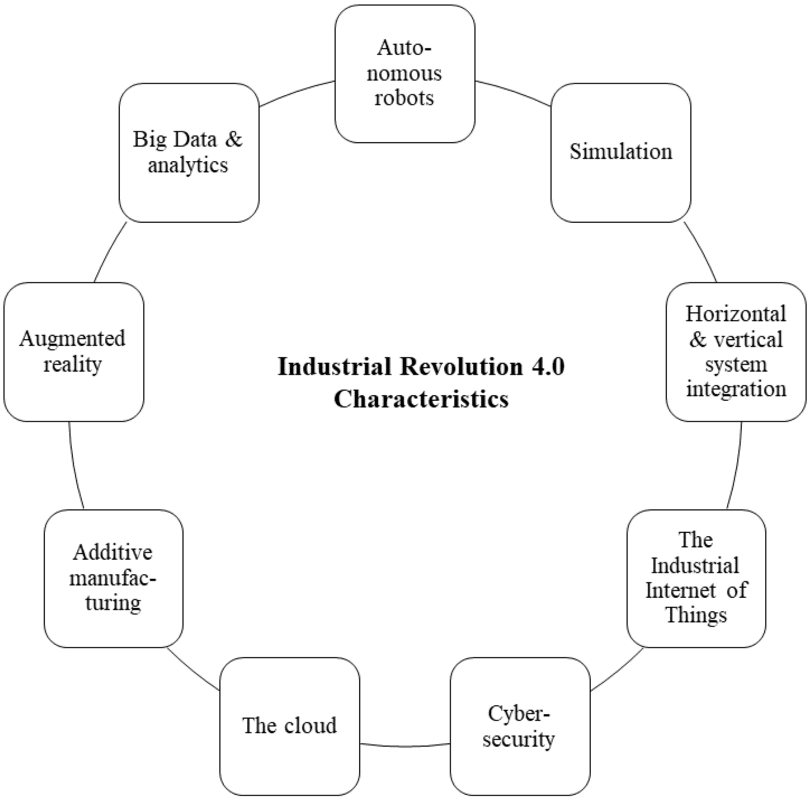
[1]*Source*: https://www.tablazat.hu/ipar-4-0 Last accessed: 28. April 2019.

technological instruments and activities and with the exploitation of the opportunities in digitalization (Nagy 2017, Davies 2015).

We can summarize the definitions of the 4th Industrial Revolution "*as a revolution enabled by application of advanced technologies (like IT) at production level to bring new values and services for customers and organization itself*" (Khan and Turowski 2016, p. 442.). This also brings quality and flexibility in production systems (Khan–Turowski 2016).

To be able to grasp this ongoing revolution, we summarize the main technologies characterizing companies that are developed with regards to Industry 4.0, this is also shown in Figure 1. The most important elements are: autonomous robots, simulation, horizontal and vertical system integration, industrial Internet of Things (Industrial IoT), cyber security, additive production, cloud-based services, augmented reality, Big Data analytics and production visualization (Rüßmann et al. 2015). In short, let us summarize the main attributes of these elements.

Figure 1 Technologies transforming industrial production



Source: own construction based on Rüßmann et al. (2015)

Regarding the first technology, autonomous robots, and the use of them in production is not new; manufacturers have been using robots to solve complex issues for decades now. However, these robots are evolving, they are becoming more cooperative, more autonomous, more flexible, and can solve more complex tasks (Rüßmann et al. 2015, Van der Elst–Williams 2017), and their adoption in industries is becoming more widespread. This is a consequence of several factors. Most importantly the cost of robotic systems (both hardware and software) has fallen considerably and is expected to decrease further, also their usage was limited but their development makes their use possible in several fields of production. Lastly, as their use needed substantial capital expenditures and the usage of specialised operatives, mostly large companies were able to adopt them, but as their cost and functionality improved, now many SMEs are also able to implement robotic systems (Strange–Zucchella 2017).

Robotization and the appearance of autonomous or cooperative robots have reshaped both production and services sectors causing substantial increase in efficiency (Nagy 2017). When designing these robots, a main aspect to be considered is that they should be able to detect the change in the operation of other robots and should be able to adapt their own operation accordingly. Robotically aided production and the efficiency increase caused by using collaborative robots are highly promising for companies. However, their adaptation is not that simple. Many obstacles might make it more difficult than the hype suggests. Among the obstacles we can mention the financing of such investment, the guarantee of enough return, or the ability to quickly adapt to dynamic developments and the fast rate of change in technology platforms themselves.

As for the second element, simulation is basically the modelling and visualization of product designs and their establishment in the manufacturing processes (Davies 2015). Similarly, to robotics, simulation is not a totally new technology, it being used in the engineering phase of products, materials, and production processes for years now, but their use is becoming more extensive, and in the future, simulations will be used in plant operations too (Rüßmann et al. 2015). Simulation with the help of virtual and augmented reality enables operators to see and compare actual production with the ideal, expected performance in real-time, thus supporting the effective operation of devices and the optimization of processes (Nagy 2017). Moreover, simulation allows operators to optimize and test the settings of machines virtually before the physical changeover, so machine setup times can be reduced while quality can be increased (Rüßmann et al. 2015, Van der Elst–Williams 2017).

The third is horizontal and vertical system integration. The collaboration of the supply and value chains can only be optimized if their systems are vertically and horizontally integrated. Nowadays the systems of suppliers, producers, and retailers or customers are integrated in a limited way, but with IR 4.0 this can be changed. Companies, departments, capabilities, and functions will be more cohesive as universal and cross-company data integration networks evolve. This enables the creation of a truly automated value chain (Rüßmann et al. 2015). Moreover, the integration enables the optimization of production and inventory planning managed across several business entities to best fit to market needs and links all involved partners within the supply chain.

The next technology to be presented briefly is the Industrial Internet of Things (Industrial IoT). IoT is a crucial if not the most crucial element of the 4th Industrial Revolution. Although this technology has existed for some time now, the concept has gained importance in recent years only, as the relevance of connectivity has been understood (Sniderman et al. 2016, Baldassare et al. 2017). Today, a growing number of physical products are equipped with sensors that can process, capture and communicate data (Strange–Zucchella 2017). These sensors enable information exchange at high speed anywhere in the world, and can retrieve real-time information about the devices' and their surroundings. This way we are also able to monitor and operate technology infrastructures over large physical distances (Kagermann 2015).

The Industrial IoT is similar, it is about integrated machines and sensors and cloud-based solutions that have unique identifiers, making it possible to monitor all the processes of the supply chain (Nagy 2017). Thanks to these identifiers we will have information regarding the use, destination and provenance of devices, so there will be no need to synchronise and coordinate information, and product flows (Strange–Zucchella 2017). With Industrial IoT, more devices can be enriched with embedded, connected and computing technologies allowing these devices to interact and communicate with each other. Moreover, this also decentralizes decision making and analytics, making real-time responding possible (Rüßmann 2015). Industrial IoT might also reduce transaction costs associated with international production and might promote a deeper international division of labour (Strange–Zucchella 2017, Buckley–Strange 2015).

With the increasing use of these technologies, the importance and relevance of cybersecurity is also growing. Along with the expansion of networks and the dynamic vertical and horizontal movements of data, the stability of networks, of confidentiality, and cybersecurity systems is indispensable (Nick 2017). These are specifically important to avoid cyber-attacks. The continuous monitoring of cybersecurity is now inevitable (Nick 2017). The need for protection of critical industrial systems and manufacturing lines is increasing rapidly, so secure and reliable communications, sophisticated identity and access management of users and machines are now essential (Rüßmann et al. 2015).

The cloud is another essential driver of today's technology change. Cloud-based services are already used by firms for analytics and some enterprise applications; however, with Industry 4.0 their use will become more widespread and they will be used for several purposes. As production-related undertakings will need an increased share of data inside and outside the company, the relevance of the cloud will rise. The cloud contributes to the continuous assurance of the availability of systems and makes possible safely sharing and having access to data between different sites of the company and between companies. Furthermore, as the performance of cloud technologies improve, the whole process will become faster and the reaction time will only be some milliseconds, so we will be able to get the necessary information right away (Rüßmann et al. 2015).

A further technology is additive manufacturing. It is also known as 3D printing. This method is called additive because it creates objects by adding successive layers instead of subtracting them (Van der Elst–Williams 2017, Janssen et al. 2014). This way, customized products can be manufactured offering construction advantages, for example with complex and lightweight designs. Decentralized and high-performance additive manufacturing systems will reduce stock on hand and transport distances but also production time can be considerably decreased (Rüßmann et al. 2015, Nagy 2017). This will also shape international business. As products designed by CAD software can be produced anywhere in the world where there is a 3D printer, manufacturing will not need to be factory specific but could be be conducted close to the end-user, resulting in savings in transportation costs and delivery times, and in a minimised risk of supply resulting in simplified and more cost-efficient value chains (Strange–Zucchella 2017).

Augmented reality (AR) is also one of the nine technologies transforming industrial production. Due to AR, products and product components have digital data transmission capabilities to better monitor, observe, and inspect product characteristics and the state of these products during manufacturing and consumption. With the use of the results, the production process or the product itself can be optimized. One practical field of the use of AR nowadays is maintenance (Nagy 2017). Yet, their utilization is still in its infancy, and in the future AR will have a much broader use to give real-time information and improve work procedures and decision making (Rüßmann et al. 2015).

When introducing the main elements and technologies forming and driving Industry 4.0, Big Data and analytics cannot be left out. Big Data and its analytics mean the collection and comprehensive evaluation of data from various sources, from production equipment through enterprise management to customer management systems. Analytics based on huge data sets have appeared in manufacturing only in the past years. Yet it can help in the optimization of production quality, in energy savings and in equipment service improvements (Rüßmann et al. 2015). Furthermore, firms will be able to monitor trends and possibilities in faraway markets without having to make big resource commitments there. They will also be able to more effectively optimize their production, supply, and distribution activities by the analysis of the vast amount of data (Strange–Zucchella 2017).

All these above-mentioned technologies bundled in Industry 4.0 are expected to have a great impact on the global economy. There are estimations that the 4^{th} Industrial Revolution can bring an annual efficiency gain in manufacturing of 6–8%. Just in Germany alone, IR 4.0 will contribute to the country's GDP by 1% per year over the next 10 years and will create 390,000 jobs. Globally, by 2020, investments in the Industrial Internet will be 25 times that of investments in 2012, and the value-added is expected to grow from 32 billion USD in 2012 to 1.3 trillion USD in 2020 (Davies 2015). Besides the figures, Industry 4.0 will have a serious impact on the way goods are produced, how companies do business, how economies operate, how markets function and on how societies react (Van der Elst–Williams 2017).

Although IR 4.0 is a global process, the readiness for it among countries and regions differs to a great extent. Employing the technologies of Industry 4.0 is critical as it can be a competitive advantage for countries that are able and willing to use these technologies, and a competitive disadvantage to those who are not. In the following section we interpret Hungary's position.

3. The 4th Industrial Revolution in Hungary and in Hungarian SMEs

When looking at Hungary's position we find that Hungary is a traditionalist country, meaning that the share of manufacturing in GDP is high but the countries readiness for the 4th Industrial Revolution is low (Dujin et al. 2014). An indicator of readiness is gross expenditure spent on R&D (GERD). The European Union has a target of 3% of GDP regarding R&D expenditure to be achieved by 2020 to improve the EU's competitiveness and readiness for the changes of this new era. Even the EU as a whole is far from reaching this goal (in 2016 GERD was 2.04 %); Hungary is among the moderate performers. The gross domestic expenditure on R&D in this country was 1.2% in 2016 (Eurostat 2017). This low value already indicates to some extent Hungary's weak readiness for Industry 4.0, yet in their study Kuruczleki et al (2016) pointed out the same by constructing their own index to measure the readiness for the 4th Industrial Revolution in the European Union. To create their own index, they used the following indicators: the above-mentioned GERD, total intramural R&D expenditure, community trademark applications, community design applications, total R&D personnel and researchers, tertiary educational attainment, ICT specialists, and digital single market. According to their analysis Hungary ranked 18th in the EU lagging behind most of the West European countries, and even behind some Central and East European country. Their analysis underlies that there is much room for improvement and that Hungary has to improve a lot to embrace the technologies and opportunities of IR 4.0.

Another way to present that Hungary has to do a lot to improve its readiness is looking at its performance regarding the EU's digital transformation index. This index has 7 dimensions: digital infrastructure, digital transformation, investments and access to finance, supply and demand of digital skills, changes in ICT start-ups environment, entrepreneurial culture and e-leadership[2]. Based on these dimensions Hungary is a modest performer in digital transformation, yet improvements can be observed in particularly in the field of investments and access to finance and entrepreneurial culture but also in digital transformation, supply and demand of digital skills and changes in ICT start-up environment. To these improvements a big contribution is made by a generally favourable investment climate that creates incentives for ICT firms (both domestic and foreign) to invest in Hungary. Yet the country's performance is rather poor in terms of e-leadership and of digital

[2] More on the EU's Digital Transformation Scoreboard: https://ec.europa.eu/growth/tools-databases/dem/monitor/scoreboard Last accessed: 28 April 2019.

infrastructure indicating that further policy efforts are needed. Regarding these fields, the country is performing far below the EU average. Digital infrastructure is a great challenge for the country; and special efforts should be made to increase the integration of ERP software by businesses and to improve internet bandwidth's quality (Probst et al. 2018).

All these pose a big challenge for the country, and to overcome it and to boost developments in the field of technological readiness, R&D & innovation, and digital transformation, several strategies have been implemented. In February 2016 a strategy for industrial development initiated by the Ministry for National Economy was adopted. This national strategy, called the Irinyi Plan, includes the most important directions for economic development for the time period of 2016-2020; moreover, this plan is a framework for the development of an Industry 4.0 strategy in all the key sectors concerned (Nick 2017, Klitou et al. 2017).

The Irinyi Plan also concretizes the Hungarian economic development strategy as a target has been set to increase the share of industrial value-added in GDP from 23.5% in 2016 to 30% in 2020. The plan names these key sectors as well. They are: the already strong automotive industry, specialized machine and vehicle production, healthcare, tourism, food industry, the IT sector, the green economy, and the sector of shared service centres (SSCs). There is a goal of improving the processing industry as well, which should be achieved using new technologies, the improvement of energy and material efficiency, and the more extensive use of Hungarian resources[3].

Based on the Irinyi Plan a national initiative was also adopted. Industry 4.0 National Technology Platform (IR 4.0 NTP) is an initiative of the Ministry for National Economy and the Hungarian Academy of Science Institute for Computer Science and Control (Klitou et al. 2017, Haidegger–Paniti 2016). The objective of this platform is to enhance manufacturing and industry transformation in the country to get Hungary ready for Industry 4.0. The expected effects of IR 4.0 NTP are to prepare the industrial sector for the 4th Industrial Revolution and adapt it to the requirements of IR 4.0, but also to boost the countries competitiveness. This initiative is an indispensable element in improving Hungary's readiness for digital transformation. The five main objectives of IR 4.0 NTP are: responding to challenges in a prompter way, fostering bold steps toward innovation, supporting the readiness of the country's economy for innovative adaptation, accelerating innovation particularly in digitisation and in production, and fostering information exchange, partnership establishment and cooperation between all actors of economy (Klitou et al. 2017).

To reach the set strategic targets several organizations were set up. These organizations help to roadmap the above objectives. One of such organizations is the Excellence in Production Informatics and Control (EPIC) (Fülep et al. 2018). The goals of EPIC are to improve innovation culture in Hungary, to speed up the innovation process, and to introduce and promote new technologies and

[3] *Source*: https://piacesprofit.hu/kkv_cegblog/ipar-4-0-az-uj-szabvany/?hf=1 Last accessed: 28 April 2019.

methodologies, also to strengthen the institute's research potential, especially in the field of Cyber Physical Systems (Nick 2017, Haidegger–Paniti 2016). Besides multinational companies, a special focus is put on the development of Hungarian SMEs. EPIC supports competitiveness increases and the strengthening of supply industry from a practical point of view. Among its objectives are: improving high-quality trainings and constructing and promoting sample solutions for the 4th Industrial Revolution (Nick 2017).

It is of key importance that all economic actors change their development systems from extensive to intensive development, and this is particularly true for SMEs, which in many cases still use traditional methods. SMEs are affected negatively by both their lack of knowledge and lack of capital, which are critical in the realization of high value-added production methods[4]. The value-added contribution of SMEs is lower than the EU average in Hungary, which lowers the country's competitiveness[5]. Yet as Table 1 presents, Hungarian SMEs make up around one quarter of the Hungarian business expenditure on R&D (BERD), which was even higher during 2011-2013 when, BERD by SMEs exceeded 30%. This indicates that despite their size, SMEs are crucial for Hungarian research and development as these enterprises are also highly committed to R&D activities and are investing in R&D. By supporting SMEs R&D activities not only show their readiness for the 4th Industrial Revolution can be improved, but also the competitiveness of the whole country.

Table 1 The share of SMEs in Business Expenditure on R&D (BERD) in Hungary, 2008–2016, % of total Hungarian BERD

Number of employees	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
10–49	9.26	9.72	9.45	11.09	11.57	12.06	11.73	8.78	9.20
50–249	8.58	11.57	15.76	18.93	19.17	20.95	16.84	15.09	13.67
SMEs total	17.83	21.30	25.21	30.02	30.74	33.02	28.57	23.87	22.86

Note: micro enterprises (enterprises of 1–9 employees) are left out from the investigation, however if we include them, the latest values exceed 25%, and between 2011–2013 the share of micro, small and medium-sized enterprises in total Hungarian BERD was close to 40%.

Source: own calculations based on Eurostat data[6]

[4] *Source*: https://evkszakkollegium.blog.hu/2016/12/06/negyedik_ipari_forradalom_avagy_tenyleg_elveszik_a_robotok_a_munkankat Last accessed: 28 April 2019.

[5] *Source*: http://www.mkt.hu/wp-content/uploads/2016/05/Lepsenyi_Istvan_05_12.pdf Last accessed: 28 April 2019.

[6] *Data source for calculations*: https://ec.europa.eu/eurostat/web/products-datasets/-/isoc_bde15ar2 Last accessed: 28 April 2019.

So, it is not surprising that IR 4.0 development plans have a special focus on SMEs and the help and support given to them. Several programs give SMEs the opportunity to implement their own development ideas, while others help them in conversion. One such program, called the Program of Modern Entrepreneurs, is organized by the Hungarian Chamber of Commerce and Industry (HCCI). This program supports the IR 4.0 developments of SMEs with professionals and consulting. Moreover, the Industry 4.0 Model Factory program was established with the aim of helping Hungarian suppliers of multinational companies to meet the growing needs of MNCs regarding the quality of products and services[7]. Five model factories were set up with government support to familiarize and prepare SMEs for the use of IR 4.0 technologies. Representatives of SMEs can see these technologies in real-life operation in functioning factories[8]. As a part of this program, participating SMEs can test their readiness for Industry 4.0 but can also find firm specific ideas to foster and motivate their aspiration to improve. In these Model Factories SMEs can meet new technologies, for instance they can find autonomous robots or the various uses of AR[9]. An important goal is to increase SME readiness for IR 4.0 so that they have better chances of becoming suppliers to multinational companies.

Besides SMEs, a special focus is put also on start-ups and their improvements. For this a special strategy was set up, the Digital Start-up Strategy[10]. The main aim of this strategy is to boost the cooperation between multinational companies and start-ups in Hungary. The synergies could be beneficial for both parties, because while multinationals have a lot of capital, they can only react to changes slowly, while start-ups are innovative enterprises that are able to react to market changes fast and effectively; however, they have lack of capital (Ritter et al. 2016).

Although, Hungary's position regarding Industry 4.0 should be improved, as it is presented in this section, there are several kinds of actions and plans that aim at this improvement. The focus is not only on multinationals, as they alone cannot bring change. There are special programs specifically dedicated to SMEs and start-ups and their development. These programs aim at the integration of SMEs and start-ups to the supply chains and value chains of MNCs by improving their readiness for technological change, and increasing their adaption ability so that they will be able to embrace the new technologies and enhance their competitiveness.

[7] *Source*: https://piacesprofit.hu/kkv_cegblog/ipar-4-0-kkv-k-is-elkezdhetik/ Last accessed: 28 April 2019.
[8] *Source*: https://piacesprofit.hu/infokom/ipar-4-0-tanuljunk-mintagyaraktol/ Last accessed: 28 April 2019.
[9] *Source*: https://piacesprofit.hu/infokom/ipar-4-0-tanuljunk-mintagyaraktol/ Last accessed: 28 April 2019.
[10] *Source*: http://www.kormany.hu/download/d/8c/e0000/Magyarorsz%C3%A1g%20Digit%C3%A1lis%20Startup%20Strat%C3%A9gi%C3%A1ja.pdf Last accessed 28. April 2019.

4. Research method and research questions

In our study we used a quantitative method: questionnaires distributed among students and Hungarian SMEs. To examine student answers, we used printed and online questionnaires to reach respondents. SMEs were directly asked to fill out the questionnaire sent out to them via e-mail. Both groups were asked to evaluate the importance of the attributes of the 4th Industrial Revolution on a 1–5 Likert scale. Data was collected directly. Students and SMEs were asked to fill in the questionnaire handed out to them. We wanted them to evaluate the following 14 characteristics that we articulated based on the Hungarian Irinyi Plan and literature review (Nick 2017, Klitou et al. 2017):

1. Visualization of production, real-time collection of data about production and resources
2. Supply chain visualization (real-time visualization of actual and expected quantity of stocks, of the place of stocks and their status, both inside the company and with the other actors of the supply chain)
3. Supply chain collaboration (sharing of production plans and stock information through electronic data exchange)
4. The optimization of supply chain, production planning and stock planning: making optimal stock, production and procurement plans to sufficiently meet market demand, doing so with the help of ERP, SCM and Advanced Planning systems
5. Predictive maintenance
6. Big Data solutions
7. Intelligent energy utilisation
8. Robot-aid production, collaborative robots
9. Modern warehouse and production logistics solution
10. Solutions supporting unique production and small-scale production
11. Internet of Things, Machine to Machine communication, autonomous robots
12. The use of augmented reality (AR) in maintenance and in remote assistance
13. 3D printing, additive production technologies
14. Fast prototype-making, involving customers to prototyping.

We used purposive sampling. We were interested in the opinions of students familiar with management, economy and finance, so we asked only the students of the University of Szeged, Faculty of Economics and Business Administration. We decided to select this group of students as we assume that their theoretical knowledge on economic processes must be above average, therefore their perception should at least be partly based on solid knowledge, also they will constitute the future entrepreneurs and economic employees of current enterprises. Moreover, we wanted to get to know the view of SMEs operating in Hungary. Hungarian SMEs were chosen assuming that they have a greater understanding of happenings in the Hungarian economy.

The sample of our study consists of 216 answers from students and 21 answers from SMEs. Table 2 contains a summary of the attributes of the respondents.

Table 2 The composition and main attributes of the research sample

STUDENTS	SMES
216 responses	21 responses
Female: 137, male: 75, no answer: 4	Field of operation: services 8, industry 11, agriculture 1, no answer 1
Working:81, owning an enterprise: 9	Years of operation ranges: some month to 31 years (average: almost 9 years)
Both BSc and MSc students (190 BSc, 23 MSc, 3 no answer)	Number of employees' ranges: 2–200 (average: 29)
Age group: 20–25; average age: 21,5	Operating mostly on the global scale

Source: own construction based on questionnaire responses

Our investigation focused on three main areas comparing SME and student perceptions on IR 4.0 related themes. The first one investigates the *perceived importance of technology challenges*. This topic is covered by Q1 detailed below. The second area is studying the *relevance of 14 IR 4.0 characteristics from the respondents' point of view*. This topic is addressed by Q2 and Q3. The third theme is comparing *respondent group alignment on the perceived importance of IR 4.0 characteristics*. This is reviewed using Q4 and Q5.

During our research, the following specific questions were raised and analysed to measure respondent perceptions:

- Q1: How important is the overall technology challenge based on the respondents' view?
- Q2/3: What are the top 3 most / least important characteristics of Industry 4.0 according to the respondents?
- Q4/5: Where is group alignment /misalignment on the perceived level of importance of the characteristics?

The analysis and results are presented in the next section.

5. Analysis and Results

We will review the summary of responses for each question with the objective of extracting the key observations and messages on the selected 3 themes.

The perceived importance of technology challenges is addressed in the first question. We were investigating the overall technology development challenge from the students' and SME's point of view (Table 3 Q1). This score represents the overall aggregated number derived from the technology characteristics taken them into account with equal weight.

On a 5-point scale, the aggregate score for students was 3.306 as a grand mean score, while for SMEs' this figure resulted to be 3.389. These numbers show that there is a difference in Industry 4.0 perceptions between students and SMEs. The impacts of Industry 4.0 are sensed more strongly and intensively by the enterprises. On the other hand, it shows that students undervalue the importance of IR 4.0 characteristics compared to companies who might be potential employers for them in the future. We can conclude that through the education process it is important to place stronger emphasis on and awareness of the importance of business technology for students to be more qualified for the demands of the job market.

Table 3 Summary of questions and responses

Q1. HOW IMPORTANT IS THE OVERALL TECHNOLOGY CHALLENGE BASED ON THE RESPONDENTS' VIEW?		
	SME	Student
Grand Mean	3.389	3.036
Q2. What ARE the top 3 most important characteristics of Industry 4.0 according to the respondents?		
Rank	SME	Student
1	Big Data	Intelligent energy utilization
2	Unique and small-scale production	Big Data
3	Supply chain optimization	Supply chain visualization
Q3. What ARE the top 3 least important characteristics of Industry 4.0 according to the respondents?		
Rank	SME	Student
1	Robots	Robots
2	3D printing, additive production	Use of AR
3	Intelligent energy utilisation	3D printing, additive production
Q4. Where is group alignment on the perceived level of importance of IR 4.0 characteristics?		
Rank	SME	Student
1	No alignment	Intelligent energy utilization
2		Robots
3		Use of AR
Q5. Where is misalignment on the perceived level of importance of IR 4.0 characteristics?		
Rank	SME	Student
1	3D printing, additive production	Supply Chain optimization
2	Use of AR	Modern warehouse and logistics
3	Internet of Things, Machine to Machine communication, autonomous robots	Prototype making

Source: own construction based on questionnaire responses

The relevance of 14 IR 4.0 characteristics from the respondents' point of view is investigated in Q2 and Q3. In the second question, that we raised during our analysis, we were interested in seeing how respondents rank the characteristics of Industry 4.0 (Table 3 Q2). Here we wanted to capture the top three most important elements based on the perception of students and SMEs.

We observed that SMEs capture the basic, traditionally developed elements of IR 4.0 such as collecting large amounts of data; being prepared for small scale customization; and working closely with partners in the supply chain to improve responsiveness while reducing cost. Students put more weight and importance on managing environment impact more closely through intelligent energy utilization. Big data was their second most important factor. Students demonstrated strong demand for visualization and clearer interpretation of the supply chain which comes as a natural desire based on their digital maturity.

As a message we can conclude that understanding integrated data management is an essential requirement for both from the students and the SMEs side. Developing business system understanding and system development concepts transferred from the educational experience is critical. A hint of ERP is not adequate as basic knowledge. Therefore, it is recommended to embed and teach more simulation, data management and visualization aspects at least as elective courses in higher education programmes.

The third question we studied focuses on the respondents' rank of the least important characteristics of Industry 4.0 (Table 3 Q3).

In both respondent groups we can observe that the newest technologies resonate with all respondents as less important factors than robotization or 3D printing or additive production. One key difference between the groups is the importance of environment impact, where SMEs considered intelligent energy utilization among the least important factors.

As a message we can articulate that the latest technology development is an area where awareness is to be raised for both students and companies. However, we can also state that this is a field where educators most likely will need to be educated as well. An option could be to purchase expertise from the practitioners and link it to relevant courses.

Respondent group alignment on the perceived importance of IR 4.0 characteristics is studied through Q4 and Q5.

With the fourth question we were looking for intra-group alignment concerning the IR 4.0 characteristics investigated (Table 3 Q4). Concerning SMEs, we found all standard deviation above 1.1, which indicates that there is no alignment and agreement across the respondents. Companies of this size face a variety of challenges which have diverse links to the 4th Industrial Revolution. Students on the other hand have similar perception and „consensus" both on the most and least important features. It is important to note, that students are to be prepared for a very diverse environment. To address that, educators must work two ways: partly to take new technologies into the classroom, but also by taking students out of the classroom to experience the full diversity of challenges.

With the fifth question we were looking for intra-group misalignment concerning the investigated IR 4.0 characteristics (Table 3 Q5).

SMEs and student have a completely different list of characteristics, where their perception is misaligned. SMEs consider 3D printing and additive production among the least important but with a high standard deviation, which indicates, that there are some companies, who are impacted and involved in their use of technologies. Interestingly students consider supply chain management as an area where the largest misalignment is experienced.

As a take-away message we can conclude that SMEs are strongly encouraged to take development opportunities offered by technology development in order to stimulate learning and develop knowledge and technology-based improvement capabilities.

6. Conclusions & Recommendations

Many aspects of current economic and social life are impacted by the 4^{th} Industrial Revolution. However, different actors play different roles and take part in the evolving developments in a variety of ways. Technological innovations can be asset intensive changes in this new era, which requires significant capital and knowledge investment for those who want to stay competitive. SMEs might not be the leading actors of these changes unless they are specifically set up to focus on driving certain innovations as start-ups.

In Hungary, there is a significant effort being made to stay competitive in the European and global environment driven by the Irinyi Plan, which serves as an economic strategy for Hungary. This strategic plan not only emphasises the importance of the development of IR 4.0 characteristics in the Hungarian environment, but also initiates the set-up of specific tools and support for SMEs to be able to stay in the game. This serves as a top-down approach for the small and medium-sized players in the Hungarian economy. The development of higher education in respect of IR 4.0 features contributes to the human resource development of the younger generation. As the world is in a period of fast change, the revitalization of higher education is increasingly important. Indeed, higher education revitalization is essential. Students absolutely thrive on technology, but not in a business context which is to be bridged by education. Large business players do develop young candidates intensively to meet challenges, but SMEs lack the knowledge and resources to capture all relevant development opportunities for the company, including young talents.

The importance of technology challenges is experienced more sensitively by SMEs, which necessitates that educators link the latest technology developments more intensively to specific business needs. Students with their high technological agility are ready and eager to face this challenge but their economic and business understanding must be developed to materialize their digital capabilities as a business competency. The relevance of 14 IR 4.0 characteristics from the respondents’ point of view shows a mixed picture when comparing students and SMEs. The list of the most important IR 4.0 characteristics for SMEs contains technological and digital capabilities which are in general practice in the developed economies. In contrast, students rank more recent innovations at the top of their list. This results in a gap which is natural and healthy. Concerning the least important characteristics, SMEs and students strongly overlap. The respondent group alignment on the perceived importance of IR 4.0 characteristics shows that SMEs have very diverse views and no alignments, while students are sensitive and responsive to the latest developments.

Education must provide stronger awareness of the importance of business technology, even if all knowledge is not internally available. New technologies need to be brought into the classroom for students and educators. Involving leading companies or other education institutes in the network can bring great value and quality improvements to the education process.

Also providing that opportunity to SMEs could bring mutual gains for all participating parties. Education that cooperates with SMEs can provide mutual benefits for students, educators, and companies, helping them to remain on top of the new waves of industrial technology developments.

Our study has a number of limitations, partly due to the low sample size of SMEs on the business investigation side. On the education side, the questionnaire was distributed only at a single university. These limitations are to be addressed in the next phase of the project. The results of this study bring many insights which validate the continuation of in-depth investigation of the impacts of IR 4.0 on SMEs in the Hungarian environment, in parallel with identifying the necessary changes demanded of our higher education partners.

References

Abonyi, J. – Miszlivetz, F. (2016): *Ipar 4.0 megoldások kialakításának feltételrendszere*. Pannon Egyetem, Veszprém, https://www.researchgate.net/publication/301892028_Ipar_40_megoldasok_kialakitasanak_feltetelrendszere Last accessed: 27. March 2019.

Baldassare, F. – Ricciardi, F. – Campo, R. (2017): The Advent of Industry 4.0 in Manufacturing Industry: Literature Review and Growth Opportunities. *Dubrovnik International Economic Meeting*, 3, 1, 632–643.

Buckley, P.J. – Strange, R. (2015): The governance of the global factory: location and control of world economic activity. *Academy of Management Perspectives*, 29, 2, 237–249.

Davies, R. (2015): *Industry 4.0 Digitalisation for productivity and growth*, EPRS Briefing, European Parliamentary Research Service, Brussels, Belgium

Dujin, A. – Geissler, C. – Horskötter, D. (2014): *Industry 4.0 The new industrial revolution – How Europe will succeed*. Roland Berger Strategy consultants, Operations Strategy Competence Center, Münich, Germany http://www.iberglobal.com/files/Roland_Berger_Industry.pdf Last accessed: 31. March 2019.

Eurostat (2017): *Smarter, greener, more inclusive? Indicators to support the Europe 2020 Strategy*. Eurostat Statistical Book. Eurostat, Luxembourg, Luxembourg

Fülep, I. – Nick, G. A. – Várgedő, T. (2018): Zászlón a digitalizáció – Ipar 4.0. *E-KÖZIGAZGATÁSI TANULMÁNYOK.* http://kozszov.org.hu/dokumentumok/UMK_2018/2/07_Ipar_4_0.pdf Last accessed: 14. April 2019.

Haidegger, G. – Paniti, I. (2016): Industry 4.0 Platform Activities in Hungary, Past – Present – Plans. *International Scientific Journal Industry 4.0*, 1, 1, 7–10.

Janssen, R. – Blankers, I. – Moolenburgh, E. – Posthumus, B. (2014): The Impact of 3-D Printing on Supply Chain Management. *TNO Transport and Mobility* The Hague, Netherlands

Kagermann, H. – Helbig, J. – Hellinger, A. – Wahlster, W. (2013): *Recommendations for implementing the strategic initiative INDUSTRIE 4.0: securing the Future of German Manufacturing Industry*. Forschungsunion Federal Ministry of Education and Research, Frankfurt/Main, Germany https://www.din.de/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf Last accessed: 28. April 2019.

Kagermann, H. (2015): Change Through Digitalization – Value Creation in the Age of Industry 4.0 In Albach, H. – Meffert, H. – Pinkwart, A. – Reichwald, R. (eds.): *Management of Permanent Change*, Springer Fachmedien Wiesbaden Publisher, Wiesbaden, Germany, 23–45. DOI: 10.1007/978-3-658-05014-6_2

Khan, A. – Turowski, K. (2016): A Perspective on Industry 4.0: From Challenges to Opportunities in Production Systems. *Proceedings on the International Conference on Internet of Things and Big Data (IoTBD 2016)*, 441–448.

Klitou, D. – Conrads, J. – Rasmussen, M. – Probst, L. – Pedersen, B. (2017): *Hungary: "IPAR 4.0 National Technology Platform"*. European Commission Digital Transformation Monitor, European Commission, Brussels, Belgium

Kuruczleki, É. – Pelle, A. – Laczi, R. – Fekete, B. (2016): The Readiness of the European Union to Embrace the Fourth Industrial Revolution. *Management (Slovenia)*, 11, 4, 327–347.

Monostori, L. (2015) Cyber-physical production systems: roots from manufacturing science and technology. *Automatisierungstechnik*, 63, 10, 766−776.

Nagy, J. (2017): *Az ipar 4.0 fogalma, összetevői és hatása az értékláncra* (The definition of Industry 4.0, its components and impacts on the supply chain), Budapesti Corvinus Egyetem Vállalatgazdaságtan Intézet, Budapest http://unipub.lib.uni-corvinus.hu/3115/1/Nagy_167.pdf Letöltve: 2018. 10. 27.

Nick, G.A. (2017): *Az Ipar 4.0 Nemzeti Technológiai Platform támogató szerepe* (The supporting role of Industry 4.0 National Technology Platform). XXIV. NMK Plenáris ülés, Balatonalmádi, 2017. szeptember 14. https://www.isoforum.hu/media/programnaptar/files/NickGabor-eloadas.pdf : Last accessed: 04. April 2019.

Probst, L. – Lefebvre, V. – Martinez-Diaz, C. – Bohn, N. U. – Klitou, D. – Conrads, J. (2018): *EU businesses go digital: Opportunities, outcomes and uptake*. Digital Transformation Scoreboard 2018, European Commission, Brussels, Belgium

Ritter, M. – Török, J. – Pongrácz, F. (2016): Hogyan használható a mesterséges intelligencia, az innovációmenedzsmentben a Negyedik Ipari Forradalomban a kis-és középvállalkozások növekedési pályára állításában? (How to apply Artificial Intelligence in Innovation Management to support the growth of small and medium businesses in the area of industry 4.0?). *Logisztika - Informatika – Menedzsment*, 1, 1, 111–127.

Rüßmann, M. – Lorent, M. – Gerbert, P. – Waldner, M. – Justus, J. – Engel, P. – Harnisch M. (2015): *Industry 4.0 The Future of Productivity and Growth in Manufacturing Industries*, Boston Consulting Group, https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx Last accessed: 04.April 2019.

Sniderman, B. – Mahto, M. – Cotteleer, M.J. (2016): *Industry 4.0 and manufacturing ecosystems – Exploring the world of connected enterprises*, Deloitte University Press. Deloitte, Westlake, USA

Strange, R. – Zucchella, A. (2017): Industry 4.0, global value chains and international business, *Multinational Business Review*, 25, 3, 174–184.

Szalavetz, A. (2017): Industry 4.0 in ’factor economies. In Galgóczi B. – Drahokoupil J (eds.): *Condemned to be Left Behind? Can Central and Eastern Europe Emerge from its Low-Wage Model?*, European Trade Union Institute, Brussels, 133–152.

Van der Elst, K. – Williams, A. (2017*): Industry 4.0: The new production paradigm and its implications for EU policy*. IN European Commission: Report of the Expert Group ‘Strategic Foresight for R&I Policy in Horizon 2020’ Background Papers. European Commission, Brussels, Belgium

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 79–92.

Managing inter-organisational dynamics – distinctive patterns?

Margit Tarjányi – Márton Vilmányi

The relationship between organizations is a regular topic of both relationship marketing, relationship management research and strategic management research. Such a constant, high degree of interest is not surprising. The identification of change patterns in relationships of organizations can help in the explanation of several business phenomena.

In our study, the results of a questionnaire survey are presented. The main research question is whether the development of cooperation behaviour between business enterprises could depend on the size of enterprises.

For this research the resource-based view and the theory of dynamic capabilities were used. The results confirm the relevant starting points of the approach to dynamic capabilities while providing further insights at some points in light of Hungarian specificities.

Keywords: dynamic capabilities, organizational relationship

1. Introduction

This study focuses on the problem of how to manage organisational change, and how to implement changes within an organisation. While traditional paradigms – using the vantage point of strategic management – ten regard the problem from the perspective of competitive forces and strategic conflicts, present day standpoints have their roots instead in a resource-based perspective, evolutionary economics and in the concept of dynamic capabilities (Teece et al. 1997). The competitive forces perspective (Porter 1993) draws attention to industry structure. The management of organisational dynamics aims to recognize change within the industrial structure and to reconfigure the organisation's position within the industrial structure. The strategic conflicts approach is based on game theory.

This approach embraces both the source and management of dynamics as a function of competition, where competitors aim to keep their rivals away from those areas that are seen as key priority areas (Shapiro 1989). The resource-based view states that profit is generated from the possession or control of valuable, rare, and unique resources that have no substitutes. As the core of dynamics, this approach focuses on the acquisition of resources, the management of knowledge and know-how, or on the management of learning as a strategic area (Teece et al. 1997). To explain the strategic dynamics of organisational change, and how such change is implemented, present day approaches rely heavily on the evolutionary economics concepts of path dependency and routine (Helfat–Peteraf 2009). In this terminology, routines refer to such organisation-specific and embedded patterns of action through which an organisation can repeatedly perform those individual tasks that are tied to its functions. Path dependency means that investments and the repertoire of routines that had been set in the past by the company restrict how the company will act in the future (Penrose 1959).

In line with the above-mentioned concepts, we can see from the point of view of dynamic capabilities that routines that are responsible for the implementation of changes in an organisation and for the management of organisational dynamics, are those routines that aim to reconfigure the product, the production process, the buyers, etc. (Winter 2002). From a dynamic capabilities approach these actions are such mechanisms that implement change, and in the background, there is a pattern of a set of linked routines that help the repeated performance of a given behaviour.

2. An approach to dynamic capabilities and their characteristics

In the management of organisational dynamics, the concept of dynamic capabilities is a very frequently used term. Some definitions refer to dynamic capabilities as a set of resources, while others as capacity, and some approaches define these capabilities as complex routines or organisational competence. The concept of capacity originates from an approach by Teece et al. (1997) and refers to individual competencies with no regard to the level at which a given action is performed (Helfat–Peteraf 2009). Helfat et al. (2007) point this contradiction out and introduce the concepts of the intentional/deliberate. In accordance with the authors' concept, capacities or routine/resource sets that had been developed in order to manage change fall into the category of dynamic capabilities. Finally, it is important to emphasise the "higher-order competence" nature of dynamic capabilities, a component that regularly recurs in its definitions (Teece 2012). The capabilities to implement change can be defined as those (hierarchically higher represented) behaviour patterns that formulate the above-mentioned stable competencies.

To draw the above-mentioned definitions to a conclusion and to agree on a fixed definition of dynamic capabilities, we shall hereinafter rely on a description by Eisenhardt–Martin (2000), who state that dynamic capabilities encompass those processes that use resources to integrate, reconfigure, gain and downsize a further set resources with the specific intent to adjust to or initiate market change.

Dynamic capabilities may be present in several forms and display some rather strong characteristics (Eisenhardt–Martin 2000). A particular form of dynamic capabilities is what has been called resource integration capabilities, such as product development capabilities. Another form of dynamic capabilities is reconfiguration capabilities that copy, transfer or reconfigure resources (typically knowledge) in a distinctive way. Furthermore, capabilities that support shared development and, thus, help to create links between several areas of an organisation in order to set up a sort of collaboration network can also be identified as a type of dynamic capability. Combinations of the above can also be viewed as dynamic capabilities. Within this framework, resources with different qualities are combined in order to reconfigure business opportunities. Dynamic capabilities encompass both alliance and acquisition routines that provide access to certain resources and exit routines that serve to terminate existing resources or resource combinations.

There may be differences in how a given dynamic capability is embedded. Dynamic capabilities function differently in moderately dynamic markets than in high-velocity markets (Eisenhardt–Martin 2000). In moderately dynamic markets, dynamic capabilities establish a well-structured process and remain stable and linear. In contrast with this, in high-velocity markets, dynamic capabilities do not embrace complex processes but rather consist of simple routines and the formation of real-time knowledge. General principles and forms of behaviour are laid down for the actors in order to focus the attention of the actors within an organisation onto important areas though the provision of sufficient reference points that help the implementation of reconfiguration (Gupta–Winter 2009).

While some studies that aim to uncover dynamic capabilities focus – as demonstrated above –on how these capabilities emerge, other studies focus on those areas that these capabilities improve.

Following this line, depending on how path dependency works in the case of dynamic capabilities, Desmond (2007) defines weak and strong dynamic capabilities. Weak forms of dynamic capabilities refer to the fact that there is familiarity with how to make use of the company's resources. In contrast to that, strong forms of dynamic capabilities do not build on experiences but use innovative experiments and improvised practices to discover new resources and fields of use.

Dynamic capabilities – as their targets are considered – can aim at a given resource or an existing capability, but also at resource-architecture (Kusnoki et al. 1998). Resource (or knowledge) architectures describe the structure of resource combinations, the method of how two or more of those components combine which exist independently and function on the markets, and as a result, are able to satisfy the needs of a set of buyers. Resource architecture goes beyond organisations and, consequently, dynamic capabilities that are tied to resource-architecture are basically dynamic capabilities that are linked to inter-organisational collaborations. Such capabilities go beyond organisational boundaries and enable the actors in an organisation, in association with other components in the architecture, to redefine knowledge for the individual components (Andersson et al. 2008).

When a summary is made for the characteristics of dynamic capabilities, it becomes clear that dynamic capabilities exist in three basic forms (Teece 2011; 2012):

(1) Sensing capabilities: the capability to identify and evaluate opportunities. Sensing capabilities encompass the observation and evaluation of phenomena and the development and evaluation of hypotheses. Such activities require either managerial insight and vision or the completion of an analytical process.

(2) Seizing capabilities: the capability to mobilize resources, tap opportunities and take hold of certain assets. Seizing capabilities are the sum of those routines and procedures that aim to tap already identified opportunities; examples include the design of business models, ensuring access to capital, and the setup of buyer/supplier relationships.

(3) Transforming capabilities: the capability to continuously renew in order to implement the changes and improvements within the currently existing working routines and capabilities. The function of transforming capabilities is to level out those changes between organisational capabilities that are triggered by seizing capabilities and, thus, support the management and maintenance of those organisational areas that accommodate the newly introduced functions and tasks.

3. Dynamic relationship capabilities

Dynamic relationship capability is a multi-dimensional phenomenon that has been studied and described by several disciplines. The phenomenon is approached from many different aspects: strategic management (Gulati 1999, Dyer-Singh 1998), inter-organisational collaboration (Möller–Halinen 1999, Äyväri–Möller 2008), dynamic capabilities (Knight et al. 2005, Vesalainen–Hakala 2014) and organisational learning and knowledge (Saeedi 2014, Csontos–Szabó 2017). Dynamic relationship (also termed network) capabilities can be seen as the sum of all those routines, processes, and patterns of action that allow an organisation to improve its relationship competence or make use of its collaborations, in other words, to set up its collaborations with external partners, optimize its relationship portfolio, and allocate its resources between its partner relationships (Gemünden et al. 1997, Ritter 1999, Ritter et al. 2002, Mitrega et al. 2012, Horváth et al. 2018).

Because of the nature of dynamic capabilities, dynamic relationship capabilities can be described from various aspects. Dynamic relationship capabilities already have two interpretations: on the one hand, the capability to manage collaboration – the way this process has been described above – is by itself a form of dynamic capability that can result in the reconfiguration of organisational capabilities; on the other hand, based on another approach, the literature identifies an aspect of dynamic relationship capabilities where the function of dynamic relationship capabilities is to renew relationship behaviour within organisations. In what follows, we shall proceed with dynamic relationship capabilities as defined by this latter meaning.

The literature provides plenty of examples of relationship competence reconfigurations. From a primarily learning and knowledge-based theory, Jaratt (2009) looks at dynamic relationship capabilities as the sum of those elements, learning patterns (learning that helps to advance, to adapt, and includes the use of knowledge) that can lead to the implementation of relationship management reconfiguration. As the outcome of her research, Jaratt concludes that the implementation of reconfiguration relating to relationship management competence is significantly influenced by organisational culture (the degree and type of learning orientation), the structure of organisational learning, and the degree of learning (which helps to advance, and is adaptive) that has been embedded during relationship management.

By taking a resource-based and an IMP[1] relationship-oriented aspect into consideration, Johnsen–Ford (2006) define the implementation of relationship competence reconfiguration as a multi-dimensional construct. From the authors' perspective, the role of dynamic relationship capability (in their own words: interaction capability) is to enable the collaborating actors to improve their relationship management. The authors describe dynamic relationship capability as the product of four capability factors: human interaction capability, technological interaction capability, organisational structure interaction capability, and cultural interaction capability. The study by Johnsen–Ford (2006) also clearly demonstrates that in the literature dynamic relationship capability is a construct that has been, either explicitly or implicitly, identified or investigated on multiple levels. These results have all contributed to the interpretation of dynamic relationship capabilities.

Roseira et al. (2013) highlight those features of dynamic relationship capability that are found on a strategic level. By starting from a dominantly IMP interaction approach, the authors emphasise the integration of relationship strategy, interactions, network pictures and organisational positioning. From the perspective of relationship strategy, the authors highlight that individual experience, background, and the formation of a cognitive opinion are just as important as the process of such social interaction which results in the formation of an organisational attitude. The authors stress that any reconfigurations that are made in the relationship strategy of an organisation are determined just as much by a reconfiguration in the network picture/the capability to reconfigure this picture than by the actual change in the situation itself. Consequently, to maintain a value-creating strategy in inter-organisational interactions, it is necessary to manage the full experience of all members within the organisation, the setting or reconfiguration of the perceived network position, and the formation of the network strategy.

Reinhartz et al. (2004) examine relationship management reconfiguration capability from a different angle: the perspective of CRM processes. The authors find that in connection with the composition/modification of relationship processes, there are three distinctive characteristics. Firstly, there are organisational and industry-specific characteristics, and the capability to manage the reconfiguration of these characteristics. Secondly, in the course of relationship process modifications, the authors highlight the significance of relationship life-cycle evaluation. And thirdly, the authors point out the management of the diverse distributions of relationship value, which stays heterogeneous between partners over time.

Finally, when research on the various implementation levels of relationship management reconfiguration is discussed, Havila and Medlin's 2012 study has to be mentioned. The authors focus on those characteristics that are related to the implementation of change in relationship management activities. From the perspective of project, relationship, and knowledge management, the authors have done research on the termination of collaborations, concluding that those behaviour

[1] Industrial Marketing and Purchasing Group

patterns that are displayed during reconfigurations are always the products of inter-organisational and extra-organisational elements, where reconfiguration is always part of a broader network. This network has to be taken into consideration. Experience that is derived from change related to the implementation of collaborations is generalized, and the implementation process is affected by the following: experience (embedded in a process or an organisation) that members in an organisation (individuals, groups) have and that is available for the completion of a certain task; the implementation of accessible external (partner) knowledge; and the capability to manage the outcomes and effects of reconfiguration on the level of personal and organisational interactions. The study distinguishes between operational and strategic levels and concludes that when reconfiguration is successfully implemented on each of these levels then, reciprocally, it is determined that reconfiguration can also be successfully implemented within the entire organisation.

4. Aim, model and methodology of the study

In our study, we have the following research question: with a change in their size, do Hungarian business units which operate in an inter-organisation market display any change in their behaviour with the aim to improving their collaborations? Results by Eisenhardt–Martin (2000) have demonstrated that depending on industrial dynamics, the structure of dynamic capabilities displays various differences. In the field of relationship capabilities, Äyväri–Möller (2000), O'Toole and McGrath (2008) and Sutton-Brady et al. (2011) have verified the existence of these differences in connection with the size of business units as well. These facts have heightened our interest into whether, in the case of Hungarian business units, the characteristics of dynamic relationship capability that are typical of business organisations change when there is a change in the size of a business unit.

To operationalize our work, our study has focused on the issue of change management in collaboration strategy. As conclusions from previous studies have indicated, dynamic relationship capabilities surface in a rather complex way. Thus, we tested our hypothesis by placing the strategic dimension of the phenomenon into focus; on this dimension, the various patterns of reconfiguration dynamics will, by all means, leave a mark. It has been the goal of our research to examine to what extent the reconfiguration capability of relationship strategy adjusts when there is a change in the size of a business unit.

During the design of our research model, we defined the strategic dimension of dynamic relationship capabilities in the following way: how can an organisation reconfigure its relationship strategy within the very network where it fulfils its role? Relying on studies by Teece (2011; 2012) and Roseira et al. (2013), we have defined the strategic dimension (strategic flexibility) of dynamic relationship capabilities on four levels:

- Perception capability: supports those solutions, tools, and routines that allow the perception of such experience as s that derive from relationship interaction.
- evaluation capability: supports those solutions, tools, and routines that, in line with experience derived from interactions, enable the evaluation of relationship investments and relationship value.
- learning capability: supports those solutions, tools, and routines that, in line with relationship value, enable the reconfiguration of the perceived position and the network picture.
- integrating capability: supports those solutions, tools, and routines that enable the integration of the reconfigured strategic elements into the practice of an organisation.

To be able to examine the successful functioning of strategic flexibility, we have made use of two target variables that fit how we have approached dynamic relationship capabilities:

- perceived suitability of relationship capabilities: the level of suitability of those routines and methods that are utilized for relationships management.
- perceived success of collaboration: the general level of satisfaction with partner relationships.

To test the research questions posed in this study, a questionnaire methodology was made use of and carried out on a nationwide representative sample between 1 December 2016 and 31 January 2017.

From the Business Units Registry, population data was generated, and with the assistance of the staff of the Hungarian Central Statistical Office (KSH), a random sample compiled to represent the target group. After retrieving data on all publicly available and operational business organisations, we sorted out self-employed businesses and those that fall into the unknown staff numbers category. Furthermore, the population data did not include those organisations that were under liquidation, faced bankruptcy or were being wound up. When we compared the outcome with the population data, we realized that the item numbers did not match the staff number category; subsequently we made use of weighing factors. The generated sample was representative.

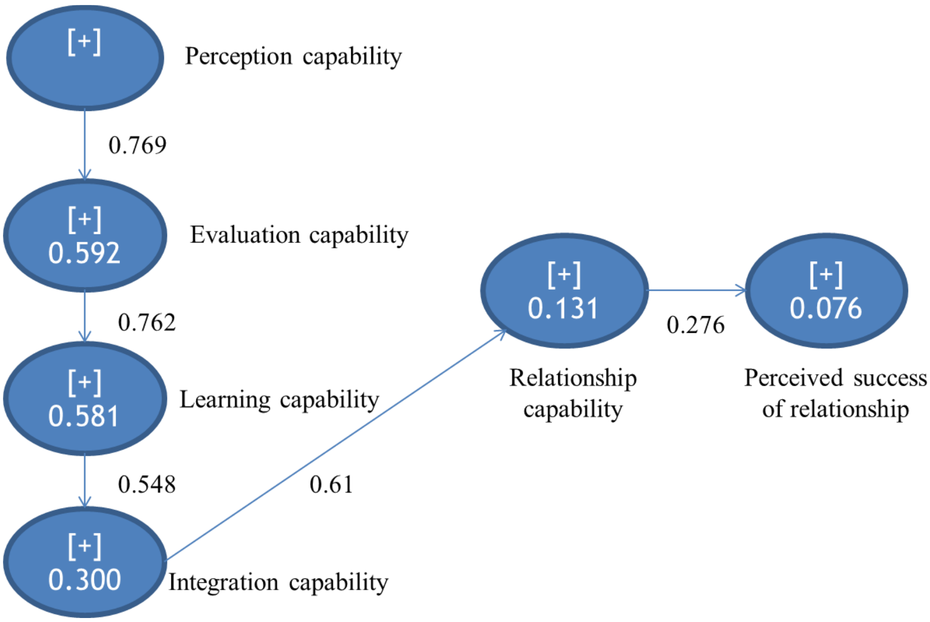
The questionnaires were sent to a thousand corporate bodies altogether, electronically, with the assistance of the staff of the Hungarian Central Statistical Office. Respondents were given two weeks to fill in the questionnaires, and it was our request that decision-makers filled in these documents. A 6-point Likert scale was used to measure the characteristics under scrutiny. 312 corporate bodies returned the questionnaires at a 30% response rate. Altogether 301 corporate bodies responded to all of our survey questions, therefore these bodies represent our research sample.

Taking the exploratory nature of this study into consideration, we have made use of PLS path analysis to examine the influential force of strategic-level flexibility factors on perceived relationship success and perceived business success (Kazár 2014).

5. Research outcomes

During the course of this study, we have analysed strategic level dynamic capabilities along four factors (sensing, evaluation, learning, and integration), while the stable relationship capability of organisations and the success of inter-organisational relationships have been measured by one factor each. For the assessment of data, we have made use of PLS path analysis. The outcomes of the assessment are presented in Figure 1.

Figure 1 Correlation between strategic flexibility in collaborations and collaboration success



Source: Authors' own work.

The outcomes presented in Figure 1 highlight three significant correlations:

– Firstly, there is a strong link between those factors that define strategic flexibility: the perception of the network picture exerts a major (0.769) influence on how the network picture is evaluated, the evaluation of the network picture exerts influence on the learning dimension (0.762), while the learning dimension exerts influence on integration (0.548). The perception of the network picture explains in 59% of respondents how the network picture is evaluated, the evaluation of the network picture explains in 58% the learning dimension, while the learning dimension explains in 30% the reconfiguration of strategy. The resulting data reflect well on the strong link between the intertwining categories of sensing, evaluation, learning, and integration.

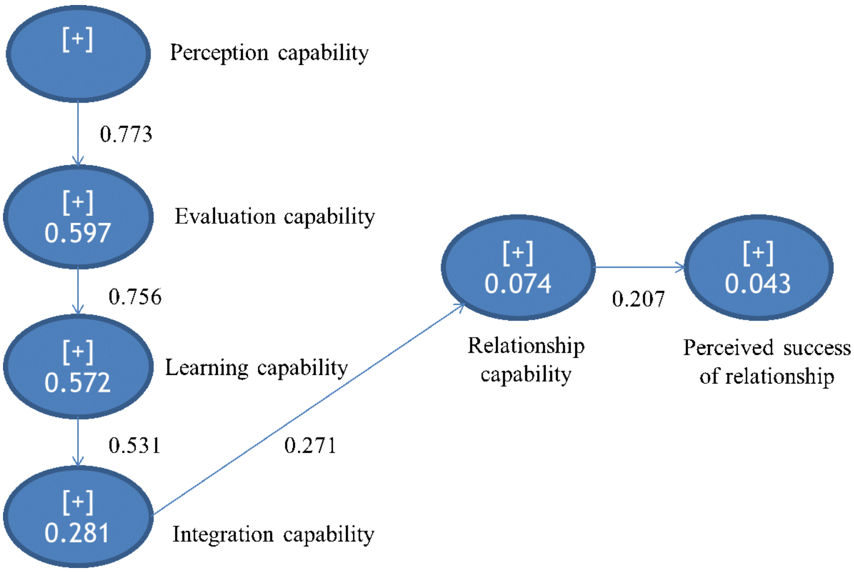
- Factors describing strategic flexibility and stable relationship capability have a weaker correlation than what is demonstrated above; the only significant one (36%) being between integration capability and stable relationship capability. However, the explanatory value of the four factors is rather low: 13% overall.
- The strategic flexibility and relationship capability of organisations affects (0.276) the success of organisations, but the explanatory value is very low: 7%.

In the second stage of our research, in line with our research question, we have divided our sample, based on their staff numbers, into two sets. This resulted in two groups: (1) businesses that employ 2–50 people and (2) businesses that employ more than 50 people. To assess the two sample groups, we have again made use of PLS path analysis. The outcomes of these assessments are presented in Figures 2 and 3.

The following correlations have been established from the research outcomes linked to businesses that employ between 2–50 people (micro and medium enterprises):

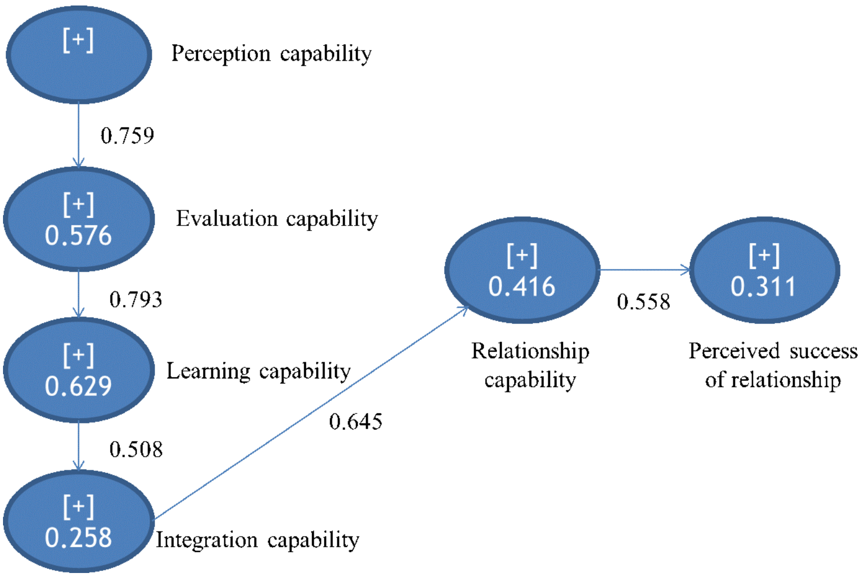
- Similar to the results derived from the entire sample spectrum, there is a strong link between those factors that define strategic flexibility. The perception of the network picture exerts a major (0.773) influence on how the network picture is evaluated, the evaluation of the network picture exerts influence on the learning dimension (0.756), while the learning dimension exerts influence on integration (0.531). The perception of the network picture explains in 59% of responses how the network picture is evaluated, the evaluation of the network picture explains in 57% the learning dimension, while the learning dimension explains in 28% the reconfiguration of strategy.
- Compared to the entire sample spectrum, there is a weaker (0.271) correlation between the factors that describe strategic flexibility and stable relationship capability. Overall, the explanatory value stays very low: 7%.
- The strategic flexibility and relationship capability of organisations affects (0.207) the success of organisations, but the explanatory value is very low: 4%.

Figure 2 Correlation between strategic flexibility in collaborations and collaboration success with businesses that employ between 2–50 people



Source: Authors' own work.

Figure 3 Correlation between strategic flexibility in collaboration and collaboration success with businesses that employ more than 50 people



Source: Authors' own work.

The following correlations have been established from the research outcomes linked to businesses that employ more than 50 people:

- Similar to the results derived from the entire sample spectrum, there is a strong (0.759; 0.793; 0.508) link between those factors that define strategic flexibility. The explanatory value is also similar to what was measured on the overall sample spectrum.
- But compared to the entire sample spectrum, there is a major difference in the correlation between factors that define strategic flexibility and stable relationship capability. Dynamic strategic capability exerts a stronger (0.645) influence on relationship capability, and the explanatory value is also higher (41%).
- Strategic flexibility and relationship capability of organisations have a stronger effect (0.207) on the success of organisations, and the explanatory value is also higher (31%).

6. Summary

Our current study has achieved two goals: first, it has made an attempt to observe how strategic level dynamic capability affects relationship capability and perceived relationship success, and, second, it has examined whether there are any changes in this model depending on the size of a business.

This study has demonstrated that, in B2B situations, dynamic strategic capability reveals itself as a clear model. The research outcomes indicate that the examined characteristics of dynamic strategic capability (perception and evaluation of the network picture, learning, and modifying strategy) significantly influence each another.

On the other hand, in B2B situations, dynamic strategic capability exerts a weak influence on relationship capability and on the perception of collaboration success. However, when the responses have been examined by taking staff numbers into consideration, we have identified a compelling distinction. In businesses that employ between 2–50 people, we have identified a pattern similar to the one that has emerged from the data related to the overall population spectrum. In businesses that employ more than 50 people, strategic flexibility exerts a stronger influence on their stable relationship capability, and, through that, on their perceived relationship success. This divergence means that in businesses that employ more than 50 people, it is clearly proven that the pattern of sensing-evaluation-learning-integrating affects the implementation of collaborative actions. On the other hand, in businesses that employ less than 50 people, the formation of collaborative behaviour does not clearly follow the model described in this study but rather occurs along a different pattern.

The research outcomes of our study partly corroborate, partly go beyond those research outcomes that had previously been demonstrated in this field of study – and in such a way that further questions arise. Firstly, our results confirm that the structure of strategic flexibility in entrepreneurships differs from the structure of strategic flexibility in larger companies. Secondly, the study discloses the layers of strategic flexibility in businesses that employ more than 50 people. Thirdly, our results raise questions for further examination on the issue of strategic flexibility structures in entrepreneurship.

On the one hand, our results expose those areas that managers of businesses that employ 50 people should address in order to continuously improve the relationship capability of their organisations. On the other hand, our results also demonstrate that such expertise is by no means universal: the adaptation of such experience to entrepreneurship is not possible, except in limited form.

References

Andersson, M – Lingren, R. – Hendfridsson (2008): Architectural knowledge in inter-organisational IT innovation. *The Journal of Strategic Information Systems*, 17, 1, 19–38.

Äyväri, A. – Möller, K. (2000): *Entrepreneurial Networking and Marketing – Dealing with People*. 16th IMP Conference, Bath, U.K. https://www.impgroup.org/uploads/papers/30.pdf Accessed: 12/11/2018.

Äyväri, A. – Möller, K. (2008): *Understanding relational and network capabilities – a critical review*. 24th IMP-conference in Uppsala, Sweden, http://www.impgroup.org/paper_view.php?viewPaper=6862 Accessed: 12/11/2018.

Csontos, R. – Szabó, Zs. R. (2017): A kapcsolatok erősségének szerepe a hálózati tanulásban. In Veresné, S. M. – Lipták, K. (szerk.) *„Mérleg és Kihívások" X. Nemzetközi Tudományos Konferencia = „Balance and Challenges" X. International Scientific Conference*: Konferenciakiadvány: A közgazdászképzés elindításának 30. évfordulója alkalmából. Miskolc-Egyetemváros, Magyarország: Miskolci Egyetem Gazdaságtudományi Kar, 68–84.

Desmond, W. N. (2007): A Modern Resource Based Approach to Unrelated Diversification. *Journal of Management Studies*, 44, 8, 1481–1502.

Dyer, J. H. – Singh, H. (1998), The relational view: Cooperate strategy and sources of interorganisational competetive advantage. *Academy of Management Review*, 23, 4, 660–679.

Eisenhardt, K. M. – Martin, J. A. (2000): Dynamic Capabilities: What are They? *Strategic Management Journal*, 21, 10/11, 1105–1121.

Gemünden, H. G. – Ritter, T. – Walter A. (1997): *Relationships and Networks in International Markets*. Pergamon, UK.

Gulati, R. (1999), Network location and learning: the influence of network resources and firm capabilities on alliance formation. *Strategic Management Journal*, 20, 5, 397–420.

Gupta, A. – Winter, S. (2009): *Dynamic Capabilities of the Firm and Strategic Change*. Druid Summer Conference.

Havila, V. – Medlin, C. J. (2012): Ending competence in business closure. *Industrial Marketing Management*, 41, 3, 413–420.

Helfat, C. E. – Peteraf, M. A. (2009): Understanding Dynamic Capabilities: Progress Along a Developmental Path. *Strategic Organisation*, 7, 1, 91–102.

Helfat, C. E. – Finkelstein, S. – Mitchell, W – Peteraf, M. – Singh, H. – Teece, D. J. – Winter, S. (2007): *Dynamic Capabilities: Understanding Strategic Change in Organisations*. Malden, MA: Blackwell.

Horváth, D. – Móricz P. – Szabó Zs. R. (2018): Üzletimodell-innováció. *Vezetéstudomány / Budapest Management Review,* 49 (6), 2–12.

Jaratt, Denise (2009): Organisational preconditions supporting relationship management capability renewal. *Journal of Strategic Marketing*, 17, 5, 365–381.

Johnsen, R. E. – Ford, D. (2006): Interaction capability development of smaller suppliers in relationship with larger customers. *Industrial Marketing Management*, 35, 8, 1002–1015.

Kazár, K. (2014): A PLS-útelemzés és alkalmazása egy márkaközösség pszichológiai érzetének vizsgálatára. *Statisztikai Szemle*, 92, 1, 33–52.

Knight, L. – Harland, C. – Walker, H. – Sutton, R. (2005), Competence requirements for managing supply in interorganizational networks. *Journal of Public Procurement*, 5, 2, 210–234.

Kusnoki, K. - Nonaka, I. - Nagata, A. (1998): Organisational Capabilities is Product Development of Japanese Firms: A Conceptual Framework and Empirical Findings. *Organisational Science*, 9, 6, 699–718.

Mitrega, M. – Forkmann, S. – Ramos, C. – Henneberg, S. C. (2012): Networking capability in business relationships – Concept and scale development. *Industrial Marketing Management*, 41, 5, 739–751.

Möller, K. K. – Halinen, A. (1999), Business Relationships and Networks: Managerial Challenge of Network Era. *Industrial Marketing Management*, 28, 5, 413–427.

O'Toole, T – McGrath, H. (2008): *Implementing a Relational Capability Framework through an SME Network.* 24th IMP-conference in Uppsala, Sweden. https://www.impgroup.org/uploads/papers/6736.pdf Accessed: 11/12/2018.

Penrose, E. (1959): *The Theory of the Growth of the Firm.* Oxford University Press, London.

Porter, M. E. (1993): *Versenystratégiák.* Akadémiai Könyvkiadó, Budapest.

Reinhartz, W. – Krafft, M. – Hoyer, W. D. (2004): The customer relationship management process: Its measurement and impact on performance. *Journal of Marketing Research*, 41, 3, 293–305.

Ritter, T. – Wilkinson, I. F. – Johnston, W. J. (2002): Measuring network competence: some international evidence. *Journal of Business & Industrial Marketing*, 17, 2/3, 119–138.

Ritter, T. (1999): The Networking Company, Antecedents for Coping with Relationships and Networks Effectively. *Industrial Marketing Management*, 28, 5, 467–479.

Roseira, C. – Brito, C. – Ford, D. (2013): Network pictures and supplier management: An empirical study. *Industrial Marketing Management*, 42, 8, 234–247.

Saeedi, R. M. (2014), *Fostering Dynamic Capabilities of SMEs – The Impact of Inward International Licensing on Absorptive Capacity and Networking Capability: A Multiple Case Study*. Linköping Studies in Science and Technology Thesis 1653, Linköping University Department of Management and Engineering.

Shapiro, C. (1989): The Theory of Business Strategy, *RAND Journal of Economics*, 20, 1, 125–137.

Sutton-Brady, C. – McGrath, H. – O'Toole, T. (2011): *The Evolution of Network Capability in an SME Context*. 27th IMP-conference, Glasgow, Scotland. https://www.impgroup.org/uploads/papers/7644.pdf Accessed: 11/12/2018.

Teece, D. J. (2011): Dynamic Capabilities: A Guide for Managers, *Ives Business Journal*, 75, 2, 29–31.

Teece, D. J. (2012): Dynamic Capabilities: Routines versus Entrepreneurial Action. *Journal of Management Studies*, 49, 8, 1395–1401.

Teece, D. J. – Pisano, G. – Shuen, A. (1997): *Dynamic Capabilities and Strategic Management. Strategic Management Journal*, 18, 7, 509–533.

Vesalainen, J. – Hakala, H. (2014), Strategic capability architecture: The role of network capability. *Industrial Marketing Management*, 43, 6, 938–950.

Winter, S. G. (2002): *Understanding Dynamic Capabilities*. A working Paper of the Reginald H. Jones Center. The Warton School University of Pennsylvania. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.202.2236&rep=rep1&type=pdf Accessed: 11/12/2018.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 93–114.

The impact of country image, country-of-origin image and consumer ethnocentrism on purchase decisions: A study about Azerbaijani food companies' entry into Hungarian Market

Ilkin Salmanli – Sevinj Omarli

The main aim of the study is to examine the impact of Azerbaijan's brand image as a country on the purchase intentions of Hungarians towards Azerbaijani products. Azerbaijani products are not on the Hungarian market yet, and this research identifies whether it would be wise for Azerbaijani food companies to launch their operations in Hungary or not, and what kind of prejudices and obstacles they would need to cope with so as to be successful. To this end, the impact of the country image of Azerbaijan, the country-of-origin of Azerbaijani products, and ethnocentrism among Hungarians on purchase intentions are analyzed to fulfil the objective of the research. It was identified that the country image of Azerbaijan in Hungary is neither highly negative nor highly positive because of the fact that Hungarians do not know much about Azerbaijan. However, values showed and proved that they are likely to have a positive attitude towards Azerbaijan as a country. As for COI, similar statistics were unearthed, and it was noticed that Hungarians do not have any objection towards Azerbaijani product COI. These two variables also confirmed, with statistical analysis of dependent variable of purchase intentions, there is a relationship between positive country image, COI and purchase intentions.

Keywords: country image, country-of-origin image, consumer ethnocentrism, purchase decisions

1. Introduction

The initial interest of this study is to look at the effect that the country image (CI), COI (COO or COI) and ethnocentrism is likely to have on the purchasing decisions of consumers towards fast-moving consumer goods. The influence of country-of-origin is about the impact of one's generalizations and perceptions on the evaluation of a particular country's products and brands (Elliot–Cameron 1994). Additionally, Elliot–Cameron (1994) state that COO is the impact of producing country's image on the consumer decision-making process, which can be influenced positively or negatively.

COO, as an extrinsic attribute, is used by buyers where they lack tangible factors. Not surprisingly, the country-of-origin of a product is of a great importance as an informative sign for consumers in making decisions (Solomon et al. 2016)

Furthermore, globalization (which is a buzz word in the modern day) has enabled companies to broaden their markets by expanding overseas, and the whole world has become much smaller than it was thought to be before and, therefore, for consumers to make a decision has turned out to be a cumbersome task. Willingness to buy domestic/foreign products is influenced by country-of-origin and quality judgement (Cai et al. 2004). As a result, a great deal of research has been conducted to identify how consumers decide when faced with such a bewildering amount of choices in market places. Plenty of similar concepts have been unearthed and taken up in the literature in this context, such as country-of-origin, country image, made-in country image, country equity, etc. A country image is defined by Martin and Eroglu (1993), “as the total of all descriptive, inferential and informational beliefs one has about a particular country. The country image can be utilized by purchasers in product evaluations when they are not so much adept at defining the true quality of country’s products before purchases”.

Having a powerful country image along with a commensurate country-of-origin image is very important nowadays, especially in terms of boosting exports of a particular country. In light of this, the research problem to be examined is about how the perception of consumers is impacted by the information they get, and why consumers give preferences to a particular product focusing on its origin. Given the nature of similar fast-moving consumer goods (FMCG), how would consumers be likely to react were they to have no information about the product and vice versa.

The aim of this study is to explore if there is a positive impact of the brand image of Azerbaijan as a country on purchase intentions in Hungary, and how the country-of-origin image of Azerbaijani products and ethnocentrism levels among Hungarians affect purchases of Azerbaijani products by Hungarians in the food industry. We are looking to find the answer for the following questions: What brand image does Azerbaijan possess among Hungarian citizens? How favorable is the opinion Hungarians have of foreign-made products? What are the reactions of Hungarians towards Azerbaijani-made products? How successfully can an Azerbaijani food manufacturing company be operating in Hungary?

2. Literature review

2.1. Evolution of the concept of COI

Country-of-origin (COO) and country-of-origin image (COI) are inextricably linked concepts. COO is exploring the effect of national origin of a certain product which would impact on consumer assessments and preferences, whilst COI is the concept which elaborates more on the particular aspects of the country to drive consumer perception and attitudes towards products from a given country (Roth–Diamantopoulos 2009).

Having been investigated for in excess of 50 years, the COI concept has undergone a series of changes in terms of literature richness during this time frame. Early period studies were carried out to study the effect of country-of-origin as a product feature, and until the year of 1982, the effect of country-of-origin had been addressed as the only variable in the studies. Over time, different product groups and counties were evaluated and studied, and simple experimental designs used as a method of exploration.

In the food industry, the investigation of COI is still on the spotlight and researchers continue to investigate the field intensively. Consumers tend to shy away from calorific and animal welfare information in case their perception causes emotional perturbation such as guilt or cognitive dissonance, and knowing the origin information can serve merely to better results or it can result in the emergence of negative emotions (Beiermann et al. 2017). Feldmann and Hamm found there were research gaps in various areas including cross-national comparisons, the influence of different types of products and food product origin, and they found local food is not perceived to be very expensive, unlike organic food, however, buyers are prone to spend a premium for local food rather than the foreign-made (Beiermann et al. 2017).

Based on an extensive literature review of main COO related publications, it is possible to draw a conclusion that the research field in question still lacks an integrative theory which could make the COO phenomenon more universal and thus better understood and utilized. However, persistence of scholars has slowly started to pay off and an overall picture of the structure of COO seems to be emerging. Researchers are also keen to identify the process of how consumers incorporate information about product COO in forming their attitudes and expressing their buying intentions.

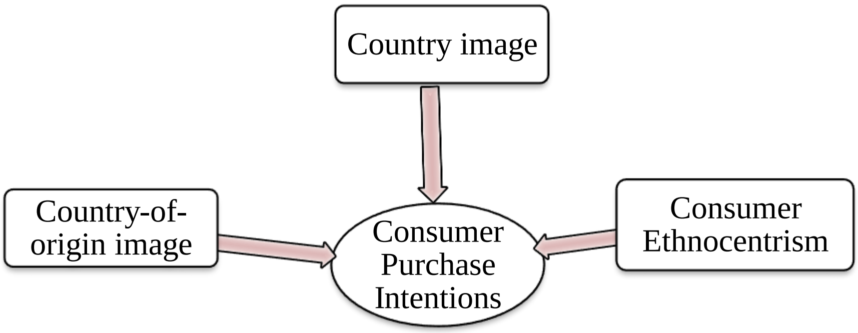
3. Theoretical Framework

The context and focus areas of this research are presented in the theoretical framework displayed in Figure 1. The case of Azerbaijani products entering Hungarian markets enables the creation of the context in which the effect relationship between country image, COI, consumer ethnocentrism, and purchase intentions is observed.

The image of a country, in the formulation of the overall image of a product, is one of the numerous extrinsic cues, such as price and brand name, and consumers are prone to approach brands based on portrayed positive or negative attitudes of a given country (Bilkey–Nes 1982). Products from developing countries are perceived to be riskier in comparison with the products from developed countries. Country image makes a considerable impact on consumer perception of product quality risk, purchase intention, and other similar valuables (Tran et al. 2017).

The image of an individual country may vary depending on the issues discussed. While there are areas where the image of each country is more positive or stronger, there may, however, be more negative perceptions than these. For example, while India is reminiscent of the exotic nature of its historical regions, its rapid development in information technology, and its qualified human power, millions of people living under the poverty line on the other hand also represent a different aspect of India's image. Another study investigated that the culture personality is a special case of country image (Gyulavári–Malota 2018) due to the medley of nations and societies living in countries.

Figure 1 The relationship between country image, COI, consumer ethnocentrism and purchase intentions



Source: Han (1989)

The concept of country image has been the subject of research especially on the basis of the perception of country-specific products. As a result of the increasing extent of globalization and international trade, Country-of-Origin Impact is highly relevant to both academics and practitioners as one of the important factors affecting purchasing decisions of users. In other words, it would be appropriate to consider the image of the country a concept that plays an increasingly important role in every aspect of the development of an overall country, which is related to the preference of commercial products.

One implication of country image is that, consumers tend to evaluate products from developed countries as better than less developed ones and, therefore, according to they use association of brand origin which is considered to be a place, region, or country to where the brand befits from the its target customers' point of view (Kilduff–Tabales 2016). Hence, it can be said that brand origin from a developed country is valued much more in comparison with less-developed or under-developed countries. For example, UK products have exclusively strong link with the United Kingdom which is their brand origin, however, they have less strong associations with an array of countries worldwide from where their parts are sourced (Eng et al. 2016). The impact of country-of-origin will be discussed in more detail in the following chapter.

COI is quite similar to brand image where associations with the origin tend to be meaningful to buyers and, consequently, buyers perceive the origin as a brand. The stereotypical judgments of the consumers concerning the countries and their products are the general perception of the quality of the products and services, and of the people of that country. In other words, consumers, by using COI to establish attitudes towards products, are stereotypically associating product categories with specific origins. Thus, it is absolutely essential to understand the consumer perceptions of the country. (Chattalas et al. 2008).

In the past, country-of-origin has solely been considered as the place where the goods were actually manufactured, but after a lot of research conducted in this field, the definition was later widened by Laroche and his colleagues to include "country of assembly, country of design, the location where the headquarters are situated, country of brand and geographic origin of a product" (Laroche et al. 2003).

According to Hong–Wyer (1989), the effects of country-of-origin on consumer behavior when consumers are given factors as country-of-origin, brand image or price of a product can be observed in two ways which include the Halo Effect and the Summary Construct.

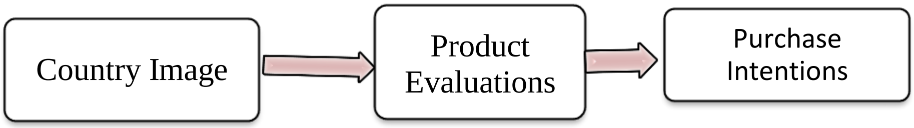
A product's country-of-origin or product-country image influences consumer evaluations of it, thereby, German cars, Japanese electronic devices, French wines are distinguished from others and evaluated differently (Roth–Diamantopoulos 2009). For instance, a variety of products manufactured in China are generally perceived as cheap and low quality and the attitude towards Chinese products is also negative in terms of safety. On the other hand, German products, particularly in the automobile industry, are accepted as being of high quality, durable, and safe which also affects overall perception of consumers in making a decision to purchase products from the country (Kerbouche et al. 2012). According to Han (1989), if the consumers are familiar with product, the COI appears in the form of summary structure, while if there is no familiarity with products, the evaluation emerges in the form of a Halo Effect.

Country-of-origin can affect consumers in three ways: cognitive, emotional, and canonical (Verlegh–Steenkamp 1999). To elaborate, consumers giving preference to products based on the development level of their manufacturing country, and making judgements over their quality show that country-of-origin impact is, to a great extent, highly associated with cognitive comprehension. Additionally, having emotional ties with a certain country, in the framework of country-of-origin, can also impact on consumer opinions about the products in a positive way or adversely in a negative way. For example, a person who is fond of France, will be more likely to possess positive attitudes towards French products while a person who detests France will be prone to avoid French products on accounts of his negative opinion of France (Veselá–Zich 2015).

As a matter of fact, a positive impact of country-of-origin on the evaluation of products is observed for developed countries. However, as Balabanis–Diamantopoulos (2011) state, it is not only because of those countries' economic power, but also their cultural and technological development that play a great role, because it is not only for reasons of economic development that consumers buy French wine, German cars or Italian clothes, there are some other factors as well. Considering halo impact, consumer decisions are formed by their perceptions about countries and when there is insufficient familiarity with products, COI directly affects the purchasing decisions of consumers (Balabanis–Diamantopoulos 2011).

Along with halo-effect, Han (1989) determined another type of cognitive process called Summary Construct which is inextricably linked to Country Image influences. The main difference between the Halo Effect and the Summary Construct is whether the buyers of certain products are aware of or familiar with their country-of-origin or not, and where consumers are informed about the product originating from another country, the Summary Construct model is activated. Summary Construct is more considerable, by having a direct effect on consumer evaluations, and this construct with an array of information consumers possess about country, brands retrieve easily upon assessment (Han 1989).

Figure 2 Halo impact, Adapted by Han (1989)



Source: Han (1989)

Figure 3 Summary Construct, Adapted by Han (1989)

Source: Han (1989)

As shown in Figures 2 and 3, consumers develop their intention to purchase directly or indirectly through the framework of Halo Effect or Summary Construct mechanism owing to the influence of their prior knowledge, attitudes, beliefs, and country image perception about a certain country.

A long history of consumers' reactions to products, and about 1,200 studies on the country and its impact on origin, show that the effect of the country-of-origin is an important issue in international marketing. Examples related to the researches are listed below:

Akira Nagashima's research in 1970 found that the image of the country-of-origin is highly affected by familiarity with a particular country, the ease of buying the country's product, and the stereotypical reputation of that country. Some products representing countries have been seen to affect the overall image of the country. For example; popular products such as Coca Cola, Chevrolet, Ford, IBM, and Sunkist favor and positively affect American goods in a great number of countries such as Japan. On the other hand, Japan's leading products such as Sony, Nikon, Toyota and Honda are positively affecting the image of Japanese products on the American market. In this case, the poor image of Japanese products is improved by customer satisfaction results (Nagashima1970).

Ahmad Zafar and his colleagues conducted a study in 2004 that revealed that the origin country played a role in evaluating low involvement products but had a weak influence. The purchase decision in this product category is of minor importance and, therefore, consumers pay little attention to signs such as country-of-origin. Consumers will prefer popular, frequently preferred, and popular products instead of habits. However, the brand is more important than the country-of-origin in evaluating low-involvement products (Ahmed et al. 2004).

In a study conducted in Lithuania, Urbonavicius and his colleagues identified the automobile as a high involvement product class, revealing that the effect of the country-of-origin on the decision to buy a car was a significant effect (Urbonavičius et al. 2007).

According to Ozturk and Cakir's findings based on the research which was done in 2015, on the basis of various cultural patterns in different countries, the country-of-origin effect varies between cultures. Evaluators and conceptual responses show that individualist cultures value their country's products more positively, while only being competitive. The collectivists, on the other hand, came to the conclusion that they evaluated their country products positively, regardless of the superiority of the product. These results show that in strategies based on country-of-origin, it is necessary to take intercultural information into consideration. In another study they conducted, they investigated how new information was influenced as a determining factor in country-of-origin assessment. This research has shown that motivation, processing goals, and the type of information affect intergenerationally intertwined country-of-origin evaluations (Ozturk–Cakir 2015).

Yunus–Rashid, who were researching Malaysian consumers on Chinese mobile phone preferences, found that "Made in China" products can be seen everywhere in international markets but still have negative effects on consumers. However, it has been observed by researchers that since China became the second largest economy after the United States, this situation has gradually changed. Ultimately, this research concludes that Chinese origin information has a positive effect on the intention of Malaysian consumers to purchase in terms of Chinese brand telephone makes (Yunus–Rashid 2016).

4. Importance of Country-of-origin in the assessment of product from consumer perspective

Country-of-origin information encompasses a trait that is external to the product itself, serving as a surrogate for quality, performance, prestige etc. which cannot be precisely measured. The studies of COO-image, perceived product quality and risk have suggested that different COO-image in terms of a country's economic development influence the perceived product quality and risks (Laroche et al. 2003).

Obviously, by triggering the perceptions of the manufacturing country and general quality of products manufactured in a certain country, the COO-image may have a positive or negative effect on the interpretation of other available information of product attributes and features (Magnusson et al, 2013). Some countries are successful in establishing a unique reputation for specific products. For example, Japan is known for technology especially in cars, cameras, and consumer electronics, France for perfumes, and Switzerland for chocolate (Aichner 2014). Furthermore, Swiss watches, French cosmetics and Argentinean beef are generally considered to be of high quality just because of their origin. Research found that the COO-image works as an information-cue regarding the quality, reliability, dependability, and value for money of the product when more specific information is not readily available (Roth–Romeo 1992). Research further concluded that in international marketing, the association between COO and perceived quality plays the foremost role in evaluating the product image even before the brand name (Magnusson et al. 2013). Similarly, the COO-image effects on a new brand have a similar role to family branding where the COO-image is generalized for the new brand (Ozturk–Cakir 2015). However, research suggested that the association between the COO-image and perceived quality can be moderated by some factors such as price and strong brand name (Aichner 2014).

According to Sharma (2011), COO-effects are also related to perceived risk, which has three dimensions: social, financial, and performance. Consumers have no favorable attitudes and lower intentions to purchase goods produced in countries with high perceived risk in terms of performance (Verleegh–Steenkamp 1999). Research has suggested that consumers perceive social and financial risk related to the products that are manufactured in a given country based on its manufacturing infrastructure, marketing sophistication and level of economic development (Sharma 2011). Consumer perceptions of perceived risk related to the product, together with perceived quality are important as they affect the consumer's choice of buying a product (Magnusson et al. 2013). The COO may be perceived as a risk evaluator, in which consumers perceive greater risk in purchasing products from countries with a poor reputation and image (Chattalas et al. 2008). On the other hand, they may seek to enhance their status by purchasing products from countries with a positive repute and image (Bilkey–Nes 1982). A research study by Sharma (2011) found that consumers in emerging markets show negative perceptions of the quality of products made in other emerging markets coupled with low purchase intentions due to higher perceived risks. Even with the recent boom of manufacturing in China and increased acceptance of Chinese products among consumers in other emerging markets such as India, consumers in Western countries have negative attitudes towards Chinese products due to quality concerns (Hamin et al. 2014).

Safety concerns, food safety issues, and harm associated with food brands are also topics of great importance nowadays. Consumers often construct product evaluations with incomplete information when they evaluate products in a retail stores and, therefore, consumers seem likely to make attribute-related inferences when country-of-origin information is disclosed in such restricted information provision environments (Tran et al. 2017). Berry et al. (2015) propose that a country-of-origin label inevitably activates general perceptions based on the specific country-of-origin, which influences product-related inferences. Specifically, inferences related to food safety, taste, and freshness are highly positive for some countries and less positive for other countries. No wonder these attributes are likely to have a corresponding effect on the purchase intentions, and COO labelling is utilized as a cue in the evaluation of food-related attributes (Tran et al. 2017).

5. Consumer ethnocentrism

The concept of ethnocentrism, introduced by William Graham in 1906, is defined as "individuals tend to view their own group as superior to others", and are reluctant to deny things that are different and admit things that are alike with other groups (Çilingir 2014). According to this concept, people become highly proud of their own ethnic and social groups' norms and symbols while they regard other groups' norms and traditions as not important as theirs. Not surprisingly, ethnocentrism plays an important role in helping social, ethnic groups maintain unity (Abdolvand et al. 2016).

Consumer ethnocentrism is the dimension of ethnocentrism in consumer behavior studies, and it depicts more of consumption habits or feelings of consumers towards consumption (Sharma–Shimp 1995). As the product is the focal point of consumer consumption, the impact of product categories on ethnocentrism level is inevitable. To illustrate, in Russia the level of ethnocentrism in the consumption of fast-moving consumer goods is very high, whilst in more durable and long-lasting products such as electronics the level is very low.

Consumer ethnocentrism is an unavoidable phenomenon, still pervasive in the developing world. Some researchers have proved positive regard towards foreign goods from consumers and some have very severe opposition towards foreign products (Acikdilli et al. 2017). For instance, the demand for foreign goods has always been very high in India and this is due to the fact that Indian consumers seek status symbols, being burdened with inferiority complex, however, when it comes to Turkey, a developing country, consumers, based on the assumption that purchasing foreign product hurts the national economy, are inclined to avoid foreign products (Acikdilli et al. 2017).

Two important concepts; "The influence of the country-of-origin" and "consumer ethnocentrism" are two concepts that are closely related to one another, even though they are handled independently (Çilingir 2014). For the non-ethnocentric consumers, the country-of-origin is based on the ethnocentric relationship of the country-of-origin, accentuating that it is not essential where the foreign products are produced, but preference for those products that are beneficial in terms of other factors. Consumer ethnocentrism is a stimulus that motivates people to favor the products of their own country. The effect of country-of-origin is a factor that may cause individuals to prejudge their perceptions and make decisions in advance. Consumers who are sensitive to the country-of-origin of the products they don't buy, don't have to be ethnocentric, whereas ethnocentric consumers have to be extremely sensitive about the country-of-origin of the products they purchase (Balabanis et al. 2001). According to Thelen, the most important factor for consumers with low ethnocentrism is the price, while the most important factor for consumers with high ethnocentrism is the concept of country-of-origin (Thelen et al. 2006).

The characteristics of consumer ethnocentrism have been described by Sharma et al. (1995) in three different terms:

1. This is due to the loss of control over the interests of the individual against the adverse consequences that one may have for himself and for others through buying imports and imported goods, owing to the excessive love and affection he has for his own country
2. Consumer ethnocentrism involves the importance of the inclination towards the rejection of foreign products. For consumers with a high level of ethnocentrism, buying a foreign product is not only economically relevant but also a moral issue. This moral issue leads consumers to buy domestic products, even if they are inferior in quality compared to foreign products. The purchase of domestic products is better, more accurate, more desirable and more nationalistic - patriotic behaviors, and the purchase of foreign products is intolerable, unwanted, and irresponsible.
3. Although the consumer in a social system thinks that the total level of ethnocentrism is the aggregation of individual tendencies, when viewed at the individual level, this represents a negative attitude towards foreign products.

As a result of research, it has been determined that many factors contribute to the formulation and development of consumer ethnocentrism. Ethnocentric tendencies, at the individual consumer level, are found in much of the individual's socialization experience. The first reason that comes to mind as a means of socialization is the family, and social extensions such as intellectuals, friends, and the mass media which influence the individual's development of ethnocentrism from early childhood. Just as a child accepts the religion that his or her family believes and follows its worship and lifestyle changes accordingly, consumer ethnocentrism begins to develop in the same way in childhood. If the level of ethnocentrism in the family is high, the child will be affected by these behaviors and the behavior against foreign products will be affected for this reason (Erdoğan–Burucuoğlu 2016).

The increase in ethnocentric trends towards indigenous products is increasing, especially in national security issues and economic crises (Lusk et al. 2006). The terrorist attack on the United States on September 11, 2001 caused considerable changes in the country and affected its population. This terrorist attack on American soil has brought a new sense of protection that is unparalleled in the history of the country. That is why Americans have a different perspective on their own country and world. These changes manifest themselves in areas ranging from voluntary work to charitable work and consumption habits. Recent surveys, after these attacks, show that the public is embracing the slogan of "Buy American goods" to an even greater extent now (Kam–Kinder 2007)

The sentiments of nationalism and patriotism are among the factors that trigger consumer ethnocentrism. The survey of Balabanis et al. (2001) on the existence of these elements showed that nationalism and patriotic feelings cause consumer ethnocentrism and these factors significantly affect consumer purchasing preferences for domestic and foreign products, According to the survey sample consumer ethnocentrism of the two countries relations with Turkey and the Czech Republic citizens were examined. While the Turks advocated increasing ethnocentrising as the level of patriotism rose, the Czechs' nationalist sentiment was found to be the most important factor causing ethnocentrism (Siamagka–Balabanis 2015).

All in all, ethnocentrism is a universal issue, and sources of past intercultural consumer behavior show us that the nationalist, patriotic, and ethnocentric tendencies of consumers influence the preference of domestic and foreign products. Positive effects are encountered when these concepts are investigated in countries where production of domestic products is limited in comparison to imported goods. Even if consumers conceive that imported products are necessary, they opt to buy domestic alternative in the interests of contributing to the country's economy, and it clearly results in the decrease of the unemployment rate and raise in the national economic level (Erdoğan–Burucuoğlu 2016).

6. CI, COI influence on decision process.

Consumers' consideration of country-of-origin when buying a product has increased the importance of studies on country-of-origin for businesses that want to evaluate every opportunity to increase their sales. In this regard, some researchers have fallen into disagreement and some have rejected other findings while agreeing that the country's origin has an effect on the purchase decision.

Johansson et al. (1994) carried out a study on the subject of selling a product of a low image country in the US. It was found that the consumer of products of Russian origin agriculture is quite satisfied and affected by the features and thus the effect of the country-of-origin effect on product sales was not observed (Johansson et al. 1994). In another study, the impact of the country's origin on luxury goods was investigated and it was seen that the brand name was the most important factor in the purchase decision, while the country name had little effect on the purchase intention

Lin–Chen (2006) have suggested that the country's origin is a major influence in Taiwanese decisions to purchase insurance and catering services. The authors asked about the country-of-origin, the product information (how much information about the consumer product) and the product dependency (how much thought before considering the consumer product) to provide insurance and catering services to Taiwanese. Correlation analysis in this regard resulted in the following findings: Country image has a positive influence on the decision to purchase original goods; product information has a positive influence on the decision to buy; product loyalty has a positive effect on the purchase decision; the country origin effect of the consumer was not affected by product dependency; the effect of product information level was not influenced by product dependency (Lin–Chen 2006).

Ozturk–Cakir (2015) found that Turkish consumers prefer Turkish products rather than foreign-made products in certain product categories. However, they also show that rich Turkish people are more inclined towards foreign-origin products. Especially, it is observed that Turks aged between 33 and 40 have a stronger tendency to purchase foreign products. Along with this, many families and families with higher education levels are more interested in foreign products (Ozturk–Cakir 2015).

Vendrell-Herrero et al. (2018) have suggested that there is no impact of country of origin for digital products in the UK. There are, in general, studies that argue that the product's country-of-origin has little or no effect on the purchasing decision. Nonetheless, other researchers have been working on proving the influence of COI on purchasing in different product categories and have found different results.

As a result, although some studies claim that the effect of the country origin has no effect on the purchasing decision, many studies have achieved results that support COI influence in purchasing products. This effect has produced different results depending on the consumers and what the products or services received are.

7. Influence of ethnocentrism on decision process

Ethnocentrism is a universal issue, and sources of past intercultural consumer behavior show us that the nationalist, patriotic, and ethnocentric tendencies of consumers influence the preference of domestic and foreign products. Positive effects are encountered when these concepts are investigated in countries where production of domestic products against imported goods is limited. Even though consumers think that imported products are necessary, they prefer to buy a domestic alternative because of the contribution of this to the country's economy. These results clearly show that the unemployment rate and the national economic level raise the ethnocentric level.

Kavak–Gumusoglu (2007) identified ethnocentrism as having a significant influence over the intention to purchase. While Turkish consumers were investigating the effects of ethnocentrism on fast food preferences, they found that consumers with high ethnocentric preference opt for Doner Kebab and consumers with low ethnocentric preference prefer McDonald's (Kavak–Gumusoglu 2007)

Erdoğan–Burucuoğlu (2016) identified the fact that consumer ethnocentrism alone has not had a major effect, and that the relationship between product evaluation and the influence of country-of-origin is also negatively affected. Research has found that among consumers with high ethnocentric and low ethnocentric levels, consumers with high ethnocentric levels are more influenced by products of the local brand than those with low ethnocentric levels in product evaluation.

In the research conducted by Garmatjuk–Parts (2015), it is shown that ethnocentric tendencies are significantly related to demographic characteristics. For example, when the level of education increases, the ethnocentric tendency has been found to decrease. In addition, it has been observed that young people and high-income individuals are less ethnocentric than older and lesser-income consumers, and that males have less ethnocentric behavior than females. Consumers with a high ethnocentric level had a negative attitude towards foreign products but a positive attitude toward domestic products. This study also concluded that people are much more willing to purchase non-domestic skin care products in Estonia and that this was not a matter of irresponsibility or erro on the part of consumers (Garmatjuk–Parts 2015).

Balogh et al. (2016) found that people with high ethnocentric inclination pay attention to where a product is produced. It is also observed that individuals with high ethnocentric inclination have a higher positive attitude that the country where the brand is from will inform the quality of the product. Another result is that people with high ethnocentric inclination respond more positively to the perception that the product quality produced in underdeveloped countries is lower than in other countries. It has been observed that consumer evaluations of domestic and foreign products differ according to ethnocentrism ratios in an established study. It is observed that individuals with high ethnocentric levels have a low level of brand loyalty during product evaluation. As a result of these reasons, consumers have exhibited negative results in purchasing foreign products in order to protect the domestic economy and business power (Balogh et al. 2016).

In Silili–Karunarathna's research, consumer ethnocentrism has been studied as one of the main factors affecting the intention to purchase on many global markets. In a study examining the intentions of Sri Lankan youth to buy domestic brands, it was concluded that the consumer has a positive effect on the intention to buy ethnocentrically, and that the demographic characteristics of the young consumers are higher than those of the low-income young consumers (Silili–Karuharathna 2014).

Mohammad Ali Abdolvand and his colleagues have explored the relationship between consumer awareness, ethnocentrism and loyalty in international brands in an investigation they conducted in Tehran, capital of Iran. In the research, it was concluded that consumer awareness is consumer ethnocentrism and consumer loyalty, and consumer ethnocentrism has an effect on consumer loyalty. In addition, there is a belief that domestic brands are better quality than foreign brands for ethnocentric consumers (Abdolvand et al. 2016).

Awdziej and his colleagues have done a research on local and foreign food product preferences among Polish consumers and they have reached the result that the ethnocentric level of the consumer increases as age progresses. Again, in the study, it has been found that the ethnocentric level of the consumer does not change according to sex, and the ethnocentric level does not decrease as the consumer grows wealthier. In addition to these results, it was observed that Polish consumers were exposed to foreign food products together with local food products (Awdziej et al. 2016)

Chiciudean and his colleagues studied Romanian consumers, the conclusion being that the CETSCALE scale was useful in determining the ethnocentric levels of consumers because it helped to determine the trends in the local product decision-making process (Chiciudea et al. 2015)

This review leads to the following hypotheses:

H1 High level of country image has a positive impact on consumers which lead to high purchase intentions

Similarly, considering positive COI causes more inclination to purchase from those countries, it inevitably has a positive impact on purchase decisions. For this reason, the next hypothesis will be as follows:

H2 High level of COI has a positive impact on consumers leading to high purchase intentions

When we look at the previous definitions and evaluations in evaluating the concept of consumer ethnocentrism, the origin of the products in consumer choice will be effective when we consider the concept of ethnocentrism only as nationalism. If we look at it in a broader sense, we can think of consumer ethnocentrism as being effective in selecting the economic, social, cultural, and social environment of the person. As mentioned before, high level of ethnocentrism leads to low level of willingness to purchase foreign products. Therefore, the next hypothesis will be as the following:

H3 High level of ethnocentrism has a corresponding negative impact on consumer which lead to low purchase intentions.

8. Methodology

8.1. Research Design

Data is gathered through the Internet to give a clear idea about Azerbaijan as a country and its leading companies which could potentially become huge exporters in the food industry. Later, based on the primary data obtained from questionnaire distributed among Hungarians, quantitative method technique is used to understand the brand image of Azerbaijan as a country, the COI of products from Azerbaijan, the ethnocentrism of Hungarian consumers and their relative impact on Hungarian consumer purchase intentions.

In order to evaluate the perception of Hungarians about Azerbaijan, country image scale by Martin and Eroglu (1993) and from Jenes (2012) is used, where the respondents are asked with 21 questions to rate the country based on the Likert scale ranging from one to seven.

Turning to the perception of country-of-origin, the scale from Pisharodi–Parameswaran (1994) was adapted and respondents were asked nine one to seven Likert questions to evaluate their perception of Azerbaijani products. The level of ethnocentrism among Hungarian citizens is analyzed by CETSCALE which was developed by Shimp–Sharma (1995), where the respondents are asked nine questions to rate the questions with 1 to 7 on the Likert scale.

The study is based on quantitative analysis by using regression analysis with 3 independent variables: country image, COI, and ethnocentrism level. The dependent variable of the study is the purchase intentions of the individuals which helps find out whether the independent variables affect the dependent variable in a positive way or not.

The validity of the first hypothesis is investigated by devoting the first part of survey to it which only contains questions about the evaluation of country image. In the second part of the survey, respondents are asked about their perception of the COI of Azerbaijani products. At the final level, respondents give their input based on the questions related to ethnocentrism and they also evaluate the dependent variable to what extent they would be willing to purchase Azerbaijani products.

All the analysis of primary data is made by using SPSS software tools.

8.2. Data Collection

In this research, the survey was distributed on the Internet by using the power of social media networks and the company email server of Deutsche Telekom. All questions were clear, and participants read and responded to each individual question. The sample size stood at 170 respondents. Females constituted 51% of all respondents whilst male respondents were slightly less at 49%. The largest age group partaking in the survey were 25–35-year-old respondents at slightly more than half the population.

9. Data Analysis

Country image scales consisted of 21 items and this scale was adapted from Martin–Eroglu (1993) and Jenes (2012). In order to check the reliability of the scales, Cronbach’s Alpha is applied. The value is 0.867 for Country image, 0.775 is for COI scales, 0.838 is for Ethnocentrism scale, 0.818 is for Purchase intention scale (own developed). The alpha coefficient for all scales js very high, suggesting that the items have relatively high internal consistency as the reliability coefficients are higher than 0.70, which is considered acceptable.

Table 1 Model Summary

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.468[a]	.219	.205	1.32099	2.248

a. Predictors: (Constant), Ethnocentrism, Country_image, Country_of_origin_image
b. Dependent Variable: Purchase_intentions

Source: SPSS output based on a compilation of authors

It is found that the adjusted R^2 of our model is .205 with the $R^2 = .219$. This means that the linear regression explains 21.9% of the variance in the data.

The Durbin-Watson d = 2.248, which is between the two critical values of 1.5 < d < 2.5. Therefore, we can assume that there is no first order linear auto-correlation in our multiple linear regression data.

Table 2 F test

ANOVA[a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	81.285	3	27.095	15.527	.000[b]
	Residual	289.672	166	1.745		
	Total	370.957	169			

a. Dependent Variable: Purchase_intentions
b. Predictors: (Constant), Ethnocentrism, Country_image, Country_of_origin_image

Source: SPSS output based on a compilation of authors

The linear regression's F-test has the null hypothesis that the model explains zero variance in the dependent variable (in other words $R^2 = 0$). The F-test is highly significant, thus we can assume that the model explains a significant amount of the variance in purchase intentions.

Table 3 Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
(Constant)	0.024	0.674		0.036	0.972	–1.307	1.355					
Country_ image	0.715	0.171	0.344	4.177	0.000	0.377	1.053	0.442	0.308	0.286	0.694	1.441
Country_ of_ origin_ image	0.369	0.176	0.173	2.096	0.038	0.021	0.717	0.367	0.161	0.144	0.690	1.448
Ethno-centrism	0.054	0.090	0.042	0.603	0.547	–0.123	0.232	0.082	0.047	0.041	0.990	1.010

a. Dependent Variable: Purchase_intentions

Source: SPSS output based on a compilation of authors

Table 3 shows the multiple linear regression estimates including the intercept and the significance levels. Subsequent analysis was run with the following statistically significant variables: country image and COI. Each variable has tolerance value in order of 0.694 and 0.690 and for both variables over 0.10, VIF value less than 10.

For this regression model F-value was significant for country image (0.00) and COI (0.033). (sig. = .00 <.05) meaning that the model was statistically significant (p= .00). Therefore, both of the independent variables, namely, CI and COI can explain purchase intention.

Regression analysis found that Ethnocentrism was not statistically significant (p-value = 0,547) and it means that we have to reject H3, that a high level of ethnocentrism has a corresponding negative impact on the consumer leading to low purchase intentions or vice versa. There is no identifiable relationship.

According to the regression analysis, the coefficient of determination is 0.344 which means that 34.4 % of the purchase intention can be explained by country image and 17.3% of purchase intention can be explained by the country-of-origin image in this regression model. The following equation for predicting customer satisfaction was obtained:

Purchase intention = $\beta 0$ + 0.re344 (Country Image) + 0.173 (COO) + ε

We accept H1 and H2 hypotheses, which a positive country image has a corresponding positive impact on consumers which leads to high purchase intentions, and positive COI has a corresponding positive impact on consumer which leads to high purchase intentions.

10. Conclusion

The main aim of this research was to explore the impact of the country image of Azerbaijan in Hungary, its products' COI in Hungarians minds, and the effect of Hungarian ethnocentrism on purchase intentions towards Azerbaijani food products. For the research part, the theory was put into the practice with the help of the data obtained through survey. It was identified that the country image of Azerbaijan in Hungary is neither highly negative nor highly positive, simply because of the fact that Hungarians do not know much about Azerbaijan. However, values showed and proved that they are likely to have a positive attitude towards Azerbaijan as a country. As for COI, similar statistics resulted, and it was noticed that Hungarians do not have any objection towards Azerbaijani product COI. These two variables also confirmed, with statistical analysis of dependent variable of purchase intentions, that there is a relationship between positive relationship country image, COI and purchase intentions.

On the other hand, the study notes that this research did not provide statistical evidence for any relationship between ethnocentrism and purchase intention. However, there is, certainly, a large amount of research in the literature both proving the positive relationship and rejecting it.

11. Contribution and implications of the research

The research contributes to the existing information on Hungarian opinion of Azerbaijan and Azerbaijani products. It is an undeniable fact that the literature provides countless pieces of academic researches, articles about CI, COO and their impact on purchase intentions, however, there was not a single piece of information specifically regarding the relationship between Hungarian consumers and Azerbaijan as a COO.

Additionally, this research, by analyzing Hungarians' perceptions of Azerbaijan in a quantitative way, paves the way for Azerbaijani food companies in terms of making their task of market analysis more straightforward. The companies previously mentioned before such as AzGranata, Azersun and Veyseloglu could harness this advantage to make their entry to Hungary. However, one obstacle will also have to be overcome in the shape of intensive and integrated marketing plan to increase awareness of both Azerbajan and its products. In order to create a proper image and position for the country, firstly the current situation should be assessed and the strategy for how to reach the determined target prepared based on the current situation. Otherwise, the goals and targets will be unsuccessful no matter how rational the strategist and the strategic plan prepared. Because any such plans will only be believable and to be relied upon as a sound basis if the current situation is taken into consideration.

12. Limitations and Further research

Due to the target group which had to consist of solely Hungarians, along with time pressure, the sample size of the survey consisted of only 170 subjects. Additionally, the respondents surveyed were largely English-speaking urban people living in Budapest. Therefore, it would not be justifiable to make widespread generalizations about the whole Hungarian population.

Furthermore, the third hypothesis could not be proven owing to the fact that data received from the survey was not significant enough to measure the influence on purchase intentions. Therefore, the research cannot say whether ethnocentrism affects the purchase intentions in a positive or negative way.

Last but not least, there is a great scarcity of respondents in the age group of 55–64, which is represented by only one respondent.

The sample set of the research is not very widespread and consists of English-speaking urban people of Hungary. For further research, it would be advisable to consider the whole of Hungary, and to hear the opinions of people living in other cities and regions.

Moreover, in the sample the age group of 55–64 is merely represented with one person only, and there is no respondent from older age groups. For this reason, it would be better if the further researches focus on older age groups, as well, in order to get more comprehensive and thorough data.

References

Abdolvand, M. A. – Farzin, M. – Asl, A. M. – Sadeghian, M. (2016): The Effect of Consumer Awareness On Consumer Ethnocentrism And Loyalty Regards To The Functions Of International Brands. *International journal of life science and Pharma Research*, Special issue, 2, 102–114.

Acikdilli, G. – Ziemnowicz, C. – Bahhouth, V. (2017): Consumer Ethnocentrism in Turkey: Ours are Better than Theirs. *Journal of International Consumer Marketing*, 1–13.

Ahmed, Z. U. – Johnson, J. P. – Yang, X. – Kheng, F. C. – Sack, T. H. – Chee, B. L. (2004): Does country of origin matter for low-involvement products?. *International Marketing Review*, 21,1, 102–120.

Aichner, T. (2014): Country-of-origin marketing: A list of typical strategies with examples. *Journal of Brand Management*, 21, 1, 81–93.

Awdziej, M. – Wlodarek, D. – Tkaczyk, J. (2016): Are elderly consumer more ethnocentric? Attitudes towards Polish and'foreign'food products. *Journal of Economics & Management*, 23, 91.

Balabanis, G. – Diamantopoulos, A. (2011): Gains and Losses from the Misperception of Brand Origin: The Role of Brand Strength and Country-of-Origin Image. *Journal of International Marketing*, 19, 2, 95–116. doi:10.1509/jimk.19.2.95

Balabanis, G. – Diamantopoulos, A. – Mueller, R. D. – Melewar, T. C. (2001): The impact of nationalism, patriotism and internationalism on consumer ethnocentric tendencies. *Journal of international business studies*, 32, 1, 157–175.

Balogh, P. – Békési, D. – Gorton, M. – Popp, J. – Lengyel, P. (2016): *Consumer willingness to pay for traditional food products*. Food Policy, 61, 176–184.

Beiermann, J. – Ritten, C. J. – Thunström, L. – Ehmke, M. (2017): Measuring the value of information – revealed preferences for country of origin information. *Journal of Behavioral and Experimental Economics*,71, 96-104. doi:10.1016/j.socec.2017.10.002

Berry, C. – Mukherjee, A. – Burton, S. – Howlett, E. (2015): A COOL effect: The direct and indirect impact of country-of-origin disclosures on purchase intentions for retail food products. *Journal of Retailing*, 91, 3, 533–542.

Bilkey, W. J. – Nes, E. (1982): *Country-of-Origin Effects on Product Evaluations*. Journal of International Business Studies, 13, 1, 89–100.

Cai, Y. – Cude, B. – Roger, S. (2004): *Country-of-Origin Effects on Consumers' Willingness To Buy Foreign Products: An Experiment in Consumer Decision Making*. Consumer Interests Annual, Volume 50, 98–105.

Chattalas, M. – Kramer, T. – Takada, H. (2008): *The impact of national stereotypes on the country of origin effect*. International Marketing Review, 25, 1, 54–74.

Chiciudean, D. – Funar, S. – Chiciudean, G. – Muresan, I. (2015): *Study Regarding the Romanian Consumer Ethnocentrism. Bulletin of University of Agricultural Sciences and Veterinary* Medicine Cluj-Napoca. Horticulture, 72, 1, 263–264.

Çilingir, Z. (2014): *Tüketici etnik kökenciliği eğilimi ölçeği (CETSCALE): İstanbul ili tüketicileri üzerine bir pilot araştırma.* Uluslararası İktisadi Ve İdari İncelemeler Dergisi, 13, 13, 209–232.

Elliott, G. R. – Cameron, R. C. (1994): *Consumer perception of product quality and the country-of-origin effect*. S.l.: S.n.

Eng, T. – Ozdemir, S. – Michelson, G. (2016): Brand origin and country of production congruity: Evidence from the UK and China. *Journal of Business Research*, 69, 12, 5703–5711.

Erdoğan, E. – Burucuoğlu, M. (2016): *"The Effect of Consumer Ethocentrism, Cosmopolitanism and Patriotism on Product Evaluations in Turkey"* International Marketing Trend Congress

Garmatjuk, K. – Parts, O. (2015): *Consumer ethnocentrism in Estonian skin care products market.* Procedia-Social and Behavioral Sciences, 213, 610–615.

Gyulavári, T. – Malota, E. (2018): *Do Perceived Culture Personality Traits Lead to a More Favourable Rating of Countries as Tourist Destinations?* Trziste/Market, 30, 1, 77–91.

Hamin, E. M. – Gurran, N. – Emlinger, A. M. (2014): Barriers to municipal climate adaptation: Examples from coastal Massachusetts' smaller cities and towns. *Journal of the American Planning Association*, 80, 2, 110–122.

Han, C. M. (1989): Country Image: Halo or Summary Construct? *Journal of Marketing Research*, 26, 2, 222. doi:10.2307/3172608

Hong, S. – Wyer, J. R. (1989): Effects of Country-of-Origin and Product-Attribute Information on Product Evaluation: An Information Processing Perspective. *Journal of Consumer Research*, 16, 2, 175. doi:10.1086/209206

Jenes, B. (2012): *"Theoretical And Practical Issues In Measuring Country Image"* Dimensions And Measurement Model Of Country Image And Country Brand, 1–288.

Johansson, J. K. – Ronkainen, I. A. – Czinkota, M. R. (1994): Negative country-of-origin effects: The case of the new Russia. *Journal of International Business Studies*, 25, 1, 157–176.

Kam, C. D. – Kinder, D. R. (2007): Terror and ethnocentrism: Foundations of American support for the war on terrorism. *Journal of Politics*, 69, 2, 320–338.

Kavak, B. – Gumusluoglu, L. (2007): *Segmenting food markets the role of ethnocentrism and lifestyle in understanding purchasing intentions*.

Kerbouche, M. – Adouka, L. – Belminoun, A. – Guenouni, H. (2012): The Country of Origin and the Consumer Behavior: How to Improve Chinese Products Brands?. *Mediterranean Journal of Social Sciences*, 3, 11, 551–559.

Kilduff, K. – Tabales, J. M. (2016): *Country Branding and its effect on the consumer in the global market*. Cuadernos De Gestión,17, 1, 83–104.

Laroche, M. – Papadopoulos, N. – Heslop, L. A. – Mourali, M. (2003): *The influence of country image structure on consumer evaluations of foreign products*. International Marketing Review, 22, 1, 96–115. doi:10.1108/02651330510581190

Lin, L. Y. – Chen, C. S. (2006): The influence of the country-of-origin image, product knowledge and product involvement on consumer purchase decisions: an empirical study of insurance and catering services in Taiwan. *Journal of consumer Marketing*, 23, 5, 248–265.

Lusk, J. L. – Brown, J. – Mark, T. – Proseku, I. – Thompson, R. – Welsh, J. (2006): *Consumer behavior, public policy, and country-of-origin labeling*. Review of agricultural economics, 28, 2, 284–292.

Magnusson, P. – Westjohn, S. A. – Semenov, A. V. – Randrianasolo, A. A. – Zdravkovic, S. (2013): The role of cultural intelligence in marketing adaptation and export performance. *Journal of Marketing Research*, 21, 4, 44–61.

Martin, I. M. – Eroglu, S. (1993): Measuring a multi-dimensional construct: Country image. *Journal of Business Research*, 28, 3, 191–210. doi:10.1016/0148-2963(93)90047-s

Nagashima, A. (1970): A Comparison of Japanese and U. S. Attitudes toward Foreign Products. *Journal of Marketing*, 34, 1, 68. doi:10.2307/1250298

Ozturk, R. G. – Çakir, H. Ö. (2015): *The role of the original country concept of the country and the effect of the market choice from the country*: *A research on Turkish consumers*. Global Media Journal: Turkish Edition, 5, 10.

Roth, K. P. – Diamantopoulos, A. (2009): Advancing the country image construct. *Journal of Business Research*, 62, 7, 726–740. doi:10.1016/j.jbusres.2008.05.014

Roth, M. S. – Romeo, J. B. (1992): Matching product catgeory and country image perceptions: A framework for managing country-of-origin effects. *Journal of international business studies*, 23, 3, 477–497.

Parameswaran, R. – Pisharodi, R. M. (1994): Facets of Country of Origin Image: An Empirical Assessment. *Journal of Advertising*, 23, 1, 43–56.

Sharma, P. (2011). Country of origin effects in developed and emerging markets: Exploring the contrasting roles of materialism and value consciousness. *Journal of International Business Studies*, 42, 2, 285–306.

Sharma, S. – Shimp, T. A. – Shin, J. (1995): Consumer ethnocentrism: A test of antecedents and moderators. *Journal of the academy of marketing science*, 23, 1, 26–37.

Siamagka, N. T. – Balabanis, G. (2015): Revisiting consumer ethnocentrism: review, reconceptualization, and empirical testing. *Journal of International Marketing*, 23, 3, 66–86.

Silili, E. P. – Karuharathna, A. C. (2014): Consumer Ethnocentrism: Tendency of Sri Lankan Youngsters. *Global Journal of Emerging Trends in e Business*, Marketing and Consumer Psychology (GJETeMCP), 1, 1, 1–15.

Solomon, M. R. – Bamossy, G. – Askegaard, S. – Hogg, M. (2016): *Consumer behaviour: A European perspective*. Harlow: Pearson.

Thelen, S. – Ford, J. B. – Honeycutt, J. E. D. (2006): The impact of regional affiliation on consumer perceptions of relationships among behavioral constructs. *Journal of Business Research*, 59, 9, 965–973.

Tran, T. T. – Moritaka, M. – Fukuda, S. (2017): Country of Origin, Price Consciousness, and Consumer Innovativeness at Food Service Outlets in Developing Markets: Empirical Evidence from Brands of Imported Beef in Vietnam. *International Journal of Marketing Studies*, 9, 3, 50. doi:10.5539/ijms.v9n3p50

Urbonavičius, S. – Dikčius, V. – Časas, R. (2007): *Automobile country of origin related associations in Lithuanian market*. Transport, 22, 3, 1–6.

Vendrell-Herrero, F. – Gomes, E. – Collinson, S. – Parry, G. – Bustinza, O. F. (2018): *Selling digital services abroad: How do extrinsic attributes influence foreign consumers' purchase intentions?*. International Business Review, 27, 1, 173–185.

Verlegh, P. W. – Steenkamp, J. E. (1999): A review and meta-analysis of country-of-origin research. *Journal of Economic Psychology*, 20, 5, 521–546. doi:10.1016/s0167-4870(99)00023-9

Veselá, J. – Zich, R. (2015): *The Country-of-Origin Effect and its Influence on Consumer's Purchasing Decision*. Acta Universitatis Agriculturae Et Silviculturae Mendelianae Brunensis, 63, 2, 667–673. doi:10.11118/actaun201563020667

Yunus, N. S. – Rashid, W. E. (2016): *The Influence of Country-of-origin on Consumer Purchase Intention: The Mobile Phones Brand from China*. Procedia Economics and Finance, 37, 343–349. doi:10.1016/s2212-5671(16)30135-6

Azerbaijan Exports 1998-2018 | Data | Chart | Calendar | Forecast | News. (n.d.). https://tradingeconomics.com/azerbaijan/exports Accessed: April 08, 2018

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 115–129.

Tactical marketing decisions of managers in times of crisis

Qatar Airways Company, Grounded Theory

Sevinj Omarli

This research study is conducted with the primary aim of exploring tactical marketing decisions of managers in times of crisis; using Qatar Airways Company as a sample. The main purpose of the study is to determine which factors impact on the tactical marketing decisions of managers and what marketing activities are applied in times of crisis. The interviews with marketing information executives were analyzed using grounded theory methodology. Data analysis in the study followed Strauss and Corbin (1990) coding processes. 4500 coding materials were analyzed in two interviews. Results indicated that four major components marketing strategy of a company, marketing informational system (MIS), environmental factors, and internal restrictions and limitations of the company - play important roles in the decision-making process of managers in the crisis period. It was determined that if a company makes tactical promotional decisions such as proposing new campaigns, placing advertisements at local websites and tourism agencies, allocating budget for promoting the website as a sales channel, they will get positive outcomes during crisis periods.

Keywords: Crisis, Tactical decision, Marketing activities, Grounded Theory

1. Introduction

Economic difficulty influences customers to change their buying behavior. Not only customers but also companies make changes in their marketing policy to provide better service and to fulfill the new customer preferences.

In crisis periods, companies make very significant adaptations in their marketing strategy. Crisis situations are an integrable characteristic of corporate behavior in the market economy. Given the need for external environmental factors; internally, managers should set up a plan for anti-crisis corporate activity and adaptation in order to stabilize the process. Rollins et al. (2014) examined B2B companies and determined that they quickly adapt to the crisis, they reallocate their marketing investments, and change marketing strategy.

2. Literature Review

Crisis may have an important impact on company performance, however its impact is not the same for every firm. Some companies consider crisis as an opportunity to strengthen their business and follow an aggressive marketing approach. According to Srinivasan et al. (2005), proactive marketing leads to superior business performance even during recession. Another author, Mirjavadi (2015) determined that the marketing strategy impacted on business performance during the recession in Iran and confirmed the role of proactive marketing in improving market and business performance. Similar results were found that companies modifying their strategies in times of crises can improve their performance.

In the table below, you can see some studies in the literature that pointed to the importance of marketing activities in a crisis period.

Table 1 The Literature on the Marketing Impacts of Crisis.

Subject	Research Topic	Source
General effects of the crisis	The impact factor of crisis on marketing in the United States., field interviews	Rollins et al. (2014)
	Tackling A Crisis Smoothly, synthesis of business practice and academic theories	Kitching et al. (2009)
	The impact of economic downturns on marketing	Rollins et al. (2014)
	The recent literature review of crisis, providing the integrative framework	Latham and Braun (2011)
	Economic crisis impacts on companies, assessment of the effects of marketing strategies on company performance in such conditions, survey, Turkey	Köksal and Özgül (2007)
	Propose the construct of proactive marketing	Srinivasan et al. (2005)
Marketing Budget	Time series analysis of advertising and promotional expenditures on firms' earnings, a differential impact of crisis	Graham and Frankenberger (2011)
	Meta-analysis of extant research on marketing expenditure during recessions.	O'Malley et al. (2011)
	Marketing's contribution to the profitability of Greek enterprises during the economic crisis	Chouliaras et al. (2015)
	The Impact of Proactive Strategies on Market Performance in Economic Downturn: The Case of Hungary	Gyulavari, Kolos (2015)
	Whether firms should spend more on research and development and advertising in recessions., large panel of US firms.	Srinivasan et al. (2011)
	Firm- and industry-level antecedents of advertising spending during economic contractions	Özturan, et al. (2014)

Source: own study

Decreasing the marketing budget is the typical reaction of companies in crisis periods. Retrenching or investing seem to be critical decisions in crisis times. Companies face pressure to decrease marketing expenditures sacrificing future sales and profits (Gyulavari–Kolos 2015). On the other hand, proactive firms have the chance to improve their competitive advantage (O'Malley et al. 2011). Contrary to this, companies can also benefit from the increasing number of Internet users and of different and innovative low-cost online advertising (Quelch–Jocz 2009).

If a manager assertively decides to apply promotion techniques such as coupons, bonuses, free samples etc. in a crisis period, it can positively affect the customer, and achieve value immediately or improve company performance. Other studies suggest that during the crisis, the best strategy for companies is to change distribution policy. In this way, companies can eliminate inefficient channels and intermediaries, and reallocate limited resources (Ang et al. 2000, Köksal–Özgül 2007) When companies change their promotional strategies during and after the crisis, sales income and market share can be increased.

Every company has constant interaction with its surroundings. In this context environmental factors affect companies quite strongly. In the decision-making process, managers should take environmental factors into consideration during the crisis. The company has a mutual and continuous relationship with its surroundings in the open system model. Managerial decisions making and behaviors should be based on the interaction model considering many factors, both internal and external. According to Omarli (2017), the manager's deep knowledge of environmental factors such as competitors, economic crisis, politics and so on are very important to achieve success in companies. Another study conducted by Graham (2004) found 46 strategic marketing decisions from 32 small businesses. The author held six in-depth interviews in designing and implementing a theoretical research framework in the strategic decision-making process. It was found that the external contextual factors are strongly affected by decision-making processes. Companies are directly impacted upon by all types of fluctuations and adjustments in general environmental factors. As examples of external contextual factors, stakeholders, competition, technology, macroeconomic indicators (financial credit. interest rate, inflation), and regulation (political-legal) etc. were cited.

On the other hand, the general environmental factors mentioned above also involve local environmental factors affecting companies and their actions. For example, during economic crisis times, customers are impacted and therefore stop purchasing. As a result, competitors face challenges in finding new customers by attacking each other's market shares, hence competition in the industry increases (Eren 2001). In this case, managers have to consider changing their decisions based on varying environments (Omarli 2017)

In this research tactical marketing decisions of managers in times of crisis were identified; using Qatar Airways Company as a sample. Qatar Airways is one of the youngest global airlines to serve all six continents. It connects more than 160 destinations on the map every day. A lot of companies like Qatar Airways were affected by crisis in 2017. The main purpose of the study is to determine which factors impact on tactical marketing decisions of managers in times of crisis and what marketing activities they apply.

3. Research Methods

Grounded theory is one of the most popular research methods in qualitative studies. It is a method which involves the progressive process of identification and integration of categories of meaning from data. It consists of two part; the process of category identification and integration part is a method and its product is a theory. Grounded theory provides us with instructions on how to identify categories, how to make a connection between categories and how to constitute relationships between them.

Grounded Theory was originally conceptualized by Strauss and Corbin in 1990 and it refers to a general systematic research methodology of data collection and analysis that uses a systematically applied set of methods to generate theory about a substantive area (Glaser 1992).

This research study is conducted with the primary aim of exploring tactical marketing decisions of managers in times of crisis; using Qatar Airways Company as a case study. To this end, two interviews with marketing information executives were analyzed using grounded theory methodology.

4. Grounded Theory (Interviewing)

4.1. Research Question

Which factors impact on tactical marketing decisions of managers and what marketing activities are applied in times of crisis?

4.2. Purpose of Selecting Interviewee

How many participants is enough? According to Seidman (2005), "enough is an interactive reflection of every step of the interview process and different for each study and each researcher". If we consider practical exigencies of time, money, and other resources, one interviewee is enough for learning how to do grounded theory.

The first interviewee was Mr. A. E., one of the marketing information executives of Qatar Airways Company in Azerbaijan. The second was Mr. S.E. who is one of the marketing managers of Qatar Airways Company in Turkey.

4.3. Interview Guide

According to Weiss (1994), "An interview guide is a listing of area to be covered in the interview along with, for each area, listing of topics or questions that together will suggest lines if inquiry". In the unstructured interview method, the researcher comes to the interview with no predefined theoretical framework, and thus no hypotheses and questions. Rather, the researcher has conversations with interviewees and generates questions in response to the interviewees' narration. The interview guide consists of two parts; problem-solving and decision-making.

4.3.1. Questions of Crisis Period

Could you give the relevant information about difficult decisions you have handled during the period of crisis?
Could you talk about issues in the company that needed to be taken into consideration during the crisis period?
Could you share with me what type of information managers need in times of crisis?
Could you give information about tools that help managers in times of crisis?

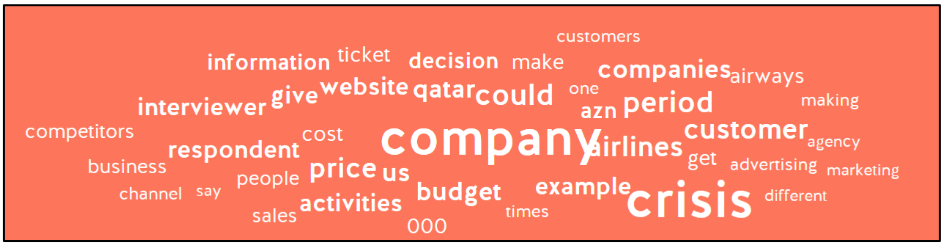
4.3.2. Questions of Making Decision

Please give me a sample of the difficulties you faced when making decisions in the crisis period?
As a manager, did you encounter any situation that changed your plans during the crisis period?
Could you please describe your relationship with employees in the company during the crisis period?
Could you describe your daily mood at the time of decision making, as a manager during the crisis?

5. Data Analysis

Data analysis in Study followed Strauss and Corbin (1990) coding processes. The tables below show frequently used words in the interviews.

Figure 1 Analyzing the most frequently used words in the interviews



Source: Own construction

Table 2 The most frequently used words in the interviews

Frequently used words	Number
Company	57
Crisis Period	40
Customer	37
Budget	35
Price	33
Activity	32
Competitors	25

Source: Own construction

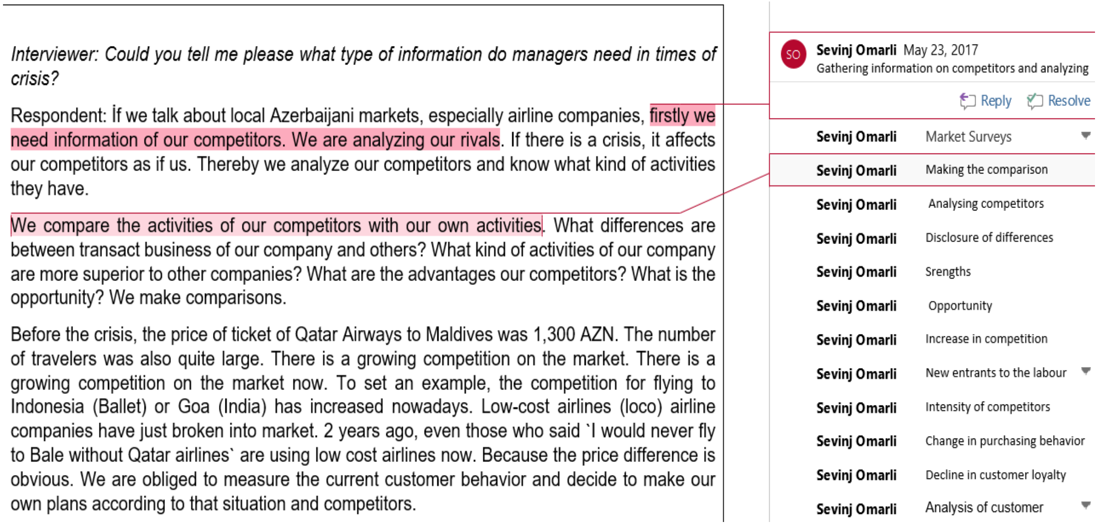
Theoretical coding refers to a systematic process used to make sense of research data by categorizing and grouping similar examples from the data. There are three commonly accepted types of theoretical coding: open/initial coding, axial coding and selective/focused coding (Charmaz 2006, Strauss–Corbin 1990).

Each step of coding enables scholars to break data down further into categories and themes to examine the relationships within the data (Birks–Mills 2012).

5.1. Coding Materials

The initial or open coding is the first step in data analysis. It is a way of identifying important words, or groups of words, in the data and then labeling them accordingly (Birks–Mills 2012).

Figure 2 Sample from the interview material


Interviewer: Could you tell me please what type of information do managers need in times of crisis?

Respondent: İf we talk about local Azerbaijani markets, especially airline companies, firstly we need information of our competitors. We are analyzing our rivals. If there is a crisis, it affects our competitors as if us. Thereby we analyze our competitors and know what kind of activities they have.

We compare the activities of our competitors with our own activities. What differences are between transact business of our company and others? What kind of activities of our company are more superior to other companies? What are the advantages our competitors? What is the opportunity? We make comparisons.

Before the crisis, the price of ticket of Qatar Airways to Maldives was 1,300 AZN. The number of travelers was also quite large. There is a growing competition on the market. There is a growing competition on the market now. To set an example, the competition for flying to Indonesia (Ballet) or Goa (India) has increased nowadays. Low-cost airlines (loco) airline companies have just broken into market. 2 years ago, even those who said `I would never fly to Bale without Qatar airlines` are using low cost airlines now. Because the price difference is obvious. We are obliged to measure the current customer behavior and decide to make our own plans according to that situation and competitors.

SO Sevinj Omarli May 23, 2017
Gathering information on competitors and analyzing
Reply Resolve

Sevinj Omarli Market Surveys
Sevinj Omarli Making the comparison
Sevinj Omarli Analysing competitors
Sevinj Omarli Disclosure of differences
Sevinj Omarli Srengths
Sevinj Omarli Opportunity
Sevinj Omarli Increase in competition
Sevinj Omarli New entrants to the labour
Sevinj Omarli Intensity of competitors
Sevinj Omarli Change in purchasing behavior
Sevinj Omarli Decline in customer loyalty
Sevinj Omarli Analysis of customer

Source: Own construction

Developing categories through the process of intermediate coding will increase the level of conceptual analysis in the developing grounded theory. At this time, the researcher may choose to select a core category that encapsulates and explains the grounded theory as a whole. Developing categories through the process of intermediate coding will increase the level of conceptual analysis in the developing grounded theory. At this time, the researcher may choose to select a core category that encapsulates and explains the grounded theory itself (Birks–Mills 2012).

Table 3 Open and Axial Coding

<table>
<tr><th>Open Coding</th><th colspan="2">Axial Coding</th></tr>
<tr><td rowspan="2">New Competitors to the Labor Market
Intensity of Competitors
Changing Purchase Behavior of Customer
Increasing Competition</td><td colspan="2">Category 1: Characteristics of the crisis period</td></tr>
<tr><td>Subcategory 1.
Consumer Issue</td><td>Subcategory 2.
Competitors Issue</td></tr>
<tr><td>Analyzing Market Surveys
Customer Analyzing</td><td colspan="2">Category 2: Marketing Research</td></tr>
<tr><td>Gathering Information from Competitors</td><td colspan="2">Category 3: Marketing intelligence</td></tr>
<tr><td rowspan="2">Considering Particular Importance to Customer Satisfaction
Creating New Target Audiences
Making Special Discount to the customer
Efforts to Increase the Degree of Loyalty
Trying to Increase the Intention to Buy
Keeping Customer Requests in the Planning Front</td><td colspan="2">Category 4: Tactical Marketing Decision</td></tr>
<tr><td colspan="2">Subcategory 1: Customer oriented activities</td></tr>
<tr><td>Proposed New Projects or Campaign
Placing Ads in Local Website
Placing Ads in Tourism Agency
Promoting Website as a Sales Channel
Using Google AdWords</td><td colspan="2">Subcategory 2: Promotional activities:</td></tr>
<tr><td>Transferring Budget
Effective Use of Budget
Decreasing Advertisement Expenditure
Reduction/ Cancellation of Marketing Activities</td><td colspan="2">Subcategory 3: Efficient using of budget:</td></tr>
<tr><td>Arrange Events with Partner Company
Arranging Business Events with Agencies
Making Joint Marketing Activity</td><td colspan="2">Category 5: Cooperation with Partner Companies:</td></tr>
<tr><td>Decline in Customer Loyalty
Failure to Reach Target
Decrease in Marketing Budget
Reduction in Advertisement Expenditure (Budget)</td><td colspan="2">Category 6: Negative result of crisis period</td></tr>
<tr><td>Increasing Incentives
Strict Business Rules
Low Decision-Making Speed
Slow Procedure
High-Price Ads</td><td colspan="2">Category 7: Challenges of difficulties</td></tr>
<tr><td>Analytical Tools - Adobe Marketing Cloud
Using Different Analytical Tools
Database (MIS)
Increasing Productivity with MIS</td><td colspan="2">Category 8: Informational Tools</td></tr>
</table>

Receiving New Email Newsletters Increasing Money Circulation Doubling the Budget in Crisis Period	Category 9: Achievement of company in crisis period
Decision-Making Becomes Constructive Individual Decision Making Group Decision Making	Category 10: Type of Decision Making
Cost Sharing Win - Win Strategy Making Comparison Disclosure of Differences Strengths Opportunity Conducting Cost Optimization Development of Cheap Sales Channels	Category 11: Marketing Strategy of Company
Restriction for Placing Adv. in the local website Restriction for Cost Issues High-Price Difference with Competitors High Decision-Making Speed of Competitors	Category 12: Restriction/ Limitation of Company
Facilities Evaluation Process Competition Evaluation Process Key Performance Indicator (KPI) Checking Process KPI Evaluation Process	Category13: Internal Evaluation Process
E-Mail Customer Segment Social Media Customer Segment Online Channel Orientation	Category 14: Type of Customers

Source: Own construction

63 coded materials are in the Open Code, 14 categories are (5 subcategories) in the Axial Code.

Selective coding is the process of choosing one category to be the core category and relating all other categories to that category. The essential idea is to develop a single storyline around which everything else is draped. There is a belief that such a core concept always exists.

5.2. Selective category emerged from Axial coded materials

Selective coding begins only after the researcher has identified a potential core variable. Subsequent data collection and coding is delimited to that which is relevant to the emerging conceptual framework (the core and those categories that relate to the core) (Holton 2010).

Table 4 Selective Coding

Selective Coding	Core Category: Tactical Decision of Manager in Crisis Period
Cooperation with Partner Companies Marketing Strategy	Category 1. Marketing Strategy of Company
Marketing Research Informational Tools Marketing intelligence	Category 2. Marketing Informational System
Advantage of Competitors Limited Budget Procedure	Category 3. Restriction and Limitation of Companies
Characteristics of the crisis period Advantage of Competitors Type of Consumer Internal Evaluation	Category 4. Environmental Factor Subcategory 4.1: Internal Factors Subcategory 4.2: External Factors
Negative result of crisis period Challenges of difficulties Achievement of company in crisis period	Category 5: Outcomes Subcategory 5.1: Negative Result Subcategory 5.2: Positive Result

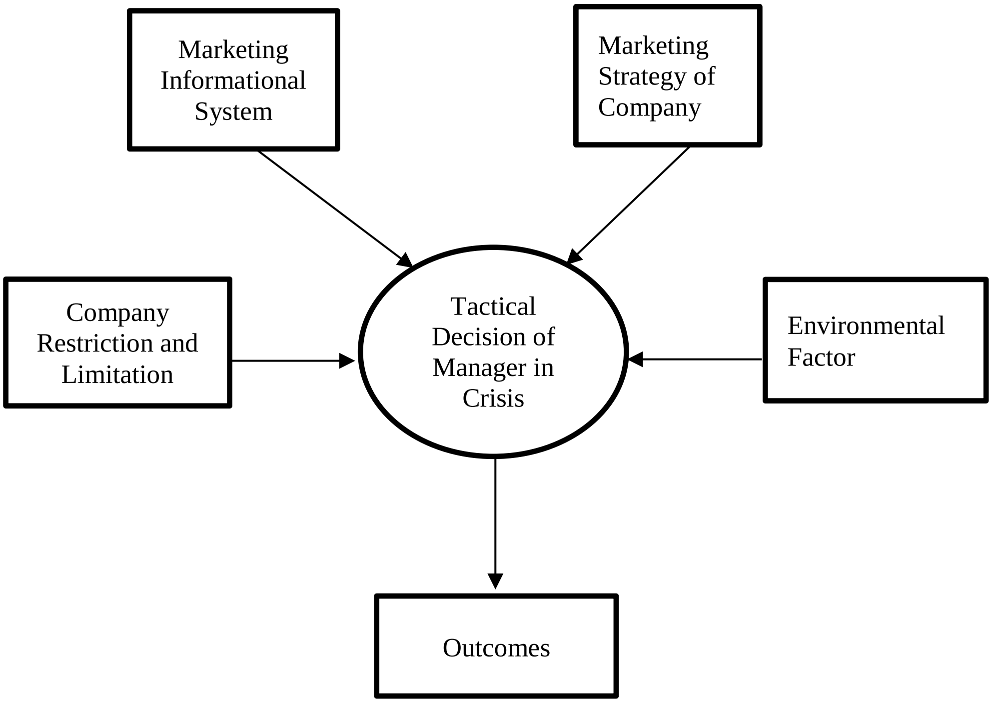
Source: Own construction

6. Findings and Results

The final product of a grounded theory study is an integrated and comprehensive grounded theory that explains a process or scheme associated with a phenomenon.

The main theme is “Tactical Decision of Managers in the crisis period”. The tactical decisions in marketing are related to the implementation of strategic decisions. They are directed towards developing divisional plans, structuring workflows, establishing distribution channels, and acquisition of resources. It is obvious that these decisions are taken at the middle level of management. During economic crisis, four major components are related to tactical decision making in marketing and especially on their marketing activities: These are; Marketing Strategy of Company, Marketing Informational System, Environmental factors, Restriction, and Limitations of Company.

Figure 3 Grounded Theory



Source: Own construction

In the crisis period, the most significant factor is environmental factors. Especially the external environment is constantly changing. Moreover, marketing managers understand the changes and know how to make a tactical marketing decision.

The competition also plays a considerable role in shaping marketing mix, market share, and market demand since the enterprise cannot have full control of this variable. During the crisis, new competitors enter the labor market and competition is increasing and getting more intensive.

> […] *To set an example, the competition for flying to Indonesia (Ballet) or Goa (India) have been increased nowadays. The low-cost airline (loco) companies have just broken into the market. Two years ago, even those who said `I would never fly to Bale without Qatar airlines` are using low-cost airlines now. Because the price difference is obvious. We are obliged to measure the current customer behavior and decide to make our own plans according to this situation and we have to consider our competitors as well.*

Marketing managers should consider internal environmental factors when making a tactical decision. For example, managers have to evaluate business facilities which increased competitiveness or to monitor and evaluate the key performance indicators. Another environmental issue is changes in purchasing behavior of customers. The financial crisis motivated customers to act more economically. In general, during financial crisis customers become more "money minded". They start to buy only necessities, switch to cheaper brands and have a more rational view of promotion.

When managers are making a tactical decision, they arrange some customer orientation activities in crisis periods. The company has considered customer requests on the planning front, giving importance to them, granting special discounts, increasing intention to buy and making efforts to increase the degree of loyalty.

Tactical Marketing Decision is the type most related to marketing strategy. In crisis periods, managers arrange events with a partner company, conduct Joint Marketing Activities, and apply Cost Sharing Strategy, as building a partnership directly with marketing companies proves to be less expensive and more successful than the alternatives from inside of the company.

> […] *I use these data and present it to some companies like "Master card" and suggest that let's do a joint marketing. Let's spend 5000 dollars and arrange the event with cost sharing, then make special discounts for the customer of Master Card. In short, we can provide opportunities which are more efficient by using such analytical information during the crisis periods.*

In the crisis periods, Qatar Airways decided to promote the website as a sales channel. This is also a marketing strategy. In that case, they developed cheap sales channels, saved the budget and performed optimization.

Another important category "Marketing Information System" (MIS) is related to the tactical decision-making process as well. The role of MIS has been described and analyzes the capability of decision-making. According to the coding result, the marketing information system constitutes marketing intelligence: how management can keep up with the new knowledge of competitors and emerging conditions. In our study, it is determined that market research is the process of collecting and analyzing data for the purposes of identifying and resolving problems. It is related to companies' marketing services and marketing opportunities, and planned and managed activities on a scientific basis to ensure efficiency in dealing with these problems and opportunities.

Qatar Airways Company has used certain technical tools such as Adobe Marketing Cloud to analyze the company data to support better tactical decision-making, identify new business opportunities and reduce costs and in this case, the company increased productivity in crisis periods.

> […] *Therefore, you can be more certain in the crisis period. The Adobe Marketing Cloud helps us to measure the effectiveness of our advertising channels. For instance, we placed a bank account on the website (www.hesab.az). It has a given price approximately 500 AZN (Azerbaijan National Currency). We can see the "search visual" by clicking on "create", "purchase create" buttons. Adobe Marketing Cloud informs us that 1,500 people out of 2000 people made search the advertisement which is placed on the hesab.az website by us and 5 people of them display purchasing behavior.*

Company restrictions and limitations influence tactical marketing decisions during economic crisis. In terms of crisis, decision-making is getting more constructive. These are the restrictions the company faced; cost issues, strict rules, high price difference with competitors, and greater decision-making speed of competitors. Moreover, some big companies face difficulties such as increasing incentives, strict business rules, low decision-making speed, slow procedure, and high-priced ads. Thereby it affects tactical decision making. They start to decrease the advertisement expenditures or reduce/cancel marketing activities to use their budget more efficiently.

The last category is "outcomes of a company". Well-planned entrepreneurial activities can increase company sales and profits by overcoming customer retention tendencies during the crisis period.

Some negative outcomes for the company are; failure to reach the target, reduction in the marketing budget, decreasing offline advertisement expenditures, and so on.

> […] *The marketing budget is reduced because your company can`t reach the target, and you have to cut off the budget of advertising or eliminate some marketing activities.*

When companies are recalculating their budgets, they mostly decide to reduce the expenses for advertising and marketing during crisis periods. But specialists in the field of communications advise that advertisement is the most important tool for bringing the biggest volumes of sales in difficult periods, and that companies must take advantage of these opportunities for subsequent periods.

Qatar Airways Company did make some tactical promotional decision. They proposed new projects and campaigns, placed ads in local websites and tourism agencies, allocated money for promoting the website as a sales channel, and so on. In this way, the company got positive outcomes, for example, instituted a new email newsletter, increased money circulation, and doubled the budget during the crisis period.

> […] *We were arranging 12 commercial events during the year. Due to the crisis, we reduced the number of events 3 times. We spent money on the last 9 activities on the website's advertising*

expenditures and promoted our website as a sales channel. When I started to work for this company in 2015, the money circulation of the website was 0.9 million dollars. I decided to make it doubled and it increased to $ 1.8 million dollars. The reason for our achievement is to cancel different activities and put money on different budgets.

7. Conclusion

In periods of economic crisis, customers are affected and therefore they stop buying, while competitors have difficulty finding new customers, attacking each other's market shares as competition in the industry gets hotter (Eren 2001). It has become imperative for managers to make and change decisions according to the ever-changing environments (Omarli 2017).

The competition also plays a considerable role in shaping marketing mix, market share, and market demand, since the enterprise cannot have the full control of this variable. During the crisis, the number of new competitors entering the labor market increases and the market becomes more intensive. In this study we analyzed Qatar Airways Company using certain technical tools such as Adobe Marketing Cloud to analyze company data, in order to support better tactical decision making, identify new business opportunities, and reduce costs, and by doing so, the company actually increased productivity in the crisis period.

We applied a qualitative research method, grounded theory, to analyze our data. In the study, we used Qatar Airways Company as a case study. During the crisis periods, exactly what kind of marketing activities are applied by managers was determined, as were the factors which have an effect on tactical marketing decisions by managers. Two interviews with marketing information executives were analyzed using grounded theory methodology. To conclude, by diminishing the negative influences of the crisis, Qatar Airways Company made marketing budgets more efficient by increasing customers' loyalty and attracting new customers by building a strong, distinct brand image. In the crisis period investigated, Qatar Airways Company also arranged events with a partner company, engaged in certain Joint Marketing Activities, and applied Cost Sharing Strategy to form a partnership with a marketing company which proved less expensive and more successful than the alternatives available within the company.

References

Ang, A.S. – Cole, R. – Lin, A.J.W. (2000). Agency Costs And Ownership Structure. The Journal Of Finance . Vol. LV, No. 1

Birks, M. – Mills, J. (2012): *Grounded Theory: A Practical Guide*. Sage Publication Asia-Pacific Pte Ltd.

Charmaz, K. (2006): *Constructing grounded theory: A practical guide through qualitative research*. London: Sage.

Chouliaras, V. – Gazepis, A. – Kargidis, T. (2015): Marketing's contribution to the profitability of Greek enterprises during the ecenomic crisis. *Procedia Economics and finance*, 19, 217–225.

Eren, E. (2001): *Yönetim ve Organizasyon, Çağdaş ve Küresel Yaklaşımlar*. Ankara: Beta Publishing, 34.

Glaser, B. G. (1992): *Emergence vs forcing: Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.

Graham, J. (2004): How do small business managers make strategic marketing decisions? A model of process. *European Journal of Marketing*, 38, 5–6, 659–674.

Graham, R. C. – Frankenberger, K. D. (2011): The earning effects of marketing communication expenditures during recessions. *Journal of Advertising*, 40, 2, 5–24.

Gyulavari, T. – Kolos, K. (2015): „*The Impact of Proactive Strategies on Market Performance in Economic Downturn: The Case of Hungary. Convergence and Divergence in the New Europe: Marketing Challenges and Issues"*. Volume: Proceedings of the 6th EMAC Regional Conference.

Holton, J. (2010) : The Coding Process and Its Challenges. *Grounded Theory Review, An International Journal*, 9, 1, 21–40.

Kitching, J. – Blackburn, R. – Smallbone, D. –Dixon, S. (2009): *Business strategies and performance during difficult economic conditions*. London, U.K.: Department for Business Innovation and Skills (BIS).

Köksal, M. H. – Özgül, E. (2007): The relationship between marketing strategies and performance in an economic crisis. *Marketing Intelligence & Planning*, 25, 4, 326–342.

Latham, S. – Braun, M. (2011): Economic recessions, strategy, and performance: a synthesis. *Journal of Strategy and Management*, 4, 2, 96–115.

Mirjavadi, S. M. (2015): The survey of proactive marketing strategies impacts on business performance during recession; Study of active companies in Kaveh Industrial Park. *Journal of Social Issues & Humanities*, 3, 1, 235–245.

O'Malley, L. – Story, V. – O'Sullivan, V. (2011): Marketing in a recession: retrench or invest? *Journal of Strategic Marketing*, 19, 3, 285–310.

Omarli, S. (2017): *Which Factors have an Impact on Managerial Decision-Making Process? An Integrated Framework*. Essays in Economics and Business Studies.

Özturan, P. – Özsomer, A. – Pieters, R. (2014): The role of market orientation in advertising spending during economic collapse: The Case of Turkey in 2001. *Journal of Marketing Research*, 51, 2, 139–152.

Quelch, J. A. – Jocz, L. E. (2009): How to market in a downturn. *Harvard business review*, 87, 4, 52–62.

Rollins, M. – Nickell, D. – Ennis, J. (2014): The impact of economic downturns on marketing. *Journal of Business Research*, 67, 1, 2727–2731.

Seidman, I. (2005): Interviewing as Qualitative Research, Third edition, 55.

Srinivasan, R. – Lilien, G.L. – Sridhar, S. (2011): Should Firms Spend More on Research and Development and Advertising During Recessions? *Journal of Marketing*, 75, 49–65.

Srinivasan, R. – Rangaswamy, A. – Lilien, G. L. (2005): Turning adversity into advantage: does proactive marketing during a recession pay off? *International Journal of Research in Marketing*, 22, 109–125.

Strauss, A. L. – Corbin, J. M. (1990): *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.

URL: http://isites.harvard.edu/fs/docs/icb.topic463254.files/October_14/Weiss.pdf Accessed: 11-12-2018

Weiss, R.S. (1994): *Learning from Strangers*. The Art and Methods of Qualitative Interview Studies. The free press. New York, 48–50.

Chapter II

Innovation

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 131–161.

Technological Change in the System of National Accounts – the Effect of New Capitalization Rules

Ilona Ida Balog

The increasing role of technological advancement in a nation's economic performance is also reflected in the recent changes to the System of National Accounts. As a result of the 2008 modifications in this macroeconomic accounting system, the expenses of research and development (R&D) are accounted as accumulation of capital and capitalized on the balance sheets. In the present research paper, I compare the strength of detectable relationships between the growth in economic performance and the different measurements of R&D at the level of nation economies. R&D is taken as stock type data from newly available statistics on the one hand and computed from flow type data on the other. The computation methods are tested with a regression using data from 2009 to 2017 for 15 European countries. The analysis may highlight not only the importance of technology, but also reveal information on the accounting treatment of R&D expense

Keywords: System of National Accounts, Asset capitalization, R&D productivity

1. Introduction

The most important change in the 2008 version of the System of National Accounts (SNA) is the capitalization of research and development (R&D) costs as intangible assets on the balance sheet of a nation's economy (EC et al. 2009). This change has a twofold relevance in the advancement of economic analysis studies. First, it highlights the ever-increasing importance of and attention paid to R&D activities as the engines of growth for modern economies. The new regulation is expected to give more opportunities to analysts to find better articulated relationships in the behaviour of such activities. Second, the capitalization of R&D is part of a lengthy process of extending the asset boundary accounted for in the national accounts. In fact, it seems to be a considerable milestone, since earlier capitalized assets all have reasonably fair and justifiable evaluation methods. In the case of R&D, however, the outcome of the investment expense is very uncertain both in magnitude and occurrence, if it happens at all. Under such circumstances, the evaluation and capitalization of an intangible asset requires more sophisticated approaches and methods on the part of economists and statisticians.

Is the capitalization of R&D in the national accounts of real help to economic analysis? The main question of this study is whether estimation results can be improved using the new stock type statistical data available due to this capitalization, instead of those calculated from flow type data in the old-fashioned way. Since the most important research question in the field of R&D from a macroeconomic point of view is the productivity growth achieved by R&D investments (Hall et al. 2009), I test the estimated values of R&D assets on an equation, which aims to find an economy's elasticity of overall productivity on the change of R&D assets.

In the second chapter I outline the asset capitalization problem, and the extension of the asset boundary from both the economist's and the statistician's point of view. In the third chapter I review the literature analysing the productivity of R&D at the macroeconomic level. Emphasis will be given to the choice and calculation of the R&D input, since that is the variable, which can be changed in the models to available statistical data. In the fourth chapter I prepare R&D productivity estimations with the help of a regression of 15 European countries with available data for the period between 2009–2017. The first two estimations will be performed with newly available statistical data of R&D assets, while the third and fourth estimation is prepared with R&D stock data calculated from the flow type R&D capital formulation data from the same data set in the way specified in earlier studies. Finally, the statistical and economic relevance of the two estimations will be compared.

2. Extension of the asset boundary

Initially macroeconomic models identified two distinct resources as factors of production and economic activity, labor and capital (Lichtenberg 1992). The interpretation of these factors was rather simple in taking into account only the tangible aspect of activities, and dividing it into a human related part in the one hand (labor) and a tool and structure related part on the other (capital). Both factors are essential for economic activity, still they behave in very different ways.

As for the tool and structure part, it was regarded as useful in the long-term in activities. These resources owned and operated by individuals and organizations and used over a longer period of time have to be planned and managed in order to achieve reasonable returns. Planning and management require measurements of value on the basis of expected future performance. Such measurements are then registered in regular points of time in a document listing the resources and their value to show the wealth, that is, the potential to produce useful goods and services for every economic unit. This document is generally called the balance sheet, where the use and the changes in the value of listed resources called assets, can be followed. Including an asset into this list is consequently called capitalization, signaling that the item is regarded as capital, or physical capital in a narrower sense (Lequiller-Blades 2014). Therefore, it was inevitable from the beginning that tools and structures have to be evaluated and accounted for as economic assets on balance sheets, that is, they have to be capitalized.

The human related part of production factors, on the other hand, proved to be much more difficult to evaluate. It is hard to decide whether these resources are utilized in a longer or a shorter run, while their productivity is highly uncertain. Both evaluation and management seemed to be almost impossible at first sight, because in this case we do not know the length of the utilization period, and also the performance of future periods can vary greatly. Moreover, it is not possible to put them onto a balance sheet, because they are inseparable from human beings, which in turn cannot be owned by economic units under the laws of modern societies respecting human rights (Lequiller-Blades 2014). The behavior of a human being will always carry more uncertainties than that of tools, machines and structures, therefore it was prudent to handle resources related to them separately, while avoiding their evaluation and capitalization in a balance sheet is also understandable.

Intangible drivers of growth and development like technological change were regarded to be exogeneous to these first-generation models (Lichtenberg 1992). This means that due to difficulties of interpretation and evaluation, economic accounting did not engage in predicting their future effect and regarded them to be out of the scope of economic calculations. This approach is understandable to a certain extent, as economists may not want to base calculations on overly unpredictable phenomena. However, the need of societies changed this situation.

Economists identified certain intangible factors of production included in those termed residual or technological change, and of which it was possible to find out at least something, even if punctual calculations were not feasible (Griliches 1998). Among these, human capital was the first, defined as the knowledge and experience of working people, which can be estimated by the means of measuring education activities (Lichtenberg, 1992). The human capital theory, building on the similarities of knowledge and physical capital, started to show the way ahead to including more types of resources in macroeconomic models and including them in calculations like endogenous factors. However, uncertainty still remained in terms of measurements, utilization period and productivity, and therefore evaluation and asset capitalization did not seem to make sense.

Likewise, a wide range of other types of intangible capital were researched, however measurement problems remained, therefore including them in capital formation and their capitalization was viewed as highly problematic (Vanoli 2005). Though the preface of the 1968 SNA prompted fast development in the modification of SNA recommendations to include balance sheets in the system, and extend the boundary of the included circle of assets (UNITED NATIONS 1968), these developments came along only partially in the 1993 modification (Vanoli 2005).

The most suitable intangible assets to be included in the SNA balance sheets and capital formation were those related to technological change. These, including R&D values, were widely researched in the preceding decades. The relationship between productivity and R&D received special attention, though mainly at the micro level (Lichtenberg 1992). This focus was understandable, if we take into account that most of the growth and the events of economies cannot be explained by the use of traditional types of resources, only by the ambiguous "residual" (Griliches 1998). On the basis of research conducted by Griliches and others, R&D expenses and assets seemed to play a major role in inducing economic growth. However, the quality of data available for such research was poor in many respects, especially in the case of R&D assets, although not only in their case.

According to the above proceedings, it was expected that in the course of the 1993 modifications of the SNA, R&D expenses would be included in capital formation and capitalized on the balance sheets, which were officially incorporated into the account system. Unfortunately, this was not to happen at that time. Debates started around the evaluation accuracy of human capital and finally its capitalization was rejected. Confidence wavered in accounting for intangibles altogether (including intangible human capital, R&D and other intangibles) and finally the experts among national accountants decided not to capitalize R&D expenses, either (Vanoli 2005). Nevertheless, some progress was made as a small group of intangibles, including software, artworks and expenses on mineral exploration were adopted as capital expenses, and therefore included on the balance sheets.

This stumbling in the process of extension of the asset boundary was corrected in the next round of modifications to the SNA. In 2008, the most important change was that capitalization of R&D assets became recommended, meaning they were now regarded as long term intentional investments in future performance. This change also brought about the change in the balance sheet that complete R&D results are now accounted for as produced assets, thus fully integrated into the logical sequence of the SNA similarly to produced tangible assets. With this step it is acknowledged also in statistics that technological change is not an exogeneous factor to the economy, as it was anticipated by the early macroeconomic models, but that technological development can to a certain extent be managed through the control of the R&D expenses which fuel it. The changes were consequently recommended in ESA 2010 (European System of Accounts), the corresponding set of rules for the statistics of the European Union (EC et al 2009, EUROSTAT 2014/a).

The need for further extensions of the asset boundary, that is, including further asset types in the balance sheet and capital formation is obvious from the perspective of economists, as stock type data are consistently used in economic analyses as variables. These variables, however, have been mostly calculated from the available statistics by simplified methods according to the limited possibilities of economists using them (e.g. calculations of Guellec and van Pottelsberghe 2001). Perhaps, economic analyses would be more informative, if the calculation of the values of R&D assets is done by statisticians applying a wider information basis. If we wish to find explanations for the changes in productivity in economies, we should seize any opportunity to calculate the value of intangibles more accurately. The importance of innovations, knowledge, communication, and many other intangibles in productivity growth has been inevitable since at least the 1980's (Vanoli 2005).

However, there are fundamental problems to asset evaluation in the case of intangible assets. By general definition, the value of an asset can be calculated at best with its future contribution to economic wealth. This value can be estimated by the markets as the average evaluation of market players or can be estimated by economists applying for example net present value calculations (Lequiller-Blades 2014). Mostly, market prices are considered to be closer to the true value, though in their absence, at cost calculations are thought to be the second-best solution for asset evaluation (EUROSTAT 2014/b). At present, uncertainty is high in every aspect of asset evaluation when applied to intangibles. The problems are listed in the next three paragraphs.

First of all, we cannot be sure of what we are measuring by economic wealth or economic growth. Some assets contribute to the well-being of people, though this performance is not measurable, and if it is not paid in money, is not regarded as economic (Hall et al. 2009). On the other hand, some measurable contributions do not really serve true development of a society, therefore measured productivity may not be in line with the development of a country or region. However, these uncertainties are equally true for tangible assets, so it should not affect the extension of the asset boundary to include more intangibles.

Second, it is often cited, especially in the case of R&D, that results from such expenses occur in a random way, therefore the positive relationship between R&D investment and productivity cannot be justified (Vanoli 2005). This argument can be countered with the statistical success of such investments. At the macroeconomic level R&D expenses are very likely to produce good results even if lots of individual projects fail. It is also worth noting that positive relationship between R&D and productivity is definitely revealed even at the micro level using only poor-quality data (Griliches 1998).

Third, the measurement of consumption of R&D is uncertain, as well. Intangibles do not have depreciation as defined in the case of tangible assets. The amortization of intangibles comes from obsolescence instead of planned wearing out. Again, this is also the case with many tangible assets (Vanoli 2005), therefore this should not be an obstacle in the way of expanding the asset boundary.

Intangible assets are difficult to mobilize, many of them do not have a market at all. In the absence of market prices, asset evaluation cannot be regarded as punctual. The secondbest estimation of evaluating them at cost often results in lower values than for those assets valued by markets. In the case of R&D evaluation, this phenomenon often occurs in connection with government investments (Griliches 1998). In case of self-production or common consumption, the additional value given by the market to the asset is missing. Though this argument is valid indeed, it is a common feature with some of the tangible assets.

The problems of asset evaluation are inevitably present in the case of R&D assets. Still, they are profoundly different from human capital in that they are separable from the human being and belong to economic units. Obsolescence can be traced, and therefore their evaluation is not much more difficult than that of tangible assets. In fact, even investment in the most conventional tangible assets has always involved some risks (Vanoli 2005).

Taking the above into account, including R&D assets on the balance sheet and in capital formation seems justifiable. It had been demanded by economists researching the topic for a long time, and the need for possible explanations of changes in economic performance will necessitate more research towards extending the asset boundary even further. Still, asset evaluation problems warn us to be cautious (Räth 2016). If the value calculated for intangible assets does not represent their real value, calculations done with them may secede from reality, and economists may once again end up being unable to provide valid explanations for the events of economies. This risk, however, has always been present in economic research. The economy cannot be described by the old models of easily interpretable tangibles any more. While these models may still be valid, their explanatory strength has certainly lessened (Griliches 1998), and therefore measuring new assets and new relationships is necessary, even if the uncertainty in using them is ever increasing.

3. Analyses on the effectiveness of R&D activities at the macroeconomic level

Research on the effect of R&D on economic performance started almost simultaneously with the creation of macroeconomic models in the 1950-ies (Griliches 1998). The basic model interpreted the residual in economic growth not accounted for when calculating the effect of conventional factors of production, as the effect of technological change. Taking this as a starting point, it was simply natural to continue thinking in the direction of finding a relationship between R&D and productivity. However, this research took place mainly at the micro level.

The macro level studies were conducted either as cross-country comparisons or temporal case studies. The first combined estimation was done by Lichtenberg only in 1992 (Lichtenberg 1992, Hall et al. 2009). The basis of research in most cases was the traditional Cobb-Douglas production function augmented with research and development as an additional type of capital. In older studies, R&D was examined together with other capital types, most importantly with conventional physical capital, but sometimes with human capital, as well. However, even R&D capital was further divided into different parts already in earlier research, as their effect on productivity was expected to be different. As for the combined effect of all R&D investments it was remarked that the effect is positive and significant (Hall et al. 2009).

The positive and significant effect of R&D on productivity was measured through very different models and approaches. The dependent variable (productivity), the independent variable of interest (R&D), and the control variables were constructed in different ways. It is remarkable that in all cases the result was a statistically significant and positive coefficient (Hall et al. 2009).

The dependent variable in these equations is a performance measurement, which can be either a productivity ratio (level estimation) or its growth rate. A productivity ratio divides a performance indicator by the amount of resources used up to produce it. The performance indicator in macroeconomic studies is most often the GDP or some other aggregated income figure (sometimes called output), while the dividing amount of resource can be one or more of the factors of production accounted for in the model. The total factor productivity, for example, takes into account both capital and labor when dividing GDP or value added, but in many cases, the number of working or working-aged people, perhaps even population size, represent the amount of resources.

The independent variable of interest, which in this case represent the R&D activity, can be a stock type amount showing the accumulated value of R&D assets, or a flow type amount as expenses or investments. In many cases the independent variable of interest is a ratio itself, dividing the R&D amount by economic performance or total investments. In these cases, the variable of interest is the share of R&D within the total of activities. The growth rate of R&D amounts also can be regarded as an independent variable. Regarding the long term and uncertain nature of R&D investments, it is worth including some lags into the regression. Often the R&D figures are those of some previous years or computed as the average of earlier years compared to the dependent variable.

The distinction between stock and flow type measurements deserve special attention. Most of the studies apply a stock type R&D amount (Guellec-van Pottelsberghe 2001, Chandra et al. 2018). The rationale behind this may be that in case of long-term assets the expenditures will have an effect in more future years, therefore probably would not give a good estimation result in a regression with one dependent variable. However, stock type data is more difficult to access. In case of R&D, it is hardly observable directly. If there are no statistics available, economists can compute an R&D stock figure from flow type data for example with the help of the perpetual inventory method, which is also widely used (Lequiller-Blades 2014, Guellec-van Pottelsberghe 2001). The problem here is that for this method we need to know the amount of investment and depreciation for all the years. Investments as flow type data are mostly available, though we may have serious problems with depreciation rates. If no better estimation is available, it is possible to use a fixed depreciation rate. However, it has to be remarked that depreciation in the case of intangible assets like R&D is obsolescence and called amortization (Vanoli 2005). Amortization means devaluations, which do not happen in equal portions in a timely manner as in the case of planned depreciation of physical assets, therefore fixed depreciation rates in case of intangible assets certainly would distort the results. This is due to a fundamental difference between tangible and intangible assets, that tangible assets do lose further utilization capacity when used, while intangibles remain equally useful for an indeterminable length of time. If they lose value, it occurs as obsolescence and not as a result of utilization, therefore it would be inappropriate to relate their service to their loss in value within a time period.

Regressions of productivity and R&D may include a series of control variables. Mainly, these are country specific variables of size or economic situation. In case of panel regressions, it is also possible to include time lag variables of the dependent variable.

In the literature, relatively few studies deal with cross-country and temporal comparisons together (Hall et al. 2009). Measurement problems and the poor availability of data may have discouraged this type of research for a long time. With the 2008 modification of the SNA, comparable stock type R&D statistical data is now available for many countries, therefore panel research will perhaps be easier in the future. Anticipating this, it is worth outlining previous research done in this specific direction.

First, Lichtenberg made a comparison between data from 74 countries all over the world in 1992. He used real GDP/working age population as an overall productivity ratio for the dependent variable of the regression calculated for the year 1985. The independent variable of interest was the share of nominal R&D investments in nominal GNP in a 25-year average between 1964 and 1989. The author examined several other factors of production, among them human capital and physical capital, and controlled for the population growth rate. A fixed depreciation rate of 0.03 was applied. The estimation was done also for the growth rate dependent variable version. The results of the non-linear least squares regression showed that both R&D elasticity and production function parameters were in positive relationship with productivity over this longer period, and their significance and strength was equally high as those of physical capital (Lichtenberg 1992).

Guellec and van Pottelsberghe prepared a panel regression for 16 OECD countries for the years 1980–1998 (Guellec and van Pottelsberghe 2001). Their dependent variable was the multi factor productivity of the industrial sector. For independent variables they used R&D stock figures calculated by the perpetual inventory method with a fixed depreciation rate of 0.15. They were interested in the coefficients of the domestic business R&D stock, the foreign capital stock and the public R&D stock, but did not calculate the strength of relationship between the overall R&D stock and productivity. R&D stocks were taken into the regression equation with a two-year lag, while their growth with a one-year lag. Productivity growth was also included with one-year lag and productivity level with a two-year lag. For control variables a business cycle effect was included and a dummy variable for Germany signaling the years before and after unification. The model was estimated with 3 stage least squares and seemingly unrelated regression estimation methods. The results of regressions showed positive elasticity for all the three types of R&D stocks.

All these results reinforce the notion that the relationship of R&D and productivity in an economy is positive. Time lags in the independent variables should be included, though there are different methods of doing this. However, the use of stock type R&D data seems to be necessary (Hall et al. 2009), even if they can be calculated only with serious distortions. The main reason for this is that productivity may be boosted by research and development costs spent in many earlier years, and therefore the accumulated value of R&D seems appropriate for estimations.

4. Comparison of R&D net asset statistics with R&D stock figures calculated from flow type data

In this paper I examine the results of a regression equation, where the dependent variable is productivity at the macro level, and the independent variable of interest is R&D. My aim is primarily to compare the performance of newly available data with the results produced by more usual data processing methods. For an accurate estimation of the effect of R&D expenses on productivity growth, longer time data is needed, as it is still not possible to fully characterize their operating pattern in the few years that have elapsed since the 2008 financial crisis.

4.1. Data and Methods

I use data of the national accounts compiled according to the rules of SNA 2008 and ESA 2010. This set of data already contains R&D expenses as capital expenditures, and R&D stock data (net assets) is also provided. The same regression is done for the net assets data given in the statistics, and for net assets data calculated by the perpetual inventory method from available R&D expenses figures. The latter method is widely used in econometric papers and statistics (Lequiller-Blades 2014). My aim is to find out whether the new rules of statistical data compilation can improve the results of regressions compared to earlier calculations.

It is not usual to examine periods shorter than 12 years in a study of R&D effects. This can be especially problematic at the country level, since effects are not likely to be observable at higher levels of aggregation, where productivity is influenced by a multitude of factors. My reason for trying to estimate a regression with macroeconomic data for only 9 years is that I was curious, whether the newly available stock type R&D statistics can produce any better estimations than the previous data under these hard conditions. The regression I apply is an OLS regression as the length of the data time series is not adequate for a panel regression. I use the GDP and R&D figures of 15 European countries for the years 2009–2017. All my data are derived from the easily available national accounts statistics of Eurostat (EUROSTAT 2019).

In my present study, the dependent variable is the growth of GDP per capita directly taken from Eurostat statistics. The independent variables of interest are different aggregated R&D data given in the national account statistics. The control variables are also taken from statistics and include a time trend and controls for size and country specifics. Undoubtedly this regression is suitable to reveal less detailed information that way than panel regressions used in previous studies (e.g. Guellec and van Pottelsberghe 2001). Still, my aim does not go further than to find results for a shorter period with the help of more accurate data.

Based upon the above principles the first estimated equation with R&D net asset data taken directly from Eurostat statistics is:

$$GrGDPc = const + \alpha_1\, GrNA\_1 + \alpha_2\, GrNA\_2 + \alpha_3\, GrNA\_3 + \beta\, LNA\_2 + \gamma_i\, X_i + e \quad (1)$$

Here the dependent variable *GrGDPc* is the growth rate of GDP/capita between two subsequent years representing the change in productivity. The regressors are the following:

- *const*: constant
- *GrNA_1*, *GrNA_2*, *GrNA_3*: the growth rates of R&D net assets statistics between two subsequent years in one, two, and three-year lags, respectively (net asset data taken directly from statistics)
- *LNA_2*: the log of R&D net assets statistics in two years lag (net asset data taken directly from statistics)

- X_i*:* four control variables: a time trend, a country index, the log of population size and the log of GDP/capita in one-year lag.

The variables of interest are *GrNA_1*, *GrNA_2*, *GrNA_3* and *LNA_2* as they show the effects of R&D on productivity. The control variables control for the time, the size, the economic development level, and other specifics of the countries.

All of these regressions were conducted including a time trend for addressing trend stationarity. Difference stationarity, however, was not handled in the first regression. The reason for this is that in the case of the relationship between productivity and R&D, it is possible to have factors which affect productivity and the value of R&D in a different way, disturbing their co-movement, but still do not influence the underlying basic connection between them. Here, for example, I would point to those research results, which were not implemented in practice within the first years of existence. Though it is supposed that they are or would be useful, their application for the time can be delayed. In these cases, their productivity effect is not discernible. Nevertheless, they do not wear off due to this delay. Productivity may drop or stagnate while R&D values are relatively high or even rising, still at an overall scale we should not accept that this means R&D is ineffectual or not in significant relationship with productivity. Their true effect in my opinion can be grasped exactly in those non-stationary co-movements, which are eliminated by difference stationarity treatment.

Despite this concern, in the second regression all variables were tested for stationarity and treated by differencing in order to achieve difference stationarity, as well. The corresponding unit root tests are included in Appendix 1, where the names of variables contain the letter "d" for differencing. Both regressions were calculated with robust (HAC) standard errors to address heteroskedasticity and autocorrelation.

The coefficient of *LNA_2* gives the component of productivity raising effects, which is exercised by existing and accumulated stock type R&D assets. In other words, *LNA_2* represents the accumulated intellectual capital, which is assumed to be effective in a two-year lag on productivity (e.g. also in Guellec and Pottelsberghe 2001). Since effects later on are also likely as the R&D assets do not wear off with utilization, it is acceptable to make calculations using the stock type figure. Valuable assets remain on the balance sheet and continue to have effects on productivity. The coefficients of *GrNA_1, GrNA_2, GrNA_3* show the productivity effect of changes in R&D assets, that is, the flow type investments in time lags. This can be interpreted as an elasticity measurement or the effect of newly-created R&D capital on productivity. It should be noted that depreciation or utilization of R&D assets calculations were done by statisticians in this case. The applied stock type data were compiled with a very detailed version of the perpetual inventory method. This method is also used by statisticians, though they have more detailed information, and do more punctual estimations than economists can do (Lequiller-Blades 2014).

The third estimated equation works with R&D stock data calculated by the perpetual inventory method from R&D capital formation statistics by the analyst, as specified by Guellec and Pottelsberghe. This method estimates R&D stock from the actual R&D expenditures and an initial value of R&D expenses with the following formula:

$$pNA = CF/(1-1/(1+(CF/CF_0)^{1/n}(1-\delta))) \quad (2)$$

In this formula *CF* is the figure of R&D capital formation in a given year, while CF_0 is the R&D capital formation of a selected initial year. The letter *n* indicates the number of years spanning the time length of calculations, finally δ is for the depreciation rate. Depreciation rate was set as 0.15 by Guellec and van Pottelsberghe in their 2001 study of R&D effect on productivity. They also concluded that their model was not sensitive to the change in the depreciation rate (Guellec-van Pottelsberghe 2001). I applied the depreciation rate given in this basis study. As data for comparison with R&D, net asset statistics is available for 7 years I set $n = 7$. CF_0 is the R&D capital formation statistic figure of 2009.

The third regression equation is put together partly from the same statistics as the first, only R&D figures are replaced by those calculated with the formula above:

$$GrGDPc=const+\alpha_1\ GrpNA\_1+\alpha_2\ GrpNA\_2+\alpha_3\ GrpNA\_3+\beta\ LpNA\_2+\gamma_i\ X_i+e \quad (3)$$

Here the dependent variable is GrGDPc again, which is the growth rate of GDP/capita between two years. The independent variables are the followings:

- *const*: constant
- *GrpNA_1, GrpNA_2, GrpNA_3*: the growth rates of R&D stocks between two subsequent years in one, two and three-year lags respectively (net asset data calculated from flow data with the perpetual inventory method)
- *LpNA_2*: the log of R&D stock in two-year lag (net asset data calculated from flow data with the perpetual inventory method)
- X_i: four control variables: a time trend, the log of population size, the log of GDP/capita with a one-year lag and a country index.

The variables of interest were changed in this third equation to *GrpNA_1, GrpNA_2, GrpNA_3* and *LpNA_2,* as they are calculated in a different way, although their meaning and interpretation remained the same. All the other variables remained unchanged.

Depreciation rate enter into this estimation as a parameter used in calculating the variables *GrpNA_1, GrpNA_2, GrpNA_3,* and *LpNA_2*. It is set as fixed all over the period and for all countries in this calculation model, since detailed information of its true value was not available. Regarding the special features of intangible assets, among them R&D stocks, counting with fixed depreciation rates can cause serious distortions in the results, as losing value in the case of R&D assets is anything but fixed in time and space. However, in case of longer time series, the approximation provided by the perpetual inventory method may be appropriate in case of missing data on R&D assets.

The fourth regression was also done with varibales calculated by the perpetual inventory method, but this time all the variables were differenced in order to address difference stationarity, similarly to the second regression. The third and fourth regressions were estimated with robust (HAC) standard errors in order to address autocorrelation and heteroskedasticity, similarly to the first two estimations.

4.2. The effect of R&D on productivity calculating with R&D net asset statistics

The results of the first regression calculation (equation (1) without differencing treatment) are summarized in *Table 1* as follows.

Table 1 Regression result of measuring the elasticity of productivity on R&D investment with R&D net asset statistics, European countries, 2009–2017

	Coefficient	***Std. Error***	***p-value***
Dependent variable: GrGDPc			
GrNA_1	0.113	0.064	0.1002
GrNA_2	0.087	0.031	0.0151
GrNA_3	– 0.005	0.069	0.9389
LNA_2	0.058	0.025	0.0362
R^2	0.587	*Adjusted* R^2	0.537
Durbin-Watson:		1.04	

Source: own construction based on Eurostat data

In the first case, calculating with R&D net asset statistics without difference stationarity treatment, the coefficient of *LNA_2* is significant at the 5% significance level and positive. The result for *GrNA_2* is also statistically significant at the 5% level and positive. This indicates that using data directly from statistics and the usage of stock type figures for estimating the effect of R&D activity on productivity is justified. The level of accumulated R&D has a positive effect on productivity growth two years later, while the effect of additional R&D investments on productivity in the next years is also likely.

All these are largely in line with the findings of previous studies. R-squared values are also consistent with those obtained earlier (around 0.5 in Guellec-van Pottelsberghe 2001). Therefore, in spite of having shorter time series, it is possible to draw similar conclusions with the help of R&D net asset statistics to those of previous studies.

The second regression calculated with the variables adjusted for difference stationarity (equation (1) with differencing treatment) is shown in *Table 2.*

Table 2 Regression result of measuring the elasticity of productivity on R&D investment with R&D net asset statistics (treated for difference stationarity), European countries, 2009–2017

	Coefficient	***Std. Error***	***p-value***
Dependent variable: ddd GrGDPc			
dd GrNA_1	– 0.107	0.101	0.3044
d GrNA_2	– 0.362	0.471	0.4553
GrNA_3	– 0.013	0.157	0.9359
dd LNA_2	0.535	1.13	0.6431
R^2	0.806	*Adjusted* R^2	0.783
Durbin-Watson:		1.745	

Source: own construction based on Eurostat data

In this case, calculating with R&D net asset statistics, the coefficient of *ddLNA_2* is not significant after differencing. The results for the growth variables *ddGrNA_1*, *dGrNA_2*, and *GrNA_3*, are also not significant statistically.

It is apparent that if difference stationarity requirements are not applied, significant results can be achieved. On the other hand, differencing the variables according to requirement gives no significant results. This confirms that other factors have significant effects on productivity, some of them affecting R&D similarly, some of them differently. In the meantime, R^2 values and the Durbin-Watson test value improved a lot compared to the first case. Also, due to the relatively short period of time examined, it is not possible to say much about the relationship of productivity and R&D under these requirements.

4.3. The effect of R&D on productivity calculating net assets by the perpetual inventory method

Results of the third regression (equation (3) without differencing treatment) are summarized in *Table 3*.

Table 3 Regression result of measuring the elasticity of productivity on R&D investment with estimated R&D stock from R&D capital formation, European countries, 2009–2017

	Coefficient	***Std. Error***	***p-value***
Dependent variable: GrGDPc			
GrperpNA_1	0.016	0.011	0.1564
GrperpNA_2	0.003	0.007	0.6583
GrperpNA_3	– 0.025	0.044	0.5814
LperpNA_2	0.044	0.033	0.1416
R^2	0.498	*Adjusted* R^2	0.438
Durbin-Watson:		0.829	

Source: own construction based on Eurostat data

The results of the third regression show that the variables of interest – *GrperpNA_1*, *GrperpNA_2*, *GrperpNA_3,* and *LperpNA_2* – do not have statistically significant and interpretable coefficient values. Furthermore, R-squared values are worse than in the first model.

The results of the fourth regression applying variables estimated by the perpetual inventory method and treated for difference stationarity are given in *Table 4.*

Table 4 Regression result of measuring the elasticity of productivity on R&D investment with estimated R&D stock from R&D capital formation (treated for difference stationarity), European countries, 2009–2017

	Coefficient	***Std. Error***	***p-value***
Dependent variable: ddd GrGDPc			
dGrperpNA_1	– 0.017	0.017	0.3420
dGrperpNA_2	– 0.014	0.051	0.7836
GrperpNA_3	– 0.022	0.092	0.8101
dLperpNA_2	– 0.039	0.253	0.8785
R^2	0.802	*Adjusted* R^2	0.778
Durbin-Watson:		1.776	

Source: own construction based on Eurostat data

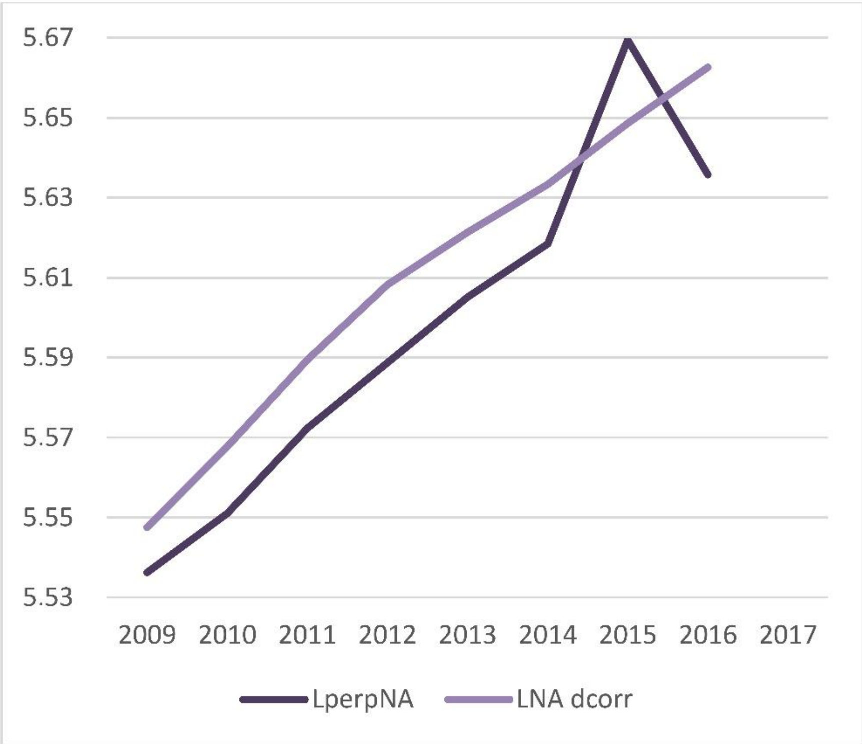
In this regression neither of the included variables have significant coefficients, again. However, due to the differencing treatment, R^2 values and the Durbin-Watson test variable are better than in the case of not controlling for difference stationarity. Differencing treatment inevitably improves the statistical features of the regression, still we cannot say anything else about the relationship of the examined phenomena other than that seemingly there is no significant co-movement at this stage.

4.4. Comparison of the R&D net assets statistics and R&D stock calculated by the perpetual inventory method

There are serious limitations in using the above calculations for evaluating the impact of R&D on productivity. The time period is too short to include the longer-term effects and also the circle of examined countries is limited. In the first years of the data series the effects of the 2008 crisis was still observable in many countries, lowering their productivity growth. All these problems imply that the model carries a high degree of uncertainty and it seems difficult to arrive at an interpretable result. For individual countries or years, it is not possible to draw relevant conclusions, no matter, which type of calculations we use.

Taking into account all these uncertainties, it seems to be a positive result for measurements using R&D net asset statistics untreated for difference stationarity, the calculated coefficient of the R&D asset level being significant and positive in relation to productivity growth. The R&D stock figures, also calculated from national accounts data with the usual perpetual inventory method, were not able to produce similar results. In order to determine the difference in their behaviour, Figure 1 shows their values aggregated for the included countries through the period examined.

Figure 1 Net asset statistics and perpetual R&D stocks in 15 European countries, in log of million euro



Source: own calculation on the basis of Eurostat data (EUROSTAT 2019). Data is shown with a shift in order to make them more comparable.

Figure 1 shows that the values of R&D net assets were more stable throughout the period than values of the calculated R&D stock. The run of *LperpNA* is much more bumpy, with an outstanding value in 2015. This is not surprising taking into account that *LperpNA* is calculated from the capital formation figures of each year, which being flow type data, show much more variability.

Also, using R&D net asset statistics treated for difference stationarity, the calculated coefficient of the R&D asset level was closer to significance than the figures calculated from flow data. This, however, does not say anything about the actual relationship of productivity and R&D, although it may be informative for the comparison of R&D data taken directly from the statistics and calculated by the perpetual inventory method by the analyst. The data of the statistics perform better than the calculated data even in the absence of interpretable results.

From the above calculations it is clear that R&D net asset statistics provide a better solution for macroeconomic estimations of the effects of R&D activities on productivity than earlier methods, which has been used in the absence of more reliable data. The difference between the two sets of estimations lies in the measurement of R&D stocks. The figures given in the statistics are compiled on the basis of more information with a more elaborate methodology than figures calculated from flow type data.

5. Conclusion

As indicated by the above analysis, R&D net asset statistics provide a better basis for future research of R&D than the previously calculated R&D stock figures. This is true not only for estimation models, which have been in use for decades, but probably, new, perhaps more simple models could also provide results, which make sense. This might well be an important step ahead in economic analysis, because research of the R&D activities at the country level has always had difficulties in obtaining statistically relevant results due to the relatively small value of R&D figures compared to GDP or total production of a country.

R&D expenses are, however, inevitably long-term investments. Their effect on productivity is observable only after a certain time lag, therefore longer time series of data will always be necessary for deeper analysis. The time span of 9 years applied here is too short to obtain an interpretable result, therefore it would be useful to conduct similar research if at least twelve-year data sets are available. It may also be useful to try more diverse estimation methods, dynamic panel regressions among others.

The case of R&D capitalization shows that extending the asset boundary is a process which is truly worthwhile. The possibility of better and faster interpretation of data is an important advantage for economists as data users, even if advancement in this field of statistics is full of difficulties. Careful consideration and innovative thinking on new methods of evaluation and possibilities of asset capitalization therefore should be welcome and continued.

References

Chandra, K. – Mansoor, A. – Sinha, K. B. (2018): The role of the state in promoting innovation efficacy: a comparative analysis of R&D spending and TFP growth in Hong Kong, Singapore, and Shenzhen. *Conference paper for the 17th International Schumpeter Society Conference,* Seoul, 2018 July.

EC – IMF – OECD – UN – IBRD (2009): *System of National Accounts 2008.* European Communities – International Monetary Fund – Organisation for Economic Co-operation and Development – United Nations – World Bank.

EUROSTAT (2014/a): *Manual on the changes between ESA 95 and ESA 2010.* Eurostat. Luxembourg

EUROSTAT (2014/b): *Manual on measuring Research and Development in ESA 2010.* Eurostat. Luxembourg

EUROSTAT (2019): European Statistics. http://ec.europa.eu/eurostat/ Accessed: February 24–25, 2019

Griliches, Z. (1998): Productivity, R&D, and the Data Constraint. In: Griliches, Z.: *R&D and Productivity: The Econometric Evidence.* University of Chicago Press, 347– 374. http://www.nber.org/books/gril98-1

Guellec, D. – van Pottelsberghe de la Potterie, B. (2001): R&D and productivity growth: Panel data analysis of 16 OECD countries. *OECD Economic Studies,* No. 33, 2001/II, 103–126.

Hall, B. H. – Mairesse, J. – Mohnen, P. (2009): Measuring the returns to R&D. *NBER Working paper,* 15622.

Lequiller, F. – Blades, D. (2014): *Understanding National Accounts.* Second Edition, OECD Publishing. http://dx.doi.org/10.1787/9789264214637-en

Lichtenberg, F. R. (1992): R&D investment and international productivity differences. *NBER Working paper*, 4161.

Räth, N. (2016): Volkswirtschaftliche Gesamtrechnungen – Reflexionen 2016. *WISTA Wirtschaft und Statistik,* 3, 96–113.

UNITED NATIONS (1968): *A System of National Accounts.* United Nations.

Vanoli, A. (2005): *A history of national accounts.* IOS Press, ISBN-10: 1586034693, ISBN-13: 978-1586034696

Appendix 1.

Dickey-Fuller tests were run for all the variables applied in this study. The variables were treated for difference stationarity by differencing the times the letter "d" is written before the name of the variable. The tests here show the performance of the variables after the treatment. In the case of GrNA_3 differencing treatment was not necessary, therefore its test is shown without such treatment.

1. dddGrGDPc
Dickey–Fuller test for d_d_d_GrGDPc
with constant and trend
model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 4, lag order = 0
estimated value of (a – 1): –0.193648
test statistic = –0.419554 [0.8616]
Unit 2, T = 4, lag order = 0
estimated value of (a – 1): –3.27577
test statistic = –4.83507 [0.1114]
Unit 3, T = 4, lag order = 0
estimated value of (a – 1): –1.22606
test statistic = –2.07901 [0.4355]
Unit 4, T = 4, lag order = 0
estimated value of (a – 1): –1.20225
test statistic = –2.14467 [0.4169]
Unit 5, T = 4, lag order = 0
estimated value of (a – 1): –1.63588
test statistic = –2.10067 [0.4344]
Unit 6, T = 4, lag order = 0
estimated value of (a – 1): –1.95584
test statistic = –2.35224 [0.3662]
Unit 7, T = 4, lag order = 0
estimated value of (a – 1): –1.12362
test statistic = –1.71123 [0.6460]
Unit 8, T = 4, lag order = 0
estimated value of (a – 1): –1.63361
test statistic = –4.84013 [0.1113]
Unit 9, T = 4, lag order = 0
estimated value of (a – 1): –1.64289
test statistic = –3.9532 [0.1799]
Unit 10, T = 4, lag order = 0
estimated value of (a – 1): –0.68602
test statistic = –1.02014 [0.7988]
Unit 11, T = 4, lag order = 0
estimated value of (a – 1): –1.82295

test statistic = –9.23072 [0.0178]
Unit 12, T = 4, lag order = 0
estimated value of (a – 1): –1.17463
test statistic = –5.98687 [0.0645]
Unit 13, T = 4, lag order = 0
estimated value of (a – 1): –1.64351
test statistic = –144.701 [0.0001]
Unit 14, T = 4, lag order = 0
estimated value of (a – 1): –1.93287
test statistic = –3.317 [0.2527]
Unit 15, T = 4, lag order = 0
estimated value of (a – 1): –1.63224
test statistic = –22.6691 [0.0007]
H0: all groups have unit root
N,T = (15,4)
Im–Pesaran–Shin t–bar = –14.0907
Choi meta–tests:
Inverse chi–square(30) = 70.0645 [0.0000]
Inverse normal test = –3.3876 [0.0004]
Logit test: t(79) = –3.96014 [0.0001]

2. ddGrNA_1
Dickey–Fuller test for d_d_GrNA_1
with constant and trend
model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 4, lag order = 0
estimated value of (a – 1): 0.610005
test statistic = 0.911672 [0.9571]
Unit 2, T = 4, lag order = 0
estimated value of (a – 1): –1.91053
test statistic = –1.86846 [0.4852]
Unit 3, T = 4, lag order = 0
estimated value of (a – 1): –1.19174
test statistic = –1.33388 [0.7418]
Unit 4, T = 4, lag order = 0
estimated value of (a – 1): –1.74239
test statistic = –6.40493 [0.0568]
Unit 5, T = 4, lag order = 0
estimated value of (a – 1): –2.07273
test statistic = –6.85839 [0.0461]
Unit 6, T = 4, lag order = 0
estimated value of (a – 1): –2.45508
test statistic = –38.2313 [0.0001]
Unit 7, T = 4, lag order = 0

estimated value of (a – 1): –1.69746
test statistic = –4.51511 [0.1288]
Unit 8, T = 4, lag order = 0
estimated value of (a – 1): –1.35296
test statistic = –2.34524 [0.3667]
Unit 9, T = 4, lag order = 0
estimated value of (a – 1): –3.40445
test statistic = –11.2602 [0.0094]
Unit 10, T = 4, lag order = 0
estimated value of (a – 1): –1.74276
test statistic = –2.56499 [0.3332]
Unit 11, T = 4, lag order = 0
estimated value of (a – 1): –1.63915
test statistic = –2.16774 [0.3884]
Unit 12, T = 4, lag order = 0
estimated value of (a – 1): –1.67623
test statistic = –3.59898 [0.2278]
Unit 13, T = 4, lag order = 0
estimated value of (a – 1): –1.8626
test statistic = –2.36419 [0.3654]
Unit 14, T = 4, lag order = 0
estimated value of (a – 1): –5.25995
test statistic = –21.8766 [0.0009]
Unit 15, T = 4, lag order = 0
estimated value of (a – 1): –1.6414
test statistic = –5.89222 [0.0699]
H0: all groups have unit root
N,T = (15,4)
Im–Pesaran–Shin t–bar = –7.35804
Choi meta–tests:
Inverse chi–square(30) = 76.3769 [0.0000]
Inverse normal test = –3.84525 [0.0001]
Logit test: t(79) = –4.41626 [0.0000]

3. dGrNA_2
Dickey–Fuller test for d_GrNA_2
with constant and trend
model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 4, lag order = 0
estimated value of (a – 1): –0.854175
test statistic = –2.33852 [0.3672]
Unit 2, T = 4, lag order = 0
estimated value of (a – 1): –1.60642
test statistic = –0.99044 [0.8010]

Unit 3, T = 4, lag order = 0
  estimated value of (a – 1): –1.16794
  test statistic = –1.81484 [0.6044]
Unit 4, T = 4, lag order = 0
  estimated value of (a – 1): –1.63114
  test statistic = –2.0494 [0.4503]
Unit 5, T = 4, lag order = 0
  estimated value of (a – 1): –2.35358
  test statistic = –196.767 [0.0001]
Unit 6, T = 4, lag order = 0
  estimated value of (a – 1): –2.53276
  test statistic = –2.15267 [0.4161]
Unit 7, T = 4, lag order = 0
  estimated value of (a – 1): –1.76964
  test statistic = –3.99068 [0.1777]
Unit 8, T = 4, lag order = 0
  estimated value of (a – 1): 1.14609
  test statistic = 2.36076 [0.9995]
Unit 9, T = 4, lag order = 0
  estimated value of (a – 1): –5.28808
  test statistic = –7.25866 [0.0384]
Unit 10, T = 4, lag order = 0
  estimated value of (a – 1): –2.62841
  test statistic = –8.6278 [0.0222]
Unit 11, T = 4, lag order = 0
  estimated value of (a – 1): –1.58793
  test statistic = –2.14455 [0.4169]
Unit 12, T = 4, lag order = 0
  estimated value of (a – 1): –1.73344
  test statistic = –3.0914 [0.2680]
Unit 13, T = 4, lag order = 0
  estimated value of (a – 1): –1.19143
  test statistic = –2.69728 [0.3107]
Unit 14, T = 4, lag order = 0
  estimated value of (a – 1): 1.02177
  test statistic = 0.260567 [0.9314]
Unit 15, T = 4, lag order = 0
  estimated value of (a – 1): –1.62597
  test statistic = –1.94437 [0.4730]
H0: all groups have unit root
N,T = (15,4)
Im–Pesaran–Shin t–bar = –15.5498
Choi meta–tests:
  Inverse chi–square(30) = 51.1795 [0.0093]

Inverse normal test = –1.19496 [0.1161]
Logit test: t(79) = –1.31793 [0.0957]

4. ddLNA_2
Dickey–Fuller test for d_d_LNA_2
with constant and trend
model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 4, lag order = 0
estimated value of (a – 1): –0.864664
test statistic = –2.29519 [0.3757]
Unit 2, T = 4, lag order = 0
estimated value of (a – 1): –1.609
test statistic = –0.997231 [0.8005]
Unit 3, T = 4, lag order = 0
estimated value of (a – 1): –1.16964
test statistic = –1.81256 [0.6044]
Unit 4, T = 4, lag order = 0
estimated value of (a – 1): –1.63304
test statistic = –2.06152 [0.4491]
Unit 5, T = 4, lag order = 0
estimated value of (a – 1): –2.37382
test statistic = –124.108 [0.0001]
Unit 6, T = 4, lag order = 0
estimated value of (a – 1): –2.54716
test statistic = –2.16458 [0.4150]
Unit 7, T = 4, lag order = 0
estimated value of (a – 1): –1.76983
test statistic = –4.00616 [0.1767]
Unit 8, T = 4, lag order = 0
estimated value of (a – 1): 1.2207
test statistic = 1.3866 [0.9724]
Unit 9, T = 4, lag order = 0
estimated value of (a – 1): –4.94891
test statistic = –8.4206 [0.0253]
Unit 10, T = 4, lag order = 0
estimated value of (a – 1): –2.625
test statistic = –8.27543 [0.0266]
Unit 11, T = 4, lag order = 0
estimated value of (a – 1): –1.5957
test statistic = –2.14848 [0.4165]
Unit 12, T = 4, lag order = 0
estimated value of (a – 1): –1.73336
test statistic = –3.0663 [0.2705]
Unit 13, T = 4, lag order = 0

estimated value of (a – 1): –1.19002
test statistic = –2.65139 [0.3206]
Unit 14, T = 4, lag order = 0
estimated value of (a – 1): 1.54343
test statistic = 0.312558 [0.9370]
Unit 15, T = 4, lag order = 0
estimated value of (a – 1): –1.6039
test statistic = –1.82626 [0.5461]
H0: all groups have unit root
N,T = (15,4)
Im–Pesaran–Shin t–bar = –10.8089
Choi meta–tests:
Inverse chi–square(30) = 51.3039 [0.0090]
Inverse normal test = –1.50048 [0.0667]
Logit test: t(79) = –1.85661 [0.0335]

5. ddddLcap (log of population size)
Dickey–Fuller test for d_d_d_d_Lcap
with constant and trend
model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 4, lag order = 0
estimated value of (a – 1): –1.84707
test statistic = –5.45565 [0.0790]
Unit 2, T = 4, lag order = 0
estimated value of (a – 1): –1.743
test statistic = –3.50544 [0.2347]
Unit 3, T = 4, lag order = 0
estimated value of (a – 1): –1.85375
test statistic = –2.52342 [0.3420]
Unit 4, T = 4, lag order = 0
estimated value of (a – 1): –2.47153
test statistic = –1.00244 [0.8001]
Unit 5, T = 4, lag order = 0
estimated value of (a – 1): –1.72874
test statistic = –2.36393 [0.3654]
Unit 6, T = 4, lag order = 0
estimated value of (a – 1): –1.48079
test statistic = –1.64201 [0.6577]
Unit 7, T = 4, lag order = 0
estimated value of (a – 1): –1.67809
test statistic = –2.29472 [0.3757]
Unit 8, T = 4, lag order = 0
estimated value of (a – 1): –1.8169
test statistic = –21.6438 [0.0009]

Unit 9, T = 4, lag order = 0
  estimated value of (a – 1): –6.31262
  test statistic = –6.74043 [0.0481]
Unit 10, T = 4, lag order = 0
  estimated value of (a – 1): –1.12536
  test statistic = –5.03574 [0.1051]
Unit 11, T = 4, lag order = 0
  estimated value of (a – 1): –2.05561
  test statistic = –7.43322 [0.0361]
Unit 12, T = 4, lag order = 0
  estimated value of (a – 1): –1.80472
  test statistic = –8.10121 [0.0282]
Unit 13, T = 4, lag order = 0
  estimated value of (a – 1): –1.32686
  test statistic = –2.10591 [0.4341]
Unit 14, T = 4, lag order = 0
  estimated value of (a – 1): –2.40826
  test statistic = –10.3771 [0.0117]
Unit 15, T = 4, lag order = 0
  estimated value of (a – 1): –1.5989
  test statistic = –1.95562 [0.4720]
H0: all groups have unit root
N,T = (15,4)
Im–Pesaran–Shin t–bar = –5.47871
Choi meta–tests:
  Inverse chi–square(30) = 65.8251 [0.0002]
  Inverse normal test = –3.6669 [0.0001]
  Logit test: t(79) = –3.90606 [0.0001]

6. dddLGDPcap_1
Dickey–Fuller test for d_d_d_LGDPcap_1
  with constant and trend
  model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 4, lag order = 0
  estimated value of (a – 1): –0.681572
  test statistic = –3.03415 [0.2736]
Unit 2, T = 4, lag order = 0
  estimated value of (a – 1): –0.545819
  test statistic = –0.315977 [0.8712]
Unit 3, T = 4, lag order = 0
  estimated value of (a – 1): –1.39503
  test statistic = –2.68737 [0.3169]
Unit 4, T = 4, lag order = 0
  estimated value of (a – 1): –1.26804

test statistic = –2.99758 [0.2828]
Unit 5, T = 4, lag order = 0
estimated value of (a – 1): –2.37594
test statistic = –15.918 [0.0037]
Unit 6, T = 4, lag order = 0
estimated value of (a – 1): –1.42619
test statistic = –2.59919 [0.3303]
Unit 7, T = 4, lag order = 0
estimated value of (a – 1): –1.20477
test statistic = –2.48281 [0.3460]
Unit 8, T = 4, lag order = 0
estimated value of (a – 1): –1.69872
test statistic = –10.5959 [0.0110]
Unit 9, T = 4, lag order = 0
estimated value of (a – 1): –1.58862
test statistic = –15.2247 [0.0043]
Unit 10, T = 4, lag order = 0
estimated value of (a – 1): –0.943855
test statistic = –2.3935 [0.3583]
Unit 11, T = 4, lag order = 0
estimated value of (a – 1): –1.81523
test statistic = –3.7205 [0.2166]
Unit 12, T = 4, lag order = 0
estimated value of (a – 1): –1.4564
test statistic = –4.21902 [0.1512]
Unit 13, T = 4, lag order = 0
estimated value of (a – 1): –1.58139
test statistic = –13.1951 [0.0062]
Unit 14, T = 4, lag order = 0
estimated value of (a – 1): –2.2206
test statistic = –3.89796 [0.1878]
Unit 15, T = 4, lag order = 0
estimated value of (a – 1): –1.70416
test statistic = –10.343 [0.0118]
H0: all groups have unit root
N,T = (15,4)
Im–Pesaran–Shin t–bar = –6.24165
Choi meta–tests:
Inverse chi–square(30) = 74.4321 [0.0000]
Inverse normal test = –4.3328 [0.0000]
Logit test: t(79) = –4.64441 [0.0000]

7. dGrperpNA_1
Dickey–Fuller test for d_GrperpNA_1

with constant and trend
model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 5, lag order = 0
estimated value of (a – 1): –0.770249
test statistic = –0.645304 [0.8727]
Unit 2, T = 5, lag order = 0
estimated value of (a – 1): –1.27991
test statistic = –2.01106 [0.4698]
Unit 3, T = 5, lag order = 0
estimated value of (a – 1): –1.57327
test statistic = –2.94216 [0.2581]
Unit 4, T = 5, lag order = 0
estimated value of (a – 1): –1.03875
test statistic = –1.40454 [0.7386]
Unit 5, T = 5, lag order = 0
estimated value of (a – 1): –2.14108
test statistic = –6.39328 [0.0281]
Unit 6, T = 5, lag order = 0
estimated value of (a – 1): –1.9878
test statistic = –78.2446 [0.0001]
Unit 7, T = 5, lag order = 0
estimated value of (a – 1): –1.75353
test statistic = –4.43943 [0.0875]
Unit 8, T = 5, lag order = 0
estimated value of (a – 1): –1.41994
test statistic = –2.77366 [0.2882]
Unit 9, T = 5, lag order = 0
estimated value of (a – 1): –2.88608
test statistic = –13.6059 [0.0015]
Unit 10, T = 5, lag order = 0
estimated value of (a – 1): –1.03759
test statistic = –1.64361 [0.6574]
Unit 11, T = 5, lag order = 0
estimated value of (a – 1): –1.68805
test statistic = –3.47471 [0.1848]
Unit 12, T = 5, lag order = 0
estimated value of (a – 1): –1.30172
test statistic = –4.51054 [0.0833]
Unit 13, T = 5, lag order = 0
estimated value of (a – 1): –1.17955
test statistic = –1.92545 [0.4970]
Unit 14, T = 5, lag order = 0
estimated value of (a – 1): –1.99468
test statistic = –2.53306 [0.3271]

Unit 15, T = 5, lag order = 0
  estimated value of (a – 1): –1.62081
  test statistic = –2.89536 [0.2641]
H0: all groups have unit root
N,T = (15,5)
Im–Pesaran–Shin t–bar = –8.62951
                    10%   5%   1%
Critical values: –3.33 –3.88 –5.72
Choi meta–tests:
  Inverse chi–square(30) = 66.4775 [0.0001]
  Inverse normal test = –3.20529 [0.0007]
  Logit test: t(79) = –3.70406 [0.0002]

8. dGrperpNA_2
Dickey–Fuller test for d_GrperpNA_2
  with constant and trend
  model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 4, lag order = 0
  estimated value of (a – 1): –1.69436
  test statistic = –2.22041 [0.4034]
Unit 2, T = 4, lag order = 0
  estimated value of (a – 1): –1.20983
  test statistic = –1.2665 [0.7640]
Unit 3, T = 4, lag order = 0
  estimated value of (a – 1): –1.60902
  test statistic = –2.12509 [0.4124]
Unit 4, T = 4, lag order = 0
  estimated value of (a – 1): –1.37717
  test statistic = –1.16123 [0.7773]
Unit 5, T = 4, lag order = 0
  estimated value of (a – 1): –1.80417
  test statistic = –3.33028 [0.2487]
Unit 6, T = 4, lag order = 0
  estimated value of (a – 1): –1.97413
  test statistic = –56.6912 [0.0001]
Unit 7, T = 4, lag order = 0
  estimated value of (a – 1): –1.75339
  test statistic = –3.13981 [0.2660]
Unit 8, T = 4, lag order = 0
  estimated value of (a – 1): –1.41575
  test statistic = –1.94318 [0.4731]
Unit 9, T = 4, lag order = 0
  estimated value of (a – 1): –3.67591
  test statistic = –15.0924 [0.0044]

Unit 10, T = 4, lag order = 0
  estimated value of (a – 1): –6.25805
  test statistic = –3.72384 [0.2165]
Unit 11, T = 4, lag order = 0
  estimated value of (a – 1): –1.71039
  test statistic = –2.6094 [0.3295]
Unit 12, T = 4, lag order = 0
  estimated value of (a – 1): –1.04355
  test statistic = –9.24633 [0.0177]
Unit 13, T = 4, lag order = 0
  estimated value of (a – 1): –0.960279
  test statistic = –1.17605 [0.7752]
Unit 14, T = 4, lag order = 0
  estimated value of (a – 1): 1.09166
  test statistic = 0.695655 [0.9515]
Unit 15, T = 4, lag order = 0
  estimated value of (a – 1): –1.66903
  test statistic = –3.8058 [0.2112]
H0: all groups have unit root
N,T = (15,4)
Im–Pesaran–Shin t–bar = –7.12239
Choi meta–tests:
  Inverse chi–square(30) = 57.8844 [0.0016]
  Inverse normal test = –2.17048 [0.0150]
  Logit test: t(79) = –2.63711 [0.0050]

9. dLperpNA_2
Dickey–Fuller test for d_LperpNA_2
  with constant and trend
  model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 5, lag order = 0
  estimated value of (a – 1): –1.16709
  test statistic = –1.66631 [0.6488]
Unit 2, T = 5, lag order = 0
  estimated value of (a – 1): –1.30804
  test statistic = –1.96712 [0.4821]
Unit 3, T = 5, lag order = 0
  estimated value of (a – 1): –1.47471
  test statistic = –2.45089 [0.3425]
Unit 4, T = 5, lag order = 0
  estimated value of (a – 1): –0.845728
  test statistic = –1.74981 [0.6227]
Unit 5, T = 5, lag order = 0
  estimated value of (a – 1): –1.65873

test statistic = –6.26341 [0.0303]
Unit 6, T = 5, lag order = 0
estimated value of (a – 1): –2.11021
test statistic = –5.61894 [0.0446]
Unit 7, T = 5, lag order = 0
estimated value of (a – 1): –1.39012
test statistic = –2.14945 [0.4335]
Unit 8, T = 5, lag order = 0
estimated value of (a – 1): –1.3989
test statistic = –2.62117 [0.3080]
Unit 9, T = 5, lag order = 0
estimated value of (a – 1): –3.15652
test statistic = –214.542 [0.0001]
Unit 10, T = 5, lag order = 0
estimated value of (a – 1): –0.869741
test statistic = –2.1646 [0.4150]
Unit 11, T = 5, lag order = 0
estimated value of (a – 1): –1.0898
test statistic = –1.64981 [0.6566]
Unit 12, T = 5, lag order = 0
estimated value of (a – 1): –1.33528
test statistic = –6.09251 [0.0332]
Unit 13, T = 5, lag order = 0
estimated value of (a – 1): –1.04581
test statistic = –2.91129 [0.2624]
Unit 14, T = 5, lag order = 0
estimated value of (a – 1): –2.0396
test statistic = –1.22642 [0.7827]
Unit 15, T = 5, lag order = 0
estimated value of (a – 1): –1.61959
test statistic = –3.03978 [0.2461]
H0: all groups have unit root
N,T = (15,5)
Im–Pesaran–Shin t–bar = –17.0743

	10%	5%	1%
Critical values:	–3.33	–3.88	–5.72

Choi meta–tests:
Inverse chi–square(30) = 56.4579 [0.0024]
Inverse normal test = –2.55848 [0.0053]
Logit test: t(79) = –2.9365 [0.0022]

10. GrNA_3
Dickey–Fuller test for GrNA_3
with constant and trend

model: (1–L)y = b0 + b1*t + (a–1)*y(–1) + e
Unit 1, T = 4, lag order = 0
estimated value of (a – 1): –1.65695
test statistic = –2.25683 [0.4025]
Unit 2, T = 4, lag order = 0
estimated value of (a – 1): –1.58412
test statistic = –1.81368 [0.6044]
Unit 3, T = 4, lag order = 0
estimated value of (a – 1): –1.45747
test statistic = –2.41304 [0.3567]
Unit 4, T = 4, lag order = 0
estimated value of (a – 1): –0.76999
test statistic = –0.901293 [0.8132]
Unit 5, T = 4, lag order = 0
estimated value of (a – 1): –2.31685
test statistic = –2.76487 [0.3052]
Unit 6, T = 4, lag order = 0
estimated value of (a – 1): –1.32404
test statistic = –1.39272 [0.7400]
Unit 7, T = 4, lag order = 0
estimated value of (a – 1): –1.7729
test statistic = –2.91476 [0.2991]
Unit 8, T = 4, lag order = 0
estimated value of (a – 1): –0.3029
test statistic = –2.78595 [0.3048]
Unit 9, T = 4, lag order = 0
estimated value of (a – 1): –3.37744
test statistic = –28.7468 [0.0003]
Unit 10, T = 4, lag order = 0
estimated value of (a – 1): –0.16541
test statistic = –0.149364 [0.8965]
Unit 11, T = 4, lag order = 0
estimated value of (a – 1): –2.5667
test statistic = –2.03964 [0.4533]
Unit 12, T = 4, lag order = 0
estimated value of (a – 1): –1.7916
test statistic = –3.0609 [0.2708]
Unit 13, T = 4, lag order = 0
estimated value of (a – 1): –1.67302
test statistic = –2.21566 [0.3948]
Unit 14, T = 4, lag order = 0
estimated value of (a – 1): –1.195
test statistic = –41.0331 [0.0001]
Unit 15, T = 4, lag order = 0

estimated value of (a – 1): –1.6793
test statistic = –2.28829 [0.3761]
H0: all groups have unit root
N,T = (15,4)
Im–Pesaran–Shin t–bar = –6.45179
Choi meta–tests:
Inverse chi–square(30) = 55.9678 [0.0028]
Inverse normal test = –1.9536 [0.0254]
Logit test: t(79) = –2.57096 [0.0060]

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 162–189.

Agriculture 4.0 in Hungary: The challenges of 4th Industrial Revolution in Hungarian agriculture within the frameworks of the Common Agricultural Policy

Sarolta Somosi – Gabriella Számfira

The impacts of the 4th industrial revolution, which is considered a "general purpose technology", do not appear only in the manufacturing sector or in the increasingly ICT-intensive services sector. Agriculture, which is using manufactured inputs in more and more automatized factories, and which relies highly on technology in its basic activities, has also been going through remarkable changes recently. Precision farming (1), optimized usage of scarce inputs (2) and so attaining a more economy and ecology-friendly (sustainable) farming system (3) are all potential positive outcomes of the technological development of Industry 4.0 taking place in agriculture.

In connection with this topic, our main aim is to determine the position of Hungary in this field, and whether there are any obstacles in this process. Both in the theoretical and empirical research conducted, we focus on the agricultural workforce, whether it is ready to keep pace with developments or if Hungarian farmers may be facing difficulties. Finally, as an outlook we examine the European Union's Common Agricultural Policy and its future changes, if it is currently or close to handling this challenge?

Keywords: agriculture 4.0, precision agriculture, Hungarian agriculture, Digital Agricultural Strategy

1. Introduction

Over the last few decades, agriculture – and of course in a broader sense natural resource management – has been facing a wide range of difficulties, challenges that require solutions beyond just the recent industrialization of farming activities. Global climate change, the necessity of resilient national and regional food systems, and the guaranteeing (establishment) of sustainable livelihoods for small-hold farmers are all related to present tensions in agriculture.

Improving productivity, effectiveness, and the competitiveness of agricultural activities can all be keys to answering the challenges in such a demanding environment. Thus, meeting the requirements of recent challenges requires complex analysis to plan possible solutions. There are different ways of addressing the challenges. In the present study we turn our focus on one of the possibilities: provided by the technological background, and the necessary technological readiness to facilitate its impacts on the productivity of agricultural systems. There are of course other paths being followed all over the world based on the circumstances, investment, and other factors that determine the overall situation of a country's or region's agricultural background. Participating in the Internal Market of the European Union, Hungary is left with only a few possibilities to address the present challenges and among them the technological readiness of the agriculture seems to be important. In the following chapters we will introduce the competitiveness problems appearing on the Internal Market, and the impact of development and technological readiness in agriculture. Precision farming (1), optimized usage of scarce inputs (2) and so reaching a more economy and ecology-friendly (sustainable) farming system (3) are all the potential positive outcomes of the technological development of Industry 4.0 taking place in agriculture.

In connection with this topic, it is necessary to gain an overview of Hungary's position in this field, whether there are any obstacles in this process, and especially what they are. Thus, the aim of the present article is to reveal *where Hungary and its agricultural sector are in the process of the necessary creation and improvement of the conditions that will determine its future competitiveness, what the potential deficiencies are, and what kind of public answers are being formulated in the limited leeway allowed by EU-membership on these challenges.* Recognizing the deficiencies in certain points shows where relevant actors should put more emphasis and effort, be they EU-level determinants or arising from the domestic field, or be they public or private investment.

We contribute to the better description of the problems or deficiencies by our empirical research. In our – by far not representative but in the meantime telling – survey we were to identify the domestic failings and challenges where, if a proper and effective coordination of public and private efforts and resources were to happen, we know, positive results would be achieved.

1.1. The structure of the study

After the introduction of our topic, in the first part of the study, the technological impacts of the 4^{th} industrial revolution (4IR) on agriculture will be described based on a literature review. Then potential parallels between the development phases of agricultural technology and industry will be shown. At their point of intersection, precision agriculture as an intermediary solution between advanced technologies of Information and Communication Technologies (ICT) and farming, ranging from livestock industry to crop production, will be discussed.

Then comes a short insight into the present heritage and importance of the agricultural sector within the Hungarian economy. Addressing the challenges requires first that they be described. The present research is an attempt to contribute to this, so the empirical research focuses on Hungarian practice.

The summary of the necessary technological development and the needs/opinions of the farmers surveyed will be compared to the present EU-level and Hungarian actions and measures, so a compliance-check is performed. The study ends with conclusions, suggestions, and outlooks about future research that may contribute to the better targeting of the needs of the actors in Hungarian agriculture.

2. Recent challenges of agriculture, and the related agricultural and natural resource management issues

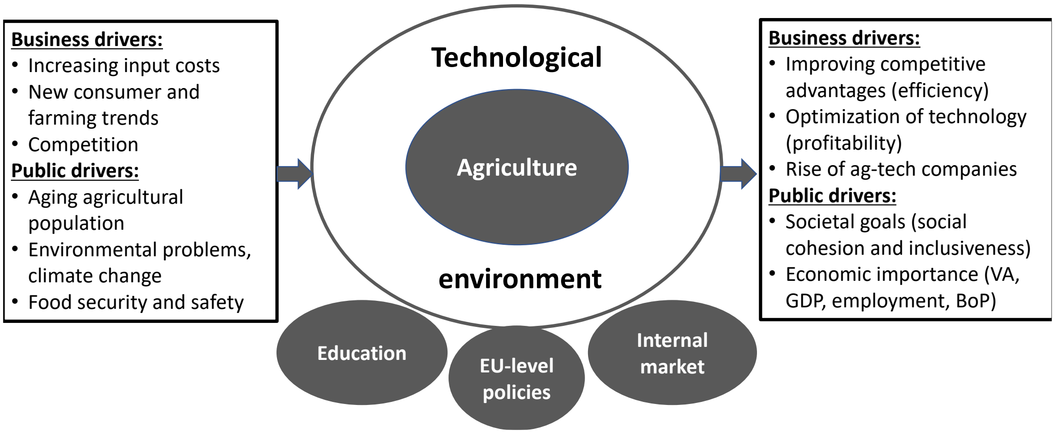
Beside acknowledging both simple and complicated problems, Turner et al. (2016) detail complex problems in agriculture, among which the frequently emphasized ones are the "usual" problems/challenges of agriculture of our times. They are also summarized in a recent report of the WGS (2018) as demographics, scarcity of natural resources, climate change, and food waste. Turner et al. (2016) characterize them as having tightly coupled components, dynamically changing behaviors, problems by which the efficiency of policies are regularly questioned, challenges where causality appears both in time and space, and where tradeoffs in the long-term and short-term interests and solutions appear. This complexity requires a different set of approaches and solutions, like the necessary interaction of actors, a systemic approach and response like agricultural intensification, improving farming systems, and increased and especially more effective resource or input integration and factor productivity.

Moreover, the original basis, the challenges, the possible solutions, and the outcomes vary country by country, by natural conditions, agricultural profile, the chosen agricultural policy, and the level of reliance on agriculture (Csáki and Jámbor 2012, Jámbor et al. 2016).

There are different ways of addressing the challenges. Among them, technological development in agriculture (the spread of automated, de-localized, and digitalized production and finally the commercialization of food) is just one direction. Other solutions can be Agroecology, which involves the application of ecological principles for the design and management of sustainable agroecosystem, so it is more in harmony with the Sustainable Development Goals (SDG) of the United Nations. Without doubting the validity of the last-mentioned solution, the present study focuses on the spread of advanced technology in farming. The spread of advanced technologies in agriculture should be considered as a possibility, rather than a constraint for actors in the market. Overall, the agricultural activities are put into a new environment, where different farming methods are/should be increasingly supported by advanced technologies (FAO 2017, WGS 2018), and where actors with different interests (business and public interests) have to react to the pushing and pulling forces appearing from different, public or private/business interests (Figure 1).

Pushing forces can be defined as pressures appearing on different levels of the market and in different orders of size or scale of operation. They can be both business or public driven pressures appearing on micro or macro levels. Meanwhile the other side of the Figure 1 summarizes those factors which appear rather as motivating (pulling) forces. These can also be a split according to their drivers – be they private or public – and according to their extent to or level at which they take place – at the micro or macro level.

Figure 1 Mapping the relevant challenges and environment of agriculture and the scales of intervention



Source: own construction

Huge differences have been mapped even between the so-called New Member States, which joined the EU in and after 2004. In their article, Jámbor et al. (2016) highlight that although original differences also occurred in agriculture after the transition to post-socialist countries, their decisions/policies in connection with agriculture had significant impacts even on their present production output, efficiency, and competitiveness. And at this point, we come to the fact that the primary sector of economies is no different from others that are traditionally considered knowledge-intensive sectors. Beyond the general comparative advantages appearing in labor and the capitalization of agriculture, the competitiveness of this sector also depends on other conditions like infrastructure, applied technology (FAO 2017) or favorable business environment, like in other fields of the economies. The development of technological background and innovative environment are also endogenous factors of growth, and finally, they appear in both national (Digital Agricultural Strategy of Hungary) and EU-level indicators (Digital Economy and Society Index, DESI of the EU) and policies. Thus, the aim of the present study is to identify the obstacles to future agricultural development and the improved competitiveness of Hungarian farmers. Furthermore, it is an attempt to collect potential solutions in addressing farmers' challenges in investing in such technologies.

3. New technological paradigm and its impacts in agriculture

The latest technological developments, the new productive assets and the appearance of a new technical and economic paradigm are all the manifestation of the so-called 4th industrial revolution (Manyika et al. 2013, Schwab 2016)[1].

Within this concept the latest technological changes are led by advanced digitalization (e.g. cyber-physical systems, cloud technology, or the Internet of Things, which contains location-based services), automation and robotization (e.g. near- autonomous machines and vehicles), 3D printing (i.e. tailor-made production becoming ever more feasible and profitable) and advanced bio- and nanotechnology (offering new materials and processes to regular industrial activities), and last but not least the use of intelligent data-based decision support systems (Pelle and Somosi 2018). The essence of these is a set of interconnected products, processes, and organizational, managerial, and business model innovations to bring fundamental changes in the entire economic system. Digital transformation in 4IR means far more than simply introducing new technological tools, machines, or solutions. Its huge value is the interconnectedness of appliances and the possibility of decision-making based on collected data and the ubiquitous computing and network infrastructure with self-configuring capability (Xu et al. 2018). These all contribute to the improvement of production in general and the transparency of processes, reveal potential malfunctions, optimize value creation and improve factor productivity (ElMaraghy et al. 2012, Xu et al. 2018). The technological environment of data collection, computing, and evaluation went through a remarkable development as well, including the potential to make data-based decisions (Babiceanu and Seker 2016). Intuitions and the experience of decision-makers are now coupled with facts and figures. The interconnectedness of things and data-based decision-making contribute to evolution and innovation in production in general.

In estimating the potential spread of advanced farming technologies, we have to mention the Digital Economy and Society Index[2] indicator, which is said to show the general digital preparedness of countries and their economic actors. In terms of overall DESI, Hungary is placed 23rd within the EU, but when the Integration of Digital Technology by Businesses is surveyed, the picture is different: Hungary obtains only 25th position, and if we see the Business digitization Index Hungary's position is even worse, only 27th in 2018[3]. From the perspective of Digital Intensity

---

[1] "Industrie 4.0" was initially introduced during the Hannover Fair in 2011. Moreover, it was officially announced in 2013 originally as a German strategic initiative to obtain a pioneering role in industries which are currently revolutionizing – at the time, only the manufacturing sector. Here the 4th refers to its rank in sequence of industrial periods from the second half of the 18th century until today, although we acknowledge that there are no clear-cut breaks between the phases of industrial development, especially in recent times.

[2] https://ec.europa.eu/digital-single-market/en/desi

[3] The Integration of digital technology indicator contains (a) 'business digitization' and (b) 'e-commerce'. From the perspective of the present study, the first is important, having five indicators (as % of firms using): electronic information sharing, Radio Frequency Identification (RFID), social media, eInvoices, and cloud solutions.

Index[4] (% of enterprises by level), in 2017 Hungary reached was 23rd. Altogether, if digital readiness is one pre-condition for the spread of advanced farming technologies, the results are far from promising.

The 4th industrial revolution is impacting the traditionally labor-intensive field of agriculture as well, furthermore, Xu et al. (2018) mention agriculture as a potential beneficiary of 4IR. The widespread use of 4IR technologies and solutions even in the primary sector underpins/confirms that they can be referred to as General Purpose Technology (GPT) (Dudley 2010). Spreading ICT usage in agriculture – similarly to other sectors – does not offer the final solutions in areas of the economy or life, but provides the tools to properly select and achieve new targets. As Bresnahan and Trajtenberg (1992, pp. 2) wrote: "as the GPT evolves and advances, it spreads throughout the economy, and in so doing it brings about and fosters generalized productivity gains."[5] Developed countries like the USA, Japan and some from Western Europe are trying to solve agricultural issues through mechanization, automation, and modernization. According to Sung (2018), the 4IR will serve as an opportunity both in time and environmentally to accelerate the spread of advanced farming technologies and thus increase the scale and overall commercialization of agriculture. Moreover, by the use of the Internet of Things (IoT), not only will farm production be able to improve, but due to time series data, the sum of earlier experiences and infrastructure, its value may even increase also.

3.1. Development stages of agricultural knowledge and technology

The first – longest in term and broadest in expansion – phase of agriculture lasted until the first third of the 20th century. It can be characterized by labor-intensive activity with a low level of productivity. It already applied some early inventions like the plough, and the mechanized agriculture also appeared, but its productivity was low[6]. A large number of small farms was common in a majority of countries (Jóri 2017).

The second phase of agricultural development started after the Second World War with several processes and research efforts made in this field. It is regularly referred to as *"the green revolution"*. Inputs of synthetic origin appeared. Nutrients, chemical fertilizers, pesticides, and even more effective machinery were implemented, so production capacity and effectiveness increased significantly

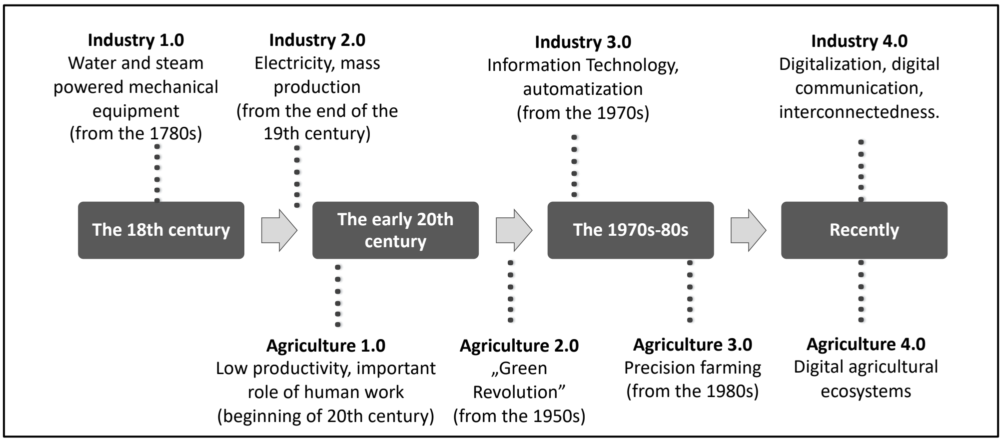
[4] The Digital Intensity Index (DII) measures the availability at firm level of 12 different digital technologies: internet for at least 50% of persons employed, recourse to ICT specialists; fast broadband (30 Mbps or above); mobile internet devices for at least 20% of persons employed; a website or homepage; a website with sophisticated functions; social media, sharing supply chain management data electronically; the use of Enterprise Resource Planning (ERP) software packages; the use of Customer Relationship Management (CRM); e-commerce web sales accounting for over 1% of total turnover, and business-to-consumer (B2C) web sales of over 10% of total web sales. The value for the index, therefore, ranges from 0 to 12.

[5] Bresnahan and Trajtenberg (1992) brought an even earlier example from agriculture. They cited Griliches's (1957) view on the method of inventing and breeding during the invention of hybrid corn seed as something close to their notion of general purpose technology.

[6] Representing this with a basic example: each farmer could produce enough food to feed about 26 people at this time (data from https://consulting.ey.com/digital-agriculture-helping-to-feed-a-growing-world/)

(Hesser 2006). Later, this period also brought genetic modification technology (GMO). In this development phase, the natural and engineering sciences went hand in hand with new management solutions – work organization, administration, control – to find applicable answers on the challenges of the time. Together, this all led to a significant increase in production – Bőgel (2018) refers to it as the "*period of the harvest*"[7] (Figure 2).

Figure 2 The development stages of industry and agriculture



Source: own construction based on Popp (2018)

In the following period, from the 1990s the impetus of inducing constantly increased productivity was deflected. Experts explain this by the law of diminishing returns. The applied inputs and farming solutions resulted in increased yields and productivity, but only up to a limit, and associated with increasing negative impacts on the environment. Thus, a special situation appeared, where developing countries were not even able to benefit from the earlier Green Revolution, meanwhile in the developed world new solutions were being planned to assist/support sustainable growth.

Digital transformation of agriculture seemed to be one possible direction, even if its spread and efficiency was not equal everywhere (Westerman et al. 2014). The appearance of digital technology as a General Purpose Technology in agriculture is witnessed from the late 80s. It brought the use of yield measuring right on machines equipped with Global Positioning System (GPS). This could even be considered as an early manifestation of decision-making based on data gathered by the help of Industry 4.0. The more precise yield maps supported the revelation/recognition of the spatial variation of yields and its natural and other causes. Modern/up-to-date and continuous yield measurement combined with GPS positioning was an important step towards computer-supported precision agriculture (Bőgel 2018).

[7] Development in this period resulted in each farmer being able to feed about 155 people.

This is why it is difficult to find an exact time when digitalized precision agriculture turned into the 4th phase of agricultural technological development. On Figure 2, Industry 4.0 and Agriculture 4.0 are positioned visibly at the same time. It should be emphasized here, that *Agriculture 4.0* or *Farming 4.0*[8], as the manifestation of impacts of the 4th industrial revolution is usually called, is not the same as the precision agriculture (PA), the one solution which is more and more based on the tools of ICT. In other words, Agriculture 4.0 is the application of digitalization processes and the evolution to Industry 4.0 within the primary sector and as such, it is more a technological environment. Meanwhile, PA, as will be introduced later, is a widely used farming method, which bridges or connects the two latest phases.

3.2. Advanced technologies and their application in agriculture

As Comparetti (2011) summarized, at the beginning of 1980s, precision agriculture, requiring GPS for sensing the position to which any measured field parameter must be geo-referenced, was implemented for the first time in the US. Mapping technology was introduced in Australia in the early 90s, and the first combine harvester mounting yield mapping technology was sold in 1997 in Europe. Since then, further improvements happened. In short, precision agriculture can be summarized as the right input, the right amount, at the right time and on the right spot.

The first wave of the precision agricultural revolution came in the form of satellite and aerial imagery, weather prediction, variable rate fertilizer application, and crop health indicators. Agriculture 3.0 was about the realization of gains deriving from early precision agriculture. The target of efficiency-based cost reductions was later replaced by profitability based on creative solutions, improved quality, or the development of new products (Jóri 2017). Meanwhile, the second wave of precision farming (that can be referred as the 4th wave of agriculture) – that is induced mostly by the earlier mentioned measures and solutions of 4IR – will aggregate the machine data for even more precise planting, topographical mapping, and soil data. So, since the new millennium, with the introduction of new technological advances, new machinery, and GPS tracking, each farmer will be able to feed about 265 people on the same unit of land by 2050.

With the use of Industry 4.0 technologies like artificial intelligence and the use of big data, and other methods, the potentials may increase further. The improved remote sensing and heading control (automatic steering), yield monitoring, diagnosis, measuring soil conditions, diagnosing harvest time, and monitoring crop health status, and the process of data originating from them are now also involved (Sung 2018). It contributes to more precise farming activities: instead of calibration of machines for

---

[8] Euractiv, 2016, Farming 4.0: The future of agriculture?, Available at https://www.euractiv.com/section/agriculture- food/infographic/farming-4-0-the-future-of-agriculture/
Euractiv, 2016, 'Farming 4.0' at the farm gates, Available at: https://www.euractiv.com/section/agriculture- food/opinion/farming-4-0-digital-technology-at-the-farm- gates/

one single but larger unit and instead of treating the whole livestock, it supports the adaptation to the spatially variable soil, water, and crop parameters within a field, and to the unique treatment of an animal (a specimen). Moreover, an additional important step is the real-timeliness of farming. Data collection and its processing happen immediately, data-based formulation of adequate decisions and delivering them are not separated by time from data collection. There are several machines, appliances, and applications that currently can communicate with each other, sometimes even without the need for human intervention (Bőgel 2018).

Nothing shows the expansion of this special market better than the numbers of sales of machines equipped with advanced and interconnected technologies. The increasing focus on farm-efficiency and productivity is expected to induce the growth of the agricultural robot market. Of the industries facing automation, agriculture could see the most benefit from robots over the next few years. And the farming robot wave, along with other new agricultural technology, could come even sooner, and with a bigger increase[9]. According to business data, the agricultural robot market is expected to grow from USD 2.75 billion in 2016 to USD 12.80 billion by 2022, at a Compound Annual Growth Rate (CAGR) of 20.71% between 2017 and 2022[10]. According to the GSA (2018), the CAGR of the global precision agriculture market will be around 12% through 2020, whilst the total market value will surpass USD 5.5 billion by then.

3.3. The benefits and potentials in precision agriculture based on Hungarian examples

Besides the subject of precision agriculture introduced earlier, its economic impact should also be emphasized. Its impacts can both be measured from profitability and from the return on investments (investment recovery) aspect. The present study focuses on researches made in Hungary since the empirical research presented later, which has also been delivered in this environment. Moreover, the focused studies presented on the potential impacts of introducing PA in arable crops production.

According to the study of the Hungarian Research Institute of Agricultural Economics (Kemény et al. 2017), the value of the overall economic impacts of transformation to PA was about 6.9% of 656 farmers/holdings reported. More precisely, *yield growth* (wheat: 7–17%, corn: 2–9%, sunflower: 6–10%) was detected (Kemény et al. 2017) together with the controlled and restrained environmental impacts also mentioned in the study by Hart and Bas-Defossez (2018).

When we come to *investment recovery*, the same research (Kemény et al. 2017) found that the investments targeting software and precision tools were recovered in farms both over and under 1000 hectares. But of course, differences were highlighted: in case of an exchange of complete machinery, smaller farms should face investment of over 500,000HUF (EUR 1,600) per hectare, meanwhile agricultural

[9] https://www.investors.com/news/farming-robot-agriculture-technology/
[10] Data from https://www.marketsandmarkets.com/Market-Reports/agricultural-robot-market-173601759.html

holdings extending over 1000 hectares have to invest around 328,000HUF (a little more than EUR 1,000) per hectare. Such investment would induce about a 6 (small farms) to 9% (holdings) increase in yearly incomes.

Another study by Lencsés (2013) also shows the differences between the potentials in investment recovery. It investigates the potentials by already differentiating by the size of the farms, their investments, and by the future prospects being them pessimistic, realistic, or optimistic. The results show that investments to introduce PA technologies are recovered only in case of medium and large agricultural holdings. These findings align with the results of Smuk et al (2009) which proved by modeling that the investment recoveries are highly dependent on, and show a strong correlation with, the size of the farm/agricultural holding. This is not a surprising result when considering that despite the advantages of Industry 4.0 for large enterprises, small and medium-sized enterprises (SMEs) often face complications in innovative processes due to the continuous development constraint in innovations and technologies (Zambon et al. 2019). This fact increases the complexity of the problem and the challenge that decision-makers face if they are willing to improve the innovative background of Hungarian agriculture in total, and not just of those bigger holdings that are already capable of improving.

4. Hungarian agriculture

As mentioned earlier, Member States (MSs) face different pressure (pushing) and challenges (pulling forces) but in the meantime, they also have limited means (lack of traditional own agricultural support system) to improve their positions. For Hungary, it would be an important task to improve especially the competitiveness of its agricultural sector. As Bojnec and Fertő (2008), Csáki and Jámbor (2013) revealed, after their accession to the EU in 2004, Hungarian farmers and the agricultural sector faced serious difficulties and remarkable reductions in the competitiveness of agricultural products both in price and quality, especially in the field of higher value-added, produced products. Altogether, there are several reasons which allowed ineffectiveness to pervade Hungarian agriculture of the present. Among other things, the following should be highlighted, since they are related to policy decisions:

(1) the traditional lack of networking and cooperation in agricultural activities (Juhász 2016),
(2) the wrong-headed agricultural policy that was not effective in exploiting the beneficial use of pre-accession aids coming from the SAPARD fund (Csáki and Jámbor 2013),
(3) the pre-accession policy targeting mostly the artificial creation of competitiveness through price and market support, that neglected measures for inducing agricultural competitiveness (Csáki and Jámbor 2012),
(4) the decision to implement the simplified Single Area Payment Scheme (SAPS) embodied in the Copenhagen Agreement in 2002

(5) and last but not least, during the first Multiannual Financial Framework Hungary contributed to as a member (the end of 2000–2006 and 2007–2013) we gave priority to direct agricultural support (60%) instead of support for rural development focused on the development of competitiveness and agricultural environment (Csáki and Jámbor 2012). Later in the present period of 2014–2020, this rate increased even further to 72%.

Besides these policy conditions, the decisions of farmers in their switch to producing higher value-added products is also determining their future. Here the technological background, professional knowledge, and the range of investment opportunities are all having remarkable impacts. With the above-mentioned unfavorable policy heritage, the changing technological environment and the *pushing and pulling drivers especially of businesses* should somehow be addressed in the near future. It would be necessary if Hungary wants to improve its position on the EU's internal market of agricultural products, and so, indirectly, conditions for Hungary's agricultural population - performance as *a driver of development*.

4.1. Hungarian agriculture's position in the economy

Besides the unquestionable importance of agriculture in food production, in producing basic raw materials for the secondary sector, and in its role in rural areas, its contribution to the whole economy can be also measured. Based on the indicators shown in *Table 1*, a balanced performance is visible.

Table 1 Basic indicators of Hungarian agriculture (as total contribution, year by year)

	2010	2011	2012	2013	2014	2015	2016	2017
Contribution to GDP (%)	3.0	3.9	3.8	3.9	4.0	3.7	3.7	3.3
Contribution to gross value added (%)	3.5	4.6	4.6	4.6	4.7	4.4	4.4	3.9
Employment (%)	4.6	4.9	5.0	4.7	4.6	4.8	5.0	5.0
Investments (%)	4.8	5.6	5.8	5.9	6.0	4.8	5	4.8
Trade balance (surplus) of agricultural products (billion HUF)	580	763	1007	992	900	889	806	915

Source: own construction based on the data of the Hungarian Central Statistical Office (2018–2019)

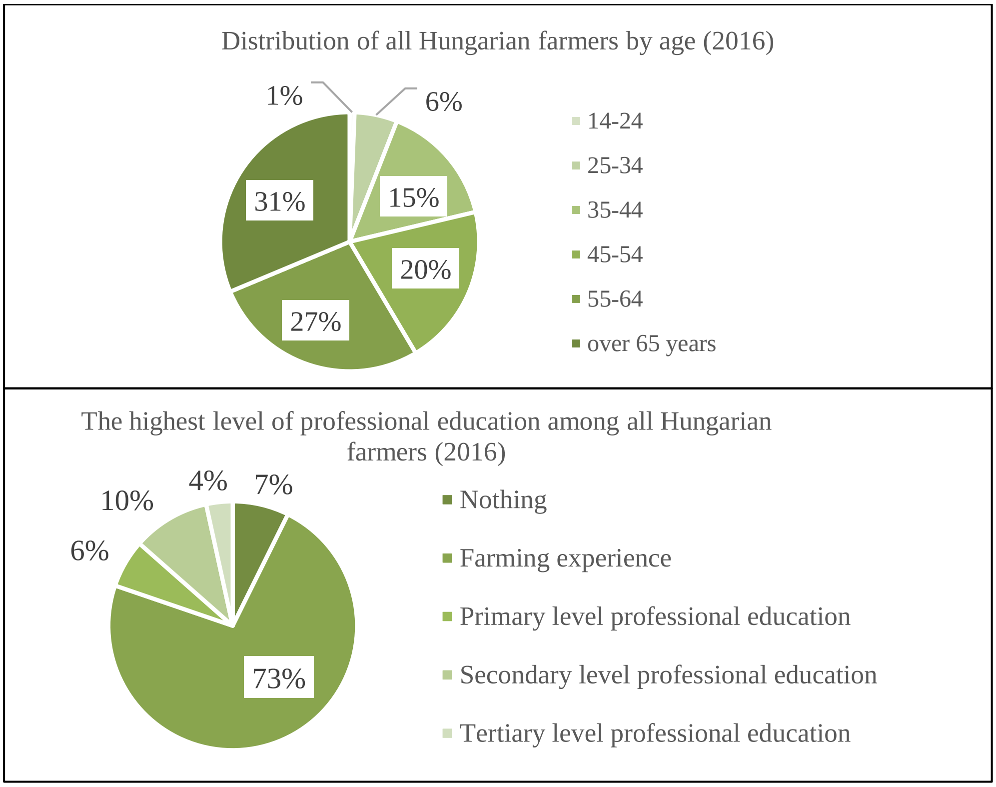
The contribution of Hungarian agriculture to GDP was 3.3% in 2017. The same indicator for the EU average was 1.2%. In agriculture's contribution to gross value-added, Hungary reached 3.3% (with only Romania, Bulgaria overtaking it in this field, and with Greece attaining the same rate), although meanwhile, the same indicator for the EU-28 was 1.3% in 2018[11].

Agricultural employment in Hungary is again over 5%, which is higher than the EU-average, but this could be a result of whitening the sector by new regulations

[11] Calculations based on Eurostat data using gross value added and income by A*10 industry breakdowns [nama_10_a10], while the definition of the agricultural industry is based on Division 01 of NACE Rev. 1.

providing easier, less complicated part-time and seasonal employment in agriculture. What is more important is the distribution of Hungarian farmers by age and by educational background (Figure 3). The Hungarian agricultural population coincides with the EU average as the majority of farmers are above 55 years. More problematic is their professional education. Compared to the rate of tertiary education graduates with respect to the total population of Hungary, which was around 18% in 2016.[12] tertiary level education is boasted by only 4% of the agricultural population. It could be a promising sign in connection with this indicator that, in the age group of 25–34 years, this rate is above 6%.

Figure 3 The distribution of Hungarian farmers by age and educational background



Source: own construction based on the data of the Hungarian Central Statistical Office (2018)

Furthermore, there is a remarkable difference in the educational background of farmers and managers of agricultural holdings. Due to a lack of recent data, only those from 2010 can be compared: 2.6% of the former group finished higher education, compared to the 44% of managers of agricultural holdings.

[12] Data is from the Hungarian Central Statistical Office.

4.2. Overall performance of Hungarian agriculture

The educational background, besides several other factors, has a huge impact on the productivity of agriculture. A complex indicator of this is the *total factor productivity (TFP)*, which according to the Eurostat shows the ratio between the change in production volumes over a given period and the corresponding change in inputs (or factors) used to produce them, and hence measures the growth in productivity over a given time span. It became a key indicator of the economic performance of agriculture and a driver of farm incomes and shows how efficiently the agricultural sector uses the resources that are available to turn inputs into outputs. A change – growth or decline – in TFP results predominantly from the change in public investments in infrastructure (irrigation, electricity, roads) and in agricultural research and development, from the changing efficiency in the use of water and plant nutrients, from the introduction of new technologies, managerial skills, etc. Although there is a general belief in the progress of technology in improving resource efficiency, the EC (2016) acknowledges that this is challenging in agriculture, as working with living organisms in outdoor conditions introduces variability and limits to growth.

Overall, between 2014 and 2016, the EU-28 experienced a 2.4% growth in TFP. In this period, post-2004 MSs (with an increase by 4.8%) narrowed the productivity gap and approached the higher TFP level of the pre-2004 MSs (which achieved a lower, 2.0% growth in TFP) of the EU. Besides the relative distance from the technology frontier, other drivers of this increase might be in many cases increasing labor productivity, but also improvements in yields (EC 2017). In the aforementioned period, Hungary achieved a 9.2% increase, meanwhile, the average annual change in TFP between 2006 and 2016 was a little bit over 1% (EC 2017). By these numbers, Hungary ranks among the low growers' group with Slovakia and the Czech Republic (EC 2016), but Baráth and Fertő (2016) also put Poland in this group. The Hungarian performance can be explained by the fact that per capita public expenditure on R&D in agriculture in 2014, the same as in 2005, was the second lowest for Hungary as member of the EU (EC 2016). It would, however, be essential to improve technological development in order to increase TFP. In the post-2004 MSs, it is important, because there is a lag behind developed MSs in technological development (Baráth and Fertő 2016).

According to the *Agriculture Performance Index* developed by Jámbor et al. (2016), Hungarian agriculture has been in the medium/average performers' group both by productivity and by efficiency sub-indicators in the examined period of 1999 to 2013. This research served as evidence to the earlier study of Csáki and Jámbor (2013), which claimed the comparison of MSs showed that those countries that made a switch to higher value-added agricultural activities (like dairy sector and fruit production) during their transition period, before their EU-accession, could later produce better performance.

Both the latest cited studies agree that Hungarian agriculture – to b more precise, its decision-makers – could not take the advantage of the opportunity to take advantage of the possibilities of the first decade of EU membership. Differences in the quantity and the quality of land/soil, in the labor force, and the capital background, all had a significant role in later differences in the agricultural performance of the countries. In Hungary, the capitalization of agriculture did not increase in the period between 1999 and 2013 (Jámbor et al. 2016). Regarding farm structure, the duality revealed was the insular presence of big, capitalized, more efficient farms surrounded by small, ineffective farms run by farmers usually of a low educational background. Furthermore, in contrast to the Czech and Polish examples, the quick implementation of privatization processes in the agri-business, resulting in significant foreign ownership of the food processing industry, to the detriment of local farmers in the early 1990s, also did not favor later agricultural development. Overall, the lack of transparent, comprehensive, and well-thought-out Hungarian agricultural policy resulted in the absence plans for the future (Potori et al. 2013, Jámbor et al. 2016) and the failure of readiness/preparedness for increased competition after accession (Csáki and Jámbor 2013).

5. Farming 4.0 in practice – an empirical research

At the beginning of 2019, we distributed a survey among Hungarian agricultural enterprises and farmers asking them about their situation, opportunities, and needs in connection with Industry 4.0 technologies. Our main questions on which the empirical research was based were as follows:

– What are the main motives/incentives of Hungarian farmers in using advanced technologies during their agricultural activities?
– Will such advanced technologies replace human interaction in the future of farming?
– What are the main obstacles and limits in the spread of technological development in Hungarian agriculture?
– What measures might support the technological development of Hungarian agriculture?

5.1. Research method and sample

Survey research using a predefined series of questions was used to collect information from farmers and managers of agricultural holdings. The questionnaire contained approximately 30 closed questions that ranged from the basic topic, like the form of their activity, their profile, and their size, to questions in connection with the research topic of the present article. With regard to the technologies used, motivations, suggested supports, etc. the respondents could mark more answers.

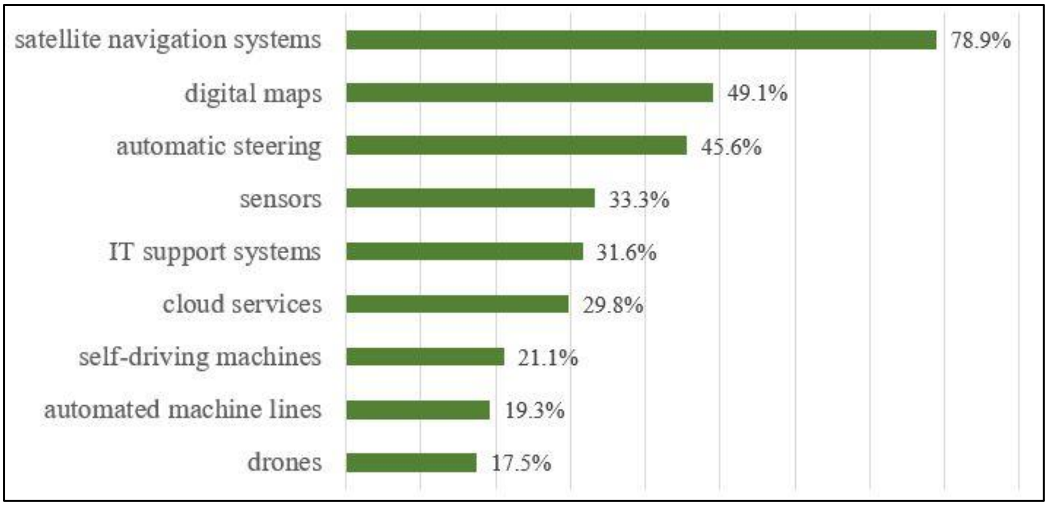
The questionnaire was sent to farmers, managers and other stakeholders primarily via the internet, social media groups, and e-mail through an advisory network of the Hungarian Chamber of Agriculture. Questionnaires were also filled out personally at professional events and agricultural forums. Altogether 134 questionnaires were filled out. During the analysis of the results, the limitations of the results became evident. The educational background of the respondents was not in accordance with what is described earlier in the present article about the Hungarian agricultural population. The distribution by age of the respondents is: 16.4% between 18–30, 23.9% between 31–40, 25.4% between 41-50, and compared to the total sample, only 34.3% were over 50 years. Altogether 81.3% (27.1 and 54.2%) of the farmers who apply advanced technologies in farming were in the below 40 years category. What is even further from the Hungarian reality – as presented in sub-chapter 3.1 – is that 78.4% of farmers in the survey had tertiary level professional education. So, *the results of the survey must be interpreted as messages containing urgent deliverables, since they come from the well-educated and younger agricultural generation of Hungary*.

5.2. Results of the research

More than half of the respondents, 56.7 percent said they did not apply the listed technologies (among others: satellite navigation systems, digital maps, automatic steering, sensors, cloud services, drones, etc.) in their business activities. The rest of the respondents, 42.4 percent, use at least one of the listed technologies in their business activities according to the distribution depicted in Figure 4.

Most of the advanced technology users implemented satellite navigation systems (78.9%), digital maps (49.1%), and automatic steering (45.6%). Approximately one-third of users applied sensors (33.3%), IT support systems (31.6%), and cloud services (29.8%), and nearly one-fifth of them used self-driving machines (21.1%), automated machine lines (19.3%), and 17.5% of them use drones.

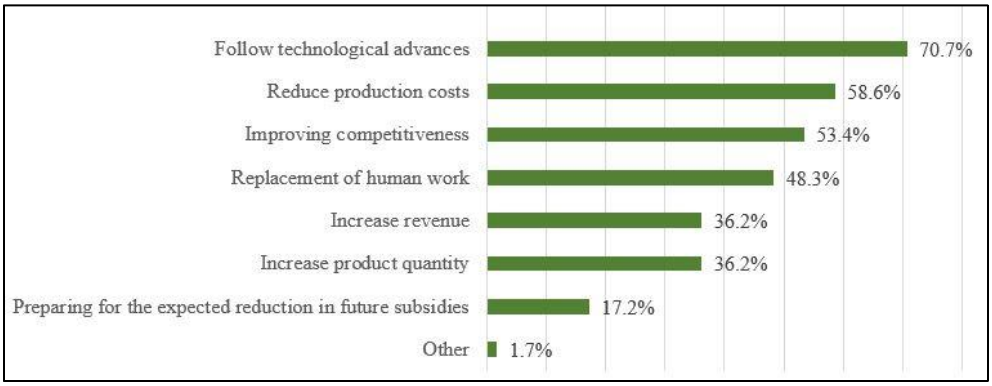
Figure 4 The distribution of applied advanced agricultural technologies



Source: Own construction based on the empirical research

One surprising result was in connection with the motivations in the implementation of advanced technology. Besides the “general” incentives of reducing the production costs (58.6%) and increase revenue (36.2%), the wish to follow the recent trends in technology was in first position (70.7%). The potential to improve micro-level competitiveness (53.4%) was also placed high among the motivations.

Figure 5 The motivations of actors using advanced technologies in agriculture



Source: Own construction based on the empirical research

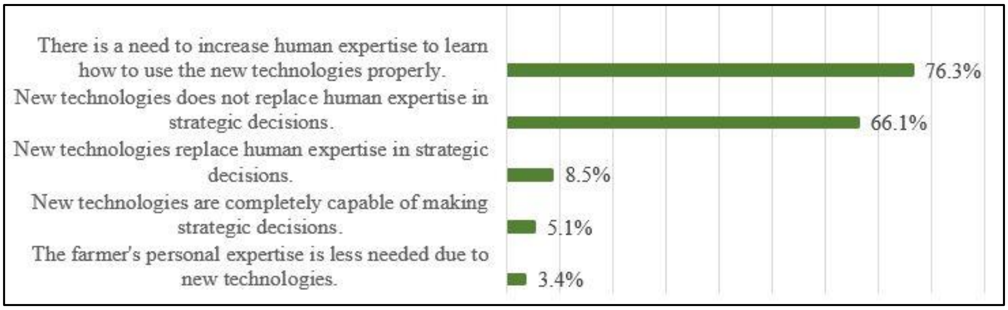
Although in sub-chapter 3.1 it was mentioned that agricultural employment increased slightly, here the intention to replace human labor (48.3%) is in line with the recent news about labor shortages in agriculture.[13] Given the aims of the present article, even the answer “preparing for the expected reduction in agricultural supports” reaching 17.2% is an important sign. In the subsequent discussion of the compliance of EU CAP with productivity and efficiency requirements, this will be further detailed.

In connection with the required knowledge in the use of advanced technologies the research brought the same results as other research about the impacts of technological advancements in other sectors (Szalavetz and Somosi 2019). Here the role of knowledge in farms using Industry 4.0 technologies has evolved and increased (Figure 6).

More than three-quarters of respondents said that there is a need to increase human expertise to learn how to use the new technologies properly (76.3%). More than two-thirds said that the use of new technologies does not replace human expertise in strategic decisions (66.1%). Therefore, the majority of responding agricultural enterprises and farms said that the role of human expertise has increased as a result of using new types of technologies, and not reduced, and only in a low proportion (8.5%), was the perception that new technologies replace human expertise, or the farmer’s personal expertise is less needed due to new technologies prevalent (3.4%).

[13] https://www.agrotrend.hu/hireink/a-mezogazdasag-a-hatalmas-munkaerohiany-erzekeny-terulete

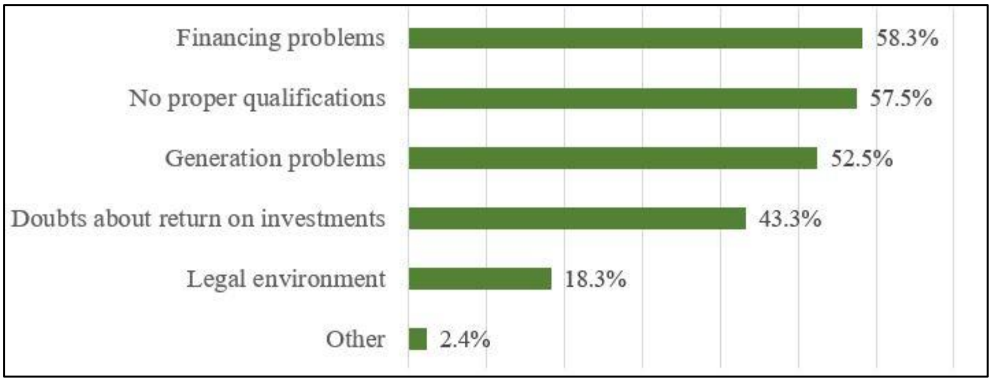
Figure 6 The role of human expertise in connection with advanced technologies



Source: Own construction based on the empirical research

Another focus of the research was the revelation of the potential obstacles of the spread of advanced technologies supporting agricultural activities in Hungary. As shown in Figure 7, financing problems (58.3%) were indicated as the first obstacle, followed closely by the lack of proper qualifications (57.5%) and generation problems (52.5%). Hence, the aging Hungarian agricultural population, which is further hampered by a lack of educational background (highlighted in sub-chapter 3.1), is visible on the micro-level after all. In addition to these three factors, there was a significant proportion of doubts about return on investment (43.3%). This is in line with the afore-mentioned studies by Smuk et al (2009), Lencsés (2013) and Kemény et al. (2017), which showed investment recoup depends on the size of farms and the rate of necessary investments, and may range only from 6-9% increase in the future incomes.

Figure 7 Obstacles to the spreading of Industry 4.0 technologies in agriculture

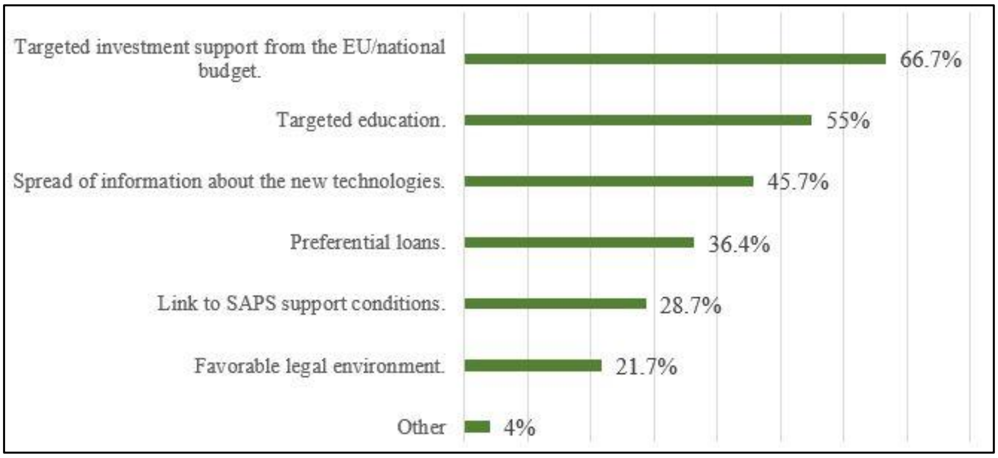


Source: Own construction based on the empirical research

According to nearly one-fifth of the respondents, the legal environment[14] may also be an obstacle to the spread of new technologies. Finally, 2.4% of respondents mentioned "other" obstacles, such as corporate culture, expensive data traffic, lack of online coverage, fragmented fields.

When the respondents arrived at naming the supporting factors of the spread of advanced technologies in agriculture (Figure 8), most of them indicated that targeted investment support from the EU/national budget (66.7%), targeted education (55%), spread of information about the new technologies (45.7%), and preferential loans (36.4%) would support its wider use. Linking SAPS support to investment in advanced technologies was mentioned by almost one-third (28.7%) of the respondents, but others especially highlighted that this would not be a favorable change in the SAPS system since it is already complicated enough to meet all its conditions. Mentioning the creation of a favorable legal environment is in accordance with the difficulties that derive from the aforementioned inadequate regulation.

Figure 8 Key factors in the spread of Industry 4.0 technologies in agriculture



Source: Own construction based on the empirical research

Here, the "Other" factors (4%) were responses such as: reducing costs, increasing purchase prices, and creating a professional, independent advisory network. Based on the answers and the deeper personal discussions three parties are likely to emerge: one would require professional assistance from the part of the authorities and the relevant ministries, the second would like to see an increase in financial support, and the last group would ideally like a combination of the two.

[14] According to the respondents' experiences from abroad, in general, it is a price increasing factor if data of soil, irrigation, and past and future potentials in yields are collected in a time series for a special unit of land. But recent Land Law regulations in Hungary – besides its other anti-market regulations – does not support the increase of value embodied in the price of land. Others mentioned the regulation of drones for agricultural purposes or water management regulations.

6. Compliance(?) of CAP and Hungarian measures with the necessary technological development

When one intends to map the potential factors of technological development in agriculture, the different levels and their impacts should all be analyzed. Governance issues, infrastructure, connectivity, data ownership, but also the emphasis on specific sectors like agriculture are – beside the increasing impact of international processes derived from the most opened markets ever – being shaped by the private sector, governments, national strategies, legal systems, but also by regulations of the EU and other international organizations as well (FAO 2017). For example, even the World Economic Forum is currently supporting agricultural transitions in 21 countries through its "New Vision for Agriculture" initiative[15].

The present study is not putting the analysis of the EU's CAP into main focus. Other studies (Gorton et al. 2009, Möllers et al. 2011, Popp and Jámbor 2015) have done that, even with an emphasis put on the impacts of the CAP on the agriculture of post-2004 MSs. Gorton et al. (2009) analyzed that the CAP is not suitable for this post-socialist region, and they even listed the reasons for that. Later Popp and Jámbor (2015) also found that the CAP was obviously hardly able to meet the challenges it faced before and during the present 2014-2020 period due to the inconsistencies between the predefined challenges and the measures proposed to meet them with respect to the territorial imbalances of the EU's agricultural sector. The ongoing challenges in conjunction with new ones (technological readiness as a tool to be able to improve competitiveness) and the design of the forthcoming MFF also raise doubts in connection with its future success.

Just to highlight some private actors among the pushing forces of business drivers we can mention service providers, actors from manufacturing and supply, participants in the Agrifood industry, research institutes and innovators, and finally the policy creators. These actors are split by their scale of operations from the farm level, through regional, national, European levels all the way to the global level (GSA 2018). Increasing input prices as a pressure on farmers are transmitted by these agricultural input suppliers. Among the incentives of business drivers to keep pace with the innovative core of agriculture, one could cite the increasing number of Ag-Tech companies providing farm management solutions and applications.

Without aiming to give an exhaustive analysis, only a few stakeholders influencing the basic factors of agricultural improvements/development through their regulatory and financing potential will be mentioned, with respect to the results of the empirical research.

6.1. Compliance of CAP with the present technological environment

Among the challenges of adopting Farming 4.0 technologies from the perspective of the farmers, we can mention the necessity of common standards (1), the ability of

[15] http://theconversation.com/the-battle-for-the-future-of-farming-what-you-need-to-know-106805

farmers to modernize (2) and the modernization of infrastructures (3) (EC 2017b). The EU and its policies have something to do with all three areas. Here the focus is put on the ability of farmers to modernize.

After the latest CAP-reform concluded in 2013, the Commission initiated an action to collect ideas on the "modernization and simplification" of the CAP. A communication was published in 2017 titled "The Future of Food and Farming", which outlined the Commissions ideas for further CAP reform to coincide with the introduction of the next multi-annual financial framework (MFF) of the EU for the period 2021–2027. The objectives of the Common Agricultural Policy for the upcoming period are:

- ensuring fair income for the actors in the sector (1);
- improving competitiveness, including by prioritizing the development of research, technology and digitalization (2);
- improving the position of farmers in the food chain (3);
- action against climate change (4);
- supporting efficient management of natural resources (5);
- preserving biodiversity (6);
- generational renewal support (7);
- developing rural areas (8); and
- promoting sustainable production of safe, nutritious food (9).

As is widely known, the CAP as % of the MFF will receive 28.5% compared to the present period's 35.3% portion (Matthews 2018). In respect to this, it should be mentioned that the role of subsidies in the future of farmers could be ambiguous. As Pechrová (2015, 16) wrote based on the Czech example, direct subsidies "can improve farms' viability, but may mitigate farmers' motivation to engage in efficient resource usage." If the criticism, that CAP preserves conditions of ineffective farming are valid, such a cut in support may contribute to the necessary endogenous improvement of productivity implemented by the farms themselves. The use of advanced technology may later supplement the lower direct financial support.

In addition, about 10 billion EUR will also be available in the Horizon Europe research program during the next MFF, although not as a direct income substitution. This will support research and innovation activities in agriculture, food, and rural development. The agricultural **European Innovation Partnership (EIP-AGRI)** combines the resources of Horizon Europe and Rural Development to contribute to the digitization of agricultural holdings in rural areas, including the spread of precision farming techniques, and thus to contribute to sustainable and competitive agriculture (EC 2018). It aims to foster competitive and sustainable farming and forestry that *"achieves more and better from less"*.

Beyond financial support, it is also essential to build a stronger innovative network as well. There are several initiatives, like the **Agricultural Knowledge and Innovation Systems (AKIS)** with which the EU aims to boost initiation and development of innovation projects, to disseminate their results as best practices to support their use as widely as possible.

Overall it is visible that the new EU agricultural policy aims to strengthen actors' capacities for innovation by taking into account the complexity (as mentioned also in the introduction part) of innovation processes. Faure et al. (2019) summarized the key innovation support services (ISS) that help actors in agriculture-related innovation. Even in this specified case, their results show that ISS depends on the phase of innovation. During the initial phases – where the majority of Hungarian farmers are –, there is a need for innovative support services (e.g. network building, financial and other support for the innovator). In the latter phases – where some best performer, Hungarian insular agricultural holdings are –, there is a need for more conventional services (e.g. training, further credit for constant development). Moreover, the assistance is needed at both farm, value chain, and territory level, and from different – national, community, public or even private – sources. Overall, given increasing input costs and aging agricultural population, it is a necessity for the whole EU to improve efficiency, mainly through innovation, but from a MS perspective, the competitiveness of their farmers within the EU's Internal Market for agricultural products is the challenge.

6.2. The compliance of the Hungarian "environment" with the requirements

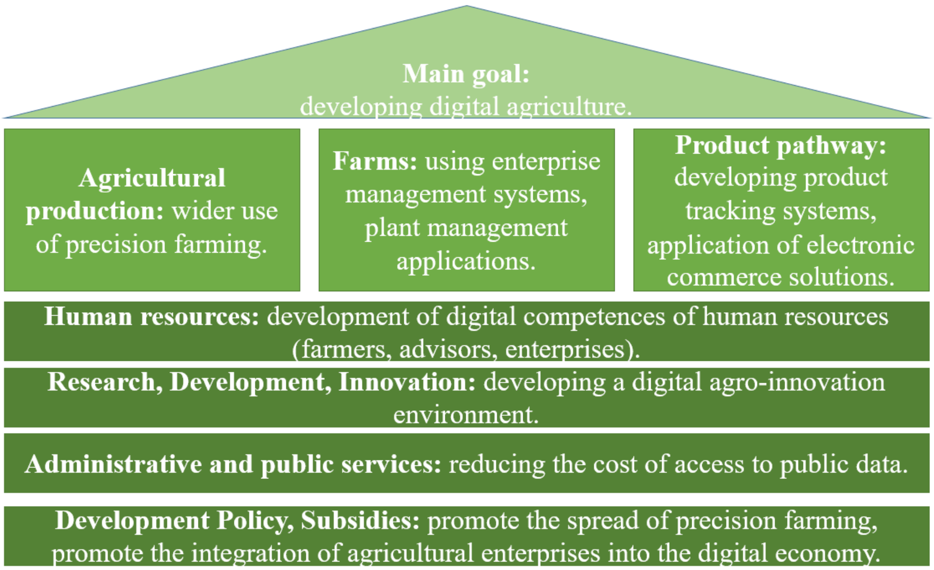
The large input suppliers (like equipment manufacturers and the agro-chemical industry) and the downstream food processing and food service firms generally have the capacity and the background to adopt and to adapt ICTs to address the needs of farmers and consumers. The farm sector, on the other hand, is characterized by the presence of SMEs, many of which may face difficulties in the adoption of ICT-based solutions (FAO 2017). Beyond EU-level actions, government policies can enhance the widespread adoption of appropriate ICTs, including the facilitation of access to "hard" (physical) and "soft" (skills and technical support) infrastructure.

In connection with the necessary and widely accepted mode of improving "soft" skills generally on a domestic level, Gunasekera et al. (2018) emphasize the opposite connection also between agriculture and education. They highlighted the potential use of IoT in agricultural education, as an incentive even to attract new, younger generations. The same impact occurs when digital games are used not only for entertainment but when they are designed for training the player in solving problems in different fields of life or the economy. As the FAO (2017) also highlights, these are simulation platforms that allow learners to experience a number of scenarios and situations, and provide solutions, having a positive impact on analytical skills, learning and recollection abilities, problem recognition and problem-solving.

With a narrower focus on agriculture, in respect of the level of innovation, and the necessary contribution of different (even national) policies and measures, it is a novelty that the Commission proposed *National CAP Strategic Plans* to be developed by the MSs for the next period. Transparently and with the involvement of relevant actors a comprehensive agricultural policy strategy should be developed at the national level in which MSs plan how their measures will contribute to the objectives of the Common Agricultural Policy (EC 2018). This is a key new instrument in the EU's legal proposal for the CAP post-2020 to shift EU agricultural policy to a more performance-based framework (Matthews 2019). Besides this new leeway in the policy with respect to the national conditions, there are no further national possibilities in agricultural policies. Independently from the direct agricultural payments, the targeted financial support of technology improvements remains the sole tool in the hand of MSs to improve agriculture's and its actors' performance.

The *Hungarian Digital Welfare Program* (mentioned later as DJP 2.0 (2017)) serves as a framework to develop the previously mentioned soft skills, the digital competencies, to promote the digital transformation of Hungarian industry and agriculture, with special regard to domestic SMEs and micro-enterprises. The aim of the program is to help businesses take advantage of the benefits of digitization to increase their competitiveness. It emphasizes the role of digitalization also within agriculture, as automatization is a basic condition of increasing productivity and efficiency. The program makes proposals for the digital development of agriculture, such as preparation of Digital Agricultural Strategy, creation of a Digital Agricultural Academy, training of digital consultants and advisors, preparing an Agricultural Data Integration Program, digital development of agricultural enterprises, all encouraging the better exploitation of potential in Hungarian agriculture. Here, it should be emphasized that agricultural advisors and consultants from the national network are supposed to provide trustworthy, independent support, impartially and independently from any input supplier, application provider or dealer. The goal of the Digital Agricultural Strategy (DAS) is to transform and develop the agricultural economy based on digital, intelligent tools and solutions. The Strategy defines three pillars and four cross-bars, as shown in Figure 9.

Figure 9. The structure of the Digital Agricultural Strategy (DAS)



Source: own construction based on Digital Agricultural Strategy (2018)

Beyond the above-mentioned programs, the government has another measure to improve the capitalization of Hungarian agricultural development. Recently the investment-friendly environment contributed to the increasing amount of outstanding credit. Furthermore, within the *Growth Loan Program* (Növekedési Hitelprogram) altogether 480 billion HUF loans were outsourced between 2013–2017, within which the majority – over 50% – has been spent on investments[16] largely directed at technological and efficiency improvements[17].

7. Conclusions

At the beginning of our study, we put emphasis on the present challenges of agriculture, with a focus on Hungarian farmers. As discussed, there are more ways to address these challenges. Here our focus was on the technological development of agriculture, from which the EU and Hungary are both aiming to get the desired outcome. In the following, the development phases of agriculture – from the technological aspect – and industry have both been introduced and compared. It is shown that their point of intersection is in so-called precision agriculture, but recently with constant technological improvements. Industry 4.0 has brought the real-time capability of reacting and making decisions easier based on data collected due to the interconnectedness of machines, so, as the WEF summarized, it resulted in the "*fusion of technologies that blurs the lines between physical, digital and biological domains*"[18] known widely as Farming 4.0.

The aim of the present study was to see how Hungary is performing in this field. Our empirical research have helped us draw conclusions and create suggestions specific to Hungarian conditions. Especially because the results of the survey had to be interpreted as messages containing urgent deliverables, since they come from the well-educated and younger agricultural generation of Hungary, which meanwhile boasts genuine insight into both the needs and possibilities. As found, there would be an *increasing need for advanced technology* in agriculture because of economic, ecological and social reasons, and although for slightly different reasons, both in the EU and Hungary.

In connection with the role of human expertise in Farming 4.0 technologies, empirical research has revealed different results. As the literature summarizes, level of knowledge already plays an important role in the spread of advanced technology in the primary sector. This has been supported by the field research as well. Compared to the threats arising in the secondary sector in connection with the impacts of Industry 4.0, *technology is not likely to replace personal knowledge and experience* in agriculture. But the required level of knowledge, the need for a higher skilled and better educated workforce, is constantly increasing.

---

[16] Data from https://nhp-hitelek-statisztika.mnb.hu/reports/powerbi/NHP?rs:embed=true

[17] https://www.agrarszektor.hu/agrarpenzek/a-2019-es-ev-is-mozgalmasnak-igerkezik-a-mezogazdasagban.12953.html

[18] https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

Having gained an insight into Hungarian conditions and the impacts of Common Agricultural Policy, it is also a possible revelation that the constraints – labor shortage, water scarcity, increasing input prices, decreasing direct agricultural subsidies – may have a bigger and more effective impact on the future development of Hungarian agriculture compared to the personal intention of catching up with the latest technological developments. The findings of the present research were different from this, but we must not forget that our research summarized the views of the “top farmers” of Hungary: younger and better-educated farmers compared to the Hungarian average agricultural population. On a national level, the constraint is to become more effective and competitive within the Internal Market. If we accept that Farming 4.0 is an adequate and widely accepted measure of this, the improvement of the educational, information, knowledge level, and awareness of the farmers is essential. The Hungarian education system thus has to focus on the improvement of overall readiness in the use of Industry 4.0 technologies with special regard to the agricultural sector as well. If it is possible, here both educational background and the acceptance of advanced technologies is lagging behind levels for the same indicators in society as a whole. Farmers should be educated to see the contribution of the use of advanced technologies not as threats, or as is the general opinion of the majority: as “a harmful difficulty”. Attracting younger generations would be of the highest importance, but here the controversial effect is also revealed. Based on the surveys, it is also found that the younger generation considers farming with advanced technologies a more attractive form of work, which traditionally is said to be hard.

Beyond appropriate education and the involvement of new generations, the adequate financial background is also a factor in advanced technologies in agriculture. The lack of financial background should be handled by a re-structured system of financial resources. It is visible that the decreasing level of future direct supports of EU-origin is going to be partially replaced by other sources, like the Horizon program. At this point our opinion considering the EU’s plans are ambiguous. On the one hand, we support the transition to an EU-level system which instead of conserving ineffective farming conditions tries to contribute to the necessary endogenous improvement of productivity implemented on farm level. But on the other hand, we fear that such a shift in the supporting system favors those farms that are already at a certain level of development, which is like the “tip of the iceberg” in the case of Hungarian agriculture. If this happens, it will increase the necessity of domestic contributions to the development of the sector.

The EU’s subsidies should be – as they partially have been – replaced by widening investment measures on the national level, be they public or private. Thus, the objectives of Digital Agricultural Strategy of Hungary should be coupled with domestic financial support. The favorable loan environment may not be enough or not be sustained in the future, but within the framework of research and development programs, the EU and Hungary are both able to support the technological advancements of farmers without harming international trade processes. In this case the “only” factor that may hinder success is the knowledge level of farmers.

Acknowledgments

This research was supported by the project nr. EFOP-3.6.2-16-2017-00007, titled Aspects on the development of intelligent, sustainable and inclusive society: social, technological, innovation networks in employment and digital economy. The project has been supported be the European Union, co-financed by the European Social Fund and the budget of Hungary.

References

Babiceanu, R. F. – Seker, R. (2016): Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *Computers in Industry*, Vol. 81., 128–137.

Baráth, L. – Fertő, I. (2016): Productivity and Convergence in European Agriculture. *Discussion Papers*, MT-DP – 2016/26, August, Hungarian Academy of Sciences.

Bojnec, Š. – Fertő, I. (2008): European Enlargement and Agro-Food Trade. *Canadian Journal of Agricultural Economics*, 56, 4, 563–579. http://dx.doi.org/10.1111/j.1744-7976.2008.00148.x.

Bőgel, Gy. (2018): Digitális transzformáció a mezőgazdaságban. (Digital Transformation in Agriculture). *Magyar Tudomány*, No. 5. doi: 10.1556/2065.179.2018.5.12

Bresnahan, T. F. – Trajtenberg, M. (1992): General Purpose Technologies: "Engines of Growth?". *NBER Working Paper Series*, No. 4148, August.

Comparetti, A. (2011): *Precision Agriculture: Past, Present and Future*. Conference Paper, September.

Csáki, Cs. – Jámbor, A. (2012): Az európai integráció hatása a közép- kelet-európai országok mezőgazdaságára. *Közgazdasági Szemle*, 59,7–8, 892–910.

Csáki, Cs. – Jámbor, A. (2013): The impact of EU accession: lessons from the agriculture of the new member states. *Post-Communist Economies*, 25, 3, 325–342. http://dx.doi.org/10.1080/14631377.2013.813139.

DJP 2.0 (2017): *A Digitális Jólét Program 2.0.* Budapest. Available from: https://digitalisjoletprogram.hu/files/40/49/404929cfc6bdda1988f927bf1af59ffa.pdf

Dudley, L. (2010): General Purpose Technologies and the Industrial Revolution. *Papers on Economics and Evolution*, 1011. Jena: Max Planck Institute of Economics.

ElMaraghy, W. – ElMaraghy, H. – Tomiyama, T. – Monostori, L. (2012): Complexity in engineering design and manufacturing. *CIRP Annals-Manufacturing Technology*, 61, 2, 793–814.

EC (2016): Productivity in EU agriculture – slowly but steadily growing. EU Agricultural Markets Briefs, No 10, December. European Commission, Retrieved 24.04.2019 from https://ec.europa.eu/agriculture/sites/agriculture/files/markets-and-prices/market-briefs/pdf/10_en.pdf

EC (2017): *CAP Context Indicators 2014-2020 – 27. Total Factor Productivity*. European Commission.

EC (2017b): Industry 4.0 in Agriculture: Focus on IoT aspects. Digital Transformation Monitor, Internal Market, Industry, Entrepreneurship and SMEs, European Commission.

EC (2018): The CAP after 2020. European Commission. https://ec.europa.eu/commission/sites/beta-political/files/budget-may2018-modernising-cap_en.pdf

EUROSTAT (2017): European Statistics. http://ec.europa.eu/eurostat/ Accessed: November 15, 2017

FAO (2017): *Information and Communication Technology (ICT) in Agriculture. Report to the G20 Agricultural Deputies*. Food and Agriculture Organization of the United Nations. Rome. ISBN 978-92-5-109979-7

Faure, G. – Knierim, A. – Koutsouris, A. – Ndah, H. T. – Audouin, S. – Zarokosta, E. – Wielinga, E. – Triomphe, B. – Mathé, S. – Temple, L. – Heanue, K. (2019): How to Strengthen Innovation Support Services in Agriculture with Regard To Multi-Stakeholder Approaches. *Journal of Innovation Economics & Management* 2019/1, 28, 145–169 ISBN 9782807392816

Gorton, M. – Hubbard, C. – Hubbard, L. (2009): The Folly of EU Policy Transfer: Why the CAP does not fit Central and Eastern Europe. *Regional Studies*, 43, 10, 1305–1317. http://dx.doi.org/10.1080/00343400802508802.

GSA (2018): *Report on Agriculture User Needs and Requirements* (GSA-MKD-AGR-UREQ-233667) European Global Navigation Satellite Systems Agency 18/10/2018

Hart, K. – Bas-Defossez, F. (2018): *CAP 2021-27: Proposals for increasing its environmental and climate ambition*. Institute for European Environmental Policy. Retrieved: 15.02.2019 from: https://ieep.eu/uploads/articles/attachments/63db952e-0825-4eb8-80fe-f88708cfd62f/NABU%20CAP%20Report%20-%20FINAL%20.pdf?v=63710723894

Hesser L. (2006): The Man Who Fed the World: Nobel Peace Prize Laureate Norman Borlaug and His Battle to End World Hunger. Dallas: Durban House

Jámbor, A. – Somai, M. – Kovács, S. (2016): Tíz év az Európai Unióban – az új tagországok agrárteljesítményei. *Közgazdasági Szemle*, 63, 3, 260–284.

Jóri, J. I. (2017): A Digitális Mezőgazdaság fejlődésének története https://agroforum.hu/szakcikkek/gepeszet/cema-a-digitalis-mezogazdasag-fejlodesenek-tortenete

Juhász, P. (2016): Farmok, technológiák, hálózatok, kooperációk – nemzetközi trendek a mezőgazdaságban. In *Kovács, K. (szerk.): Földből élők – Polarizáció a Magyar vidéken*. Argumentum, ISBN 978-963-446-773-1

Kemény, G. – Lámfalusi, I. – Molnár, A. (eds); Gaál, M., Kiss, A., Péter, K., Sulyok, D., Takácsné György, K., Domán, Cs., Illés, I., Keményné Horváth, Zs. (2017): *A precíziós szántóföldi növénytermesztés összehasonlító vizsgálata (Comparative study of precision arable crop production)* Agrárgazdasági Könyvek, Agrárgazdasági Kutató Intézet, Budapest. ISBN 978-963-491-601-7

Lencsés, E. (2013): A precíziós (helyspecifikus) növénytermelés gazdasági értékelése. (Economic Analysis of Precision Crop Production. Dissertation) PhD-értekezés (SZIE GSZDI). 179 p.

Manyika, J. – Chui, M. – Bughin, J. – Dobbs, R. – Bisson, P. – Marrs, A. (2013): *Disruptive technologies: Advances that will transform life, business, and the global economy*. San Francisco, CA: McKinsey Global Institute.

Matthews, A. (2018): The CAP in the 2021–2027 MFF Negotiations. *Intereconomics*, 53, 306–311. DOI: 10.1007/s10272-018-0773-0

Matthews, A. (2019): Introducing a development policy perspective into CAP Strategic Plans. *Trinity Economics Papers*, 0319, Trinity College Dublin, Department of Economics.

Möllers, J. – Buchenrieder, G. – Csáki, Cs. (eds.) (2011): Structural Change in Agricul- ture and Rural Livelihoods: Policy Implications for the New Member States of the European Union. *Studies on the Agricultural and Food Sector in Central and Eastern Europe*, Vol. 61. IAMO, http://www.iamo.de/fileadmin/documents/sr_vol61.pdf .

Pechrová, M. (2015): The effect of subsidies on the efficiency of farms in the Liberecký region. *Ekonomická revue – Central European Review of Economic Issues*, 18, 15–24. doi: 10.7327/cerei.2015.03.02

Pelle, A. – Somosi, S. (2018): Possible Challenges for EU-level Industrial Policy: Where Do Potentials for Policy Improvement in Central an Eastern European Countries Lie? In: *Progressive Industrial Policy, Journal für Entwicklungspolitik*, 34, 3–4, 143–172.

Popp, J. – Erdei, E. – Oláh, J. (2018): A precíziós gazdálkodás kilátásai Magyarországon. (The Outlooks of Precision Farming in Hungary) *International Journal of Engineering and Management Sciences*, 3, 1, 133–147.

Popp, J. – Jámbor, A. (2015): How consistent is the new Common Agricultural Policy with the challenges it faces? *Society and Economy*, 37, 2, 224–242.

Potori, N. – Kovacs, M. – Vasary, V. (2013): The Common Agricultural Policy 2014–2020: An impact assessment of the new system of direct payments in Hungary. *Studies in Agricultural Economics*, 115, 3, 118–123. http://dx.doi.org/10.7896/j.1318.

Schwab, K. (2016): *The Fourth Industrial Revolution*. Geneva: World Economic Forum.

Smuk, N. – Milics, G. – Salamon, L. – Neményi, M. (2009): A precíziós gazdálkodás beruházásainak megtérülése. (The Investment Recovery of the Switch to Precision Agriculture) *Gazdálkodás*, 53, 3, 246–253.

Sung, J. (2018): The Fourth Industrial Revolution and Precision Agriculture. In *Hussmann, S. (ed.): Automation in Agriculture*, *Securing Food Supplies for Future Generations*. IntechOpen, DOI: 10.5772/intechopen.69016

Szalavetz, A. – Somosi, S. (2019): Ipar 4.0 technológiák és a magyarországi fejlődés–felzárkózás hajtóerőinek megváltozása – gazdaságpolitikai tanulságok. (Advanced technologies of Industry 4.0 and the altering engines of Hungarian development – lessons for economic policy) *Külgazdaság*, 63, 3–4, 66–93.

Turner, B. L. – Menendez, H. M. – Gates, R. – Tedeschi, L. O. – Atzori, A. S. (2016): System Dynamics Modeling for Agricultural and Natural Resource Management Issues: Review of Some Past Cases and Forecasting Future Roles. *Resources*, 5, 4; doi:10.3390/resources5040040

Westerman, G. – Bonnet, D. – McAfee, A. (2014): *Leading Digital*. Boston: Harvard Business Review Press

WGS (2018): *Agriculture 4.0 – The Future Of Farming Technology*. World Government Summit, February.

Xu, L. D. – Xu, E. L. – Li, L. (2018): Industry 4.0: state of the art and future trends. *International Journal of Production Research*, 56, 8, 2941–2962.

Zambon, I. – Cecchini, M. – Egidi, G. – Saporito, M. G. – Colantoni, A. (2019): Revolution 4.0: Industry vs. Agriculture in a Future Development of SMEs. *Processe*s, 7, 36; doi: 10.3390/pr7010036 www.mdpi.com/journal/processes

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 190–209.

4th Industrial Revolution: The middle income trap, technological advancement and socio-economic development nexus

Timothy Yaw Acheampong

The Middle-Income-Trap (MIT) concept has received considerable attention among development practitioners and economists in recent times due to its associated lower socioeconomic development and negative welfare consequences such as increasing poverty and inequality in affected economies. Of the many factors proposed for breaking out of the MIT, technology has been singled out as absolutely essential, based on the hypothesis that MIT countries have lower technological development and lower socio-economic development compared to high income countries. Although the relationship between technology and socioeconomic development has been established, existing studies have not utilized the latest Global Innovation Index (GII) to examine this relationship. In the current era, the "4th Industrial Revolution" offers a higher potential for human development through rapid technological development. The Sustainable Development Goals (SDGs) also recognize innovation and technology as critical for ending all forms of poverty and inequality. As a result, there is the need to examine the MIT hypothesis by answering the question: do countries with higher levels of technological development also have higher socio-economic development? A cross-sectional research design was utilized. Quantitative analyses GII scores and selected socioeconomic indicators for 126 countries corroborate the MIT hypothesis that countries with higher innovation and technological development also have higher income levels.

Keywords: 4th Industrial Revolution, Technological Progress, Middle Income Trap, Global Innovation Index, Human Capital, Sustainable Development Goals

1. Introduction

As the world enters the 4th Industrial Revolution (or Industry 4.0), the important role of technological progress in the socioeconomic development status of countries has taken center stage in economics and development discourse (See UNCTAD 2017, United Nations 2017, World Bank 2019). The importance of technological advancement to economic growth and development have featured in economics literature at least since the 1950s with the introduction of Solow's growth model (Mankiw–Taylor 2014), and currently it is argued that technological change is an extremely important factor if not the main factor, in economic growth (Bajmócy–Gébert 2014). The recognition of the importance of technological advancement to economic growth and development has resulted in policy makers across globe adopting various innovation policies over the years making the term 'innovation policy' very popular in the last two decades (Edler–Fagerberg 2017). It is also widely accepted that innovation can help address global challenges and affect various socioeconomic situations (Edquist 2014). Thus, the 2016 edition of '*The Global Information Technology Report'* posits that a key feature of the Fourth Industrial

Revolution is that, "the future holds an even higher potential for human development as the full effects of new technologies such as the Internet of Things, artificial intelligence, 3-D Printing, energy storage, and quantum computing unfold" (Samans–Hanouz 2016).

In spite of its potential for improved socioeconomic development, the most recent World Development Report 2019 has cautioned that, some countries are likely to benefit more from technological development than other countries. For instance Samans–Hanouz (2016) points out that Information and Communication Technologies (ICTs) are the backbone of this Fourth Industrial Revolution, and that countries and businesses that embrace these developments as well as anticipate challenges, and deal with them in a strategic way, are more likely to prosper, while those that do not are more likely fall behind (Samans–Hanouz 2016). Similarly, *Information Economy Report 2017* notes that "the world is on the cusp of a new digital era…This has major implications for the implementation of the 2030 Agenda for Sustainable Development, presenting significant opportunities, but also challenges, for developing countries" (UNCTAD 2017).

As part of global efforts to ensure that all countries and their citizens benefit from the 4th Industrial Revolution and the opportunities technological progress and innovation offer, the Sustainable Development Goals (SDGs), which was adopted by world leaders at the 2015 United Nations General Assembly, has set – among other aims – specific targets for governments to increase investments in science, technology, and innovation, and also to ensure that everyone has access to ICTs (UN 2015, UN 2017). In spite of these efforts, Nikoloski (2016) posits that, there is a very refined technological gap that currently exists between developed countries on the one hand and developing countries on the other, with the gap widening to the detriment of developing countries. According to Nikoloski, the "developed countries have a monopoly on the sources of technological development and export of modern equipment and technology while developing countries are technologically dependent on developed countries (2016, p. 48). Empirical evidence from studies on the Middle Income Trap (MIT) – a phenomena whereby when countries enter the middle income bracket, they are unable to progress to high income status – suggests, that technological advancement is a critical factor in escaping the trap (Glawe–Wagner 2016). Out of 101 middle-income countries in 1960, only 13 were able to escape the MIT by the year 2008, and Glawe–Wagner (2016) have cited several empirical studies which indicate that the few countries that were able to escape the MIT were those that have moved up the technological development ladder from being only consumers of technology to also becoming producers and exporters of technology. Meanwhile the MIT is of concern to development practitioners and economists because empirical evidence suggests that countries that get stuck in the middle income bracket have higher levels of poverty, inequality, and lower scores on other socio-economic indicators (Glawe–Wagner 2016).

In spite of the MIT hypothesis that higher income countries have higher levels of technological and socioeconomic development, empirical studies are yet to examine this relationship using the most recent data. This justifies the need to examine the relationship between national levels of technological progress and socio-economic development indicators using the most recent data. This paper seeks to provide new insights into the technology and socioeconomic development nexus by answering the following questions: What is the relationship between the level of socio-economic development of countries as measured by GDP, per capita income, and the Human Development Index (HDI) and their level of technological progress, as measured by the most recent Global Innovation Index (GII) 2018; also, is there a significant difference between the GII stores of high income and middle income countries; is there equal participation by countries in different income groups in the 4th Industrial Revolution. In this study, participation in the 4th industrial revolution is as measured by trade in ICT goods, and the proportion of people using the internet by the respective income levels of these countries? The results of these analyses provides new insights into the existing literature on the countries that are likely to benefit from the 4th Industrial Revolution, as well as the nexus between technological progress, socioeconomic development, and the MIT. The next section provides an overview of the relevant theoretical and conceptual issues including the theoretical foundations of technology in economic growth, the MIT hypothesis concerning the role of technology in socioeconomic development, as well as the nexus between the 4th industrial revolution, technological development, and the socioeconomic development status of countries. This is followed by discussions on the methodology and key findings.

2. Theoretical and Conceptual Issues

2.1. Theoretical Foundations of Technology and Innovation in Economic Growth and MIT Hypothesis

The theoretical foundations linking technological progress to the economic growth of nations can be traced to classical economists such as Adam Smith, David Ricardo, Thomas Malthus, and much later Frank Ramsey, Frank Knight, and Joseph Schumpeter, among others (Barro–Sala-i-Martin 2004). However, contemporary theorising on the importance of technology to the economic growth and productivity of nations can be traced to the seminal work of Solow in the 1950s (Bajmócy–Gébert 2014, Mankiw–Taylor 2014). According to Gill and Kharas, economists started to "unpack the technological black box of the Solow growth model" after the pioneering work by Romer in the year 1986, Lucas in the year 1988, and a decade later by Aghion and Howitt in 1996 (Gill–Kharas 2015). For instance, building on Solow's growth model, the endogenous growth theory proposed by Robert Lucas and Paul Romer in the late 1980s (Mankiw–Taylor 2014) posited that investment in a nation's human capital will be a key driver in economic growth rather than trade, because human capital is likely to lead to increases in technology, which in turn would help promote efficiency and increases in productivity (Mankiw–Taylor 2014). This assumption is captured by Mankiw–Taylor in the model below:

Determinants of Economic Growth Model

$$Y=\beta f(K,L,H,N)$$

Where;

Y – the output (GDP) of a country is dependent on the following:
β(beta) – the rate of technological progress;
K – the quantity of physical capital;
L – the quantity of labour;
H – the quantity of human capital; and
N – the quantity of natural resources.

Based on the model above, developed countries have high levels of physical and human capital and in a production function analysis, this would explain their high levels of output per person whereas the opposite is true for low income developing countries (Mankiw–Taylor 2014, Todaro–Smith 2015). Whereas Solow's growth model and the endogenous growth theories seemed to better explain the phenomenon of "club convergence" and also pointed to the importance of technological progress in disparities between the economic development status of different countries, Gill–Kharas (2015) posit that these models were only successful in addressing growth problems in high income and low income countries; however, neither of those two frameworks were satisfactory in understanding and addressing the nature of economic growth challenges in middle-income countries. This gave rise to the notion of MIT which Glawe–Wagner (2016) observed is a relatively new phenomenon, conceptually. Thus, according to Gill and Kharas who introduce the term middle income trap' in a 2007 World Bank Report *titled 'An East Asian Renaissance, Ideas for Economic Growth',* the MIT concept emerged due to the inability of the existing economic growth theories to satisfactorily inform development policy in middle income countries (Gill–Kharas 2015).

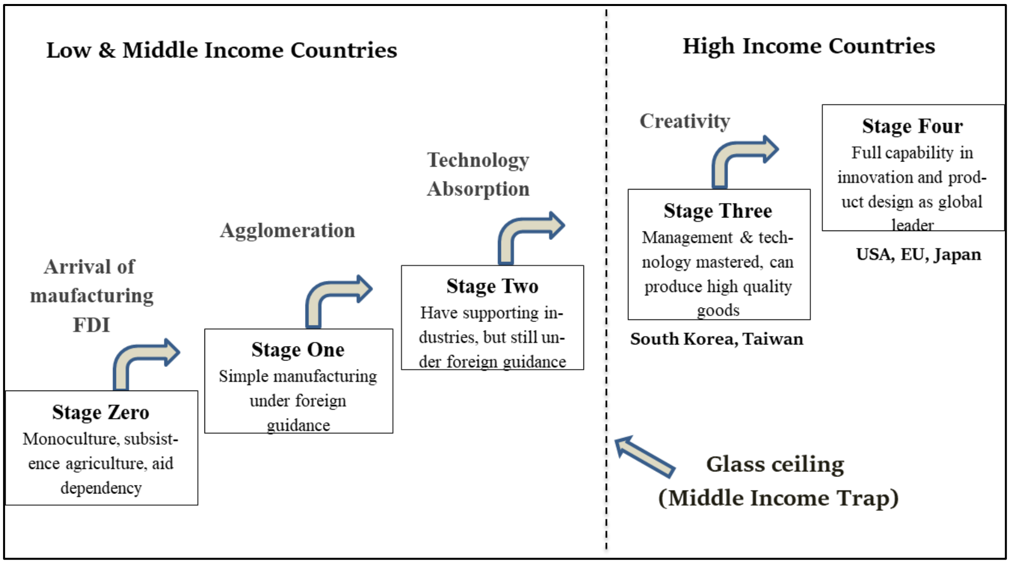
Insights from various economic growth theories have been used to explain the poverty trap in developing countries and also to justify the need for increased investments in human capital. For instance, Mankiw–Taylor (2014) posit that 'education – investment in human capital – is at least as important as investment in physical capital for a country's long-run economic success' (p. 487). Todaro–Smith (2015) also posit that health and education are inputs into the national production function in their role as components of human capital, meaning productive investments embodied in persons; however, improvements in health and education are also important development goals in their own right. To underscore this point, UNDP (2016) states that 'human capital is an asset, and differences in educational attainment prevent poor people from becoming part of the high-productivity growth process' (p. 12). Similarly, World Bank (2019) points out that 'delivered well, education – and the human capital it creates – has many benefits for economies and for societies as a whole. For individuals, education promotes employment, earnings, and health. It raises pride and opens new horizons. For societies, it drives long-term economic growth, reduces poverty, spurs innovation, strengthens institutions, and fosters social cohesion' (p. 11).

The collective impact of human capital and technological advancement on economic growth is based on studies which analyse mechanisms known as the 'productivity channel' (Li–Wang 2018). The core argument of this approach is that higher levels of human capital increase a country's ability to innovate and/or to adapt to existing technologies. Brue–Grant (2013) have observed that human capital in the form of the entrepreneur is central to a key process in economic change and the introduction of innovations – which can be defined as changes in the methods of supplying commodities, such as introducing new goods or new methods of production. Innovation can also be distinguished from invention in that, an invention becomes an innovation only when it is applied to industrial processes and this transformational process requires people with exceptional abilities who seize opportunities that others are oblivious to or who create opportunities through their own daring and imagination (Brue–Grant 2013). According to Edler–Fagerberg (2017) it was the founding father of innovation theory, Josef Schumpeter, who introduced the distinction between invention (a novel idea for how to do things) and innovation (carrying it out into practice). The justification for this distinction was based on the realization that, what matters economically and societally is not the idea itself but its adoption and subsequent exploitation in the economic and social system (Edler–Fagerberg 2017). Meanwhile the processes of invention, innovation, the adoption, and exploitation of technology are dependent on the human capital base of a nation. Therefore, the "productivity channel," approach, argues that differences in growth rates across countries largely arise from differences in levels of human capital in those countries (Li–Wang 2018).

2.2. Technology, human capital, and Middle-Income-Trap Hypothesis

Li–Wang (2018) as well as Wang et al. (2018) have recently investigated the nexus between human capital and the MIT. According to Li–Wang (2018) there is a more recent work by Vandenbussche et al. in 2006 which measures the role of human capital and technology in economic growth, in which the contribution of human capital to growth has both a level effect as well as a composition effect through the "productivity channel". Based on this model, the productivity-enhancing impact of human capital depends on not only its level but also, controlling for its level, the composition of skilled human capital and the country's position relative to the technological; therefore, skilled human capital is more important for countries that are closer to the technological frontier (Li–Wang 2018). It is in this regard, that empirical studies on the MIT such as Eichengreen et al. (2013) have concluded that the MIT is 'less likely in countries where the population has a relatively high level of secondary and tertiary education and where high-technology products account for a relatively large share of exports (Glawe–Wagner 2016). Therefore, in order for countries to break out of or avoid the MIT, they must move up the technological ladder as depicted in the stages of the catching up process in Ohno's MIT Model (Figure 1).

Figure 1 Middle Income Trap Hypothesis – Stages of the catching up in Ohno's Model



Source: Author's construct adapted from Ohno (2009)

Ohno's MIT model indicates that, the countries that have escaped the MIT are those that have mastered technology, which Nikoloski (2016) defines as the sum of knowledge about procedures and processes not only in manufacturing but also in other spheres of social life, and have full capability in innovation and product design. Based on this model, the process of escaping or breaking out of the MIT is described as a catch-up process in which development is viewed as a linear process in which countries must move from one stage to another in order to develop; however, empirical evidence suggests that transitioning from one stage to another is not as smooth as many countries have remained in the poverty trap and middle income bracket for several years (Glawe–Wagner 2016). Similar to Ohno's model, Todaro and Smith had earlier observed that, "technology transfer is critical to more rapid growth, competing internationally, and beginning to catch up with advanced countries" (2015)

2.3. Technology, the 4th Industrial Revolution, and the SDGs

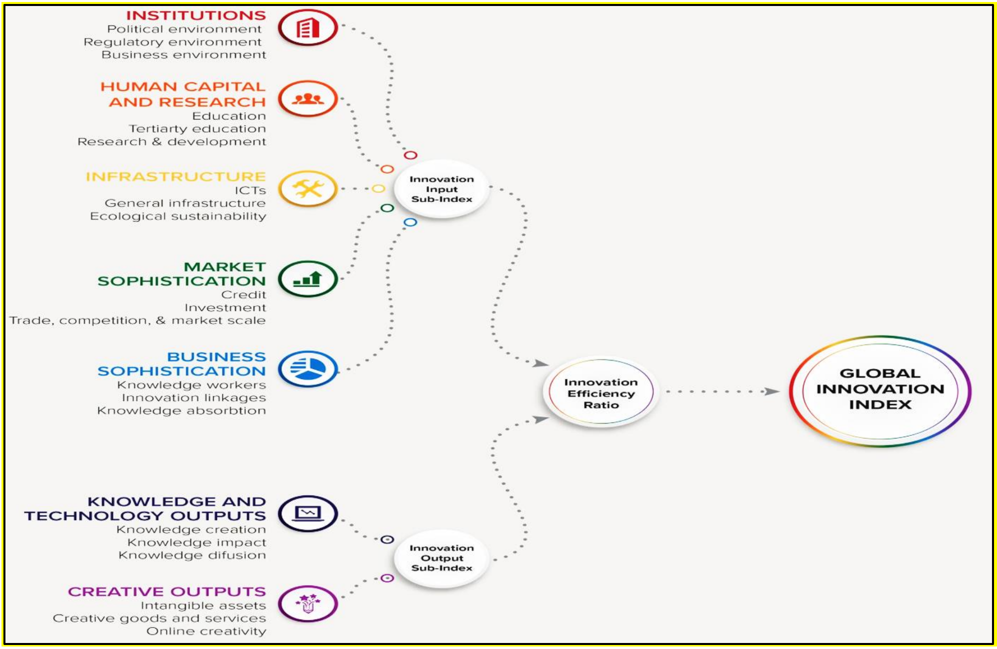
The academic literature related to the developments of the so-called 'Fourth Industrial Revolution" is still relatively new (Glawe–Wagner 2018) thus there are different conceptions as to the scope of this revolution. A distinguishing feature between the 4th and previous industrial revolutions is the recognition of the need for more skilled and knowledgeable workers (World Bank 2019). Nikoloski (2016) summarises the key features of the previous industrial revolutions as follows: the 1st industrial revolution which was accompanied by the steam engine resulted in replacement of some of the physical effort with machines; the 2nd industrial revolution or simply automation changed man and human development not only in the execution of physical operations, but also in the performance of certain mental operations; the third industrial or technological revolution referred to as the electronic revolution brought about a transistor whose application enabled the development of computers or computers and microprocessors. The emergence of the 4th Industrial Revolution and the growing importance of technology in life and business means that all types of jobs (including low-skill ones) require more advanced cognitive skills; therefore, a basic level of human capital, such as literacy and numeracy, is needed for economic survival (World Bank 2019). According to Samans–Hanouz (2016) ICTs are the backbone of this 4th Industrial Revolution, and the countries and businesses that embrace these developments, anticipate challenges, and deal with them in a strategic way are more likely to prosper, while those that do not will more likely fall behind. In this regard, the World Bank (2019), points out that developing countries need to increase investments in human capital and technological capabilities.

As part of global efforts to ensure everybody benefits from the opportunities that the 4th Industrial Revolution offers, the SDGs have several targets and indicators for countries to achieve. Specifically target 17.6 of SDG 17 requires nations to "Enhance North-South, South-South and triangular regional and international cooperation on and access to science, technology and innovation and enhance knowledge-sharing on mutually agreed terms, including through improved coordination among existing mechanisms, in particular at the United Nations level, and through a global technology facilitation mechanism" (UN 2015, UN 2017). Additionally, SDG target 17.8 also has the aim to "Fully operationalize the technology bank and science, technology and innovation capacity-building mechanism for least developed countries by 2017 and enhance the use of enabling technology, in particular information and communications technology" (UN 2015, UN 2017). The global indicator for measuring progress on SDG 17.8 for instance, is the 'Proportion of individuals using the Internet' in various countries (UN 2017). Similarly, the SDG 9.c has the aim to "Significantly increase access to information and communications technology and strive to provide universal and affordable access to the Internet in least developed countries by 2020" (UN 2017).

In addition to ensuring that everyone has access to the available ICTs, the SDGs further emphasize the importance of technology to poverty reduction, and economic growth. For instance the SDG on No Poverty has target 1.4 which seeks to "ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology" by the year 2030 (UN 2017). According to the UN (2017) SDG 8.2 also aims to "Achieve higher levels of economic productivity through diversification, technological upgrading and innovation, including through a focus on high-value added and labor-intensive sectors". Furthermore, SDG 9 also has the following targets and indicators: 9.5 Enhance scientific research, upgrade the technological capabilities of industrial sectors in all countries, in particular developing countries, including, by 2030, encouraging innovation and substantially increasing the number of research and development workers per 1 million people and public and private research and development spending. An indicator for this target is research and development expenditure as a proportion of GDP (UN 2017). The targets and indicators highlighted above, give an indication of the pledge that world leaders have made to ensure that as many people as possible benefit from the 4th Industrial Revolution through the SDGs.

According to UNCTAD (2017) information technology has provided opportunities for businesses and countries to improve productivity across all sectors and to build new sectors. In spite of this potential of ICTs provided by the 4th Industrial Revelotion, Nikoloski posits that developed countries have a monopoly on the sources of technological development and export of modern equipment and technology while developing countries are technologically dependent on developed countries (Nikoloski 2016). If such a trend persists, then it has the potential to inhibit the ability of developing countries to benefit from the 4th Industrial Revolution and also achieving the SDGs. Furthermore, this trend would also inhibit countries from breaking out of or avoiding the MIT as indicated earlier. It is in view of of the potential welfare consequences of the development challenge discussed above that this study focuses on examing the implications of the technological development status of countries using the Global Innovation Index.

Figure 2 Framework of the Global Innovation Index 2018



Source: Cornell University, INSEAD & WIPO (2018); Dutta et al. (2018, 16)

The Global Innovation Index (GII) is a composite measure of seven (7) indicators, also referred to as pillars, that are used to measure the technological progress and level of innovation of countries (See Figure 2). Five (5) of the 7 pillars constitute the Innovation Input Sub-Index, comprised of the following elements of the national economy that enable innovative activities: (*1) Institutions, (2) Human capital and research, (3) Infrastructure, (4) Market sophistication,* and *(5) Business sophistication.* The other two pillars constitute the Innovation Output Sub-Index, which provides information about outputs that are the results of innovative activities within the economy: *(6) Knowledge and technology outputs* and *(7) Creative outputs.* Each pillar is divided into sub-pillars and each sub-pillar is composed of individual indicators (80 in total in 2018). Sub-pillar scores are then calculated as the weighted average of individual indicators; pillar scores are calculated as the weighted average of sub-pillar scores and the overall GII Score is computed by taking a the simple average of the Input and Output Sub-Index scores (Dutta et al. 2018). According to Cornell University et al. (2018), GII gathers data from more than 30 sources, covering a large spectrum of innovation drivers and results, and the framework is revised every year to improve the way innovation is measured. Although the relationship between technology and economic growth has been established, existing studies have not utilised the latest GII to explore this relation. The relationship between GII and HDI is also yet to be analysed in literature.

3. Methodology

The study utilized a cross-sectional research design. Cross-sectional studies entail the collection of data on more than one case (usually many more than one), at a single point in time, with the objective of providing a snapshot of a given phenomenon (Babbie 2008, Walliman 2006). Walliman (2006) points out that the advantage of using cross-sectional design is that it allows for patterns of association between variables to be examined in order to detect associations; in addition, causal influences can also be inferred. As indicated earlier, the objective of the paper was to investigate the relationship between the level of socio-economic development of countries as measured by GDP, per capita income, and the Human Development Index (HDI), and their level of technological progress as measured by the most recent Global Innovation Index (GII) 2018. The socioeconomic development indicators adopted for this study were selected because they are the most widely used and accepted. A total of 126 countries were selected based on the availability of data sets used to compute the GII.

The study relied on secondary data sources and all indicators analyzed were based on official 2017 figures. The data sources included the World Development Indicators database (World Bank 2019), Human Development Index database (UNDP 2018) the Global Innovation Index 2018 database (Cornell University et al. 2018), the World Economic Outlook database (IMF 2018; 2019), and UNCTADStats. Various statistical analyses including descriptives, t-tests, and Pearson’s product moment correlation were conducted to answer the research questions. Descriptive statistics were used to describe the distribution of the study countries by their income levels for the various variables studied. The t-tests were used to examine whether there is a significant difference between the mean GII scores of high income and middle-income countries. The correlation analysis was also conducted to examine whether there is significant relationship between the level of technological development of countries and their respective socio-economic development indicators. In order to determine if the participation of various countries with different income in the 4^{th} Industrial Revolution is equal, their trade in ICT goods and persons using the internet were cross tabulated. The focus of this paper on analyzing the income level of countries is based on the MIT hypothesis and empirical literature that suggest there is a gap in the level of technological and socioeconomic development of the high-income countries, on one hand, and the middle and low income countries on the other hand. These assumptions formed the basis of the various analysis used to answer the research questions. The key findings of the various analyses are discussed next.

4. Findings and discussions

4.1. The sample countries

A descriptive analysis of the 126 countries analysed in this study indicates that most of the countries were from the Europe and Central Asia region (37%) followed by Sub-Saharan Africa (19%) and then the Latin America and Caribbean region (14%); North America (2%) and South Asia (4%) had the fewest countries respectively (Table 1). As indicated earlier, these countries were selected due to the availability of complete data sets. Although the GII data does not cover all countries, the number of countries in this study represents about 65% of the countries in the world. The countries analysed also represent close to 90% of the world's population and 97% of global GDP (Cornell University et al. 2018, IMF 2019).

Table 1 Distribution of countries analyzed by regions and percent of world GDP 2017

Regions	Frequency	Countries analysed (%)	2017 GDP share in Purchasing power parity (%)
East Asia & Pacific	15	11.9	30.5
Europe & Central Asia	46	36.5	21.2
Latin America & Caribbean	18	14.3	6.0
Middle East & North Africa	16	12.7	5.4
North America	2	1.6	16.6
South Asia	5	4.0	9.3
Sub-Saharan Africa	24	19.0	7.5
Total	126	100.0	96.6

Source: Author's construct based World Bank Classifications and GII (2018), and IMF (2019).

A crosstabulation of the GNI per capita of the income levels of various countries against the World Bank income group classifications indicates that, most of the countries analyzed were in the middle-income group (59 representing 47%) whereas 39% and 13% of the countries were in the high income and low income groups, respectively. The World Bank's classification distinguishes between four income categories based on the real per capita gross national income (GNI) calculated on the basis of the Atlas method and is the most widely used indicator. Since analysis of the MIT requires the comparison of middle income against high income countries, the sample size of each of the groups which are above 30 satisfies the required assumptions to make statistical comparisons among the GII scores of the two income groups using correlation analysis and t-tests.

Table 2 Distribution of countries analyzed by income groups

Income Group	World Bank Threshold ($)	Frequency	Percent
High Income	> 12,055	50	39.68
Upper Middle Income	3,896 – 12,055	34	26.98
Lower Middle Income	996 – 3,895	25	19.84
Low Income	< 995	17	13.49
Total Countries Analyzed		126	100.00

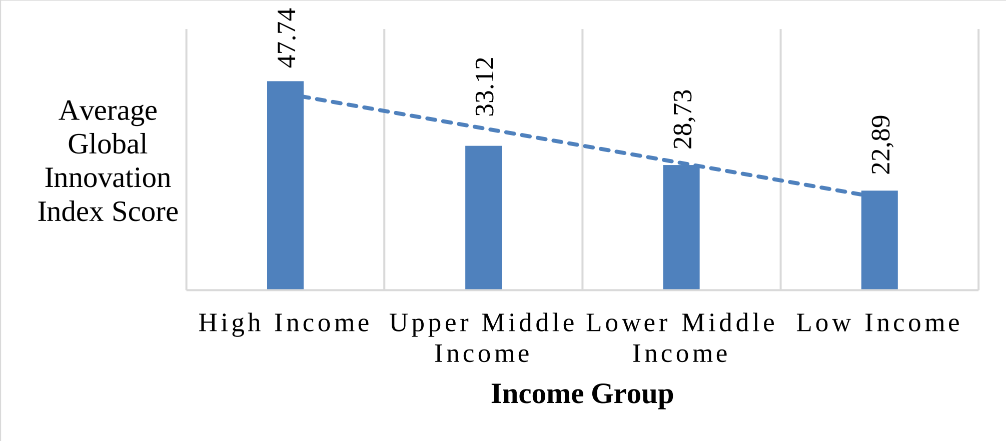
Source: Author's construct based on World Bank classifications and GII 2018 data

4.2. Findings on the relationship between Global Innovation Index scores and income of countries

In order to investigate the existing literature on the nexus between the level of technological progress and output of countries as theorised by literature on the MIT, a correlation analysis was conducted on the GII scores of the countries against both their output as measured by GDP and the GNI per capita used by the World Bank to categorise countries into different income brackets. The study finds a significant relationship between the technological progress and output of countries. The GII scores positively correlated with both the GNI per capita (r= .836) and GDP per capita in both current prices (r = .780) and purchasing power parity (r = .696). The correlation between the GII and GDP was also statistically significant and positive (r = .347); however, it was weaker than the correlation between the GII and various measures of income per capita. It should be noted that the GNI per capita is the indicator that is used to classify countries by income groups. The implications of this finding is that countries in the higher income also have higher GII scores and vice versa, which is consistent with the MIT hypothesis that high income countries are higher up the technology ladder.

An independent-samples t-test was also conducted to compare GII Scores for the high and middle income countries. There was a significant difference beween the GII scores for the high countries (M = 47.74, SD = 10.10) and middle income countries (M = 31.26, SD = 6.04); t (77) = 10.11, p = .00, two-tailed). The magnitude of the differences in the means (mean difference = 16.49, 95% CI: 13.24 to 19.73) was very large (eta squared = . 488). The eta square value means that about 50% of the variance in the GII scores of the the high income and middle income countries can be explained by the income levels of the respective countries.

Figure 3 Distribution of Global Innovation Index scores by income groups



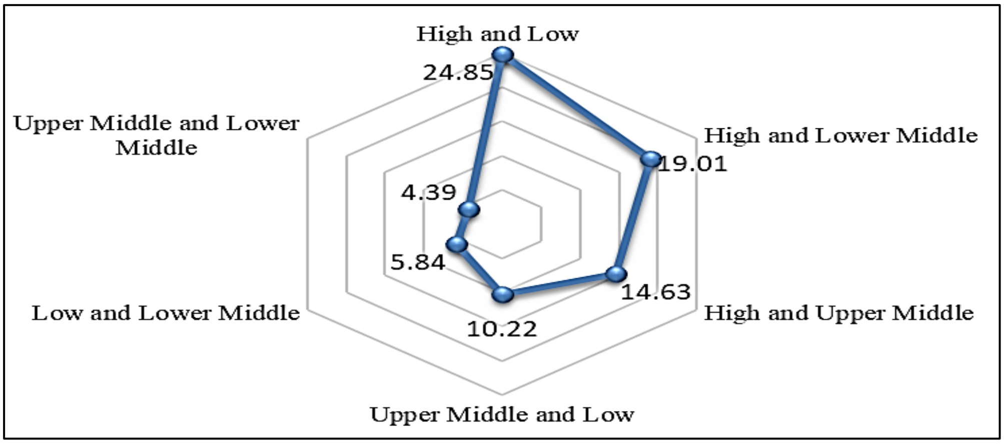
Source: Author's construct based on GII 2018 data and World Bank classification.

Another independent samples t-test was also conducted to determine if there was a significant difference between the GII scores of the high income countries and the lower middle income countries, and also between the high income countries and the uppper middle income countries. In both cases the study finds a significant difference between the GII scores of the high income countries and middle income countries. However, the gap was wider between high income countries (mean difference = 19.01, 95% CI: 15.37 to 22.65, t(72)=10.42, p = .00, two-tailed) and the lower middle income countries as compare with the gap between the high income countries and the upper middle income countries (mean difference = 14.63, 95% CI: 11.19 to 18.06, t (79) = 8.46, p = .00, two-tailed). In the case of the difference between the GII scores of the high income and lower middle income countries the eta square was .593, indicating that about 60% of the difference in the gap of GII could be explained by the income levels of the countries. For the difference between the high income and upper middle income countries the eta square was .476 slightly, lower that the eta square when all the middle income countries are combined. The highest gap in GII scores is between the high income and low income countries (See Figure 4)

These findings corroborate the earlier correlation analysis which indicates that the higher the income level of a country the higher their technological development. The findings are also consistent with the assertion of Nikoloski (2016) that there is a very refined technological gap that currently exists between developed countries on the one hand and developing countries on the other. Although Nikoloski

(2016) indicates that the technolgical gap is widening to the detriment of developing countries, the focus of this study was just to provide a snapshot of the current situation with respect to the relationship between the technological development of countries and their income levels. Consistent with the Global Innovation Index 2018 Report, the analysis of the GII scores against the income levels of countries indicate that the middle income countries performed worse than the high income countries on the GII with the exception of China. The Report notes that with the single exception of China, which is an upper-middle income economy, there has been a stable group of high-income economies that composes the top 25 of the GII suggesting a form club convergence as theorised Nikoloski (2016) and other MIT literature (Glawe–Wagner 2016).

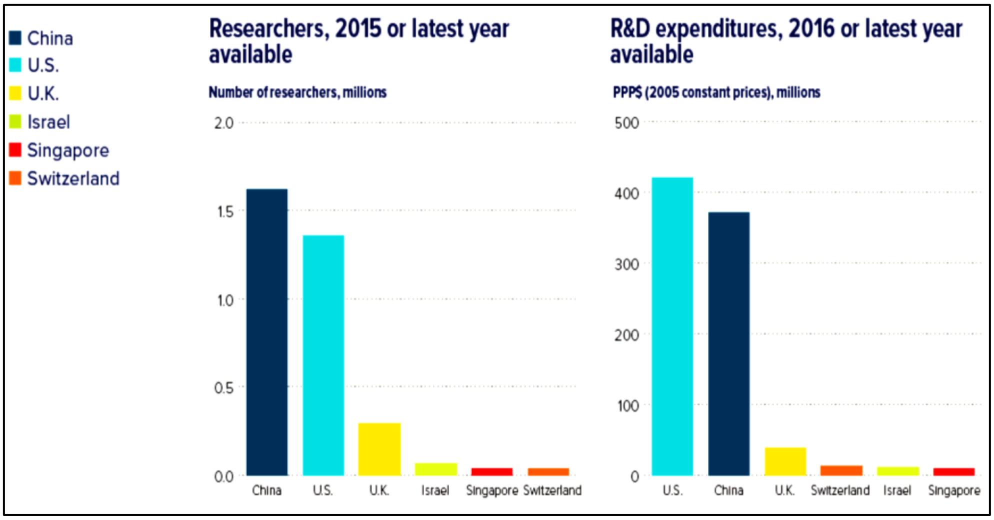
Figure 4 Gaps in Global Innovation Index scores by income groups



Source: Authour's calculations

An analysis of the GII rankings indicates that China was ranked 17th overall, performing better than more than almost 70% of the high income countries. A trend analysis of the GII rankings indicates that China entered the top 25 group in 2016 and has consistently moved up in the rankings to reach 17th place in 2018. According to the GII 2018 Report, China has been able to make this rise as a result of improvements in global R&D companies, high-tech imports, the quality of its scientific publications, and tertiary enrolment. Furthermore, the Report indicates that, China's score in knowledge and technology outputs continues to be above that of the top 10 group average. According to the GII Report, China's rapid rise in the GII rankings indicates how other middle-income can possibly bridge the technology gap. China currently has the highest number of researchers per 1 million people and was second only to the United States in terms of research and development expenditure (See Figure 5). As indicated earlier, these 2 indicators have be identfied in the SDGs. Specifically, the aim of target 9.5 of the SDGs is for countries to enhance scientific research, upgrade the technological capabilities of industrial sectors in all countries, in particular developing countries, including, by 2030, encouraging innovation and substantially increasing the number of research and development workers per 1 million people and public and private research and development spending and an indicator for measuring the progress of this target is the research and development expenditure as a proportion of GDP (UN 2017).

Figure 5 Comparison of China's GII/SDG indicators against selected high income countries



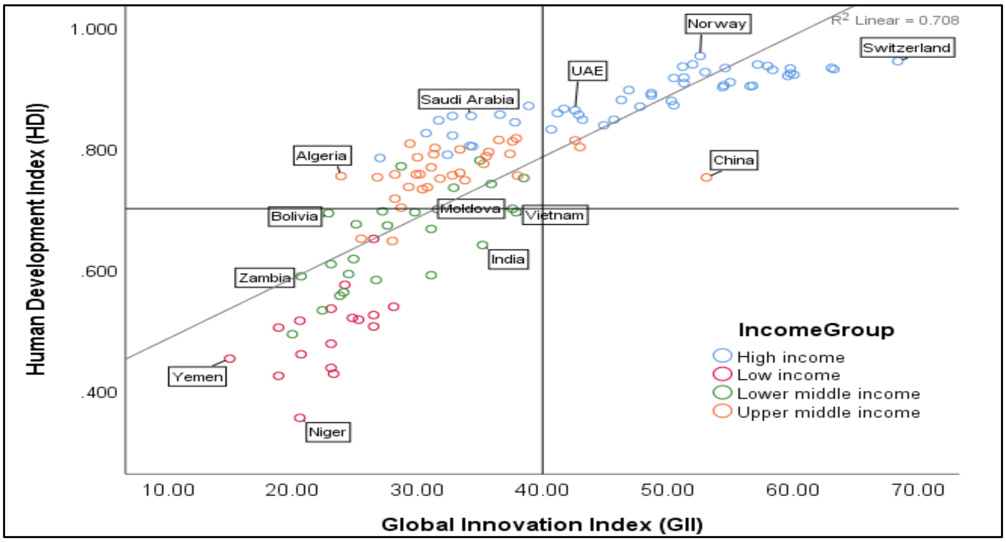
Source: Global Innovation Index Report (2018) and Dutta et al. (2018)

The findings on China's performance on the GII gives an indication of how the achievement of some of the SDG targets may improve the technological progress of countries and particular help low and middle income countries to avoid the MIT. This finding also gives credence to the earlier observations including that of the World Bank (2019) that developing countries may miss out on the benefits of the 4th Industrial Revolution, and also miss out on various SDG targets, if they do not make the necessary investments in human capital, science, innovation, research and development. Thus, the World Bank (2019) points out that, innovation will continue to accelerate, but developing countries will need to take rapid action to ensure they can compete in the economy of the future. Although China is still a middle income country, empirical studies indicate that investments in human capital and technological progress played a a key role in the the 5 East Asian countries (Japan, Taiwan, Singapore, South Korea, Hong Kong) that were able to break out of the MIT (Kanchoochat 2015).

4.3. Findings on the nexus between technological progress and Human Development Index

In order to investigate existing MIT literature claiming that countries with higher levels of technological development also have higher socio-economic development levels, a correlation analysis was conducted on the GII and HDI scores of the 126 countries The result of Spearman's correlation test of .842 indicates a strong significant positive relationship between the HDI and GII scores of countries. The implication of this result is that, countries with higher levels of technological advancement also tend to have higher HDIs. Furthermore, the coefficient of determination indicates that the GII scores helps to explain about 71% (.842 x .842 x 100) of the variance in HDI of different countries. As already discuss in section 4.3 there is also a strong positive correlation between the GNI per capita of countries and GII scores. This finding justifies the need for countries to invest in technological development not only for avoiding the MIT but also for improving the well-being of their citizens.

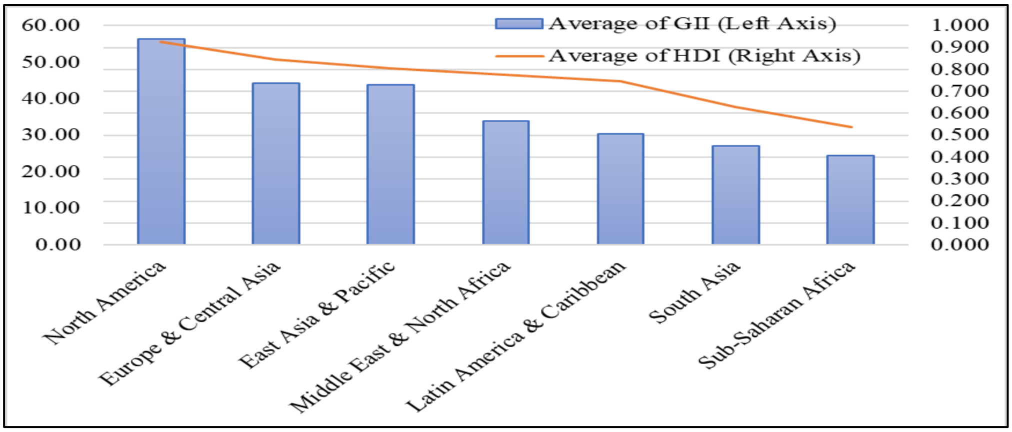
Figure 6 The relationship between the GII, HDI, and income groups of 126 Countries



Source: Authors construct based on data from World Bank (2018) and UNDP (2018) and GII.

In terms of geographical regions, the study finds that North America had the highest average GII scores as well as the highest average HDI scores followed by Europe and Central Asia (See Figure 7). The Sub-Saharan Africa region had the lowest average GII scores as well as the lowest HDI scores. The regions with higher GII scores also had higher HDI scores. These findings give an indication of the geographical location of countries with the lowest technological development as well as lower human development indicators. The findings also give an indication of the location of countries that are likely to miss out on the benefits of the 4th Industrial Revolution.

Figure 7 The relationship between the GII, HDI, by geographical regions

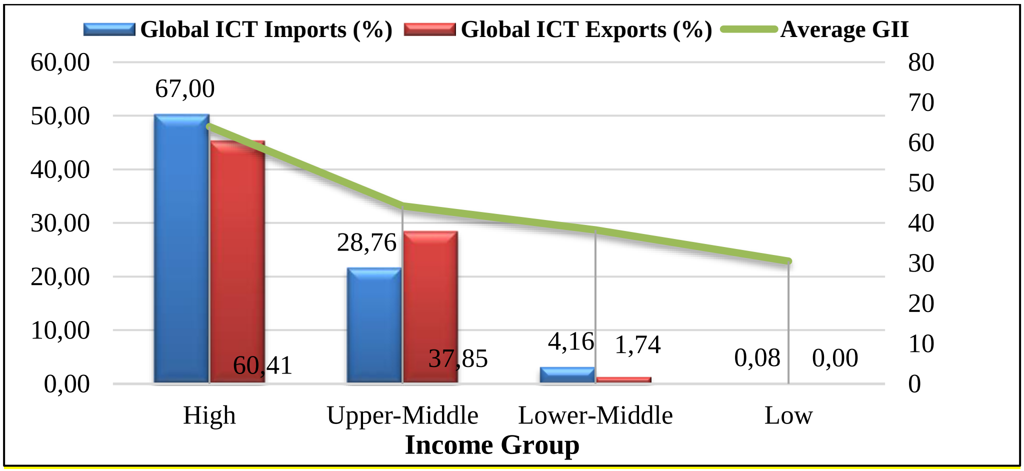


Source: Author's construction

4.4. Findings on the level of participation in the 4th Industrial Revolution by income groups

As indicated earlier, information and communication technologies (ICTs) are the backbone of the Fourth Industrial Revolution, and countries and businesses that embrace these developments, anticipate challenges, and deal with them in a strategic way are more likely to prosper, while those that do not will more likely fall behind (Samans–Hanouz 2016). In this regard, the study analsed data on global ICT Goods trade of countries by their income to access the extent to which countries are positioned to benefit from the 4th Industrial Revolution. The study finds great disparities between the high income and lower income countries (see Figure 8).

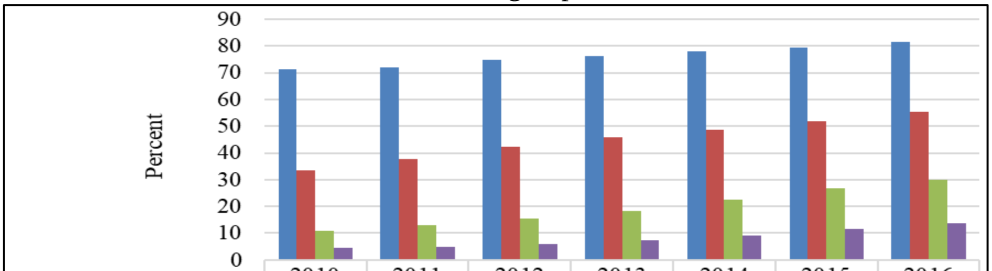
Figure 8 Correlation between GII, global ICT trade and income level of countries



Source: Author's Construction based on data from UNCTADStats and GII 2018 data.

An analysis of the most recent data on bilateral trade goods flows of countries by income level indicates that high income countries account for the majority of both global ICT imports (67%) and exports (60%) whereas middle income countries accounted for about 40%. Low income countries accounted for less than 1% of both global ICT imports and exports. There was also a positive correlation between the bilateral trade in ICT goods the and the GII scores (See Figure 8). These findings support the observation of Nikoloski (2016), who posits that developed countries have a monopoly on the sources of technological development and export of modern equipment and technology while developing countries are technologically dependent on developed countries.

Figure 9 Trend of individuals using the internet 2010–2016 (% of population) by income group



	2010	2011	2012	2013	2014	2015	2016
High income	71	72	75	76	78	79	82
Upper middle income	33	38	42	46	49	52	55
Lower middle income	11	13	16	18	23	27	30
Low income	4	5	6	7	9	12	14

Source: Author's construct based on World Bank (2018)

A trend analysis of the proportion of people using the internet in various countries from 2010 to 2016 also indicates that lower income countries lag behind the high-income countries (See Figure 9). This trend gives an indication of the extent of participation by countries in the 4th Industrial Revolution according to their income levels, and is consistent with the observation of UNCTAD (2017) that there is a "digital divide" between the rich and the poor, as developed countries massively buy goods or services from the Internet, while less than 5 per cent do so in most developing countries. As indicated earlier, the SDGs have several targets aimed at bridging the digital divide. SDG 9.c has the aim to "Significantly increase access to information and communications technology and strive to provide universal and affordable access to the Internet in least developed countries by 2020" (UN 2017). SDG target 17.8 also has the aim to "Fully operationalize the technology bank and science, technology and innovation capacity-building mechanism for least developed countries by 2017 and enhance the use of enabling technology, in particular information and communications technology" (UN 2015, UN 2017). The global indicator for measuring progress on SDG 17.8 is for instance, the 'Proportion of individuals using the Internet' in various countries (UN 2017). The findings of this study clearly indicate that the lower income countries are far behind on this indicator and this has implications for the middle income trap.

5. Conclusion and Recommendations

The study finds a significant difference between the level of technological progress among countries in different income brackets. Higher levels of technological advancement were finds to be a strong determinant of the socioeconomic development status of countries. Countries with higher GII scores were finds to have higher output, higher income per capita, and higher human development. There was a strong positive correlation between the GII scores of countries and their HDI (.842), GNI per capita ($r = .836$), as well their GDP per capita in both current prices ($r = .780$) and purchasing power parity ($r = .696$). The findings of this study are consistent with the MIT hypothesis that countries with higher levels of technological development also have higher incomes and higher socio-economic development. The findings of this study also corroborate suggestions that countries with higher technological progress are likely to benefit from the 4th Industrial Revolution. For instance, higher income countries appear to be benefitting more as evidenced by their trade in ICT goods and the percentage of people with access to the internet. The findings of this study also underscore the need for policy makers to pay attention to the observation of UNCTAD (2017) that a major global concern is the prevalence of various "digital divides" between the rich and the poor and that, as the old ones remain, new ones are emerging.

China's stand-out performance on the GII among middle income countries and rise in GII rankings, which has been attributed to investments in human capital, science, technology, and innovation, gives an indication of how developing countries could possibly avoid or break out of the so-called middle income trap (MIT), since the hypothesis is that countries that get stuck in the MIT are those that are lower down the technological ladder. The findings also give credence to the global calls on policy makers and governments as captured in the SDGs and the World Development Report for increased investments in research and innovation as well as health and education, which are the building blocks of human capital – the drivers of innovation – particularly for developing countries, if they are to harness the benefits of technology and mitigate its adverse disruptions which are an inevitable by-product of the 4th Industrial Revolution. Although the GII does not cover all countries, this paper has used the most recent data to confirm the technological advancement, socioeconomic development, and MIT nexus and emphasizes the need for governments to invest in human capital and technological development. The weakness of a cross-sectional study is that it does not allow for explanations and understanding causal processes that occur over time; it only provides a snapshot of the prevailing phenomena. The focus of this study was just to use current data to understand the relationship between technological development and socioeconomic development, but not to infer causality. Therefore, more in-depth time series analyses is needed to understand the mechanics of the potential role of innovation and technology in helping countries to avoid and break out of the MIT. Regression models could also be used in future studies to investigate causality and the impact of technology on socioeconomic development. The author together with Udvari Beata (PhD) and Associate Professor of Economics (name missing) at the University of Szeged in Hungary are currently undertaking in-depth studies on the determinants of the MIT and key drivers for escaping the MIT. Nevertheless, this paper contributes to the existing MIT literature by providing an empirical situation analysis of the relationship between the level of innovation of countries and their socioeconomic development status using the most recent data. Whiles the GII is the index currently available that captures the multi-dimensional facets of innovation and provides the tools that can assist in tailoring policies to promote long-term output growth, improved productivity, and job growth, the methodology and scope of the GII also needs further examination. Again, further studies could build on this study by investigating the implications of the SDGs on the MIT and participation in the 4th Industrial Revolution.

References

Babbie, E. (2008): *The Basics of Social Research.* Belmont: Sage Publications. Thomson Wadsworth.

Bajmócy, Z. – Gébert, J. (2014): The outlines of innovation policy in the capability approach. *Technology in Society*, 38, 93–102. Elsevier Ltd.

Barro, R. J. – Sala-i-Martin, X. (2004): *Economic Growth second edition:* Cambridge, London: The MIT Press.

Brue, S. L. –Grant, R.R. (2013): *The Evolution of Economic Thought Eighth Edition.* Mason: Cengage Learning, South-Western.

Cornell University – INSEAD – WIPO (2018): *The Global Innovation Index 2018: Energizing the World with Innovation*. Ithaca, Fontainebleau, and Geneva.

Dutta, S. – Lanvin, B. – Wunsch-Vincent, S. (2018): *Global Innovation Index Report, Energizing the world with innovation.* Geneva: Cornell University, INSEAD, and World Intellectual Property Organization.

Edler, J. – Fagerberg, J. (2017): Innovation policy: what, why, and how. *Oxford Review of Economic Policy*, 33, 1, 2–23.

Eichengreen, B. – Park, D. – Shin, K. (2013): *Growth Slowdowns Redux: New Evidence on the Middle Income Trap.* Cambridge: National Bureau of Economic Research.

Edquist, C. (2014): Striving towards a Holistic Innovation Policy in European countries – But linearity still prevails! *STI Policy Review*, 5, 2, 1–19.

Gill, I. – Kharas, H. (2015): *The Middle-Income Trap Turns Ten. Policy Research Working Paper No.7403.* Washington DC: World Bank.

Glawe, L. – Wagner, H. (2016): *The middle-income trap - definitions, theories and countries concerned: a literature survey.* Hagen: Munich Personal RePEc Archive.

Glawe, L. – Wagner, H. (2018). *The Middle-Income Trap 2.0, the Increasing Role of Human Capital in the Age of Automation and Implications for Developing Asia.* CEAMeS Discussion Paper No. 15/2018. Hagen: Center for East Asia Macro-economic Studies.

Li, T. – Wang, Y. (2018): Growth channels of human capital: A Chinese panel data study. *China Economic Review*, 51, 51, 309–322.

IMF (2018): *World Economic Outlook. Washington DC; International Monetary Fund.*

IMF (2019): *World Economic Outlook. Washington DC; International Monetary Fund.*

Kanchoochat, V. (2015): *The Middle-income Trap and East Asian Miracle Lessons Rethinking Development Strategies after the Financial Crisis Making the Case for Policy Space*. 1, 55–66. New York and Geneva: United Nations.

Mankiw, N. G. – Taylor, M. P. (2014): *Economics, 3rd Edition.* Hampshire: Cengage Learning

Nikoloski, K. (2016): Technology and Economic Development: Retrospective. *Journal of Process Management – New Technologies (JPMNT)*, 4, 4, 45–50.Ohno, K., 2009. Avoiding the middle-income trap: renovating industrial policy formulation in Vietnam. *ASEAN Economic Bulletin*, 26, 1, 25–43.

Samans, R. – Hanouz, M. D. (2016): Preface. In S. Baller, S. Dutta, & B. Lanvin, *The Global Information Technology Report 2016, Innovating in the Digital Economy* (p. v). Geneva: World Economic Forum and INSEAD.

Todaro, M. P. – Smith, S. C. (2015): *Economic Development 12th Edition.* Upper Sadle River, New Jersey: Pearson.

UN (2015): *Transforming Our World: The 2030 Agenda for Sustainable Development.* New York: United Nations.

UN (2017): *Resolution adopted by the General Assembly on 6 July 2017, A/RES/71/313: Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development.* New York. United Nations

UNCTAD (2017): *Information Economy Report 2017: Digitalization, Trade and Development.* Geneva. United Nations

UNDP (2018): *Human Development Indices and Indicators 2018 Statistical Update.* New York: United Nations Development Programme.

Walliman, N. (2006): *Social Research Methods.* London, Thousand Oaks, New Delhi: Sage Publications

Wang, L. – Li, M. – Abbey, C. – Rozelle, S. (2018): Human Capital and the Middle Income Trap: How Many of China's Youth are going to High School? *The Developing Economies*, 56, 2, 82–103.

World Bank (2019): *World Development Report 2019: Changing the world of work.* Washington DC: World Bank.

World Bank (2018): *World Development Indicators.* Washington DC: World Bank.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 210–222.

The effect of changing information technology on doctor–patient communication

Éva Málovics – Beáta Kincsesné Vajda

Information technology is changing rapidly worldwide and likewise in healthcare. The digital revolution will have a great impact on how physicians interact with patients. The Internet has become a powerful healthcare tool for many people. This will have an important effect on doctor-patient communication, and the transformation of these relationships has become a fundamental question. Due to these changes, an e-patient movement has evolved. This process requires a significant change in the attitude of patients as well as physicians. After a literature research, in order to explore how participants in the Hungarian healthcare system stand with this issue, we conducted in-depth interviews with six highly educated patients with serious illness and with four doctors. We used the grounded theory method to discover the characteristics in the doctor-patient communication. Our results show that some elements of becoming an e-patient have appeared on the patient side, but the habitual barriers of paternalistic doctor-patient communication overwrite these efforts on both sides.

Keywords: Information technology, communication, e-patient

1. Introduction

In a networked society, the digitization of information and communication changes the way people communicate and co-operate. E-shopping, e-banking, e-learning, and e-health are becoming increasingly important in our everyday lives. The letter 'e' mainly means 'electronic', but it involves much more than just a specific technology and refers to a series of far-reaching social and cultural changes. When it comes to e-health and e-patient, it's not just about the opportunities offered by the Internet and social media. It is also about changing the health care system, and about the opportunities that this entails in making health care more efficient and more humane.

According to Meskó (2016), the e-patient model means opening up a conservative system of health care to innovations. Why is this system so conservative? There is a consensus that this is a very high risk area. "The stakes in medicine are very high; when you apply new technology, you play with the lives of people" (Mesko 2016). But do they not play with the same stakes in our barely functioning, unworkable health care system?

In recent years, innovation in the area of healthcare robotics has shown significant growth; robotic nurses and surgical robots have been created to assist medical jobs. This poses several questions, including ethical ones, to answer. Will robots replace nurses? Can machines give the warmth humans are capable of? Do machines objectify care receivers? (Stahl–Coeckelbergh 2016)

The lack of financial viability and the crisis in health care, as well as the problems of doctor-patient communication are in fact world phenomena. In his book, Le Fanu (2008) argues that due to the unparalleled development of modern medicine in the 20th century, health care systems face four paradoxes:

1. Disillusioned doctors, despite the success of modern medicine.
2. Worried healthy people – surveys show that the proportion of people who worry about their health has also increased.
3. The growing popularity of alternative medicine.
4. The lack of financial viability in health care.

In this paper we try to investigate the topic of becoming an e-patient in all its complexity, and to explore the factors that contribute to the fact that many find this to be a good solution to the current problems of healthcare and doctor - patient communication, as well as those factors that hinder the realization of this concept in the Hungarian health care setting.

2. Innovations in the health care industry

Technological changes and new virtual devices may greatly influence the functioning of medicine. The way Internet and social network sites operate, basically determines individuals' attitudes towards healing, doctors, and patients. Finding and sharing information has become easier than ever, and a large number of online patient forums point to this need. Patients are no longer satisfied with the information provided by doctors, and they are also interested in the symptoms of those with similar diseases or symptoms, which on the one hand can increase anxiety and on the other possibly counteract the passivity of interaction with a doctor. Properly selected pages and groups increase the possibility of self-management. In rheumatology, for example, there are solutions that use social media to promote self-management. In addition, methods (applications) with which, for example, self-evaluation or sending information remotely to the doctor becomes easier, are constantly developing. Online solutions offer many benefits not only to users but to doctors, too, in healing as well as in learning and training. Despite the many benefits of using digital devices, they have drawbacks in the healing process as well: if meetings and personal consultations are replaced by digital interactions and information transmissions, the potential in human relationships – the expression of empathy, the interpretation of non-verbal signals in conversation – will be degraded, and also, there is a need for new regulations on the adequacy of the processing of personal data in the digital space. The isolation effect of electronic devices seems to be compensated for by new forms of communication and connections; the task being to create the right boundaries for this (Berenbaum 2018). In the extension of the physical limits of care, it is important to consider how much both medical staff and patients are willing to move away from the traditional physician-patient relationship (McCabe–Timmins 2016). Our previous research suggests that there is a growing demand for patient-centred care based on partnership (Kincsesné 2014).

Nowadays, the individual data collected by wearable sensors or activity trackers (e.g. step counter, sleep monitor) provide an opportunity for certain groups in society to monitor and improve their fitness, and hence their health. The ability of the individual to be responsible for - and actually an administrator of - the improvement of their own health is called "self-quantification", which can be realized here through a system that helps the individual to gather relevant information for self-reflection and self-awareness. Related research has looked at the effectiveness of activity trackers from many aspects, and, according to some results, systems that are easy and enjoyable to use - making it easier and more enjoyable to improve fitness and health - are more likely to be recommended to acquaintances, and the likelihood of this is increased by the perception that the device has a positive effect on behaviour and by gamification as well. Social effects can also act paradoxically in this case – encouraging each other or engaging in common challenges may increase, while social pressure may reduce motivation (Ilhan–Henkel 2018).

Wearable technology (WT) not only includes activity trackers and smartphones. These only represent one category that works independently or is connected to other devices in order to convey information. In the other category, there are tools that record specific activity or measures (such as a heart rate monitors worn around the chest) that transmit data to some other primary device. A third category, now unspecified, consists of so-called smart textiles, the scope of which makes them a key future innovation. Wearable devices, after their appearance, almost immediately entered the area of health care, but there are still problems with the separation of healthcare-related and general use and with the clarification of classification for clinical use (Godfrey et al. 2018).

Technological solutions play an important role not only in prevention but also in patient care. Nisar–Shafiq (2018) in their research found that new generations of healthcare staff are eager to use the Internet and online social networks to communicate health-related information. They refer to prior research in which about sixty percent of physicians and sixty-five percent of nurses admitted using social network sites for professional purposes; and add that patients also feel that accessing information this way facilitates the understanding of treatment methods and thus promote better health outcomes. Therefore, it is becoming increasingly common for health information to be obtained from the Internet rather than from physicians. As a result, hospitals around the world have started using social media to communicate with patients. In developing countries, this type of information access is particularly important, as people in underdeveloped areas have easier access to information from professionals working in developed countries.

3. The effect of technological development on the doctor-patient relationship

Professionals generally draw attention to the fact that new technologies present both opportunities and dangers, which is also true for changes in access to health data. New technologies provide an opportunity to solve old problems in a new way, optimize communication processes, change old roles and hierarchies, improve the efficiency and quality of health services, and enable people to rethink / reconsider how they treat their health and illness (Meskó 2016).

The prevalence of blogs, webpages, customizable search interfaces and forums dealing with health care, diseases and prevention have led to more conscious consumers who access a significantly bigger mass of information, and for whom it is easier to get advice from other users. While doctor-patient relationships were previously characterized by the fact that all knowledge and information was in the hands of the physician, now the appearance of Internet has resulted in a shift, so that patients in an initiating role can even start a dialogue on their optimal health solution (PWC 2018).

Many doctors have a very negative opinion on patients using Google to answer healthcare questions, which is probably due to a variety of reasons. Doctors we interviewed justified rejecting 'internet' patients by claiming that this avenue resulted in patients receiving misinformation, and that it is intrinsically time-consuming to answer such questions. There were no mentions, however, of the fact that information received this way may be useful not only for patients but for doctors too. Knowledge sharing has at least two sides; on the one hand, knowledge is power, so if we share it, we seem to share our power. On the other side, knowledge is the only good that multiplies with sharing, and thus, its quality can improve (Davenport–Prusak 1998).On the basis of this, it would surely seem worth sharing knowledge with each other. However, for healthcare providers who are accustomed to hierarchy and to a monopoly of knowledge, it is clearly difficult to lose authority, despite it possibly leading to benefits such as better quality knowledge and service, greater consumer satisfaction, and adherence to health advice.

So, patients have more information about their own condition which leads to a completely new doctor-patient relationship: they may become partners. However, not only the increasing amount of available information plays a role in the transformation of medicine, but also technological development. With the development of smartphones and wearable smart watches, this awareness has entered a new level: consumers are not only able to read about the topics they are interested in, but they can also analyse easily available, objective and measurable data about their physical condition (Chiauzzi et al. 2015). Sooner or later, the usual medical visit will not begin with a personal meeting of the physician-patient; instead, the first step will be an e-mail exchange by which patients sends their sleeping curves, cardiogram, blood oxygen level or even blood count recorded for weeks on their smart devices. These devices nowadays got within reach; it is no extraordinary anymore to see touchscreen gadgets on one's arm or wrist that records data twenty-four hours a day. And all this comes with a completely new patient-doctor relationship, as the information will be available to both sides, so that they will face each other as partners. More and more consumers are turning to dr. Google with the aim of prevention or self-diagnosis - this process cannot be stopped. There is a need for reliable, smart-tech, wearable or always available consumer applications and solutions (Wu-Luo (2019).

According to Fox (2008), those people who feel they have a lot at stake (e.g. users living with a disability or a chronic disease) are more likely to engage intensely with online resources. The author refers to a research in which 75% of e-patients with a chronic condition stated their last health search affected a decision about how to treat an illness or condition and also calls our attention to the fact that the impact of an online search is proven to be more useful than harmful. Being an experienced e-patient in his words also means posting technical advice online about how to manage a certain disease as well as advising people about how to communicate with health care staff.

The emerging world of e-patients, according to Ferguson–Frydman (2004) can be characterized by the following phenomena:

– clinicians underestimate the benefits and overestimate the risks of the cases when patients use online health resources
– online support groups have become a significant source of information and emotional support, especially for those living with a serious or a rare disease, sometimes having an even more important role than medical staff
– net friendliness of physicians have become an important aspect of health-care quality.

The aforementioned trends are still highly disputed today; health care systems are known to be conservative and slowly changing, and many of the service providers do not want to accept these changes. This is a natural consequence of change; the first reaction of people as well as organizations, due to their basic need for homeostasis, is resistance to new things (Atkinson et al. 2002).

This situation is further complicated by the fact that healthcare-related work usually involves very high levels of risk, but risk perception of the parties varies greatly, as service providers are professionals, whereas patients are lay. Thus, thoughtful risk communication would be needed in many areas. Healthcare providers often perceive high-risk and high-responsibility solutions as barriers to new opportunities offered by technology (Yarbrough–Smith 2007).

4. Methodology and sample

Our research question is whether new directions that are predicted and considered important by technological changes can be discovered in physician-patient communications. One side of communication is represented by the patient: by his or her information, and where he/she obtains knowledge about the illness and its cure. The other party is the doctor and what he/she thinks about the changes in communication with patients in the light of the present technological changes.

Due to the complexity of our research subject, the qualitative part of our empirical research is made up of in-depth interviews, analysed by the grounded theory method. This method have several trends; in our research, we have used the improved version provided by Strauss and Corbin, according to which concepts already used in the literature may be used during the coding process, and both inductive and deductive logic are allowed. The validity of the research results is ensured by the fact that two researchers independently analyse the transcripts of the interview and then negotiate their analysis results. With a thorough study of the transcripts, the researchers search for codes that are relevant to the research question in the text, and note all thoughts and ideas that emerge in the form of memos. After studying the first results they form categories from related concepts and explore the interrelationships between categories and search for key categories for answering the research question. Based on the relationships between the categories, the results are outlined in a research model (Charmaz 2013).

The analysis of our interviews was carried out in a constructivist interpretative approach: we studied the reports, intentions, actions of the subjects, and tried to follow the direction from the data. (Horváth–Mitev 2015) According to the interpretive approach, people are not only passively reacting to external reality, but through their inner perceptions and thoughts they examine the outside world, thus actively creating their own reality. It follows from the above that we do not intend to test hypotheses with the grounded theory method, but it is possible to formulate statements about how the subjects of the research interpret reality (Horváth–Mitev 2015).

Sampling was done with a snowball method, and we also used the quality sampling of the Grounded theory method, which means that the data that are already being prepared during the analysis of the interviews also direct us to who and what to ask. We have conducted interviews with 4 doctors and 6 patients. The grounded theory method does not require extensive interview structures, and we conducted the interviews along our research questions.

During sampling, our goal was to ask highly educated patients with severe illness because we supposed that the severity of the illness would provide sufficient motivation for getting information on possible therapies on the internet and that they had the necessary conditions to do so due to their level of education. Table 1 and 2 shows the data of our research subjects.

Table 1 Qualitative sample data – patients

Code	Sex	Age group	Disease
P1	Female	55–64 years	breast cancer
P2	Female	45–54 years	breast cancer
P3	Female	65–74 years	diabetes, hypertension, stroke
P4	Male	65–74 years	bladder cancer
P5	Female	45–54 years	CMT
P6	Female	75–84 years	skin cancer

Source: Own construction

Table 2 Qualitative sample data – doctors

Code	Sex	Age group	Specialization
D1	Male	65–74 years	abdominal surgeon specialist
D2	Female	55–64 years	oncologist specialist
D3	Male	25–34 years	internal medicine specialist
D4	Male	25–34 years	internal medicine specialist

Source: Own construction

5. Qualitative results

The results of the analysis of our in-depth interviews are presented on the basis of the most important categories and interrelations that were gained from the transcripts. Due to the difference of topics, we present the results of doctor- and patient interviews in a separate way.

5.1. Results of doctor interviews

Overall, we found that none of the doctors we interviewed heard about the concept of e-patient, they did not know what it could mean. Some of them created their own version of it but nothing about what it really means. Although they seem to consider doctor–patient communication important, they are satisfied with its paternalistic style. Internet use of patients is rather negatively perceived and they connect it to hypochondria. When we clarified the concept, they were somewhat sceptic about it and state e-patients are not present in their practice.

Regarding the detailed results, we found five important categories related to the research question to emerge.

First, they do not know the concept of the e-patient (*"I haven't met it, but it may exist"* (D4)) The question thus arises: why hadn't they met such a thing when many patients surely read about their illness, and possibly about therapies on the Internet? Why doesn't patient knowledge figure in consultations? We believe that it is due to the paternalist culture in which both patients and doctors are socialized in our culture. Several doctors have tried to highlight the issue of lay patients discussing their therapies with the example of car mechanics. However, we consider the example of car repair is inadequate to use in this situation, as patients know their body better than a car. People may have greater knowledge about health, certain diseases that have occurred in their family, and if there is a problem, many are very motivated to get information – which usually does not happen with cars.

Second, patients who are informed or search for information on the Internet and wish to participate in decisions are evaluated negatively. They consider lifestyle issues an important area in which patients should be informed, but overall, they state therapy decisions should be made by doctors; *"there are cases where there is a standard treatment, but the patient arbitrarily thinks he is not taking it now because he read this and this and all that, so unfortunately, we see this, for example in chemotherapy treatment, that patients read something bad about it and does not start it despite that he should, so from this point of view, reading about it is negative"* (D4)

In her TED talk, Heffernan (2012) states that health care is very conservative, it is changing too slowly, and proven new procedures are not entering therapies promptly enough, costing lives. According to Heffernan, what has changed is that today this information no longer reaches only doctors, but is also available to a large number of patients, often sooner than to doctors. Consequently, information asymmetry has been reduced, but patients can only discuss this new information with friends on- and offline, not with doctors. Their paternalistic style has not changed, and in part because of this, trust and satisfaction in physician-patient communication has reduced. The negative effects of the Internet were emphasized by all interviewees. However, no-one talked about the possibilities of channelling patient energy spent in information searches to improving the quality of the cure. *“There will be more and more people spending more time searching for information on their own health issues on the Internet, but the issue of how competent they are depends on and how much false information and stupidity can be found there”* (D3).

Third, there are generational differences among doctors. Younger doctors believe that they are in a more symmetrical partnership with patients than older doctors. *“They do not like the "smart patient", so to speak, but during my professional practice I spent in Kecskemét, there was a younger doctor, a specialist, he was 30 years old, he was a gentleman, and he explicitly encouraged patients to search for information, and to ask anything”* (D3). Generational differences are also perceived in connection with ICT use.*”... in the electronic space that doctors have, there are such pop-ups as in the browser, which reminds the doctor of what information is incomplete (...) I cannot imagine that an older general practitioner deals with it, for a younger, obviously, this might be a motivation.”* (D3).

Fourth, doctor-patient communication is considered to be important, but has many barriers. One of the aspects assumes that patients are incompetent, therefore, they only see what is easy to notice. *“They only see or experience the way their dentist talk, how he approaches the situation, how he treats them, so he can be a good professional, but if meanwhile he ‘kicks’ patients, they won’t come back to him, as they don’t have the competence to see whether the work done was any good”* (D3). During their university studies, our subjects didn’t perceive communication as having much importance. moreover, organizational hierarchy and lack of time favors one-way communication: *“in inpatient institutions there is a very strong hierarchy within the medical community… doctors in lower positions are entrusted with the care of the patients, and senior doctors only supervise them. ... They only meet the patient effectively during visits, and there is the expectation that there should be silence, everyone should lie well, answer 3–4 questions they receive… there is little time, we are moving on… we have a department, with let’s say 4 wards, 6 people in every ward, and the visit should take place in 1 hour”* (D3). According to our subjects, this type of hierarchy sometimes creates humiliating situations for younger doctors, as older doctors correct them in front of patients, thus, they receive negative feedback in public. As we go up in the hierarchy, the control distribution becomes more asymmetric and communication is increasingly one-way. Strong hierarchy worsens communication in many respects, the powers of non-doctors are severely limited in terms of communication, the are not confident enough to provide information, and those in high positions are unwilling to do so, while “small fry doctors” are overwhelmed with this. Money is a tool that improves communication; with patients who pay, two-way communication is more frequent. The ever-growing role of technical innovations could theoretically also improve doctor-patient communication, but doctors we interviewed saw this in different ways. Since they introduced the so-called EESZT, there has been a lot of disruption, some of our subjects were grateful that it was not working, others have seen many benefits. As for the role of communication in adherence, its difference from the concept of compliance was mentioned: *"I think there is compliance as a professional term, maybe I think it is outdated. and now they are trying to introduce adherence instead. It does not mean that the doctor will tell you what to do, and the patient will obey, but that together, there is an adherence between them, which means mutual trust (...) Both parties are trying to find the solution that is acceptable to both of them ... I think the doctor who communicates properly with his patients will get this adherence… roughly, this is typically what is dependent on communication"* (D3). We can also see that although our subject knows the words, but cannot understand fully what adherence means; he is talking about communication tricks to make the patient accept the selected therapy willingly: *"…one learns communication tricks that can make the patient feel a little better, and otherwise these are not great tricks, let's just say a nice comment, smile, saying please, thank you, and so on (...) so if this adherence is created, and that is why they seek it, the majority of doctors will strive for it"* (D3).

Finally, the role of new technological trends and robotization is acknowledged and considered to improve the relationship between doctors and patients.

5.2. Results of patient interviews

According to the results of the interviews we conducted with patients, we could not find a "real" e-patient. The characteristics of the e-patient were best approached by one of our interviewees, looking for doctors with whom she could talk about the therapeutic options and with whom she could participate in decisions. However, she does not share her experiences on the Internet, although anyone can contact her in person, and she likes to share her knowledge. She fears the Internet because of the possibility of negative opinions, envy, because the therapies she chose have cost her a lot.

Based on the information retrieved from the interviews, we derived nine relevant categories.

First, the importance of searching information about the therapy is acknowledged, although our subjects use different methods for that purpose; problems of information provided by different doctors also emerged. Our respondents were informed from different sources about their illness and possible therapies. Although we were deliberately looking for well-educated patients (there were health care workers - assistants - and a biologist - genetic researcher and medical researcher - among them), it is interesting that this did not significantly affect being an e-patient, although everyone had "e-patient" attitudes, but none could be considered an e-patient as described in the literature. However, all the interviewees searched for information themselves according to their means: *"I got the paper, I went home, I turned on the Internet, I checked that it was a two-degree, so-so-centimeter, so I was informed from the Internet, as despite all the pleasant atmosphere and compassion there was no time to explain it to me there how much I am in trouble, but it was understandable"*(P3). Our subject, who is a researcher, read studies of her illness in the original literature and took it to the doctors who were not always willing to take it into account –

therefore, she changed her doctors. Patient-specific initial confidence was replaced by a series of disappointments due to rigid, incomplete information, and quickly enforced decisions. Patients did not have time to think, nor did they dare to ask for it even before serious decisions: *"... The doctor said that I should have chemo and there are two types, the stronger one, either I go through with that, or a milder type that actually doesn't let the hair fall, but actually that the former is safer. And then I said if I could bear it, let's have stronger one, I didn't have more info about it all, it would have been a good idea to have a booklet about it because it was serious there… communication is always distorted, there was a decision to immediately put everything right there and then, no questions come to your mind... you don't think there can be lasting side effects, it turned out that the picture is not so bright...* (P3). The various specialists advised our subjects quite differently, which also had a negative impact on them. Sometimes it was only the solution that was announced. Most of the patients accepted incomplete information because of their gratitude for quick treatment, and they didn't dare ask for more in most cases. Also, in many cases, doctors used one-way communication in giving information that was hard to understand, and different specialists happened to communicate different prospects, depending on patient responses.

Second, patients misunderstand doctors in several cases, which frequently comes up later, because they do not dare to question them, or had no chance to do so. This is a typical drawback of one-way communication, which is often used in healthcare.

"... And then they sent me off telling me it is ok, and I became very happy about that, but they sent me to oncology too. And I went there, why not, and there came the bad news, the doctor told me there that the picture is not so clear, I should have chemotherapy" (P3).

Third, most respondents perceived the lack of time in communication, it was not clear for them how much time they have for the consultation, and this contributed to the one-way communication: *"... The doctor was in a hurry, but told me what was needed. And it was a very big slap when he said it was going to be an operation and then I hoped for a moment that they were mistaken or there was only an inflammation (...) then he said it is going to be mastectomy* (P3).

Fourth, at many specialist departments, patients wait for hours in crowded, airless places, in extreme cases leading to people collapsing. It so happens that little individual sensitivity is accepted, although the patient remarked: "I only see the present of the doctor-patient communication, but the future is the big question mark. I see the present as being very overcrowded (…) This affects me even more, because my blood pressure is very high and I take the prescribed medication in vain. During my last four visits I got sick… I usually go for an appointment, but for some reason they are always disrupted" (P6).

Fifth, paternalism, due to its relevance in our culture, is actually instilled in the behavior of patients: "I don't have doubts, why should I? I'm not the doctor, if he thinks this is okay for me, then he must be right" (P2).

Sixth, most patients feel empathy with doctors because they see that there are many patients, and all our subjects used their relationships within the system, hoping to get better treatment, so they thought they had to tolerate incomplete communication. They didn't need to wait for long, so they're very grateful and feel they have no right to self-advocate.

Seventh, every interviewee turned to their doctor with confidence and trust when their illness was confirmed. *"I trust him in that he will do everything that is meaningful and what is reasonable and (...) and will do so with the necessary attention"* (P6).

Eighth, the importance of kindness and attention is acknowledged. Our respondents, according to their stories, met with lots of kindness and attentiveness from health professionals.

Ninth, disappointment also appeared: all of our subjects reported abusive treatment, abuse of the patient's vulnerable position, and denial of basic, legitimate requests, as well as brutal honesty and lack of empathy: *"The nurse said this was an emergency department, not a luxury hotel and I didn't get my nose drops out of my bag, though I couldn't breathe"* (P1)

6. Summary and conclusions

As a result of technical changes, the risk perception of patients and healthcare professionals is also changing. Patients who are informed about their disease and its therapy from the Internet go to the doctors with some knowledge, while doctors offer them the protocol, even in cases in which they are already aware that it is being challenged professionally. Patients assume a special status with their knowledge of their symptoms, which has led to the notion of "lay expert" in the literature. If they are able to share this knowledge with their physician, actively engage in the development of their health, and share their knowledge and experience in online patient communities, they become e-patients.

According to our research, the concept of e-patient seems rather idealistic and impracticable in a Hungarian setting, and patients are convinced that only paternalistic communication is possible in the field of health services, so they seem to accept the doctor's decision on therapy. Our respondents first usually turned to their doctors with confidence and were partly satisfied with the information, but as they began to get in touch with newer health professionals, more and more deficiencies were discovered. It seems that the initial trust in the doctor is disrupted by other information. As the patient's knowledge is not included in the communication, but information asymmetry between the two sides increases, it causes loss of trust and dissatisfaction in the patients.

During therapy, the majority of those interviewed do not dare to discuss their knowledge, doubts, and questions with their physician. They are under stress, they are embarrassed, they don't know how much time they can spend with the doctor. Using the functional model of communication, the patient detects a very asymmetric control distribution favoring doctors, with both positive and negative manifestations of affection. Affiliations would be important, but medicine has changed in a mechanical direction, which has brought some disadvantages for both sides. It is important for patients to know how much time they have at the consultation and what to ask and communicate in a few minutes. Our conclusion is that although the respondents are actively addressing the issue of their recovery, doctor-patient communication takes place in the traditional paternalistic style, one of the reasons being the patient's attitudes, an hence patient education would be needed to increase the effectiveness of clinical service.

Acknowledgements

This research was supported by the EU-funded Hungarian grant EFOP-3.6.1-16-2016-00008.

References

Atkinson, R. L. – Atkinson, R. C. – Smith, E. E. – Bem, D. J. – Nolen-Hoeksema, S. (2002): *Pszichológia*. Osiris Kiadó, Budapest.

Berenbaum, F. (2018): E-Health, social media, and rheumatology: Can they get along? *Joint Bone Spine*, 85, 3, 265–266.

Charmaz, K. (2013): Lehorgonyzott elmélet. In: *Kvalitatív kutatási olvasókönyv - Szavak, képek jelentés.* Bodor P. (szerk.) Budapest: L'Harmattan Kiadó.

Chiauzzi, E. – Rodarte, C. – DasMahapatra, P. (2015): Patient-centered activity monitoring in the self-management of chronic health conditions. BMC Medicine 13:77 DOI 10.1186/s12916-015-0319-2

Ferguson, T. – Frydman, G. (2004): The first generation of e-patients: These new medical colleagues could provide sustainable healthcare solutions. *British Medical Journal*, 328, 7449, 1148–1149.

Davenport, TH, – Prusak, L. (1998): Working knowledge: How organizations manage what they know. Harvard Business Press.

Fox, S. (2008): The Engaged E-Patient Population. https://www.pewinternet.org/2008/08/26/the-engaged-e-patient-population/ Accessed at 30 April 2019.

Godfrey, A. – Hetherington, V. – Shum, H. – Bonato, P. – Lovell, N. H. – Stuart, S. (2018): From A to Z: Wearable technology explained. *Maturitas*, 113, 40–47.

Heffernan, M (2012): Dare to disagree. TED talk, https://www.ted.com/talks/margaret_heffernan_dare_to_disagree?language=en

Ilhan, A. – Henkel, M. (2018): 10.000 Steps a Day for Health? User-based Evaluation of Wearable Activity Trackers. *Proceedings of the 51st Hawaii International Conference on System Sciences*, 3376–3385.

Horváth, D. – Mitev, A. (2015): Alternatív kvalitatív kutatási kézikönyv. Alinea Kiadó, Budapest.

Le Fanu, J. (2008): *Az orvostudomány önkritikája*. Typotex, Budapest.
Kincsesné, V. B. (2014): Az egészségügyi ellátás minőségének SERVQUAL szempontú mérése. *Marketing és Menedzsment*, 48, 1, 47–58.
McCabe, C. – Timmins, F. (2016): Embracing healthcare technology – What is the way forward for nurse education? *Nurse Education in Practice*, 21, 104–106.
Meskó, B. (2016): *Az orvoslás jövője*. HVG Kiadó Zrt, Budapest.
Nisar, S. – Shafiq, M. (2018): Framework for utilisation of social media in Pakistan's healthcare sector. Technology in Society, https://doi.org/10.1016/j.techsoc.2018.09.003
PWC (2018): Top health industry issues of 2018: A year for resilience and uncertainty. https://www.pwc.com/us/en/health-industries/assets/pwc-health-research-institute-top-health-industry-issues-of-2018-report.pdf Accessed at 5 May 2019.
Stahl, B. C. – Coeckelbergh, M. (2016): Ethics of healthcare robotics: Towards responsible research and innovation. *Robotics and Autonomous Systems*, 86: 152–161.
Wu, M. – Lou, J. (2019): Wearable Technology Applications in Healthcare: A Literature Review. *ONline Journal of Nursing Informatics Contributors*, 23, 3, available at http://www.himss.org/ojni
Yarbrough, A. K. – Smith, T. B. (2007): Technology Acceptance among Physicians: A New Take on TAM. *Medical Care Research and Review*, 64, 6, 650–672.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 223–239.

The Sharing Economy – An opportunity for individuals to become entrepreneurs

Elena Ştiubea

The phenomenon of the sharing economy is a relatively new one, emerging as a result of the development of the technology field, but also of the changes in people's consumption behaviour. The collaborative economy is a way to stimulate entrepreneurship and to open new horizons for ordinary people. This type of economy can stimulate people with different levels of training and bring them into the economic market much more easily, thus contributing to economic development. Access to the collaborative economy is driven by the development of technologies, while through the development of sharing economics new value is created. This article sets out to highlight the way in which the sharing economy can make ordinary people successful entrepreneurs, the opportunities in such a situation, and the impediments encountered in such an economy.

Keywords: sharing economy; entrepreneurs; changes; benefits.

1. Introduction

Entrepreneurial activity has a huge impact in terms of its contribution to economic growth, employment, social inclusion, and stimulating the economy etc (Hatos et al. 2015). Over the last few years, the world economy has gone through a series of changes. The economic crisis in 2008, as well as the rise of the Internet and everything related to technology and the digital world has led consumers to adopt a new attitude in relation to their consumption habits and to the way they spend their money. The collaborative economy seeks to share the resources/properties that people hold in order to save resources as well as creating added value.

Botsman (2015) defines the collaborative economy as "an economic system of decentralized networks and marketplaces that unlocks the value of underused assets by matching needs and haves, in ways that bypass traditional middlemen."

Small entrepreneurs on such a market are no others than individuals or providers who supply goods and services on this market. At the same time, it is interesting to find out whether the entrepreneurial aspirations of consumers are being explored.

The collaborative economy, through its mechanisms helps maintain a less polluted environment, and also contributes to economic growth. Helped by technology, the collaborative economy establishes a link between people-goods-services in a more efficient way, creating new business opportunities and giving people the chance to become entrepreneurs as well.

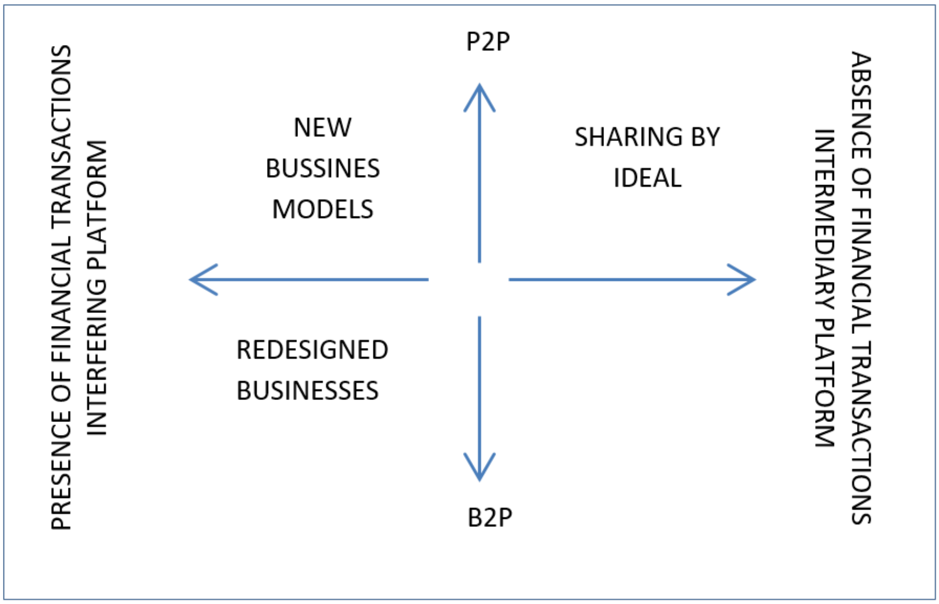
This type of economy has evolved rapidly thanks to its low barrier to entry, flexibility, convenience, and minimal regulations, and will continue to grow because it is based on technology, and technology is increasingly present and indispensable in people's lives.

In the first place, the collaborative economy may initially mean only a source of additional income, but this mechanism can be developed so that ordinary people can become small entrepreneurs with great chances of growing. The economy is changing rapidly and governments should support this type of economic development and come to support new types of business generated by the collaborative economy.

In all its manifestations, the collaborative economy has the ability to create the right environment in which people can grow their own business much more easily than by means of the classical economy. The mere fact that you have an affinity for a particular good, can lead to successful entrepreneurship.

The two models of sharing: access and transfer were identified in the two market structures: P2P and B2P. No B2P market structure was identified when there was the absence of financial transactions. Therefore, three archetypes were identified: new business models, redesigned businesses, and sharing by ideal (Petrini et al. 2017) (Figure 1).

Figure 1 Typology of sharing economy



Source: Petrini, M. and Freitas, C. S. de. and Silveira, L. M. da S. (2017)

2. Collaborative economics and entrepreneurship

2.1. Entrepreneurship and the entrepreneur

Entrepreneurship is a much discussed topic, almost anyone will be open to discuss or express an opinion on this topic. Sexton–Bowman (1991) said that entrepreneurship is a general approach to management that begins with identifying opportunities and culminates with their exploitation. This environment offers you plenty of advantages. It is an innovative, thrilling, dynamic, creative environment with potential to create the desired product/ service. Besides the advantages of creating your own business, it puts you face to face with certain risks and threats, and overcoming them requires a lot of work, involvement, determination, responsibility, and entrepreneurial spirit. We can consider entrepreneurship as a permanent domain where the entrepreneur plays a decisive role. Entrepreneurship means the capacity and ability (differing from one person to another, from one company to another) to coordinate the re-allocation (Cadar–Badulescu 2017). The entrepreneur theme is the idea that entrepreneurship involves individuals with unique personality characteristics and abilities (Gartner 1990). The entrepreneur is a person possessing abilities to solve crisis, a leader, motivated to undertake risks. Marshall suggests that the qualities associated with a good entrepreneur are rare and limited, being "so great and so numerous that very few people can exhibit them all in a very high degree" (Burnett cited in Cadar–Badulescu 2017). At the same time, entrepreneurship is an activity that has implications on more people, not just the entrepreneur. He cannot act alone because he needs collaborators to be able to achieve his own vision of the business he has initiated. Entrepreneurship involves entrepreneurs mastering several qualities such as: vision, intelligence and creativity, determination, responsibility, problem-solver, gut-feeling, positive thinking, professional ethics, and not least passion for their own business (Constantin 2014).

The innovative spirit of the entrepreneur in all its forms is very important when it comes to entrepreneurship in the sharing economy. This is an area in continuous flux and transformation, so the way in which the entrepreneur manages to capitalize and materialize his knowledge and abilities means a plus of value added to the business.

The entrepreneurial spirit is a starting point with a beneficial impact on the economy and on all of society. Through their activity, new labor markets and opportunities are created, which in turn leads to higher incomes and to social benefit in general. Through these dynamics and through the advantages that it creates, this area has a unique charm.

The importance of entrepreneurship is given by a series of reasons:

- Entrepreneurs create new business: By creating new business in the form of goods and services, new jobs are being created and that has a beneficial effect on the economy.
- Entrepreneurs also create social change: Through their unique offerings of goods, entrepreneurs indirectly support freedom by reducing dependence on

old technology systems. Overall, this thing leads to an improvement in the quality of life, a greater moral and economic freedom.

– Community Development: entrepreneurs regularly support entrepreneurial ventures by other similar individuals. They also invest in community projects and provide financial support to charities. Some famous entrepreneurs such as Bill Gates, have used their money for financing noble causes, be it education or health.
– The Role of States: Regulations have an essential role in stimulating entrepreneurship.

These regulations require a fine balancing act from regulatory authorities.

Unregulated entrepreneurship can lead to undesired social results, including unfair market practices, financial crisis, and corruption. Paradoxically, a large number of entrepreneurs could lead to fiercer competition and the loss of career choices for individuals. Also, in these conditions, aspiratory levels usually also rise. Given the variability of success in entrepreneurship, the case of having too many entrepreneurs can also lead to income inequality, weakening social cohesion. Meanwhile, the interesting interaction between entrepreneurship and economic development is providing key contributions and inferences to policy makers, development institutes, business owners, change agencies, and charity.

If we understand the advantages and disadvantages, a balanced approach to entrepreneurship will certainly have a positive impact on the economy and society. The entrepreneur has an attitude and behavior based on receptivity to new information and people; he makes independent decisions, sees opportunities in rapidly changing and uncertain economic environment, he is persistent, available to assimilate technical knowledge, has outstanding personality, leadership, and management abilities (Johnson 2001).

However, as a basic definition of entrepreneurship, it's a bit limiting. The more modern definition is also about transforming the world by solving big problems. Like initiating social change, creating an innovative product, or presenting a new life-changing solution (Ferreira 2018).

2.2. The entrepreneurship through sharing economy

Sharing tends to be a communal act that links us to other people. (Belk 2010). One of the first books on this topic was "What's mine is yours", in which Botsman and Rogers described in 2010 how the collaborative economy changes our lives. Botsman defines the collaborative economy as "an economic system of decentralized networks and marketplaces that unlocks the value of underused assets by matching needs and haves, in ways that bypass traditional middlemen" (Botsman–Rogers 2010).

Entrepreneurship is the symbol of business tenacity and achievement; it is a vital source of change in all facets of society, and its integration into the collaborative economy can be a real success (Pahuja–Sanjeev 2015). Collaborative consumption is people coordinating the acquisition and distribution of a resource for a fee or other compensation (Belk 2013).

In order to become an entrepreneur, in a traditional economy one has to perform certain steps. The first step to be accomplished is to identify the domain in which one wishes to be active, in fact, one must find a niche in the respective domain. It must be chosen in accordance with the skills and qualifications that the person who wants to become an entrepreneur has. A second step would be to conduct market studies that would result in the profitability of of activity in that particular domain. Education in the chosen field and then the actual construction of the business would be another stage that is required. The business world is extremely competitive, and the development of technologies makes finding a new or innovative niche that is not already exploited increasingly difficult. Collaborative economics helps people who want to start a business, offering them a variety of opportunities to rethink their business ideas. Thus, the process of becoming an entrepreneur, in the sharing economy, is accelerated by the fact that there is no need to identify the field in which to become active. It is enough to have a good idea and be willing to share it in order to receive a benefit. Thus the possibility of becoming an entrepreneur is much greater, also overcoming the significant barrier of requiring access to a lot of money to start a business. By comparing the path to becoming an entrepreneur in a traditional economy with that experienced in the sharing economy, it can be observed that in the sharing economy everything happens faster and with less financial resources, due to the lack of a complex legislation in this field.

Unlike traditional entrepreneurship, integrated entrepreneurship in the collaborative economy is much more accessible. This is primarily due to the fact that in order to initiate such entrepreneurship, one only needs to possess a good and be willing to share it, and to share a good or anything else with a person or a group of persons involves finding partners for teams, projects, and entrepreneurship (Bouncken–Reuschl 2018).

These things mean much more than sharing something with someone, it actually means social interaction, creativity, exchanges of ideas and knowledge, representing the perfect climate for what entrepreneurship means. Entrepreneurial performance is improved through the collaborative economy system by individuals working together, exchanging information, and thinking about the efficiency of their work. This type of cooperation is based on values such as trust, transparency, and loyalty. In the absence of these values, it is very difficult to exist and to develop an entrepreneurial collaboration. All concepts, which mean the collaborative economy, have a positive aspect. However, there are often divergences, gaps in communication, intentional or unintentional leakage, which will affect the development of entrepreneurship and the respective community.

Being an entrepreneur means constantly being in search of new business opportunities. In the long run, entrepreneurship can be a problem due to the fact that different resources are needed in this activity and, in different countries, some resources begin to deplete. Although, at a national level, they are trying to solve the problem, financial barriers often arise and there is no concrete solution. The specificity of the collaborative economy is to share resources. This aspect can help society, first of all by reducing costs.

Consequently, this type of economy can be applied to reduce costs but also to increase long term business efficiency. The collaborative economy is the perfect framework that promotes opportunities for entrepreneurs (Huarng 2018).

Collaborative consumption is rising today. One major factor contributing to this is the generalization of the internet and easier ways of accessing it. The Internet has brought new ways of sharing but it also facilitates older forms of sharing on a larger scale.

The phenomenon of collaborative consumption helps the stimulation of economic activity even in areas where it has been limited or almost non-existent. At the same time, new horizons and opportunities are being opened up to ordinary people through which they, themselves can create their own business and become entrepreneurs, mainly because the investment may be insignificant at first. An example of this might be in the gastronomy sector: the cook can use his own kitchen as a cooking space.

There are some firms that are trying to develop their own business models, so they put their web platforms at the disposal of their partners to assist foreign entrepreneurs and ease their transactions.

Moreover, these platforms can represent the market for entrepreneurs and they do not need to search for potential customers anymore (Perren–Grauerholz 2015).

This type of consumption is gaining more and more ground. People find it easy to enter into such a market as entrepreneurs mainly because they can use the skills and all the assets they possess.

Collaboration consumption, estimated at over $ 3.5 billion in 2013 (Geron 2013) allows income to go directly into people's pockets (Perren–Grauerholz 2015).

Walsh (2011) says Collaborative Consumption was one of the ten ideas that will change the world in the next few years. It can be said that sharing economy is based on exchange. Even if these exchanges have existed since antiquity, so, long before the emergence of the concept of sharing economy, these exchanges were carried out by traditional, face-to-face methods and took place in a limited framework, especially because of the large geographic space. In addition, some shifts have a temporary character and their activity is limited to a niche market.

Opportunities in the collaborative economy are related to the fact that people can use the unused capacity of tangible, as well as intangible goods, adding value and guiding people to the field in which to open their own business. Almost everyone has a tangible or intangible asset that they are underutilizing, and that can mean the chance to receive a financial benefit. It connects distribution networks for both individuals and assets. Another opportunity is related to the fact that they activate technologies such as the Internet, which helps the interaction between people at even considerable distances from each other. It encourages meaningful and trustworthy interactions, and embraces openness, inclusiveness, and common ownership.

With the emergence, development, and easy access to these technologies, traditional consumer markets have increased their spectrum, from limited economic activities, through the geographic area and the way they interact, to economic activities that are now taking place on a global scale. All these transformations are having a positive impact from the economic, social, and environmental, etc. points of view. Peer-to-peer businesses lead to the expansion of entrepreneurship and innovation. If the idea of sharing economy did not exist on the market, many of those who are today entrepreneurs would not exist. In this collaborative environment, individuals have used their own skills and properties they own to create a business that would not have started in a different situation. Probably, for many, at first, it meant only a way of earning supplementary income, but, eventually it was the road to successful entrepreneurship. Through the collaborative economy the poorly used assets are optimized, bringing value and creating financial benefits. This category can include both tangible assets such as apartments, houses, cars, and intangible ones, for example skills, knowledge. Their optimization is possible through digital platforms that give people access and at the same time their efficiency. Table 1 describes the overview of assets and actors in the sharing economy (Hyup 2016).

Table 1 Overview of assets and actors in the sharing economy

Sharing Asset	Example	Actor: International	Actor: Korean
Tangible	Transportation	Zipcar, Uber, Zilok.com, Lyft, Car2Go	Socar, Greencar
	Property	Airbnb, DestNearMe	Cozaza, Bnbhero
	Utility/Food	NeighborGoods.com, thredUP.com, EatWith	Libpia, Bookoob, CoLux, Kiple
Intangible: Service	Professional	Innocentive, oDesk	Zipbab, PumasiPower
	Personal	Landshare.net, TaskRabbit	
Intangible: Financial	Zipbab, PumasiPower	Kickstarter, Indiegogo, gofundme, CircleUP, crowdfunder	Tumblbug, Ssiatfunding
	P2P Lending P2P Lending	KIVA, LendingClub, Prosper	Popfunding

Source: Hyup, R. T. (2016)

Thanks to technology, access to these assets has become much easier, and this has led to the growth and development of the sharing economy, which in turn promotes entrepreneurial behavior. In the sharing economy everyone has to win. The customer receives what they are looking for quickly, cheaply in an authentic way, and the one who provides the services earns money. The opportunities offered by this new type of economy also extend in the financial field. Companies like Lending Club demonstrate how you can bypass traditional banking to benefit both investors and customers in this innovative new business environment. There are still regulations, of course, but in general, the process is much less complicated than setting up a bank, credit union, or traditional lending company to meet people's financial needs.

An opportunity of sharing economy is given by the fact that it is based on the exchange economy, and it is much easier for ordinary people to access such a market to create their own business.

3. Tourism and the Sharing economy

3.1. Airbnb – general aspects

The entire tourism industry is in a state of continuous transformation, and is therefore adaptable to the new. Thus, tourism cannot remain unchanged by the new trend in the economy: the sharing economy. In a broad sense, sharing can be anything to which access is enabled through pooling of resources, products or services (Hyup 2016). The sharing economy, even if it is a fairly new concept on the market, has seen a remarkable climb lately and shows all the signs of a having a remarkable future. The new business models based on sharing economy and new technologies give a fresh look at the tourism scene, giving people new options when it comes to the place they want to travel, where they want to stay, and what they will do during their stay and beyond. All these are aspects that give tourism a new look, but at the same time a challenge in terms of regulation.

The sharing economy has had a positive impact on tourism, but there are voices claiming it also has a negative impact. The positive aspects refer to the fact that it offers people easier access to a wide range of services, which by traditional methods they could not have accessed. Critics, on the other hand, claim that the sharing economy provides unfair competition, reduces job security, avoids taxes and poses a threat to safety, health and disability compliance standards (EPRS 2015). Looking beyond the criticism, the collaborative economy creates the perfect framework for anyone that wants to open a business, especially in tourism. What facilitates the efforts of people wishing to open such a business are online platforms. Through these, small entrepreneurs can find their clients, and thus do not have to advertise by other means, usually more expensive. With the flexibility it offers, a growing number of tourists take advantage of the benefits of this new trend. Some tourists appreciate these platforms for their personalised approach, authenticity and contact with local citizens (EPRS 2015). According to the OECD, sharing companies may bring tourists to destinations that were previously less popular. For example, a 2016 study by the Observatoire Valaisin du Tourisme looking at the impact of Airbnb on tourism in Switzerland found that Airbnb had expanded the city-break niche market in some Swiss cities where the high cost of hotel stays had previously hindered some tourists from staying there (EPRS 2015).

Criticism of sharing economy refers primarily to legislation. At present there are no serious regulations. Those who come to attack this type of economy say that lack of regulation can lead to illegality, evasion of various taxes, illicit business, unfair competition, and disruption of the resident population. At the same time, if we refer to accommodation services, one might add to the negatives with the fact that the owners of houses that provide tourists do not need to hire personnel, and this contributes to the increase of the unemployment rate.

Tourism is not just fun trips, but also business trips, attending various events, visiting relatives or even traveling for medical purposes. The travel and tourism industry is generally viewed as consisting of largely six segments including airlines, lodging, car rental, cruise, rail, and travel packaging (Hong 2018). That is why it is important to capitalize on each of these tourism-related segments, which have as a consequence the sustainable growth of the economy.

Table 2 Impact of the sharing economy on the different stages of the tourism customer journey

Pre-trip (booking)	Journey	Destination	Post-trip
Accommodation	Land-based passenger Transportation	Sporting, recreational and entertainment activities	Review platforms
Traditional	Share the cost of travel	Equipment rental	Exchange of views and recommendations on the quality of tourist services
House swap between two individuals	Taxi services	Tours and tourist activities	
House sitting		Workshops and classes	
Work in exchange for accommodation		Food and drink services	
Hotel booking resale platforms between individuals		Gastronomic activities	
Review platforms		Vehicle rental between individuals (bicycles, boats, cars, etc.)	
Exchange of information between users			

Source: Ferrer, R. (2018)

In recent years, with the increase of income, in most of the emerging world, there has been an increase in tourism and all its aspects. The sustained and healthy growth of this industry is driven by the increase in travel numbers. A large proportion of them is booked through online platforms, which tells us that sharing economy has made a contribution to this. In 2015, online revenue accounted for 21.6% of global sales, and that number is expected to grow to 27.7% in 2017 (Hong 2018). Tourism is a very influential area, and everything that happens on the market: economic fluctuations, political changes, natural disasters, and changing consumer habits change the tourism scene. Additional changes in the industry such as evolving consumer mindsets, technologies, and new platforms may all transform travel in the future (Hong 2018) (Figure 6).

Through the sharing economy, the places and activities less known by tourists are becoming more attractive. The travel industry focuses mainly on major hotels and activities suggested by various tour operators. Instead, if you organize your own trip yourself, by using the services provided by the sharing economy, it is very possible to have a unique experience and also have the opportunity to do various activities, and to get to know the locals, i.e. enjoying the beauty and the places your trip offers. In the figure below we can see how the sharing economy is changing the way in which some tourism services are provided and has generated new ways of travelling (Ferrer 2018).

3.2. The case of the Airbnb platform can be used to exemplify the sharing economy model

Regarding services, in the case of sharing economy, they cover a very large range from accommodation to finances. In the field of accommodation a very good example for sharing economy are collaborative platforms. Acquiera (2017) said that the platform economy forms a second core of the sharing economy. They also defined the platform economy as a set of initiatives that mediated decentralized exchanges between peers through digital platforms.

Airbnb is an online platform set up in 2008 that provides tourists and others, homes that owners own but are unoccupied for rent, usually for short periods of time. In other words, we can say that this platform responds to the accommodation needs of people and also creates a connection between them. The company, founded in 2008, has grown exponentially in the past few years, and by now it lists over 1.5 million properties, with a presence in over 190 countries and 34,000 cities (Quattrone et al. 2016l). Renters can obtain accommodation at lower prices from Airbnb than from hotels in most cities (Permalink 2013). As a result, Airbnb offers a double advantage, on the one hand, for those who can earn extra income, and on the other hand, tourists who can rent at lower costs than the hotel accommodation costs.

This platform is attracting a growing number of tourists, but there are not many studies that explain why they choose to stay in the airbnb platforms. This is primarily due to the fact that Airbnb is a new phenomenon on the market. Many studies have explored tourists' hotel choices, typically by having respondents rate the importance of different hotel attributes (Guttentag 2017). The choice of the hotel is determined by factors such as location, cleanliness, services offered, price, security, room comfort. At the opposite end, there are studies that provide information on why people choose to use the Airbnb service (Guttentag 2017). Guttentag (2015) proposed three key appeals – price, household amenities, and authenticity; Tussyadiah (2015) found peer-to-peer short-term rental users were motivated by three factors – sustainability, community, and economic benefits; and Nowak et al. (2015) found the top reasons Airbnb users chose the service were "cheaper price," "location," and "authentic experience" (Guttentag 2015).

Table 3 Descriptive statistics for the motivations to choose Airbnb

Dimension (as originally proposed) Motivation	M	SD	N
Price			
For its comparatively low cost			
Functional attributes			
For the convenient location	4.99	0.99	841
For the access to household amenities	4.70	1.31	840
For the homely feel	4.41	1.30	842
For the large amount of space	4.13	1.39	843
To receive useful local information and tips from my host	3.90	1.44	842
Unique and local authenticity			
To have an authentic local experience	4.46	1.29	841
To have a unique (non-standardized) experience	4.36	1.34	841
To stay in a non-touristy neighborhood	4.33	1.41	844
To interact with host, locals	3.46	1.51	839
Novelty			
I thought the experience would be exciting	4.06	1.30	841
To do something new and different	4.04	1.39	840
I thought the experience would be unpredictable	2.63	1.27	843
Bragging rights			
To have an experience I could tell my friends/family about	3.40	1.40	841
Sharing economy ethos			
I prefer the philosophy of Airbnb	3.91	1.34	839
I wanted the money I spent to go to locals	3.69	1.36	843
Staying with Airbnb is environmentally friendly	3.25	1.31	842

Source: Guttentag, D. A. (2016)

Table 3 summarizes the results of a poll taken to see why people choose the Aibnb platform. All items were measured on a six-point Likert scale ranging from 1="Strongly disagree" to 6="Strongly agree." As can be observed, respondents on average agreed with nearly all of the proposed motivations (with 3.5 as the mathematical midpoint of the six-point scale) (Guttentag 2016).

Thanks to tourist choice for this type of accommodation, ordinary people are given the opportunity to form their own business. This business model is relatively simple, giving the individuals the confidence to get involved in this field by creating their own business. The business model offered by this service, which is an integrated part of the sharing economy, requires only three elements: the host, the lists and the guests.

Hosts list their property on the platform and establish their own nightly, weekly or monthly price, and offer accommodation to guests. Listings – properties – may include entire apartments, private rooms or shared rooms. The 'host' of a space may be living there at the time of the rental, as with a typical bed and breakfast (B&B), or may be absent, possibly away on vacation or even operating the space as a permanent rental. Airbnb claims that 57% of its listed spaces are entire apartments and homes, 41% are private rooms, and 2% are shared rooms (Guttentag 2015)

In September 2016, Airbnb was valued at more than $30 billion, making it the second most valuable start-up in the world, after Uber (Farrell–Bensinger 2016).

Airbnb was generally expected to outperform budget hotels/motels, underperform upscale hotels, and have mixed outcomes versus mid-range hotels, displaying – but not completely – consistency with the concept of disruptive innovation (Guttentag–Smith 2017).

Airbnb is operating amidst continuous development on the market, and many sources seem to predict that the sharing economy and companies like Airbnb are here to stay (Rusli et al. 2014). Thus, there is a need for the cities and other locations where this type of accommodation works, to prepare, adapt, and find innovative ways to stimulate business in the area. Even if cities need to be prepared and find new opportunities to do so, they must comply with existing regulations and not allow abuses to occur. This means that the competent authorities must be just as innovative in order to continue to adapt in an unprecedented and quickly changing economy and world, especially as the sharing economy, in all its aspects, not just in terms of Airbnb, is controversial and at the same time flawed (Mehmed 2016). Hirshon et al. (2015) explains the lack of regulations through the rapid evolution of these new business models, and the absence of any clear precedent in the regulatory process. Given this context, each city has to determine who will handle this and this takes time. At the same time, we could point to the self-regulation of the market that often emerges as a natural by-product of economic exchange and has a long history of success. On the one hand, we can talk about "partial self-regulation", which refers to situations in which private parties are responsible only for rulemaking, while enforcement is the domain of either public bodies or the market. On the other hand, "full self-regulation" occurs when the industry engages in both rulemaking and enforcement (Cohen–Sundararajan 2017).

The different market inefficiencies for peer-to-peer transactions will require different entities to act as partners in any self-regulatory solution. If this approach is based on principles such as trust and transparency, it could provide timely solutions for regulating this field. Policymakers, however, should take account of the effectiveness of these platforms and eliminate unnecessary requirements and protectionist rules that benefit primarily the current operators (Edelman–Geradin 2016).

Schor (2017) finds that the site that has been most successful at creating new social ties is Airbnb (Frenkena–Schor 2017). Airbnb is growing at a rapid rate. It is attracting hosts and guests with alluring experiences that make them fall in love with the company, while about 2 million people stay in Airbnb rentals across the world during an average night in 2018 (Airbnb Statistics 2019). This means economic growth, because through this type of accommodation, even people who do not have large incomes can afford to travel. Among the benefits of this accommodation system are the promotion of less known places for tourists, and the interaction of the tourists with the locals, giving tourists a stay during which they can enjoy the authenticity of the place. Research conducted by the Morgan Stanley Research (2015) shows that over 50% of Airbnb users in 2016 gladly used it to replace a traditional hotel stay. Consequently Airbnb is creating a massive profile, and setting new trends across the world.

4. Conclusion

The sharing economy has its roots in the exchange and the willingness of the individual to share. If we refer to exchange, this can range from a simple exchange of ideas and opinions to those more property-related. Through the sharing economy, values such as trust and transparency are created, and at the same time barriers between people are broken down, new intercultural relationships being created. One can say that one of the factors that determine the sharing economy is the desire of people to connect and interrelate.

In a traditional economy, the process of becoming an entrepreneur involves observing a set of steps mentioned in this paper, but the collaborative economy comes to mitigate this process and bring opportunities to people who want to open a business.

This interaction between people and the way they work together can be endangered by a lack of fairness, by violating various social norms, but also by non-involvement. In addition to what has been said above, a danger or a barrier to sharing the economy may be the desire for people's intimacy. According to Durkheim (1964), the desire for intimacy with others is the most important determinant of human behavior. Everything that means belongingness and building a sense of community acts as a driver for sharing, whereas on the other hand, possessiveness, attachment, and independence inhibit and limit sharing (Belk–Llamas 2011).

Nowadays, consumption has taken on a very large scale, and the idea of sharing when it comes to consumption in any way would bring major benefits both at the individual level and at societyal level. Looking at the macro level, the sharing economy has positive consequences from the economic, social, and environmental point of view. At a personal level sharing goods / resources (money, time, space) may mean, at first, an additional source of income, and as it develops, it may even be a chance for individuals to become entrepreneurs. Sharing is a new contemporary subject, and the choice of sharing instead of monopolization can bring major benefits to a world which, it can be said, still accentuates individualism. The new trend, meaning sharing economy, contributes to collective welfare.

Sharing models are used in a growing number of areas, such as research, transport, communication, and not least in tourism. The tourist sector has followed this trend and reaped a number of advantages.

The collaborative economy is a phenomenon that will continue to develop over the coming years, and perhaps even become predominant. Thus, sharing economy will be included in hybrid economics where multiple modes of exchange (Scaraboto 2015) exist.

Concluding, it can be said that Airbnb has begun to heavily infuse the nature of business in the hotel industry. This is largely due to innovation, the development of technologies, and the very easy access of all people to the internet. Habbel, the founder of Voyat, an e-commerce optimization platform, pointed to the fact that "hotels need to steal basic things out of the Airbnb and online travel agency playbooks, like making their webpages load quickly on mobile devices and reducing the number of clicks it takes to finalize a transaction." (O'Neil 2017). Botsman–Rogers (2011) said the most important reasons why people use alternative solutions of sharing economy are the economic and social characteristics.

References

Acquiera, A. – Daudigeos, T. – Pinkse, J. (2017): Promises and paradoxes of the sharing economy: An organizing framework. *Technological Forecasting and Social Change.* https://www.researchgate.net/publication/318603405_Promises_and_paradoxes_of_the_sharing_economy_An_organizing_framework Accesed: February 24, 2019.

Airbnb Statistics (2019): *Airbnb by the Numbers: Usage, Demographics, and Revenue Growth.* https://muchneeded.com/airbnb-statistics/ Accessed: February 25, 2012

Belk, R. (2010): Sharing. *The Journal of Consumer Research*, 36, 5, 715–734. http://www.cs.cmu.edu/~jhm/Readings/612649.pdf Accesed: February 10, 2019

Belk, R. (2013): You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research*, 67, 1595–1600. https://www.researchgate.net/publication/262490610_You_are_what_you_can_access_Sharing_and_collaborative_consumption_online Accesed: February 27, 2019

Belk, R. W. – Llamas, R. (2011): The nature and effects of sharing in consumer behavior. In Mick, D. G. – Pettigrew, S. – Pechmann, C. – Ozanne, J. L. (eds): *Transformative consumer research for personal and collective well-being*. Routledge, New York, 625–646.

Botsman, R. (2015): *Defining the sharing economy*. http://www.fastcoexist.com/3046119/defining-the-sharing-economy-what-is-collaborative-consumption-and-what-isnt Accessed: February 28, 2019.

Botsman, R. – Rogers, R. (2010): *What's mine is yours: The rise of collaborative consumption*. New York, NY, Collins.

Botsman, R. – Rogers, R. (2011): *What's mine is yours: how collaborative consumption is changing the way we live*. London, Collins.

Bouncken, R. B. – Reuschl, A. J. (2018): Coworking-spaces: how a phenomenon of the sharing economy builds a novel trend for the workplace and for entrepreneurship. *Review of Managerial Science*, 12, 1, 317–334.

Cadar, O. – Badulescu, D. (2017): Entrepreneur, Entrepreneurship and Intrapreneurship. A Literature Review. *The Annals of the University of Oradea, Economic Sciences*, 2, 24, 658–664.

Cohen, M. – Sundararajan, A. (2017): *Self-Regulation and Innovation in the Peer-to-Peer Sharing Economy*. University of Chicago Law Review Online, 82, 1, 116–133.

Constantin, A. R. (2014): *Antreprenoriat*. Editura Pro Universitaria, Bucureşti.

Durkheim, E. (1964): *The rules of sociological method*. London: Free Press.

Edelman, B. G. – Geradin, D. (2016): Efficiencies and regulatory shortcuts: how should we regulate companies like Airbnb and Uber? *Stanford Technology Law Review*, 19, 293–328.

European Parliamentary Research Service (2015): *Tourism and the sharing economy*. EPRS. http://www.europarl.europa.eu/RegData/etudes/BRIE/2017/595897/EPRS_BRI(2017)595897_EN.pdf Accesed: February 27, 2019

Farrell, M. – Bensinger, G. (2016): Airbnb's Funding Round Led by Google Capital, *The Wall Street Journal*.

Ferrer, R. (2018): The sharing economy and tourism. *Caixa Bank Research*, 7, 34–35.

Ferreira, N. M. (2018): *What is Entrepreneurship? Entrepreneur Definition and Meaning*. https://www.oberlo.com/blog/what-is-entrepreneurship Accesed: February 10, 2019.

Frenkena, K. – Schor, J. (2017): Putting the sharing economy into perspective. *Environmental Innovation and Societal Transitions*, 23, 3–10.

Gartner W. B. (1990): What are we talking about when we talk about entrepreneurship? *Journal of Business Venturing*, 5, 1, 15–28.

Geron, T. (2013): *Airbnb and the unstoppable rise of the share economy*. https://www.forbes.com/sites/tomiogeron/2013/01/23/airbnb-and-the-unstoppable-rise-of-the-share-economy/#121f9139aae3 Accessed: February 28, 2019.

Guttentag, D. A. (2016): *Why tourists choose Airbnb: A motivation-based segmentation study underpinned by innovation concepts*. Waterloo, University of Waterloo.

Guttentag, D. A. (2017): *Assessing Airbnb as a disruptive innovation relative to hotels: Substitution and comparative performance expectations*. Travel and Tourism Research Association: Advancing Tourism Research Globally, 19. https://scholarworks.umass.edu/ttra/2017/Academic_Papers_Visual/19 Accesed: 20 February, 2019.

Guttentag, D. (2015): Airbnb: Disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*, 18, 12, 1192–1217.

Guttentag, D. A. – Smith, S. L. (2017): Assessing Airbnb as a disruptive innovation relative to hotels: Substitution and comparative performance expectations. *International Journal of Hospitality Management*, 64, 1–10. https://doi.org/10.1016/j.ijhm.2017.02.003 Accesed: 20 February, 2019.

Hatos, A. – Hatos, R. – Bădulescu, A. – Bădulescu, D. (2015): Are Risk Attitudes and Individualism Predictors of Entrepreneurship in Romania? *Amfiteatru Economic*, 17, 38, 148–161.

Hirshon, L. – Jones, M. – Levin, D. – McCarthy, K. – Morano, B. – Simon, S. – Rainwater, B. (2015): *Cities, the sharing economy, and what's next*. The National League of Cities.

Hong, J. (2018): Rise of the Sharing Economy and the Future of Travel and Tourism Industry, *Journal of Hotel & Business Management*, 7, 2 https://www.longdom.org/open-access/rise-of-the-sharing-economy-and-the-future-of-travel-and-tourism-industry-2169-0286-1000180.pdf Accesed: 03.01.2019.

Huarng, K. H. (2018): Entrepreneurship for Long-term Care in Sharing Economy. *International Entrepreneurship and Management Journal*, 14, 1, 97–104.

Hyup Roh, Tae. (2016): The Sharing Economy: Business Cases of Social Enterprises Using Collaborative Networks. *Procedia Computer Science*, 91, 502–511.

Johnson, D. (2001): What is innovation and entrepreneurship: lessons for larger organizations. *Industrial and Commercial Training*, 33, 4, 135–140.

Mehmed, N. R. (2016): Airbnb and the Sharing Economy: Policy Implications for Local Governments. *SPNHA Review*, 12, 1, Article 6 http://scholarworks.gvsu.edu/spnhareview/vol12/iss1/6 Accesed: February 15, 2019.

Morgan Stanley Research (2015): Global insight: Who will Airbnb hurt more – Hotels or OTAs? *Morgan Stanley Reseach*, Internet, Lodging, Leisure and Hotels.

O'Neil, S. (2017): Hotels may try to counter Airbnb's rise with bets on tech.

Pahuja, A. – Sanjeev, R. (2015): *Introduction to Entrepreneurship*. https://www.researchgate.net/publication/301659818_Introduction_to_Entrepreneurship Accesed: February 28, 2019.

Permalink (2013): Airbnb vs hotels: A price comparison. http://priceonomics.com/hotels/ Accesed: February 27, 2019.

Perren, R. – Grauerholz, L. (2015): Collaborative Consumption. *International Encyclopedia of the Social & Behavioral Sciences*, 139–144. https://www.researchgate.net/publication/303960052_Collaborative_Consumption Accesed: February 27, 2019.

Petrini, M. – Freitas, C. S. de. – Silveira, L. M. da S. (2017): A Proposal for a Typology of Sharing Economy. *Revista de Administração Mackenzie*, 18, 5, 39–62.

Quattrone, G. – Proserpio, D. – Quercia, D. – Capra, L. – Musolesi, M. (2016): *Who Benefits from the "Sharing" Economy of Airbnb?*, WWW '16 Proceedings of the 25th International Conference on World Wide Web, 1385–1394. https://www.ucl.ac.uk/~ucfamus/papers/www16_airbnb.pdf

Rusli, E. – Macmillan, D. – Spector, M. (2014): Airbnb Is in Advanced Talks to Raise Funds at a $10 Billion Valuation. *The Wall Street Journal*. http://online.wsj.com/articles/SB10001424052702303802104579451022670668410 Accesed: February 03, 2019.

Scaraboto, D. (2015): Selling, Sharing, and Everything in Between: The Hybrid Economies of Collaborative Networks. *Journal of Consumer Research*, 42, 1, 152–176.

Sexton, D. L. – Bowman Upton, N. B. (1991): *Entrepreneurship: Creativity and Growth*. MacMillan.

Walsh, B. (2011): 10 ideas that will change the world, *Today's Smart Choice: Don't Own. Share*. Times Magazine 1–4. http://content.time.com/time/specials/packages/article/0,28804,2059521_2059717_2059710,00.html

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 240–256.

A Game changer: exploring and exploiting cloud computing

Péter Füzes – Roland Zs. Szabó – Zoltán Gódor

Even Fortune 500 firms can disappear quickly because many of them fall into an exploitation trap. Examples show that certain organizations – that manifest ambidextrous features – can avoid the exploitation trap. But, how do they manage it? The Authors tracked the digital transformation of a Fortune 500 company in order to understand the strategic and organizational challenges and solutions to becoming resilient and prosperous. A cutting-edge example is shown in the paper about how an industry leader can exploit traditional industry while exploring and exploiting new ones and their markets at the same time. The Authors found that shaping an industry needs strong top-down leadership, and strong alignment between markets, strategy, and the configuration of the firm.

Keywords: ambidexterity, change management, cloud computing, exploration, exploitation

1. Introduction

Digital transformation is reshaping every industry. Today is the era of the fourth industrial revolution. Exploration of new ideas and territories has never been so popular, but exploiting existing businesses is also crucial. The trade-off between exploration and exploitation has been known for 25 years (March 1991), and the real challenge of management is to find the right balance between the two (Raisch et al. 2009)

Many firms are exploratory in the early stages of their lifecycles, but become exploiters in the later stages (Hortoványi 2012, Szabó 2014). As they become incumbent, they adapt themselves to the norms and rules of an industry which in turn decreases the future ability of adaptation. That is the reason why so many firms fall into an exploitation trap and are unable to get out of it. Only a few firms are able to remain exploratory and as such, shape an industry. These ones are those who are able to remain resilient. These firms also transform along with the new era and become ambidextrous.

Ambidextrous organizations are not only able to find the right balance between exploration and exploitation, but even more, they are able to drive internal organizational changes effectively. As an impact of the fourth industrial revolution, the average life cycle of products is becoming shorter (Xu et al. 2018). In this environment, finding the right balance between exploration and exploitation can be even more difficult, therefore understanding how the firm can become ambidextrous is a current area of interest.

In this paper, a cutting-edge example shows how an industry leader can exploit the traditional on-premise IT industry and large customers market and explore and exploit the new cloud computing industry and SME market.

2. Theoretical background

2.1. The exploitation trap

As industries evolve many of them become predictable and unchangeable. The adaptation to the environment is very important, but the constant adaptation to existing factors decreases future ability of adaptation. According to Burgelman (1991) the change carried out in the strategy is always smaller than the change happening in the environment, therefore the alterations mainly concern the peripheries of the strategy, and there is no change to the core areas. Hence during reorientation, companies facing competition tend rather to strengthen their already existing activities instead of looking for new ones. This adaptation paradox leads to exploitation trap for many successful incumbent companies.

Child (1972) pointed out that the view which says that the organizational structure is unambiguously determined by the environmental factors, technological level, and other external factors is not correct. The decision makers of the company actively contribute to the manipulation of their own environment, in order to achieve the goals, they have set. They either ignore the changes happening in the environment, or they alter the organization. In this interpretation, the proactive behavior of the company is determined by the leader or dominant coalition.

Contradicting Child’s view, Burgelman (1991) represents the view that strategy is based on the current technology, economic, and cultural factors, and adapting to these, while the task of the leader is to create such a strategy that enables the organization to attain further success. Therefore, the organizational structure defines the competencies of the organization and determines its aims. The strategy consists of technical, economic, and cultural regulations. These regulations serve the purpose of maintaining the character of the organization. He uses the theory of population ecology for strategy building. During selection, the participants on different levels perceive strategy differently, therefore, variations appear. The objectives set in the strategy cannot be achieved without internal selection systems.

Thus, the primary task of top management is managing the administrative tools (strategic planning, control system, incentive systems), developing cultural (behavioral norms) mechanisms and selection systems.

However, exploitation can be an effective strategy short term. Oh et al. suggested that in high tech industries (on-premise and cloud software belongs to this category) exploitation of existing markets with improved products is a more effective strategy to achieve sales growth than bringing new products to existing markets, and or targeting new markets with improved products (Oh et al. 2015).

2.2. Managing change

Change is a continuous phenomenon both nowadays and throughout history, but the pace of change seems to be accelerating. Change is an unavoidable result of innovations, whose effect and impact are often unimaginable and underestimated by many people, included those individuals and organizations, too, from whom the innovation derives. Managers want to govern this process better and more proactively, but there are still several unanswered questions (Schendel–Hitt 2007):

- How can and has to be change consciously (actively) managed, while one enterprise innovates, and perceives the innovations in the industry?
- How can the effect of innovations be tracked (e.g.: in case of organizational structure and business model)?
- What are the primary tasks in the preparation of the enterprise for the changes?
- What change forms are reasonable and effective?
- What obstacles might change run into and how can these obstacles be avoided or how can we overcome them?

Managers and employees at different levels of the hierarchy can have different view of the necessity of the change and its substance. Lower level managers' understanding of the situation and their emotions about the change can be dissimilar to top managers' perceptions and intention (Vuori–Huy 2016).

Change management is a consciously managed activity, during which the enterprise gets from one configuration to another. The recognition of strategic changes, and finding an adequate answer to these, present the members of the change management team with an especially hard task. The corporate environment presupposes the continuous revision of strategy and operations, which has a significant effect on the stakeholders of the organization. During the change management processes, the proper combination of strategies, the creation of the favorable reception of changes and the delivery of the results are critical factors.

The start and the maintenance of change is not an easy task, because for this the (artificial) maintenance of creative tension is needed in the organization. In order to maintain creative tension, vision has to be utilized, learning has to be directed, and planning has to be given power (Mintzberg et al. 1998). Hindering factors in recognition of the necessity of changes and in the creation of a sense of urgency (Kotter 1999):

- the absence of a major and visible problem or crisis,
- too much happy talk from senior management,
- low overall performance standards,
- performance measurement system focusing on wrong metrics,
- abundant resources,
- operating in silos with organizational structures that force employees to focus on narrow functional goals, and the underestimation of the power of denial, which turns a blind eye to problems
- not aware of how suppliers and customers actually view performance,
- low confrontational culture.

According to Clemmer (1995) changing and managing are precluding concepts, and changes don't have to be controlled manually, but the frameworks have to be set, along which change is to proceed. Change can be ignored, resisted, reacted to, exploited, or induced, and the necessary frameworks and configurations have to be developed accordingly.

During change, it is important, that it is very difficult to change everything at the same time, and it is not advisable either. Based on the recommendation of Mintzberg et al. (1998) we look for the best among the new and keep the most useful among the old. The change strategy of Dickhout et al. (1995) is much more pragmatic than this general recommendation:

– Evolutionary/institutional building: line managers direct the continuous change,
– Jolt and refocus: change of the management is necessary,
– Follow the leader: cutting the side-activities in order to have fast results,
– Multifront focus: fast results stabilize the organization, which can be followed by a multifront focus, changing many factors at the same time,
– Systematic redesign: ad hoc workgroups, but planned change,
– Unit-level mobilizing: the incorporation of the ideas of middle management and the workers.
– Changes can be induced top-down or bottom-up. Example for the top-down induced change is the drama of Tichy–Sherman (1993) in three acts, during which the prologue is the development of the new global playing field, and the acts are the processes of the organization: (1) awakening, (2) envisioning and (3) rearchitecting. The epilogue refers to the stability of changes that history repeats itself.

Kotter (1995: 61, 1999) gives more detailed guidance for the implementation of top-down changes:

– Establish a Sense of Urgency
– Form a Powerful Guiding Coalition
– Create a Vision
– Communicate that Vision
– Empower Others to Act on the Vision
– Plan for and Create Short-Term Wins
– Consolidate Improvements and Keep the Momentum for Change Moving
– Institutionalize the New Approaches

Beer et al. (1990) examined, why change programs aren't productive. They found the problem in starting change from too high above. Successful changes were typically started by a local manager, which was supported by the top management in order to achieve success. The successful elements were spread throughout the whole firm:

– the common diagnosis of business problems helps the commitment to change,
– common vision,
– consensus and resources,
– expansion of revitalization (as a possibility),
– cultivation,
– monitor the revitalization and correct the mistakes.

a.) Ambidexterity

A company can be successful in its existing operational areas and can exploit them. When solving crises, the successful embracement of new possibilities has a key role, without necessarily causing the destruction of existing areas. Companies meet a lot of “creative destruction’ (Schumpeter, 1942) ideas during their explorative activities; however the real challenge for them is not the pure implementation of these ideas, but the successful running and construction of the existing and new fields at the same time. Summarizing the concept of ambidexterity, it ensures success for a company in its existing fields (exploiting) and in its new business fields (exploring) at the same time.

The topic of ambidextrous organizations is becoming increasingly popular among researchers who deal with strategy. The key question is its joint treatment of efficiency (exploitation) and effectiveness (exploration) (Tushman and O'Reilly (1996; 2002), O'Reilly–Tushman 2004, Raisch et al. 2009, Gibson–Birkinshaw 2004).

The ambidextrous organizations are able to manage successfully their existing activities and new products, services, and processes at the same time. The ambidextrousness can be realized in several organizational structures, in functional, cross-functional, spinout or ambidextrous structures, too (O'Reilly–Tushman 2004). Organizational ambidexterity has a positive impact on overall corporate performance (O'Reilly–Tushman 2013, Junni et al. 2013), on revenue growth (Lee et al. 2003, Venkatraman et al. 2006), on innovation (Burgers et al. 2009, Tushman et al. 2010) and can improve the company’s survivability (Kauppila 2010, Yu–Khessina 2012).

The majority of enterprises struggle to find the balance between efficiency and innovation. The enterprises can gain efficiency in the short term if they replace their costly and unforeseeable activities with cheap routine processes. Though this exchange is extremely dangerous because the organization loses its long term ability to adapt. The more routine processes there are, the less flexible the organization will be. Therefore, sometimes based on strategic consideration, disturbance needs to be created artificially in the organization wishing to maintain creative tension (Raisch et al. 2009).

The trigger of creative tension might be the open business model in which the innovations come from inside as well as from outside of the traditional organizational boarders. At the same time, there is the possibility to spin off those innovations that are not realizable in the parent organization, but are viable/profitable otherwise (Chesbrough 2002; 2006).

b.) Exploiting: The traditional way of using IT

The traditional way of using Information Technology (IT) systems within companies was 'buy and build your own'. The company purchases and installs hardware and software elements to run the business applications required to support the company's business. Hardware elements include servers and storage; software elements include operating systems, security solutions, database, and middleware software and business applications. The hardware is usually installed within the premises of the company in the server room(s), hence the name of the model: on-premise (on-prem) computing.

c.) Exploring: The new way of using IT

The development of high-speed networks, virtualization and other distributed software solutions at the beginning of the 21st century allowed companies to use IT differently. Companies don't need to have servers on their premises anymore; they can use IT services from remote servers as a utility. It is not necessary to store the data or run the business applications by themselves; they can use external providers for that and access the services through the internet. This model is called 'Cloud computing' (Sultan–van de Bunt-Kokhuis 2012).

Cloud computing is a very broad term and includes different services. Usually, 3 major service levels are differentiated in cloud computing (DaSilva et al. 2013):

- Infrastructure as a service (IaaS)
- Platform as a service (PaaS)
- Software as a service (SaaS)

The expected benefits of cloud computing are cost reduction, reduction of ICT employees, increased flexibility, increased mobility and information access, and less ITC focus overall (Caldarelli et al. 2017).

Cloud models have become very popular over the last few years. In 2017, the total size of the public cloud computing market was around 130 billion U.S. dollars worldwide. (www.statista.com, 2018). New vendors have appeared on the market and grown their business significantly. Some of the leading cloud providers: Amazon (Amazon Web Services), Salesforce.com, Workday, Dropbox.

The traditional on-prem software companies have also realized the potential in cloud computing and started to re-position themselves as cloud providers. Microsoft, SAP, IBM, and Oracle are all actively investing in cloud solutions and growing their cloud business.

d.) Research gap and research questions

The IT Services and the Software industries are predictable. The malleability of the IT services industry is less than that of the Software industry, but even less than the malleability of the Internet Software and Services industry. The challenge of the Internet Software and Services industry is its unpredictability (Reeves et al. 2012).

Exploiting predictable markets and industries are essential for the incumbents, but it may lead to exploitation trap. As the industry declines, these firms could disappear. How can an incumbent firm escape from exploitation trap?

Exploring new markets and industries may contribute to the longevity of the firm, but can be painful and costly in the short term. In many cases, it is even uncertain. How can an established firm drive industry changes by exploring and exploiting new markets and industries?

Only a few firms are shaping the cloud computing industry, those who are resilient and powerful enough. These firms are also transforming in the digital era, how do they do it?

3. Qualitative research methodology

We tracked the digital transformation of a Fortune 500 company in order to understand the strategic and organizational challenges and solutions of the phenomenon. The company subject to this research is a multinational IT company, and has been a major player of the global on-prem software market, with customers and subsidiaries in a large number of countries.

The key target market for the company is the enterprise market, large customers from banking, telecommunication, manufacturing, retail, education, healthcare, and public sector. Small and medium enterprises (SMEs) also constitute a target segment for the company; however, the majority of the revenues are generated with enterprise customers.

We used qualitative research methodology to answer the main questions of the research. We used the multi-level approach in order to get a better insight of different stakeholders in the company, and wanted to get to know a detailed opinion of every interviewee about the changes in the company.

The essence of in-depth interviews is that they test the answers and statements of interviewees with specific parts of the interview (Solt 1998). We did not compare the interviews with each other, but we tested the conclusions of each answer with other interviews, and we developed the system of opinions. Based on Solt's (1998) guidelines, we did not set up a hypothesis about the interviews, and we did not insert the answers in our existing schemes to avoid processing errors.

We prepared 19 in-depth interviews with experts to answer the research questions in sufficient detail. In the selection of interviewees, we tried to characterize the employees by the following aspects: position, line of business, and territory.

Considering the positions of the interviewees, we covered many levels of the organizational structure:

- Cloud Programs Leader;
- Consulting Director;
- Digital Champion;
- Director;

- Finance Director;
- Sales Manager;
- 4 Sales Representatives;
- Senior Director, Public Sector;
- 2 Senior Vice Presidents;
- 3 Vice Presidents;

We interviewed employees from various lines of business, such as application sales, cloud digital, cloud customer success, finance, consulting, public sector business development and technology sales.

Most of our interviewees were responsible for ECEMEA (East and Central Europe – Middle East – Africa) region (4 people), but there were some employees from other territories, like EMEA (Europe – Middle East – Africa), Hungary (4), Hungary and Slovakia (1), MEA (Middle East – Africa) and North Africa – Levant.

We conducted semi-structured interviews, using the following opening questions:

- What do they think about the new way of using IT: cloud computing?
- How good are the new products in cloud?
- How important are the traditional on-prem products?
- How has the environment changed over the years?
- Which are the main economic, social, and technological trends?
- What are the characteristics of market competition?
- How do the main competitors act?
- What are the main strategic directions of the company?
- How did the structure of the company change because of the new way of thinking?
- What are the motivational factors for maintaining the changes in the company?

We collected and analyzed data in a parallel, iterative way (Miles–Huberman 1984), hence we recorded the interviews, and took notes during them to ensure complete data collection. After the interviews, we coded the answers based on contingency theory:

- Environment: PESTEL (political, economic, social, technological, environmental, legal), Porter's six forces (competition, new entrants, buyers, suppliers, substitutes, complementary products), SWOT (strengths, weaknesses, opportunities, threats);
- Strategy: value proposition, biggest challenges, growth directions, competitive strategy, first mover strategy, strategic alliances, learning, and development;
- Structure: changes in structural factors (hierarchy, coordination, collaboration);
- Behavior: personal and organizational motivational factors;
- Performance: control processes, performance measurement.
- The code structure ensured comprehensive analysis of the interviews. We were able to understand the changes through contingency theory because of the interaction of its factors.

We also collected and analyzed data from secondary sources:

– annual reports of the company between 2010–2017
– public speeches available on-line of leading executives of the company in the 2010–2017 period

4. Results

a.) Exploring a new industry: Cloud computing in the company portfolio

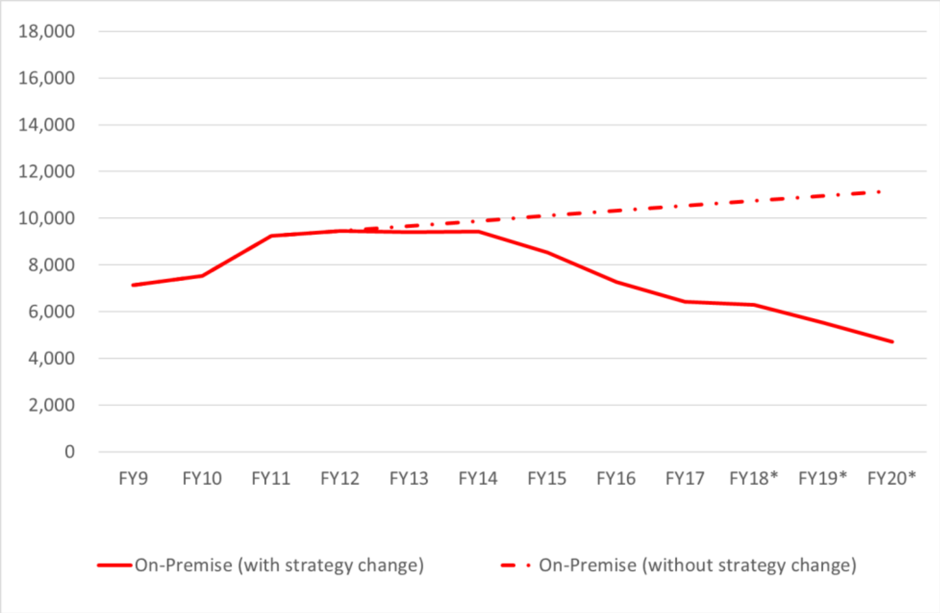
Initially, the company was not amongst the pioneers of the cloud market, therefore, it didn't have the first mover's advantage. Around 2010 when other vendors started to grow their cloud business and cloud become 'hype', the company had strategic choices to make: Invest or not into the cloud business.

Although it seems reasonable for a traditional software vendor to follow the market trend and move to the new cloud market segment, such a tack raises potential problems as well.

– New cloud products could disrupt the existing high margin on-prem business, and cannibalize on-prem revenues. If the overall market does not grow, cloud solutions will take business from on-prem. Shouldn't the company focus on keeping its strong on-prem position, and compete with the new cloud rivals?
– To offer cloud products which compete with the company's own on-prem portfolio could confuse customers.
– In most cases, perpetual on-prem software is paid when purchased. This provides the company with strong cash flow. In contrast, cloud solutions are paid over time. The average estimated length of a cloud contract is 3 years; which means the revenue and the cash flow will be distributed over 3 years. However, the investment cost associated with developing the cloud product and building a data center occurs at the beginning. While the cost of the physical data center will appear in P&L as depreciation cost over the years, the cloud software development cost hits the P&L expense line immediately. This means that mid-term profitability and cash flow of the company will be negatively impacted by selling cloud vs selling on-prem.
– In 2010 cloud was not a proven technology; it could well have been a dead-end direction. Redirecting investment from on-prem portfolio to cloud was a risky decision. The existing need of the mainstream customers was not cloud, but on-prem.

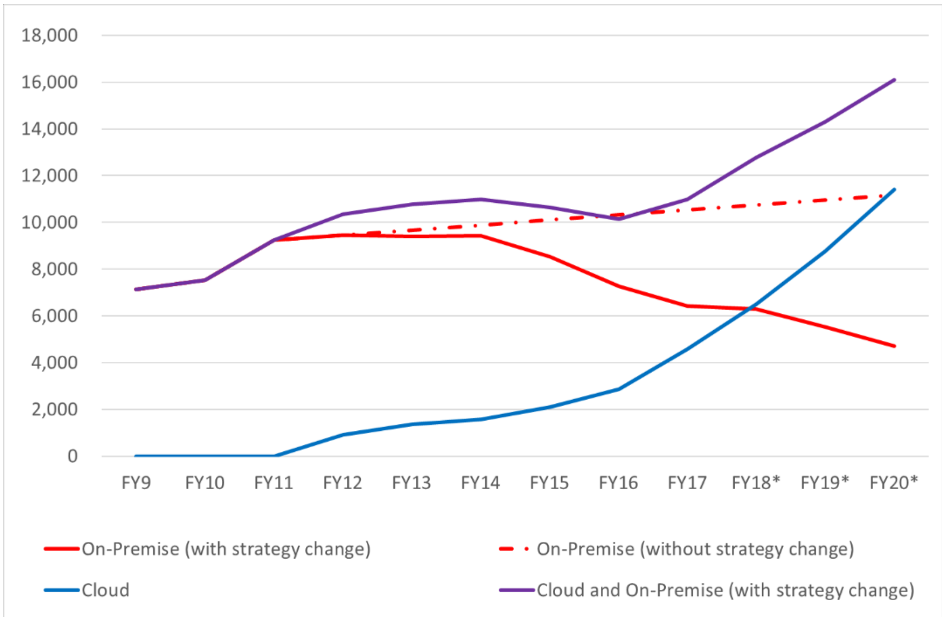
There could be two scenarios for the company regarding revenues with regard to selling on-prem products (Figure 1). On the one hand, cloud products of their own and other companies could cannibalize on-prem revenues. On the other hand, the company could maintain a slow growth of on-prem revenues, if top management focuses resources accordingly.

Figure 1 Scenarios of the on-prem revenues of the company (million USD)



Source: Company's financial report and company's forecasts based on IDC

Figure 2 Revenues of the company, in the case of escaping exploitation trap by exploration and exploitation



Source: Company's financial report and company's forecasts based on IDC

In the case of the company changing strategy, and cloud solutions are growing, the revenue from the on-prem and the cloud solutions combined will be significantly higher than the on-prem revenues without strategy change. This helps the company to escape from exploitation trap (Figure 2).

b.) Top management drove strategy change

By offering cloud services, the company has expanded to a different market segment. On the cloud market customer needs are different, as are the competitors. Sales cycles are shorter; the average contract size is smaller. Even the buyers from the customer side are different: cloud solutions often being purchased by the non-IT department.

The company subject of this research, with experience and success in selling large and complex on-prem solutions primarily to the IT departments of customers, faced a challenge. It had to deal with smaller size contracts, less complex solutions, work with different customer departments, and to compete with different firms than before.

Parallel to the growing cloud business, the company had to keep the on-prem business running to ensure stable revenue flow. When the cloud business started to become significant, the company had two options to structure its revenue generating sales force:

- Establish a new business unit to exclusively sell the cloud solution, and leave the on-prem sales within the existing sales business unit
- Add the cloud product to the sales portfolio of the existing sales business unit.

According to the 16th interviewee: “there was a debate in terms of this choice, which was a mental question.” The company selected the second option, and did not create a separate business unit for the cloud business. Instead, it added the cloud product to the portfolio of the existing sales force, and allowed and motivated the sales team to sell both cloud and on-prem products to customers.

In order to provide support to the sales force, the company created a different team of product and business development experts to focus exclusively on cloud business.

As the interviews with company employees revealed, the readiness of the market for cloud solution was not and is still not homogenous. There are differences based on geography, customer segment (enterprise vs. SMEs) and industry. There is broad adoption of cloud solution in the USA, Western Europe and in the Middle East, while customers in Eastern Europe and Africa are moving more slowly to the new way of IT.

SMEs are more open to moving to the cloud than large customers because they do not have resources to build and manage their own IT system. Some industry segments, such as Public Sector customers are more concerned with security problems and data residency issues than others.

The company was not the first mover to the cloud market, and it has consequences for the company on developed markets where cloud competitors already have a strong presence. In countries where the market is less developed, and cloud adoption less mature, that disadvantage is not as significant: other cloud vendors do not yet have significant business either.

- But there are many competitive advantages for the company that the interviewees mentioned:
- data centers are all over the world;
- complex, integrated, end-to-end solutions in the cloud;
- meet the requirements of the local markets and able to handle small countries better;
- one of the most secure companies;
- a strong sales force and next-generation sales approach;
- strong back-office and customer care system;

However, the company's strategy is to actively promote cloud across all geographies, customer segments, and industries. In some segments, the company responds to customer needs (pull mode), and in some segments, the company drives the customers to the cloud (push mode). This approach requires 'evangelization'; convincing customers that they should not invest in on-prem solutions today because that will not provide them with a modern, agile, and flexible IT system in the future.

According to the 8th interviewee: "the company is in push mode on the market, the attitude of top management is that every second we try to sell on-prem solutions is wasted. If the customer definitely wants to buy on-prem solutions, we give them to cloud solutions too, as a present." But the 6th interviewee said that "it seems irrational how the company is acting because there might be some on-prem markets which cannot be replaced with cloud solutions".

c.) Aligning the organization to the new strategy: Changes in field sales

The change from on-prem to cloud required major changes in the field sales organization. Sales reps had to learn about the new products, the new way of selling these products, and the new competition. Such significant change always requires management intervention and drive.

As described earlier, in many cases the company was not 'pulled' by customers to sell cloud products, but 'pushed' the cloud solutions on customers in anticipation of changing customer needs in the mid-term. Therefore, the drive for change was not bottom-up but top down. The top management of the company realized the need for change in strategy earlier then the lower levels did. Top management started to drive the change, and used various management methods to drive the change through the organization:

– Communication – top and later middle management focused its external and internal communication on cloud messages. What is the company strategy in the cloud and why is it important for customers?
– Financial incentives – the commission system for sales reps was changed to motivate them to sell cloud products. Multipliers were implemented in commission calculation, and sales reps started to earn two to four times more commission by selling cloud products then on-prem.
– Training – the company provided several days of training each year for the sales force, and these training sessions were only focusing on the new cloud portfolio. According to the 9th interviewee, "training is the platform to communicate a clear and strong direction for the company and its employees", but the 7th interviewee thinks that "these training are 3-day-long brainwashing events to sell cloud solutions".
– Re-defined and simplified internal processes to support the cloud business.

Using the above-listed methods led to a fast change in the mindset of the sales force, and as a result, the proportion of cloud revenues grew at a fast pace. This was a significant success, taking into consideration that tens of thousands of sales, support and operational employees who were involved in this change process. The company's internal culture and policies were crucial to this fast change and enabled this major change to occur in a relatively short time.

The company streamlined and centralized internal processes at the beginning of the 2000's. For example, 15 years ago the commission plans for sales reps were designed and prepared by individual country operations of the company. Country managers and country finance directors had the power to design renumeration plans based on their priority. For example, some renumeration plans included several KPIs, and some were based only on revenue generated by the sales rep. As a result of the streamlining and centralization efforts, the renumeration plans today are designed by the headquarters team and produced by shared service centers globally.

All sales reps in similar positions have the same renumeration scheme (they get a commission in the same way) across all countries where the company operates. This centralized system allowed the company to change the commission plan for all sales reps from one fiscal year to another to prefer cloud sales: there was no need to convince middle management and country managers why this change was necessary. Moreover, there wasn't a complicated process to work with local finance and HR operations in dozens of countries to change the commission plans. Based on top management decision, the shared service centers produced the new commission plans for all sales reps globally without interaction of middle management and country operations.

This streamlined operation made it possible for top management of the company to drive deep changes quickly and effectively across the large global organization.

d.) Exploring and exploiting new target segment: SME customers

The company traditionally was dominating the high-end of the market, focusing on large customers. With the growth of the cloud business, the mid and small size customers also become large potential segment for the company. Cloud solutions bring several benefits for SMEs, such as reduced opportunity cost, reduction of in-house ITC sunk cost, and scalability which improves business agility (Ross–Blumenstein 2015). Due to lack of their own IT staff and free cash available for CAPEX investment, SME customers have a large demand for cloud solutions. The company made a decision to build a large sales team, focusing only on SME customers across Europe, Middle East, and African territory. It announced to hire 1400 new sales representatives to address the SME segment.

The traditional sales force – working mainly with high-end large customers – was a field-based sales force: sales representatives were located across the territory to be able to regularly meet customers and interact with them. The field sales model is effective for large customers and large deals, but expensive. This model won't work effectively for SMEs.

The company decided to adopt a different model for the sales unit targeting SMEs: created a couple of telesales centers across EMEA. In each of those centers, there are several hundred sales representatives working with SME customers using modern ways of remote communication (telephone, email, chat, video calls, social media), supported by the latest technology. The sales reps in those centers are able to deliver live demos to customers and present proposals from thousand kilometers away.

The profile of the sales reps in the telesales centers is different to that of field sales. In field sales, reps have several years of experience (sometimes 10+ years); in the telesales centers, many of the reps are new graduates from university. Young and dynamic telesales reps don't have a problem with mindset change from on-prem to the cloud – most of them started to work in the cloud world.

As the 14th interviewee said, "we build a new-generation sales organization with hiring young freshly graduated people, who can act as digital marketing campaign agents", and the 15th interviewee confirm the trend of the changing ways of communication: "in 2014, only 30% of communication with customers took place by phone, in 2016 it's 80% and in the future it could be 100%".

5. Summary

Exploitation trap used to be a common pitfall for many firms. Nowadays, the digital transformation is reshaping every industry. However, it was uncertain, why some firms are able to lead this transformation and take full advantage of it, as well as avoid exploitation trap.

We examined the following research questions: How can a firm escape exploitation trap? How can an established firm can drive industry change by exploring and exploiting cloud computing?

In order to answer the research questions, a major player in the cloud computing industry was selected and closely monitored through a multi-level approach in order to track down its digital transformation.

The case study highlighted that the following actions are needed to avoid exploitation trap:

- exploring a new industry
- top management driven strategy change
- aligning the organization to the new strategy
- exploring and exploiting new target segments

This transformation was enabled by

- a change-oriented organizational culture: in the past 15 years a major organizational change was carried out in every 2–3 years,
- strong leadership: clear vision and direction
- strong and aligned support systems: structural and rewards systems

References

Beer, M – Eisenstat R. A. – Sceptor, B. (1990): Why Change Programs Don't Produce Change. *Harvard Business Review*, 6, 158–166.

Burgelman, R. A. (1991): Intraorganizational Ecology of Strategy Making and Organizational Adaption: Theory and Field Research. *Organizational Science*, 2, 239–262.

Burgers, J. H. – Jansen, J. J. – Van Den Bosch, F. A. – Volberda, H. W. (2009): Structural differentiation and corporate venturing: The moderating role of formal and informal integration mechanisms. *Journal of Business Venturing*, 24, 206–220.

Caldarelli, A. – Ferri, L. –Maffei, M. (2017): Expected benefits and perceived risks of cloud computing: an investigation within an Italian setting. *Technology Analysis & Strategic Management*, 29, 167–180.

Chesbrough, W. (2002): *Open Innovation: The new imperative for creating and profiting from technology*. Harvard Business School Press, Boston, MA.

Chesbrough, W. (2006): *Open business models: How to thrive in the new innovation landscape*. Harvard Business School Press, Boston, MA.

Child, J. (1972): Organizational structure, environment and performance: the role of strategic choice. *Sociology*, 6, 2–22.

Clemmer, J. (1995): Pathways to Performance: *A Guide to Transforming Yourself, Your Team, and Your Organization*. Macmillan Canada, Toronto.

DaSilva, C. M. – Trkman, P. – Desouza, K. – Lindic, J. (2013): Disruptive technologies: a business model perspective on cloud computing. *Technology Analysis & Strategic Management*, 25, 1161–1173.

Dickhout, R. – Denham, M. – Blackwell, N. (1995): *Designing Change Programs: Thant Won't Cost You Your Job*. The McKinsey Quartely, 4, 101–116.

Gibson, C.B. – Birkinshaw, J. (2004): The antecedents, consequences, and mediating role of organizational ambidexterity. *Academy of management Journal*, 47, 2, 209–226.

Hortoványi, L. (2012): *Entrepreneurial Management*. AULA, Budapest.

Junni, P. – Sarala, R. M. –Taras, V. – Tarba, S. Y. (2013): Organizational ambidexterity and performance: A meta-analysis. *The Academy of Management Perspectives*, 27, 299–312.

Kauppila, O. P. (2010): Creating ambidexterity by integrating and balancing structurally separate interorganizational partnerships. *Strategic organization*, 8, 283–312.

Kotter, J. P. (1995): Leading Change: Why Transformation Efforts Fail. *Harvard Business Review*, 2, 59–67.

Kotter, J. P. (1999): *Változások irányítása*. Kossuth Kiadó, Budapest.

Lee, J. – Lee, J. – Lee, H. (2003): Exploration and exploitation in the presence of network externalities. *Management Science*, 49, 553–570.

March, J. G. (1991): Exploration and Exploitation. *Organization Science*, 2, 1, 71–87.

Miles, M. B. – Huberman, M. A. (1984): *Qualitative Data Analysis: A Sourcebook of New Methods.* California, SAGE publications Inc.

Mintzberg, H. – Ahlstrand, B. – Lampel, J. (1998): *Strategy Safari*, Prentice Hall, London.

Oh, C. – Cho, Y. – Kim W. (2015): The effect of a firm's strategic innovation decisions on its market performance. *Technology Analysis & Strategic Management*, 27, 39–53.

O'Reilly, C. A. – Tushman, M. L. (2004): The Ambidextrous Organization. *Harvard Business Review*, 82, 4, 74–81.

O'Reilly, C. A. – Tushman, M. L. (2013): Organizational ambidexterity: Past, present, and future. *The Academy of Management Perspectives*, 27, 324–338.

Raisch, S. – Birkinshaw, J. – Probst, G. – Tushman, M. (2009): Organizational Ambidexterity: Balancing Exploitation and Exploration for Sustained Performance. *Organization Science*, 20, 4, 685–695.

Reeves, M. – Claire, L. – Phillip, T. (2012): Your strategy needs a strategy. *Harvard Busienss Review*, 9, 76–83.

Ross, P. K. – Blumenstein, M. (2015): Cloud computing as a facilitator of SME entrepreneurship. *Technology Analysis & Strategic Management*, 27, 87–101.

Schendel D. E. – Hitt M. A. (2007): Introduction to volume 1. *Strategic Entrepreneurship Journal*, 1, 1, 1–6.

Schumpeter, J. A. (1942): *Capitalism, socialism and democracy*. New York City, NY, USA, Harper and Brothers.

Solt, O. (1998): Interjúzni muszáj. In: *Méltóságot mindenkinek*, Beszélő, Budapest, 29–48.

Sultan, N. – Van De Bunt-Kokhuis, S. (2012): Organisational culture and cloud computing: coping with a disruptive innovation. *Technology Analysis & Strategic Management*, 24, 167–179.

Szabó, Zs. R. (2014): *Strategic Adaptation, Ambidexterity and Competitiveness*. USA: LAP - Lambert Academic Publishing.

Tichy, N. M. – Sherman, S. (1993): *Control Your Destiny or Someone Else Will: How Jack Welsch Is Making General Electric the World's Most Competitive Corporation*. Doubleday, New York.

Tushman, M. L. – O'Reilly, C. A. (1996): Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review*, 38, 4, 12–18.

Tushman, M. L. – O'Reilly, C. A. (2002): *Winning through Innovation: A Practical Guide to Leading Organizational Change and Renewal*. Harvard Business School Press, Boston

Tushman, M. – Smith, W. K. – Wood, R. C. – Westerman, G. – O'reilly, C. (2010): Organizational designs and innovation streams. *Industrial and Corporate Change*, 19, 1331–1366.

Venkatraman, N. – Lee, C. H. – Iyer, B. (2006): *Strategic ambidexterity and sales growth: A longitudinal test in the software sector*. Paper presented at the Annual Meetings of the Academy of Management, Honolulu, Hawaii, Citeseer.

Vuori, T. O. – Huy, Q. N. (2016): Distributed attention and shared emotions in the innovation process: How Nokia lost the smartphone battle. *Administrative Science Quarterly*, 61, 9–51.

https://www.statista.com/statistics/510350/worldwide-public-cloud-computing/ Accessed 24.04.18.

Xu, L. D. – Xu, E. L. –Li, L. (2018): Industry 4.0: state of the art and future trends. *International Journal of Production Research*, 56, 2941–2962.

Yu, G. J. – Khessina, O. (2012): *The role of exploration in firm survival in the worldwide optical library market,* 1990–1998. Annual Meetings of the Academy of Management.

Chapter III

Finance

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 258–274.

Has Macroeconomic Imbalance Procedure managed to reduce imbalances and reduce economic crises in EU countries?

Marcell Zoltán Végh

The European Union has suffered a prolonged crisis episode due to the global financial crisis of 2008–2009, followed by an economic crisis and a sovereign debt crisis in various Member States. Robust, pre-crisis economic growth has failed to recover ever since and levels of unemployment have remained high, thus economic performance is struggling to reach pre-crisis levels in what is called periphery countries. Common institutions have been improved through several important changes in terms of fiscal and monetary policies as well, resulting in a preferable, more stable economic structure. To ensure balanced economic growth, the European Commission has launched a monitoring system containing 14 indicators and a corrective operation (Macroeconomic Imbalance Procedure) which aims to reduce economic imbalances in Member States. This tool, complementing the regulations of the Stability and Growth Pact, may help to reduce the evolution of further crises and to establish a more sustainable economic growth rate. However, Member States do not react the same way.

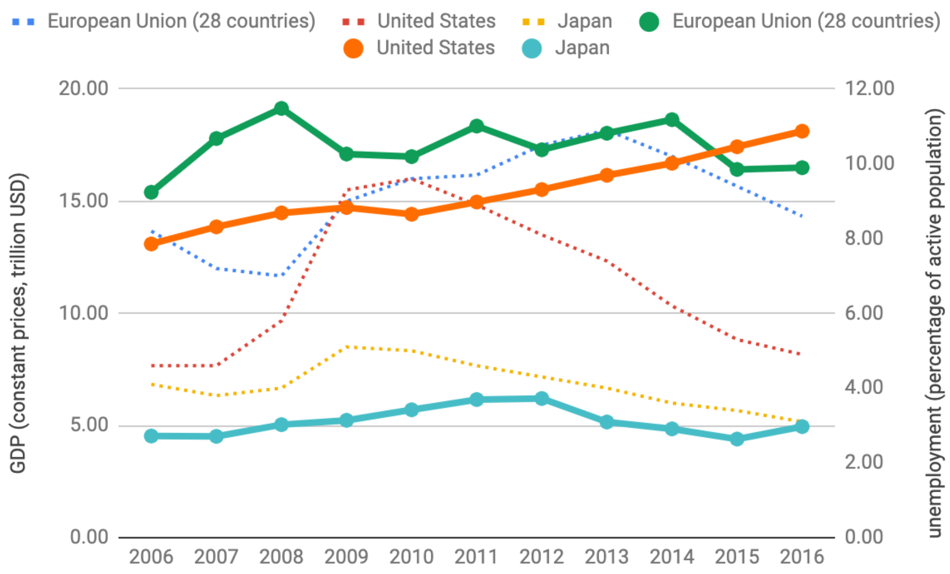
In this study, Macroeconomic Imbalance Scorecard data is examined to establish a connection between imbalances and economic growth. The study investigates whether all 14 indicators have relevance, or whether some indicators could be eliminated due to correlation within the data set. Then it also aims to identify those indicators which have greater relevance to estimate the probability of a crisis, in order to describe which imbalances lead to higher probability of a crisis event in the short term. Fighting these imbalances with the tools of the MIP could safeguard economic growth in the EU. In order to achieve all this, cross-correlation and logistic regression methods are suggested. In the future, having extended time-series database will probably allow the running of even more elaborate statistic examinations and achieving more complex results.

Keywords: economic crisis, eurozone, economic imbalance, recovery

1. Introduction

The global financial crisis of 2008–2009 has had a devastating effect on European economic integration, causing a severe, longer recession than in other developed economic regions, and leading to bigger setbacks in the field of economic growth and unemployment (Figure 1). There are plenty of explanations of why the EU's recession became more severe (for details, see Végh 2019), ranging from Optimum Currency Area theories, through institutional weaknesses, to fiscal and monetary crisis management failures. What is widely regarded as a fact is that the group of periphery countries (the terminology of which has been widely accepted in the economics literature as a distinction between core and periphery countries, see e.g. Mangone et al. 2016, Pelle et al. 2019), whose competitiveness weakened and current account problems were aggravated during the decade preceding the eruption of the crisis, became interconnected in terms of bond market risks and, at one point, the threat of government defaults became imminent. Fortunately, only two Member States (henceforth: MSs) (Greece in 2012, and Cyprus in 2013) suffered a direct default, but several other MSs (namely, Portugal, Spain, and Ireland – therefore often labelled as PIGS countries) had to undergo international financial aid programs disposed of by the European Stability Mechanism and its predecessor institutions. Meanwhile, in terms of monetary policy, the European Central Bank also had to take an active role in crisis management: with several well-targeted programs (such as the Securities Market Program and the Outright Monetary Transactions) it placed direct pressure on secondary bond markets in order to lower default risk and to ensure the financing of MSs in peril. In the meantime, common EU institutions went through a gradual development: Stability and Growth Pact has been reformed in several steps (Six Pack regulation in 2011, Two Pack in 2013 and Fiscal Compact also in 2013), followed by the gradual formation of the Banking Union, initiated in 2012.

Figure 1 GDP (constant prices, trillion USD – left axis, thick) and unemployment (percentage of active population – right axis, dotted) in advanced economic regions, 2006–2016



Source: own edition based on Eurostat (2019) and World Bank (2019)

In this study, I examine trends of macroeconomic imbalances, which is a new concept brought to life by the Six Pack regulation, aiming to identify macroeconomic disturbances. I investigate the possibility to estimate the probability of whether imbalance phenomena led to the evolution of further economic crises. First, I aim to reduce the number of indicators using cross-correlation method within the data set, then try to identify the relevant imbalance indicators which estimate the probability of crisis events. In order to achieve that, I describe macroeconomic imbalances, review their background and discover their limitations, and I also revise some relevant studies aiming to measure imbalances and their effect. Finally, I present my own statistical analysis using logistic regression on a time series database consisting of each currently known member state (EU-28), based on previous achievements. I expect to reinforce the role of macroeconomic imbalance indicators in the prediction of economic crises within the European economic integration.

2. New era of economic governance

The recent reform package of the Stability and Growth Pact called Six Pack was ratified and accepted in December 2011, following a full year of negotiations. The legislation consisting of 6 fiscal rules was created in order to increase MSs' fiscal discipline without further modification of the existing SGP regulations. Furthermore, Six Pack introduces a macroeconomic prediction and monitoring system called Macroeconomic Imbalance Procedure (henceforth: MIP), which enables common institutions to monitor MSs' fiscal discipline and their other macroeconomic tendencies. The press conference (EC 2011A) introducing the Pack's five regulations and one directive emphasizes that in December 2011, 23 of 27 MSs were sanctioned by Excessive Deficit Procedure (henceforth: EDP), which definitely justifies the reinforcement of fiscal discipline. After the ratification of the Six Pack legislation, immediate sanctions come into force in case an MS does not implement measures to improve fiscal position, in case of excessive deficit, or in case of not implementing ECOFIN recommendations. Policy makers believed that new sanctions would result in more respect towards fiscal discipline regulations: volumes of these sanctions are assessed jointly by the Commission and the ECOFIN, and they come into force automatically, except due to a contrary majority vote of MSs. Budgetary deficit rules are followed by regulation concerning public debt: the new rules render EDP automatic if a government's debt volume exceeding 60% of GDP does not decrease by 1/20 of the surplus' volume per year (EC 2011A). Regarding the fact that the majority of MSs were under EDP by the time Six Pack was accepted, this implies a severe, prompt deleveraging process and widespread fiscal consolidation to most MSs, immediately blocking fiscal stimulus programs in the whole economic integration (Végh 2019). However, this viewpoint has been widely criticized by theoretical economists because the calming and stabilizing effect of the debt brake rule is not proven to be more significant than the lack of government interventions' destabilizing effect on bond markets.

Six Pack's innovations linked to the preventive arm of SGP are mostly concerning the system of macroeconomic imbalances. The preceding years clearly showed that not only unstable fiscal position can threaten the stability of the economic integration, but other imbalances as well. In order to increase fiscal discipline, Six Pack introduces automatic barriers to maximize expansion of government spending in countries not achieving their medium-term objectives; moreover, the regulation establishes the maximized possible fine under EDP at a value of 0.2% of the MSs' GDP. This fine can be imposed after continuous disregard of necessary budgetary corrections, first as a non-interest-bearing deposit, later converting into an interest-bearing fine. In the field of macroeconomic imbalances, the regulation also introduces a new procedure, called Excessive Imbalance Procedure (henceforth: EIP) (EC 2011B). In addition, SGP's preventive arms were rendered as part of the European Semester to ensure the European Commission's (henceforth: EC) ability to monitor interconnected effects of MSs' budgets. This way, the EC can articulate direct recommendations concerning MSs' budgets (through Stability and Convergence Reports), which is a key innovation of the reform package. The corrective arms of the SGP also became stricter: EDP was rendered automatic in case of budget deficit exceeding 3%. In case any MSs' budget deficit is regarded as excessive, ECOFIN is entitled to formulate recommendations, deadlines to the desired results and also possible fines (Holler–Reiss 2011).

The aim of EIP is that economic phenomena causing disturbance and instability among MSs can be identified by the Commission and ECOFIN in due time, in order to launch preventive measures. These undesired imbalances can take various forms such as a sudden increase of real estate prices, current account deficits or surpluses, or chronic indebtedness of the private sector (for the indicator scoreboard, see Table 1).[1] In practice, all this takes place in a way that the corrective arm of the SGP can be activated if common institutions launch EIP against an MS. In these cases, EC obliges the MS at issue to present a blueprint of countermeasures. In the case of it being eligible, this blueprint is accepted by the Commission with the following logic: not adhering to the blueprint will entail a smaller amount of non-interest-bearing deposit as fine (maximized in 0.1% of MS's GDP), which can be later converted into an interest-bearing fine.

[1] In some cases, different thresholds were established for non-Eurozone MSs. In regard of ongoing development of economic relations, the set of indicators and the thresholds can be changed in the future, or even the volume of indicators can be increased.

Table 1 List of MIP indicators and the indicative thresholds (MIP scoreboard)

Type	Indicator	Threshold
external imbalances and competitiveness	3-year average of the current account balance as a percentage of GDP (CAB)	6% and –4%
	net international investment position as a percentage of GDP (NIIP)	–35%
	5-year percent change of export market shares measured in values (EMS)	–6%
	3-year percent change in nominal unit labor cost (NULC)	9% for euro area countries and +12% for non-euro area countries
	3-year percent change in real effective exchange rates (REER)	–/+5% for euro area countries and –/+11% for non-euro area countries
internal imbalances	private sector debt (consolidated) as a percentage of GDP (PSD)	133%
	private sector credit flow (consolidated) as a percentage of GDP (PSCF)	15%
	year-on-year percentage change in deflated house prices (HIP)	6%
	public sector debt as a percentage of GDP (GGD)	60%
	year-on-year percent change in total financial liabilities of the financial sector (TFSL)	16.5%
social indicators	3-year average of the unemployment rate (UR)	10%
	3-year change of the activity rate (AR)	–0.2 pp
	3-year change of the long-term unemployment rate (LTUR)	0.5 pp
	3-year change of the youth unemployment rate (YUR)	2 pp

Source: author’s creation based on EC (2011B)

3. A critical approach to MIP

There are various theoretical and practical concerns in relation to the development of economic governance which need to be answered by the EU institutions. While it can be observed that the monitoring of MSs' macroeconomically risky performance is based on an increasingly detailed and specified set of indicators, the complexity of the detailed rules eventually makes the economic performance of the MSs less transparent (Kiss 2010). Therefore, MSs' political actors are less accountable to their voters and to the public. In addition, the behavior of compliance with the numerical limits imposed by the common institutions does not fundamentally reflect fiscal compliance habits and practices of MSs. In political cultures prone to opportunism, automatic tracking of numerical fiscal rules (such as complying with MIP regulations) can occur at the expense of other, unregulated, but important and practical indicators (Farkas 2012, Bánfi 2018); other factors determining long-term competitiveness (such as education, health, or digital literacy) may easily suffer damage in such cases (Domonkos et al. 2017). As a result, in multiple cases, competitive disadvantages will be revealed only in the coming decades, mostly in the fields of absence of investment and growing current account deficit. In other, less opportunistic, rule-abiding political cultures, respect of fiscal discipline standards can also be achieved by respecting competitiveness indicators as well. In other words, indicator-based standards do not reflect the qualitative indicators of governance, and the time-inconsistency problem of economic policy-making may push short-term interests of a non-responsible political elite forward. As a result, we can conclude that the culture of numerical rule-following does not serve the objectives of cohesion policy as a common policy since it does not directly contribute to real convergence between MSs. The EU institutions should therefore – in addition to quantitative indicators – introduce qualitative analysis and monitoring of structural reforms in order to achieve a greater emphasis on convergence objectives (Halmai 2015, Kengyel 2016, Boros 2017). However, this may prove impossible without giving up additional fields of sovereignty at the MS level or by ceasing asymmetric information-based demagogic political culture (Halmai 2018). It can also be stated that one MS's fiscal discipline is fundamentally dependent on the internal commitment of its political elite while the evolution of the common fiscal and monetary rules reveals the heterogeneity of this commitment (Csaba 2018).

The definite advantages of the new economic governance practices are that the 'one-size-fits-all' approach is getting more distant (though it does prescribe strong numerical constraints), and the implementation and suspension of imbalance procedures is now a multi-staged and sophisticated process, with the participation of the MSs and the common institutions. 'One-size-fits-all' crisis management was one of the frequent criticisms concerning EU crisis management after the eruption of the global economic crisis that began in 2008 (Blyth 2013, Győrffy 2013, Pisani–Ferry 2014), which set similar structural reforms as a condition of financial assistance, disregarding economic, institutional, infrastructural, and cultural characteristics of the specific MSs. The process of adopting reform packages was difficult in many cases (e.g. Greece or Portugal) and reforms only moderately gained political legitimacy or social support. Although MIP allows for a much more sophisticated approach to fiscal discipline, social support of these reforms still remains questionable. However, the Fiscal Compact debt brake rules continue to operate under the 'one-size-fits-all' principle, i.e. they require a common debt relief process for over-indebted MSs, thus losing leeway for dealing with their MS-specific economic issues (Kertész 2014).

Another aspect is that, although the 'no-bailout-clause' is the primary guarantee of MSs' fiscal discipline, in a default situation, EU-level financial assistance institutions such as the European Stability Mechanism significantly expand the fiscal opportunities of MSs (be it due to their own irresponsibility or for reasons beyond their control). However, applying for financial assistance requires a partial suspension of economic sovereignty, which may reduce the tendency to fiscal indiscipline. However, the first historic attempt to overcome the EDP in 2005 by loosening its sanctions was driven by the MSs with the most economic power (Germany and France) while that series of events presents a worrying precedent (Authers 2013). The power of regulation to enforce fiscal discipline can be dealt with by a similar, loosening legislative process. On the other hand, the slow, deliberate rule-making procedure provides an opportunity to incorporate debates, experiences, and theoretical considerations on this topic in the development of the fiscal discipline regulatory area.

4. How to measure macroeconomic imbalances?

Measuring macroeconomic imbalances and launching procedures based on them implies a new era in terms of economic coordination among the MSs of the EU, even though the effectiveness of these efforts can be questioned. There are several methods to measure MSs' aggregated position based on macroeconomic imbalances, also linking imbalance indicators and indicators of economic growth. In this section, three methods will be discussed from the wide range of economic literature. First, Csortos–Szalai (2014) establish an early-warning system for ten Central and Eastern European Countries in order to predict financial crises based on macroeconomic imbalance indicators. They aim to find the strongest set of indicators predicting a crisis by adding together successful predictions, also nuancing with false alarms and misses (by introducing an 'average-noise-to-signal' approach). The authors use annual Eurostat data, defining a crisis event when a country's GDP year-on-year change diverges from the country's trend by the sample's standard deviation (–1.68%) but they use only 6 macroeconomic imbalance indicators with standardization by 'gap measures' (thus examining Credit-to-GDP gap, Credit growth gap, Investment gap, Real exchange rate gap, Capital flows gap, Global variable gap). They also measure the strength of the predictions on three-time horizons (1 year, 2 years and 3 years). Their key findings are that (1) prediction indicators are crucial to preventing the build-up of significant imbalances, (2) largest downturns are preceded by credit boom, investment boom and severe capital flows as well, and (3) the gap of the global variable, the real exchange rate gap and the capital flow gap were proven the most efficient predictors. The authors conclude by stating that these prediction indicators can be a powerful tool for central banks. Since the European Commission is also aiming to improve the MIP set of indicators further, this type of analyses has great importance.

Another article (Domonkos et al. 2017) provides a slightly different approach to investigating the effects of imbalances: first, they did not use all 14 MIP indicators either but applied autocorrelation calculations to reduce the number of relevant indicators from 14 to 11. The authors here also defined crisis events in a simpler way, running the tests when year-on-year GDP change was lower than –2.5%, –2% and –1.5% indicating an MS is suffering an economic crisis. Moreover, they also used annual data, and ran the calculations on the same time horizon as Csortos and Szalai (2014) (i.e. 1, 2 and 3 years), also using linear regression to show the connection between the variables. However, Domonkos et al. (2017) introduced a factor analysis to create complex indicators predicting the crisis with the easier explanation. The best descriptive factors were named 'Labour–Capital Nexus', 'Competitiveness and Catch-up Effect' and 'Real Estate Bubble', pointing out that factor analysis can further simplify the possible prediction of a crisis event. The authors also emphasize that such early warning systems need further research because false signals (be they positive or negative) can imply great societal and economic costs for MSs, possibly also harming other countries since MSs' economies are greatly interconnected in the EU (Pelle 2018).

Another composite-indicator-creating method was introduced by Bobeva–Atanasov (2017) who define the Integral Macroeconomic Imbalance Indicator (IMII), aiming to compare the level of imbalances between the countries and groups of countries in a simple way. IMII indicator is a standardized average on the imbalances compared to the threshold, observed in a specific MS or group of countries, also based on annual data. This less statistically sophisticated methodology has clear advantages and disadvantages: country data are easily comparable, and also country groups such as hierarchical clusters can be analyzed. On the other hand, measurement of an imbalance is distorted by many factors such as mild and serious imbalances evening out each other, or relative excess of an imbalance value is subject to the original threshold estimation. Despite all this, as the authors clearly point out, an accumulation of imbalances can be clearly shown, peaking 3–4 years after the eruption of the crisis, transforming into a mildly descending tendency. Still not surprisingly, periphery MSs accumulate higher imbalances while countries performing well during the years of crisis obtain lower scores. All these results imply that further observation of imbalances and their effects is necessary, and new models can be developed to interpret crisis procedures better. A summary of the previous methodological approaches is given in Table 2.

Table 2 Methodological summary of presented imbalance measurement methods

publication	Csortos and Szalai (2014)	Domonkos et al. (2017)	Bobeva and Atanasov (2017)
scope of investigation	10 CEE member states	EU28 countries	EU28 countries
number of indicators	6 macroeconomic imbalance indicators with standardization by ‘gap measures’	reduced to 11 utilizing autocorrelation methods	all 14 indicators
crisis event defined as	year-on-year GDP change exceeds from the country’s trend by the sample’s standard deviation (–1.68%)	year-on-year GDP change being lower than –2.5%, –2% and –1.5%	year-on-year GDP change below zero
methodology	‘average-noise-to-signal’ approach	factor analysis (creating 3 compound factors) and linear regression	compound indicator (Integral Macroeconomic Imbalance Indicator, IMII) based on standardized means of imbalance indicators
time horizon	1 year, 2 years and 3 years	1 year, 2 years and 3 years	none

Source: author’s own creation

5. Statistical analysis

In this section, my aim is to develop the aforementioned methods further, in a way that utilizes the advantages of previously discussed imbalance measurement approaches. To begin with, I make a comparison between the two imbalance procedures (EDP and MIP) by comparing the results. Since measuring the effectiveness of these procedures would probably be overambitious, we examine MIP scoreboard further: at first, I aim to reduce the number of necessary indicators to have a more accessible way to measure MSs imbalance performance, then we carry on with the observation of which indicators have the strongest meaning in terms of estimating the probability of economic crises. I will present results of a logistic regression model on a time series database of each currently known member states (EU-28), to have a clearer view on the usefulness and effectiveness of MIP indicators.

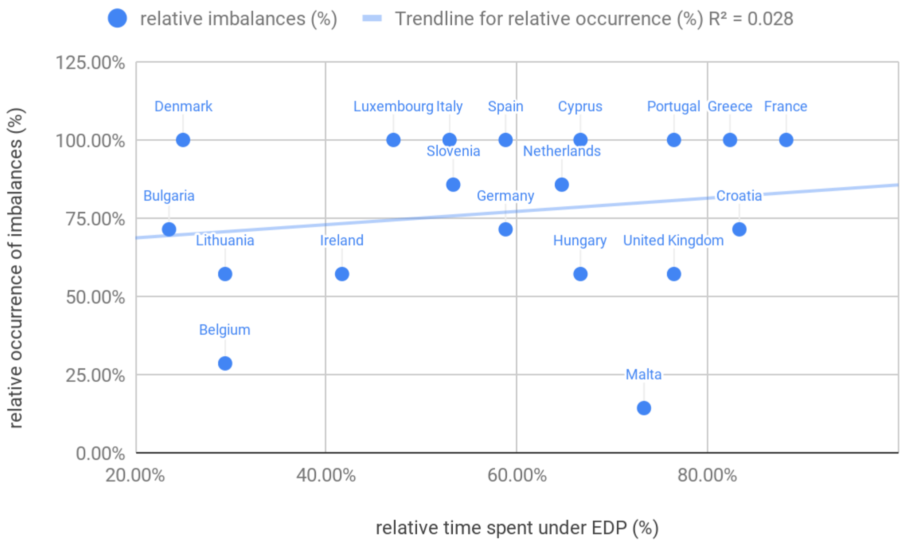
5.1. Comparison between financial discipline procedures

In order to examine fiscal discipline procedures, it is worth comparing the lessons of EDP and MIP. Analyzing the two procedures, it is obvious to try to establish a pattern of fiscal discipline (if it can be identified), for certain countries or groups of countries. Since EDPs became widespread in 2010 (striking 25 MSs out of 27), it is difficult to deduce any conclusion on MSs' behavior concerning fiscal discipline from this indicator. Thus, on the one hand, the number of years spent by a specific MS under EDP between 2002 and 2018 can be defined as a relevant indicator, such as the number of years spent under the EDP before the crisis (from 2002 to 2007). However, both indicators should be interpreted correctly, proportionately to the number of years spent in EU membership. Examining these indicators, we find that, concerning the relative time spent under the EDP, among the worst performing countries we find MSs from the core countries (e.g. France which was 15 years under EDP out of the 17 years since 2002, with a 'relative time spent under EDP' ratio of 88.24%), recently joined countries (such as Croatia and Poland, with 83.88% and 80% ratios, respectively) and countries from the periphery (e.g. Greece or Portugal, under EDP for 14 years and 13 years, expressed as a ratio of 82.35% and 76.47%, respectively). During the entire time period (2002–2018), no EDP was initiated against Estonia or Sweden, and only one against Luxembourg, in 2010, which was resolved within a year. A more relevant indicator of fiscal discipline may be the relative time spent under EDP in the pre-crisis time period (2002–2007): nevertheless here, among the worst performers, we also find core MSs (Germany) and newcomers (Poland, Malta, Hungary and Slovakia). Among the MSs unaffected by the EDP in this period, we can find peripheral countries (Spain and Ireland), core countries (Austria, Belgium, Denmark, Finland, Luxembourg, and Sweden), as well as new entrants (Slovenia, Czech Republic, Latvia, Lithuania, and Estonia). So, based on this data, it is hard to establish a trivial pattern of fiscal discipline along aforementioned country groups. However, the data can be examined further.

The macroeconomic imbalance indicators date from 2011 and the procedures indicating imbalances or excessive imbalances naturally reflect the lessons of the Eurozone crisis phenomena. Even though no strict EIP has been launched against any of the MSs up to the present, during the seven years' period of 2012 to 2018, some imbalances have been registered in 9 MSs: the Eurozone periphery countries, and Bulgaria, France and Sweden in addition. Only 8 MSs did not experience any imbalance in this time period; these countries are partially core countries (e.g. Austria and Luxembourg), and partially recently joined Central and Eastern European countries (e.g. Czech Republic, and Estonia). However, by limiting the investigation on excessive imbalances, the overall picture is more subtle: we can see from the examination of the figures that there are persistent excessive imbalances mostly in the periphery countries (with the exception of Ireland and Spain where, within 1 or 2 years, the excessive imbalances were resolved). Excessive imbalances within the core countries have only endured in France and for only 3 years. The Commission has a direct influence on the economic policies of the MSs rescued by the ESM's assistance programs; however, as the data show, imbalances could persist in the long run.

To progress further, we can compare the relative time spent under EDP and the relative time spent with macroeconomic imbalances in all the MSs concerned. If we spot these phenomena on a scatter plot (Figure 2), a trend line can be drawn although with high variance and weak correlation. Nevertheless, the data suggests that MSs where no imbalances occurred in the MIP procedures have been affected by an EDP of only 36.81% on average. In the case of MSs where some imbalances occurred over the seven-year period, this ratio is 54.91%. When examining excessive imbalances, the difference is even more striking: while MSs not subject to such a procedure (a total of 17) spent an average of 42.17% of their EU membership in the EDP process while for MSs experiencing excessive imbalances this ratio will rise to an average of 61.44%.

Figure 2 Relative occurrence of imbalances and relative time spent under EDP



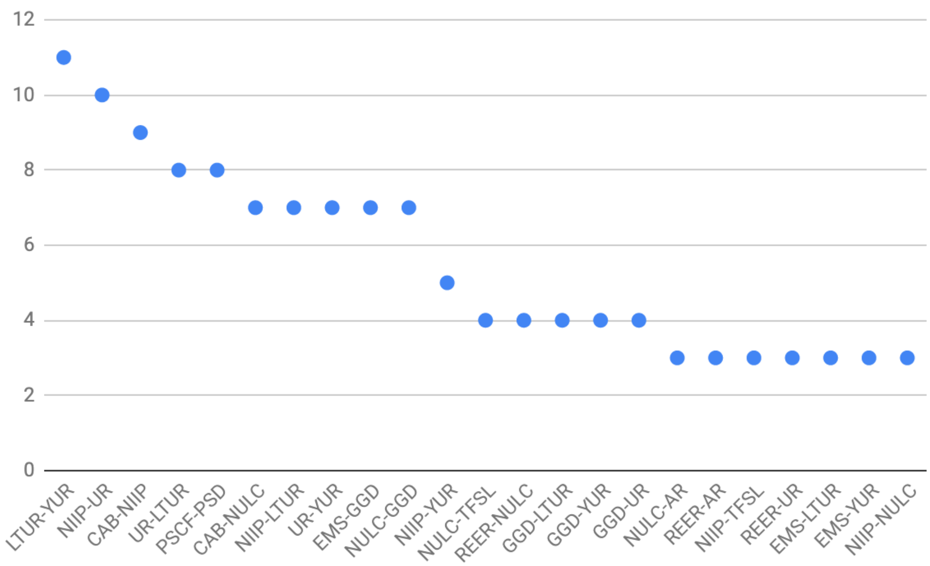
Source: own calculations based on EC (2019)

The above figures suggest that there is a slight overlap between countries sanctioned by EDP and the MIP. This can be interpreted such that the two procedures mainly sanction similar countries, i.e. the disciplinary effect of economic governance – although the years of crisis have had a significant impact – can be considered at least half-sided. However, it can be highlighted that the Commission is now in a better position to monitor macroeconomic imbalances and influence MSs' budgets with the aid of the European Semester and the ESM than before 2012. Nevertheless, this only nuances the fact that there is an overlap between pre-crisis and post-crisis disciplinary behavior, which shows that fiscal discipline primarily remains a matter for MSs' willingness (see also Végh 2019). All this implies that the importance of deficit and imbalance procedures should not be overestimated; MS governments and central banks should improve to recognize and handle their own imbalances in order to prevent future crises.

5.2. Elimination of correlating MIP indicators

Before measuring MIP data, first I examine whether all 14 indicators are relevant in the field of our investigation. To this end, I observed imbalance data for all 28 MSs between 2002 and 2017, using Pearson's correlation in order to identify any kind of (positive or negative) correlation between the variables in all 16 data sets. To describe the results: I have found 179 correlating indicator pairs within the complete data set at a 5% significance level (and a total of 98 correlating pairs at a 1% significance level). I have presented the most often correlating indicator pairs on a scatter plot, as shown in Figure 3. Not surprisingly, indicators related to unemployment show significant correlation in many cases, meaning that the various unemployment indicators (as described in Table 1) change similarly. Nominal unit labor cost also shows significant correlation with the previously mentioned unemployment indicators, which makes sense from an economic point of view: rising unemployment often coincides with decreasing labor costs. The indicator of net international investment position is also frequently observed among the correlation pairs (explained typically by current account balance, which is also a logical deduction). As a result of these observations, to progress further, we decided to omit 4 indicators from the data set: net international investment position (NIIP), Long-term unemployment rate (LTUR), Youth unemployment rate (YUR), and Nominal unit labor cost (NULC). By this step, we can focus on only 10 indicators and construct a better regression model.

Figure 3 Correlating MIP indicator pairs (Pearson-correlation) (number, 2002–2017)



Source: own calculations based on Eurostat (2019). Correlation is significant at the 0.05 level (2-tailed)

5.3. Which indicators estimate the probability of economic crises the best?

To test the predictive power of MIP indicators, I have chosen to set up a binary logistic regression model, which is very useful to estimate the probability of membership of two categorical outcomes. As an independent variable, we have chosen MIP data of the 10 indicators (reduction of relevant number of indicators is similar to the approach of Domonkos et al. 2017), from the time period 2002-2017, which is the broadest dataset currently available. We also used a standardization procedure, similar to Bobeva and Atanasov's (2017) method: when there is no imbalance, the value is set to 0, when there is an imbalance, the value is set by the following equation:

Function 1. Standardization of MIP indicators

$$N_i^j = \frac{I_i^j}{I_{threshold}}$$

Source: author's own creation

where N_i^j is the normalized value of imbalance, I_i^j is the value of MIP indicator and $I_{threshold}$ is the threshold value of the imbalance variable. For the dependent variable, I have chosen a binary indicator whether in that specific year, the GDP output gap was above the sample's current yearly average ('no crisis') or below ('crisis'), similarly to Csortos–Szalai (2014)'s approach. I also claim that it is useful to calculate and compare results for multiple time horizons, namely for t=0 year, 1 year and 2 years, also similarly to Csortos–Szalai (2014) and Domonkos et al. (2017). However, the binary logistic regression approach is completely my own, and this examination brings results in the field of MIP indicators' predictive power forecasting economic crisis. The function I applied is the following

Function 2. Binary logistic regression function of MIP indicators

$$Y_n = \beta_0 + \beta_1 * CAB_n + \beta_2 * REER_n + \beta_3 * EMS_n + \beta_4 * HIP_n + \beta_5 * PSCF_n + \beta_6 * PSD_n + \beta_7 * GGD_n + \beta_8 * UR_n + \beta_9 * TFSL_n + \beta_{10} * AR_n + \varepsilon_n$$

Source: own edition

where Y_n is the outcome, βn is the regression coefficient of the corresponding indicator value (MIP variable) and $\varepsilon$is the error of the estimation[2]. Results are shown in Table 3, where Exp(B) values are beta-coefficients of the regression analysis – the model is capable of estimating the occurrence of crisis events correctly in 72.1% of occasions in the t=0 time horizon. However, it has not shown significant results or

[2] Stationarity was tested by Augmented Dickey-Fuller (ADF) test, on significance level of 1%. Results have shown that within the database, neither time series is stationary.

successful estimation rates on the other time horizons examined. In the t=0 analysis, beta-coefficients are only to be claimed relevant on a significance level of 5%, therefore multiple indicators have been omitted from the final regression function due to too high significance values, such as real effective exchange rate (REER), public sector debt (PSD) and change in total financial liabilities (TFSL).

Table 3 Results of the binary logistic regression analysis

Results of logistic regression

		B	S.E.	Wald	Df	significance	Exp(B)
Step 1a	01CAB	0.093	0.029	10.430	1	0.001	1.097
	03REER	0.017	0.022	0.607	1	0.436	1.017
	04EMS	–0.015	0.008	3.945	1	0.047	0.985
	06HIP	–0.061	0.021	8.627	1	0.003	0.941
	07PSCF	–0.048	0.019	6.183	1	0.013	0.953
	08PSD	0.005	0.002	5.002	1	0.025	1.005
	09GGD	–0.005	0.005	1.354	1	0.244	0.995
	10UR	0.130	0.039	11.054	1	0.001	1.139
	11TFSL	–0.001	0.015	0.002	1	0.964	0.999
	12AR	–0.214	0.099	4.656	1	0.031	0.807
	constant	–1.090	0.552	3.897	1	0.048	0.336

Source: own calculations with SPSS v21

Results of the remaining indicators (as seen in Function 3) can be interpreted in the following way:

- current account balance (CAB) indicator's aggravation of 1 percentage point over the threshold results in the increase of probability of crisis event by 9.7%;
- private sector debt in the ratio of GDP (PSD) indicator's escalation of 1 percentage point over the threshold leads to increase crisis event's probability by 0.5%;
- unemployment rate (UR) indicator's upsurge of 1 percentage point over the threshold results in the increase of probability of crisis event by 13.9%.

Trivially, the model has statistical limitations. Change of export market share (EMS) indicators aggravation of 1 percentage point over the threshold obviously does not decrease the chance of occurrence of a crisis event in an economic sense, but from a statistical point of view it suggests that this indicator – similarly to the other indicators with beta-coefficients between 0 and 1, such as change in deflated house prices (HIP), private sector credit flow as a percentage of GDP (PSCF) and activity rate (AR) – predicts economic crises with lesser accuracy then indicators with beta coefficients over 1, at least based on the best-fitting regression model. Results can be displayed as show in Function 3. [3]

Function 3 Results of the binary logistic regression analysis

$$P\ (crisis) = 0.336 + 1.097\,\beta_{CAB} + 0.985\,\beta_{EMS} + 0.941\,\beta_{HIP} + 0.953\,\beta_{PSCF} + 1.005\,\beta_{PSD} + 1.139\,\beta_{UR} + 0.807\beta_{AR}$$

Source: own calculations with SPSS v21

6. Summary

The aim of this article has been to bring additional value to the Macroeconomic Imbalance Scorecard in order to predict and prevent future economic crises within the EU, based on historical data in the 2002–2017 time period. To achieve this, first I introduced the economic governance mechanism currently in force in the EU, i.e. the Macroeconomic Imbalances Procedures as a supplement of the Stability and Growth Pact's regulations. Since 2011, the European Commission has been monitoring the progression of 14 specific indicators in each member state in order to predict economic crises, restrain harmful cross-border effects, and propose corrective measures. This complex data set is examined and supervised by the European Commission under the European Semester. The structure of the scoreboard is not fixed in the sense that both the number of indicators and the values of the thresholds are up to debate; however, economic literature examining the effects of these indicators are not at all abundant.

The previous fact led to the urge to examine the existing statistical models with the aim to construct a new one, with the latest time series data. While trying to utilize the results of previous investigations, I came up with a new model, using binary logistic regression to estimate the probability of economic crises in multiple time horizons (t=0 year, 1 year, 2 years). Prior to that, I eliminated 4 indicators due to cross-correlation, and standardized the MIP indicator values, expressing them in proportion of the excess compared to the indicator threshold. I have managed to construct a model with a 72.1% success rate of estimation of crisis events, which highlighted that current account balance (CAB), private sector debt (PSD) and unemployment rate (UR) are the best-fitting indicators to estimate the occurrence of a crisis event (defined as a larger output gap in the current year then the EU-28 country group's average) on the t=0 time horizon (other time horizons did not bring significant results). MS-level and Commission-level economic policy-makers should be urged to focus on these

[3] Special thanks to Anita Pelle, András London and Éva Kuruczleki for their invaluable contribution.

indicators in particular; nevertheless, our results also show that other forms of quantitative analysis should be advised as well since fiscal compliance is still mostly based on an MSs' own internal commitment to the fiscal discipline rules. However, further investigations with different methodological approaches should be presented in order to gain a deeper understanding of crisis dynamics within the EU.

References

Authers, J. (2012): *Europe's Financial Crisis: A Short Guide to How the Euro Fell into Crisis and the Consequences for the World*. Pearson Education, New Jersey.

Bánfi, T. (2018): Az euro (eurozóna) a hibás konstrukció, vagy a tagországok az alkalmatlan tagok? *Hitelintézeti Szemle*, 17, 1, 2018. március, 137–152.

Blyth, M. (2013): *Austerity: History of a Dangerous Idea*. Oxford University Press, Oxford.

Bobeva, D. – Atanasov, A. (2017): Macroeconomic Imbalances in Euro- and Non-Euro Area Member States. *Godishnik na UNSS*, 2017, 1, 25–37.

Boros, E. (2017): Endogén egyensúlytalanságok egységes valutaövezetekben. *Hitelintézeti Szemle*, 16, 2, 2017. június, 86–116.

Csaba, L. (2018): Tőkepiaci unió vagy szabadságharc? *Közgazdasági Szemle*, 65, 5, 2018. május, 484–498.

Csortos, O. – Szalai, Z. (2014): Early Warning Indicators: Financial and Macroeconomic Imbalances in Central and Eastern European Countries. *MNB Working Papers*, 2014/2, 1–40.

Domonkos, T. – Ostrihoň, F. – Šikulová, I. – Širaňová, M. (2017): Analysing the Relevance of the MIP Scoreboard's Indicators. *National Institute Economic Review*, National Institute of Economic and Social Research, 239, 1, 32–52.

EC (2011A): *EU Economic Governance "Six-Pack" Enters into Force*. European Commission, Brussels.

EC (2011B): *Conclusions – 24/25 March 2011*. European Council, Brussels.

EC (2019): Excessive deficit procedures - Overview. https://ec.europa.eu/info/business-economy-euro/economic-and-fiscal-policy-coordination/eu-economic-governance-monitoring-prevention-correction/stability-and-growth-pact/corrective-arm-excessive-deficit-procedure/excessive-deficit-procedures-overview_en. Accessed: April 5, 2019.

Eurostat (2019): Your key to European Statistics. https://ec.europa.eu/eurostat/data/database Accessed: April 5, 2019.

Farkas, B. (2012): A világgazdasági válság hatása az Európai Unió régi és új kohéziós tagállamaiban. *Pénzügyi Szemle*, 57, 2012/1, 52–68.

Győrffy, D. (2013): Válságkezelés Európában, a gazdaságpolitika depolitizálásának kilátásai. *Pénzügyi Szemle*, 58, 2, 123 –135.

Halmai, P. (2015): Strukturális reformok és növekedési potenciál az Európai Unióban. *Pénzügyi Szemle*, 60, 2015/4, 520–535.

Halmai, P. (2018): Az európai növekedési modell kifulladása. *Közgazdasági Szemle*, 65, 2, 2018. február, 122–160.

Holler, J. – Reiss, L. (2011): What to Expect from the Latest Reform of the Stability and Growth Pact. *Monetary Policy & The Economy*, 11, Q4, 85–94.

Kengyel, Á. (2016): Mennyiségi lazítás és költségvetési fegyelem, avagy a strukturális reformok és a beruházások ösztönzésének dilemmái az Európai Unióban. *Közgazdasági Szemle*, 63, 6, 2016. június, 715–722.

Kertész, K. (2014): Alternatívák a monetáris politika és a Stabilitási és Növekedési Egyezmény megújítására az eurózónában. *Pénzügyi Szemle*, 59, 2014/3, 386–402.

Kiss, G. (2010): Az európai válságkezelés tapasztalatai: a gazdaságpolitikai koordináció megújítása. *MNB Szemle*, 2010. október, 40–46.

Mangone, J. M. – Laffan, B. – Schweiger, C. (2016): *Core-periphery Relations in the European Union: Power and Conflict in a Dualist Political Economy*. Routledge, New York.

Pelle, A. (2018): Az Európai Unión belüli versenyképesség-problematikáról és annak lehetséges kezeléséről. *Európai Tükör*, 2018/3, 45–65.

Pelle, A. – London, A. – Kuruczleki, É. (2019): A Dynamic Complex System of Clubs Comprised by Countries Performing a Variety of Capitalism. *Forum for Social Economics*, online, DOI: 10.1080/07360932.2019.1601121

Pisani-Ferry, J. (2014): *The Euro Crisis and its Aftermath*. Oxford University Press, Oxford.

Végh, M. Z. (2019): A 2008-as válság hozadékai az EU számára: gazdaságpolitikai koordinációból gazdasági kormányzás. *Európai Tükör*, 22, 2, 7–25.

World Bank (2019): Open Data. https://data.worldbank.org/

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 275–285.

Post-crisis trends in taxation – twilight or survival of the models?

Ábel Czékus

We examine corporate tax regulation of several Member States of the single market. The goal is to describe practices that considerably deviate from the mainstream in this regulatory field and potentially endanger the proper functioning of the single market. In particular, we focus on the conformity of the national legislation with the common competition policy providing a level playing field for doing business. Secondly, we desire to identify regulatory outputs that impede a deeper cooperation between the Member States.

In our research we lean on the comparison of national particularities; we identify practices that make jurisdictions attractive for taxation purposes, especially regarding research and development. However, we acknowledge the considerable approximation of national legislations that has taken place in recent years and highlight the role of the European Commission in this process. Finally, we consider the taxation of dividends as this is the main income type serving repatriation.

Keywords: corporate tax, harmful practices, single market

1. Introduction

Ten years have passed since the outbreak of the biggest crisis after the great recession of 1929–1933. The 2008 crisis resulted in shrinking fiscal latitude and serious social consequences throughout Europe. Member States of the European Economic Area, on the other hand, faced significant challenges in regulatory policies. Apart from the loosening of the implementation of the common competition policy, as the cornerstone of European integration, several novelties evolved in the field of corporate taxation as well. We see a dualistic approach in this: Member States struggle to screen their sovereign rights, on the one hand, and, on the other hand, more emphasis is put on European Union level initiatives enshrining the integrity of the single market. We put this duality into the focus of our research.

We concede the importance of corporate taxation in developed economies as this is one the traditional sources of national budgets. During the golden age of corporate taxation – in the sixties and seventies of the twentieth century – higher than 40% tax rates were common in Europe. During these decades corporate tax was mainly considered a revenue source; as a side note we recall that this period overlaps with the evolution of the welfare state. However, globalisation has brought considerable change in corporate taxation as well. Liberalisation of international trade, increased mobility of capital, the development of financial intermediaries, and the emergence of developing economies all contributed to the alteration of taxation in the European region. As a result, the share of corporate tax in budgetary revenues today is not crucial in developed economies, usually representing approximately 10% of total budgetary revenues. Considering this, we argue that currently corporate taxation rather serves socio-economic goals: it may actuate research and innovation, back the organic development of a sector, contribute to employment and the creation of new workplaces, and, in a broader sense, can increase the attractiveness of the whole economy. In open economies like Member States of the European single market are, the latter is of great importance, although meanwhile we admit the pure fiscal implication of the tax in several economies.

In the confederation of Member States the proper functioning of the single market ensures the integrity of the economic union. In the reading of taxation this means that tax measures might not divert the flow of investments. In other words, states have to refrain from applying tax incentives that artificially separate the place of economic activity and payment of taxes. Tax planning echoes a global phenomenon and reflects “the importance of cross-country spillovers in analyzing corporate tax reform” (IMF 2019, introduction) since allocation of taxation rights, base erosion, and profit shifting, the race to the bottom still create considerable tension on national tax systems. Highlighting these thoughts, we survey those measures that potentially endanger the above-mentioned considerations or may lead to significant deviation from the general patterns. In the paper we encompass the main characteristics of the Irish, Swiss, Dutch, and Luxembourg regulation.

The structure of the paper is as follows. In the first section we give a broad overview of corporate taxation; we review its history, main characteristics, and objectives. In the second section we provide a regulatory-scientific description of the above-mentioned jurisdictions, including the taxation of dividends.

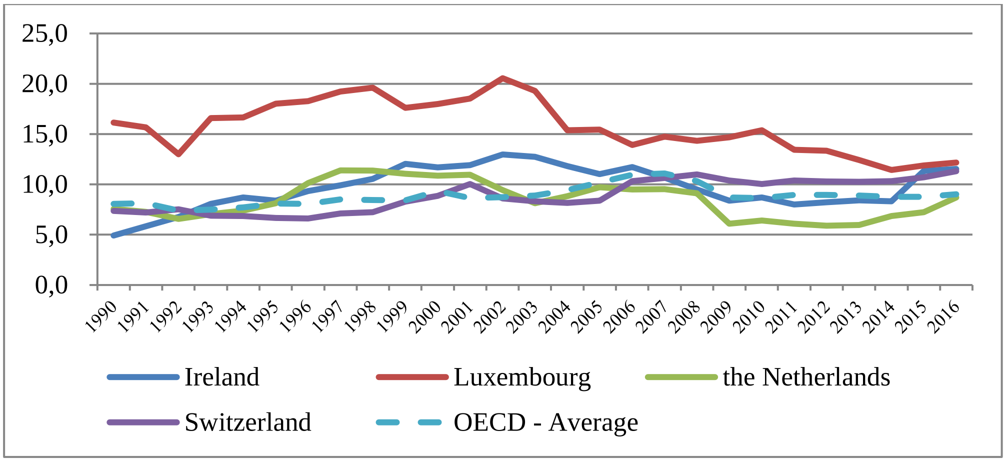
2. Corporate taxation – an overview

According to Bardopoulos (2012) the history of taxation goes back to ancient times with considerable regulatory improvement of the ancient Egyptians and Greeks. Taxes of the initial times served strategic and building goals. Later on, in the nineteenth century, the fundaments of modern income taxation were laid down in the United Kingdom, later on in the United States and in the rest of the developed world. The main novelty of these taxes was the taxation of income realised by the persons in the given jurisdiction i.e. it is being imposed on income-creating business activity.

Although corporate tax rates have decreased significantly during the last decades its share within the total revenues has not fallen significantly over the last two decades. According to the Organisation for Economic Cooperation and Development (OECD) the corporate tax/gross domestic product ratio remained below 3%, however, trends in the post-1990’s figures are hard to extrapolate. For example, observing again the most developed countries, the share increased from a stable two-and-a-half percentage value but in the years of the millennium the indicator even passed 3%. The total tax burden, however, shows a diverse picture: it is higher than the OECD average in the Netherlands and Luxembourg, but lower in Ireland and Switzerland (OECD 2018).

A more explanatory indicator, the share of corporate tax within the total tax revenues shows a slightly detailed picture. The share of corporate tax in the most developed countries comes in at only 9% while in developing countries it exceeds 15% in 2016, for instance. However, no organic development can be seen in this indicator (Figure 1). In the case of the economies studied, we rather see a convergence in corporate tax revenues than a decrease; it is also worth mentioning that the average share of corporate tax revenues in the surveyed countries has risen by 1.9 percentage points. Mainly the high Luxembourg value contributed to this effect, but Ireland and Switzerland produced an above-OECD-average outcome, too.

Figure 1 Share of corporate tax within the total tax revenues in selected economies, 1990–2016 (%)



Source: OECD (2019)

The increase in corporate tax within the budgetary revenues, however, is not followed from an increase in tax rates, as the combined corporate tax rate[1] has considerably lowered in Member States surveyed (Table 1). In Ireland, for example, the rate was almost halved in less than two decades, and in the Netherlands lowered by 10%. The increase in total corporate tax revenues therefore can only be attributed to a rise economic activity. From the data available, however, it cannot be deducted whether this is due to growing domestic business output, profit shifting or the broadening tax base (Hines 2005).

[1] We agree that corporate taxation is predominantly related to central budgets. According to Blöchliger and Petzold (2009) tax share of sub-central governments was around 17 percent, half of the resources spent. This increased the dependence of sub-central expenditures on central budgets.

Table 1 Statutory tax rate change in surveyed countries (%)

	2000	2008	2016
Ireland	24	12,5	12,5
Luxembourg	37,5	29,6	29,2
the Netherlands	35	25,5	25
Switzerland	24,9	21,2	21,1

Source: OECD (2019)

Avi-Yonah (2005) identifies three goals of taxation. It includes the fiscal implication of tax collection, the redistributive function, and the shaping of entrepreneurial economic activity. In the economic literature there is a deep debate on the nature of taxes and their impact on business activity. Devereux–Sørensen (2006) argue that the aim of corporate taxation is to gather all income generating activities under the taxable incomes, as solely taxing private individuals' income could lead to non-taxation of several items. Authors also admit that enterprises contribute by this means to common charges, including benefits to infrastructure and human resources. Nowadays fiscal impetus on research, development, and innovation can be considered the most important accessory goal of the corporate tax system. This usually takes place in the form of tax allowance, i.e. a bunch of expenses interrelated to R+D+I can be deducted against an income tax base. We do concede there are further accessory considerations served by a well designed corporate tax system (employment, social development, environmental aspects). From these items, we are mainly dealing with allowances relating to the development of a knowledge-based economy.

We also note the importance of corporate social responsibility (KPMG 2016) in corporate taxation, even if this is generally not embedded into statutory requisites. This feature usually represents a code of conduct on voluntary disclosure of taxation trends, but, broadly speaking, it represents all the countenance an enterpriseplaces outside the framework of legislative compliance.

3. Corporate tax regimes of Member States studied

In this section we overview specialities of the corporate tax regimes of Ireland, Luxembourg, the Netherlands, and Switzerland. We highlight characteristics of the national legislation with the aim of giving a broader insight. As a general rule the single market Member States[2] studied apply principles in corporate taxation consistent with the Model Tax Convention on Income and on Capital issued by the Organisation for Economic Cooperation and Development. Therefore, besides featuring national tax legislation we return to the provisions of this model and Member States' relating to international harmonization. Finally, we examine dividend distribution rules of the Member States mentioned.

[2]We acknowledge that the Swiss Confederation is not part of the single market but is deeply integrated into it through bilateral agreements.

3.1. Overview of corporate tax incentives

We assert that the economies researched impose corporate tax on resident companies' worldwide income. The worldwide income encompasses incomes of entities doing business abroad through permanent establishment or branch, in line with the generally accepted international income taxation principles. Furthermore, companies having a place of management or control in the given jurisdiction are also considered to be resident for corporate tax purposes.

These provisions are, however, subject to double tax conventions, where they exist, deductions or credits granted by jurisdictions. Deductibility reveals the maturity of the tax system and reflects legislative notions regarding economic development. The states studied provide a wide range of deductible costs and expenses against the tax base; these might serve development and financial goals (e.g. start-ups, R+D, royalty payments, and interest, respectively). States also offer taxpayers the possibility to credit taxes, mainly in the field of R+D and investments (PwC 2018). Tax incentives might considerably decrease enterprises' tax liability.

As referred to in the previous chapter we see in numerous jurisdictions that deductions can be made against the corporate tax base with the clear aim of supporting research and development. This instrument is obviously shaped for multinational companies appropriating considerable expense on R+D, however, it can even lead to tax planning. In the Netherlands, for example, an innovation box was created in which, after the approval of the competent authority, 7% tax rate applies to qualified intellectual assets. This advantage is available regarding only non-marketing assets and for groups of companies with revenues over EUR 250 million, and certain further conditions have to be met. On the other hand, a wide range of incomes is considered as qualified and enterprises are allowed to further 30% surplus enforcement. Furthermore, a separate R+D incentive (credit) exists in the Netherlands whereby nearly half of the eligible expenses can be deducted. In Switzerland there is no downright R+D corporate tax incentive, partly because of the confederal state system. Generally, tax incentives are available for activities that support steady economic and social development. Research and innovation is backed by tax incentives in Ireland as well. Thanks to the Knowledge Development Box, an interim tool for supporting knowledge-intensive business activity, eligible companies may deduct certain expenses up to an effective tax rate of 6.25%. Software-related expenses count as qualified together with other high-tech innovation costs (EY 2018). Furthermore, a tax credit is also available in Ireland, whereby one quarter of the research and development expenditures can be credited. In practice, the 25% is accumulated with the normal corporate income tax rate, meaning 37.5% allowance can be realised. Luxembourg made considerable changes to its relevant intellectual property legislation, harmonising it with the latest OECD and EU standards on nexus approach. In fact, patents and software are eligible for the exemption. Thanks to this, the place of effective taxation corresponds to the taxable activity, an 80% exemption is applicable on qualified incomes and taxing at a rate of 5.2%. It is worth mentioning that other support (loans) is available for enterprises engaged in research and development up to the entire amount (PwC 2018).

We argue that country-specific tax deductions in some cases result in considerable tax relief (Table 2). For instance, in Switzerland 8.5% central governmental tax rate is applicable and an ambulatory local rate imposed by the cantons. The two level combined average tax burden therefore amounts to 22.9%, meanwhile the effective tax rate is 3.4 percentage points lower. Similarly, nearly 10% difference is seen in the innovation friendly Netherlands. While it may be true in the case of several Member States of the European Economic Area that there is an even bigger edge between the statutory and effective corporate income tax rates, these fall outside the scope of the current study.

Table 2 Corporate income tax rates in surveyed countries, 2017 (%)

Tax rate	Central governmental, statutory	Regional/municipal, statutory	Effective tax rate
Ireland	12.5	-	11.8
Luxembourg	20.3	6.8	24.5
the Netherlands	25	-	23
Switzerland	8.5	14.4	19.5

Source: OECD (2019)

The difference between the statutory and the effective tax rates implicates further alleviation for taxpayers. One of the leading Member States is Luxembourg, its magnetism originating in its investment funds. Investment funds, for example, are subject to relatively loose conditions on setup and are exempt of corporate tax, on the one hand, and, on the other hand, some investments (private wealth management enterprises – *Société de gestion du patrimoine familial*) may benefit from extraordinary tax relief (close to zero percent), including exemption from withholding tax as well (EY 2018). Luxembourg also operates venture capital instruments (*Société d'investissementen capital à risques*), targeted to foreign residents, exempting mainly passive incomes and capital gains (PwC 2018).

The tax treatment of funds is, however, not as beneficial in Ireland as it is in Luxembourg (EY 2018). Newly (in 2017) introduced legislation is intended to overcome the non-taxation of real estate funds by burdening the payments with 40% withholding tax. Similarly, offshore funds are taxable as well, subject to several conditions (Member States of the European Economic Area, tax convention in force, etc).

Regarding corporate tax (and, in a broader sense, direct taxes) there has been no far-reaching harmonisation at a supranational level and corporate taxation to some extent embodies the sovereignty of Member States. In the early years of the millennium, the European Union initiated a legislative process with the clear aim of tackling harmful tax competition within the common market, admitting the "need for coordinated action at European level to tackle harmful tax competition in order to help achieve certain objectives such as reducing the continuing distortions in the single market" (CoC 1997). In the framework of the Code of Conduct, Member States committed themselves to cutting back harmful tax practices and refraining from introducing such measures. We also admit the importance of the Anti Tax Avoidance Directive (adopted in 2016) which "creates a minimum level of protection against corporate tax avoidance throughout the EU" (EC 2019a). The Directive is intended to address behaviours that lead to non- or under taxation in the single market and prescribed measures on controlled foreign companies, new switchover rules, rules on exit taxation and interest limitation, and a comprehensive anti-abuse rule. This set of measures serves national budgetary interests, on the one hand, and, on the other hand, the European Commission – enshrining the uniformity of the single market – intends to preclude the adoption of diverse national acts in this area.

3.2. International interaction of corporate tax systems

According to the recommendation of the OECD, Model Tax Convention states are encouraged to dissolve juridical double taxation, both of income and capital. The Model encompasses direct taxes at all governmental levels; this providing a channel for undistorted business activity. The concept of business activity is defined by domestic law (Rust 2011).

As we referred to in the previous chapter, states impose taxes on enterprises' worldwide income. The motivation behind this is to prevent non-taxation of an income earned by a resident. Tax conventions, on the other hand, share taxing rights of incomes, i.e. align corporate tax (and personal income tax) systems of the contracting parties with the clear aim of avoiding double taxation. Although in the single market, harmonisation is far advanced in the area of indirect taxes, direct taxation has remained a national competence therefore tax conventions settle double taxation between European single market Member States as well.

Two methods of tax reliefs exist for the avoidance of double taxation, both being observed in various EEA tax conventions. The exemption method disregards incomes that are taxed in the partner jurisdiction, but they might be considered in the calculation of other obligations (see progressive exemption). Passive incomes (dividends, royalties, interests) might be excluded from relief and taxed according to the provisions of a tax credit.

Crediting is a rifer method of eliminating double taxation. Thanks to this kind of relief resident taxpayers may credit taxes already paid in the partner jurisdiction. This method serves better contracting parties' budgetary interests, however, poses an administrative burden.

3.3. Taxation of dividends

In this subsection we describe the main characteristics of the dividend taxation in Member States studied. Enterprises repatriate profits by this income type, and therefore its taxation also affects the single market. We feature national acts and the Parent-Subsidiary Directive of the European Union.

According to the legislative practice dividends are taxable in the country where the recipient holds residence for tax purposes. This main rule, however, does not prevent states from taxing dividend incomes in the contracting state where the payee holds residence for tax purposes, i.e. from imposing withholding tax. In Luxembourg, for example, dividends are taxed at the normal rate of corporate tax but exemption is available. A criterion for this is that the dividend payer is subject to the Parent-Subsidiary Directive (other eligibility criteria based on the domestic law widens the scope); while on the payee side there is also a widespread criteria, however in brief, we note that local branches of companies resident in one of the Member States of the European Economic Area Member meet the prescriptions for exemption. In the Republic of Ireland, dividends from other contractual jurisdictions are also taxed on the normal corporate tax rate (currently 12.5%) provided the beneficiary owns at least 5% of the capital in the dividend payer entity. Inland dividend payments are tax free. If, however, the dividend is paid by an Irish company in which less than five persons exert more than 35% of the controlling rights, surtax is payable in some elements of the incomes (PwC 2018).

In the Netherlands dividends are also tax free if the participation reaches at least 5% in the distributor of the dividends and the share does not qualify for portfolio investment. Furthermore, as a main rule there is no withholding tax on dividends paid, subject to notification to the competent authority. In the Swiss Confederation, dividends are generally taxable according to the normal rates of corporate tax but if the beneficial owner of the dividends received holds at least 10% of the shares, it may be eligible for tax exemption. On the other hand, dividend distributions are subject to a 35% withholding tax (PwC 2018, EY 2018).

Table 3 Contractual withholding tax rates between Member States studied, (%)

	Ireland	Luxembourg	the Netherlands	Switzerland
Ireland	0	15; 0/5	0/15	0
Luxembourg	15; 0/5	15/0	2.5/15	15; 0/5
the Netherlands	0/15	2.5/15	0/15	0/15
Switzerland	0	15; 0/5	0/15	35

Source: own research based on PwC (2018)

Note: 1. for Luxembourg the first figure stands for portfolio investments, the second for substantial holdings. 2. between the Member States of the European Union no withholding tax is applicable.

Contrary to corporate income in the case of dividends (and other types of passive incomes) contracting states may collect withholding taxes. It results in a two-way interaction between jurisdictions: the state of the entity distributing dividends is entitled as well to tax the income. This kind of division of taxation rights may lead to friction between the contracting parties, leading the OECD to elaborate recommendations in the Model Tax Convention. Article 10 of the Model as a main rule assigns taxation rights to the state where the recipient is resident for tax purposes. However, the recommendation states that jurisdiction where the distributor is resident may also tax these incomes but at no more than 5% if the beneficial owner (parent company) holds at least 25% of the shares in the distributor. The aim of this limitation is to "avoid recurrent taxation and to facilitate international investment" (OECD 2014).

As seen from the previous paragraphs undistributed taxation rights may lead to legaldisputes between Member States, and in parallel to double taxation of dividend payments. For this reason – and to reduce the distortion of the single market – in the 1990's the European Commission took the initiative and proposed the Parent-Subsidiary Directive (PSD). The Directive's goal is twofold: to eliminate double taxation of the income, on the one hand, and, on the other hand, to cease withholding taxation on dividend payment (EY 2009). In the 2003 amendment of the Directive, considerable amendments were included whereby the scope of covered entities was extended (EC 2019b). Among other legal forms European companies, the newly established institutions were covered by the Parent-Subsidiary Directive. The amendment resulted in a loosening of conditions relating to shares the parents needed to hold in the subsidiary for being eligible for the benefits of the Directive. Currently 10% is the shareholding minimum. The third novelty served the avoidance of double taxation in the case of chain of companies paying dividends by imputing taxes paid by the successive entity. Ashta (2006), however, lists the issues where further legislation would be welcome but these concernmainly national mismatches in tax legislation.

4. Conclusion

The scope of our paper was to show some of the practices that represent special treatment for resident companies. We have given an overview of corporate taxation and dividend withholding taxation, respectively.

We targeted the notion of the single market. Although Switzerland is not a Member State of the European Union, it does participate in the single market and therefore its legislation also affects the flow of capital. An emphasis was put on innovation regimes as this is the bottleneck of corporate tax incentives. We showed tax treatment of the Dutch innovation box, the Irish Knowledge Development Box and the Luxembourger intellectual property regime. With a general overview of the Parent-Subsidiary Directive, it has been shown that thanks to gradual reform, withholding tax on dividends have been abolished and steps taken towards further elimination of double taxation. We consider these acts as developments towards a less fragmented single market.

Summarising, we conclude that there have been considerable attempts on the European Union level to tackle harmful tax competition possibly posed by Member State taxation practices. One of the most prominent examples of this is the Code of Conduct whereby major distortion of the single market has been staved off.

Acknowledgments

SUPPORTED BY THE UNKP-18-3 NEW NATIONAL EXCELLENCE PROGRAM OF THE MINISTRY OF HUMAN CAPACITIES

References

Ashta, A. (2006): The Parent Subsidiary Directive. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=940967 Accessed 12/03/2019

Avi-Yonah, R. S. (2005): The Three Goals of Taxation. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=796776. Accessed 05/03/2019

Bardopoulos, A. M. (2012): eCommerce and the Effects of Technology on Taxation. *Law, Governance and Technology Series* 22, DOI 10.1007/978-3-319-15449-7_2. https://www.springer.com/cda/content/document/cda_downloaddocument/9783319154480-c2.pdf?SGWID=0-0-45-1500580-p177235827. Accessed 04/03/2019

Blöchliger, H. – Petzold, O. (2009): Taxes and grants: on the revenue mix of sub-central governments. Organisation for Economic Cooperation and Development. https://www.oecd.org/tax/federalism/42783028.pdf. Accessed 27/07/2019

CoC (1997): Conclusions of the ECOFIN Council Meeting on 1 December 1997 concerning taxation policy - Resolution of the Council and the Representatives of the Governments of the Member States, meeting within the Council of 1 December 1997 on a code of conduct for business taxation - Taxation of saving. *Official Journal of the European Communities*, 98/C 2/01.

Devereux, M. P. – Sørensen, P. B. (2006): The corporate income tax: international trends and options for fundamental reform. Commission of the European Communities Directorate-General on Economic Affairs Workshop on Corporation Income Tax Competition and Coordination in the European Union, held in Brussels September 25, 2006. Accessed 05/03/2019

EC (2019a): The Anti Tax Avoidance Directive. https://ec.europa.eu/taxation_customs/business/company-tax/anti-tax-avoidance-package/anti-tax-avoidance-directive_en. Accessed 30/04/2019

EC (2019b): Parent companies and their subsidiaries in the European Union. https://ec.europa.eu/taxation_customs/business/company-tax/parent-companies-their-subsidiaries-eu-union_en. Accessed 12/03/2019

EY (2009): Implementation of the amended Parent-Subsidiary Directive. A study covering the 27 European Union Member States. https://www.ey.com/Publication/vwLUAssets/EY_study_on_the_PSD_implementation_in_the_EU/$FILE/EY-PSD-implementation.pdf. Accessed 12/03/2019

EY (2018): Worldwide Corporate Tax Guide. Ernst and Young.

Hines, J. R. Jr. (2005): Ross School of Business Paper No. 1026. Available at SSRN:https://ssrn.com/abstract=891233. Accessed 05/03/2019

IMF (2019): IMFPolicy Paper. Corporate Taxation in the Global Economy. International Monetary Fund.https://www.imf.org/en/Publications/Policy-Papers/Issues/2019/03/08/Corporate-Taxation-in-the-Global-Economy-46650. Accessed: 19/07/2019

KPMG (2016): A look inside tax departments worldwide and how they are evolving. Summary report: Global Tax Benchmarking Survey 2016.

OECD (2014): Model Tax Convention on Income and on Capital. Condensed Version. Organisation for Economic Cooperation and Development, Paris.

OECD (2018): OECD Country Reports. Organisation for Economic Cooperation and Development . https://www.oecd.org.Accessed 27/07/2019

OECD (2019): Corporate Tax Statistics Database. Organisation for Economic Cooperation and Development. http://www.oecd.org/tax/beps/corporate-tax-statistics-database.htm. Accessed 04/03/2019

PwC (2018): Worldwide Tax Summaries. Corporate and individual tax information at your fingertips. PriceWaterhouseCoopers.

Rust, A. (2011): "Business" and "Business Profits". In: Guglielmo Maisto (ed.): *The Meaning of "Enterprise", "Business" and "Business Profits" under Tax Treaties and EU tax law*, 85–106.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 286–302.

Inflation Forecasting in Developing Economies Using SARMA Models: The Case of Ghana

Senanu Kwasi Klutse

Due in part to a weakening link between monetary aggregates and the inflation variable, monetary policy authorities are specifically targeting the price level. This has made it more necessary for inflation forecasting to be considered in the conducting of monetary policy. This paper presents different approaches to forecasting inflation in developing economies using Seasonal Autoregressive Moving Average (SARMA) models. SARMA models were employed because they capture seasonal components of the inflation variable which other univariate (ARMA) models cannot effectively capture.

The analysis showed that the complex approach is good for forecasting inflation in both the short-term (1 year) and the very short-term (2 months). The results further showed that the models can capture policy changes only if they occur in the in-sample period. This feature makes it somewhat suitable for policy authorities who will want to know the direction of the inflation variable after policy decisions have been made, and can help them make future policy decisions.

Keywords: Inflation, Seasonality, Forecasting

1. Introduction

There is a widely held belief among macroeconomists that there is a long-run relationship between the growth rate of money supply and the growth rate of prices (inflation). This belief forms the basis of monetary policy making in most central banks and hence its extraordinary importance for the conduct of public policy. Its importance makes it one of the most commonly tested hypotheses in economics.

Linkages between money and inflation became increasingly important during the recent financial crisis. The long-run relationship between money and inflation is almost surely not linear, and the short-run dynamics may disguise the long-run relationship, confusing tests for this relationship. As these variables involve interaction between various economic variables, it raises the possibility that the correspondence between them may be both non-linear and time varying (Binner et al. 2010).

If indeed there is a dynamic, long run relationship between money supply and increase in prices, then it is a reasonable proposition that the near-term growth of money supply might have a predictive power for inflation. Various studies have explored this relationship. The argument here is that deregulation, financial innovation and other factors have led to recurrent instability in the relationship between various monetary aggregates and other nominal variables – inflation (Binner et al. 2010). With the adoption of the Inflation Targeting (IT) framework, more countries have shifted towards targeting the price level specifically in conducting monetary policy. Some economists have however expressed notes of caution on the importance of money, stating that money will regain an important place in the conversation of economists (Binner et al. 2010).

In order to target the price level effectively, there is the need to properly manage inflation expectations, which includes staying within the inflation target that has been set. This can be achieved in part by effectively forecasting inflation to help monetary authorities set policy rates to guide current inflation to its intended levels.

Inflation forecasting is a very important input in monetary policy making, even though others do not consider forecasting models as a useful guide for monetary policy (Fisher et al. 2002). Apart from providing an input to monetary policy deliberations, inflation forecasts also play a role in the macroeconomic policy debate. By informing the public about likely trends, inflation forecasts can influence expectations and can therefore serve as a nominal anchor in the wage-bargaining process and nominal fixed contracts like rents or interest rates (Moser et al. 2007).

This study adds to the various inflation forecasting models that exist, by taking note of the seasonality that exists in the inflation variable in developing economies, using Ghana as a case study. Since the target variable is the inflation variable, this study adopts the use of univariate models. This was in part achieved by adding a Seasonal Autoregressive (SAR) and a Seasonal Moving Average (SMA) term to produce higher order Autoregressive Moving Average (ARMA) models with non-linear restrictions. This model was then used to forecast inflation using three approaches: a simple (naïve) approach, an intermediate approach and a complex approach. As a control experiment, a normal ARMA model was also used to forecast inflation using the same approaches mentioned earlier.

The analysis showed that the SARMA model had a superior forecasting ability than the normal ARMA model. In the case of the SARMA model, the complex approach provided the best forecast ability, as a combination of the food and non-food component of inflation in Ghana provided the best out of sample result. For the ARMA models the intermediate approach, which uses the forecast values of the overall CPI, provided the best forecast ability. This conclusion was reached by first looking at various pieces of literature on inflation forecasting. Secondly, we delve into the methods of analysis in this study and finally analysed the data using the proposed methods – SARMA and ARMA forecasting.

2. Literature Review

A simple Phillips curve, which uses a single measure of economic slack such as unemployment to predict future inflation, is probably the most common econometric basis of inflation forecasting. The usefulness of the Phillips curve as a means of predicting inflation has, however, been questioned by several authors.

Focusing on the one-year-ahead forecast horizon, Atkeson and Oharian (2001), argue that unemployment – based Phillips curve models and generalized Phillips curve models can do no better than a naïve model which says that inflation over the coming year is expected to be the same as inflation over the past year (Fisher et al. 2002). Cecchetti et al. (2000) considered inflation prediction with individual indicators, including unemployment, and argue that none of these gives reliable inflation forecasts. Stock and Watson (2003, 2004) considered prediction of inflation in each of the G7 countries using a large number of possible models. Each model had a single predictor (plus lagged inflation). They found that most of the models they considered give larger out-of-sample root mean square prediction error than a simple time series forecast based on fitting an auto regression to the inflation variable.

In recent years, researchers have, however, made substantial progress in forecasting inflation using large datasets (i.e., a large number of predictive variables), but where the information in these different variables is combined in a judicious way that avoids the estimation of a large number of unrestricted parameters (Wright, 2009). For instance, Fisher et al. (2002), in accessing 'when we can forecast inflation', focused on the ability to forecast the magnitude of inflation in the CPI, CPI less food and energy component (Core CPI), and the Personal Consumption Expenditures (PCE) deflator over the 1985 to 2000 sample period. They found that the forecasting model based on core PCE, improve forecasting significantly relative to the naïve models (simple univariate) in the 1993–2000 period. However, periods of low inflation volatility and periods after regime shifts favour the naïve model. The relatively poor performance of the Philips curve model reflects its inability to forecast the magnitude of inflation accurately.

Some studies used multivariate models and compared their predictive powers to determine which was better in forecasting inflation (Stock and Watson 1999, 2001). Hassani et al. (2018) used this approach to compare professional forecasts to academic forecasts. They found that professional forecasts are good at short-term forecasts whereas academic forecasts are good at long-term forecasts. They went further to even determine causality between the two. Stock and Watson (2016) had discovered earlier that multivariate estimates of trend inflation are similar to the univariate estimates of trend inflation. They computed trend inflation using core inflation, i.e. inflation excluding food inflation and energy.

Moser et al. (2007) applied factor models proposed by Stock and Watson as well as Vector Auto-Regression (VAR) and Auto-Regression Integrated Moving Average (ARIMA) models to generate 12-month out of sample forecasts of Austrian HICP inflation and its sub-indices. According to them factor models possess the highest predictive accuracy for several sub-indices, and predictive accuracy can be further improved by combining the information contained in factor and VAR models for some indices. They favoured the aggregation of sub-indices forecasts over a forecast of headline inflation itself.

Other studies have also tried to forecast inflation in developing economies. Mohammed et al. (2015), in studying the efficacy of the inflation variable in Nigeria, favoured the use of neural networks to forecast inflation just like Binner et al. (2010). Others also forecasted the inflation variable by using ARIMA, Seasonal Auto-Regression Integrated Moving Average (SARIMA) and Vector Error Correction

Model (VECM), and since the SARIMA model performed better, they concluded that the inflation variable has some level of seasonality.

Some studies also used models new to macroeconomics in forecasting inflation. Binner et al. (2010) used two non-linear techniques, namely recurrent neural networks and Kernel recursive least square regression – techniques that are new to macroeconomics. They then compared the two models to forecasts from naive random walk model. The best models according to them were the non-linear autoregressive models based on kernel methods. Balcilar et al. (2017) also used Vector Autoregressive Fractionally Integrated Moving Average (VARFIMA) and compared it to a standard ARIMA and VAR model. They found that their model outperformed these other models used in inflation forecasting. Other models such as the Moment Estimation Through Aggregation (META) where found to compare favourably with alternative univariate and multivariate models as well as those by professional forecasters (Sbrana et al. 2017).

The conclusion here is that the best predictive performance is obtained by constructing forecasts from a very large number of models and simply averaging these forecast values (Stock and Watson, 2003, 2004, 2008, 2010, Fisher et al. 2002). This gives the best predictive performance of inflation and that it is remarkably consistent across sub periods and across countries. Stock and Watson (2003, 2004) explored other methods for pooling the different forecasts, but found that none does better than simply averaging them, i.e., giving them all equal weights.

The studies above have varying conclusions on whether univariate or multivariate models are superior in terms of predicting the inflation variable. None of the studies above considered seasonality in the inflation variable in their forecasting models, which is a very important characteristic of inflation in developing countries. Those that came close to studying seasonality in the inflation variable – as mentioned above – only used the SARIMA model to test for seasonality based on its forecast performance. Developing countries like Ghana, are predominantly agrarian in nature and rely significantly on rainfall for irrigation. These factors – agriculture and rainfall – are seasonal in nature. As a result, this study will take into consideration the above-mentioned forecasting models to determine the appropriate model for forecasting inflation in developing countries. The focus will be on SARMA models and the aggregation of sub-indices to forecast headline inflation in developing countries (Ghana) considering also the seasonality in the component of the inflation variable.

3. Methodology

If we are in a stable monetary regime and expect the regime to persist, then it makes sense for policy makers to pay attention to inflation forecasting. This study will attempt to forecast inflation using three different methods – a method that focuses on headline inflation (simple model), a method that focuses on the overall (combined) CPI to forecast headline inflation (intermediate model) and a method that will use the CPI of food and non-food inflation to determine a forecast for headline inflation (complex model). The Seasonal Auto Regressive Moving Average (SARMA) model will be used in all cases. As a control, this study will use an Auto Regressive Moving Average (ARMA) model which does not include seasonal components to test the efficiency of the proposed SARMA model.

The simple model will forecast year-on-year headline inflation using a Seasonal Autoregressive Moving Average (SARMA) model. The SARMA model is based on a series of past behaviours only. This model is able to capture rich dynamics, both seasonal and non-seasonal. The forecast model is shown in equation (1) below.

$$f_{t,s} = \sum_{i=1}^{p} a_i f_{t,s-i} + \sum_{j=1}^{q} b_j u_{t+s-j} \quad (1)$$

where $f_{t,s} = Y_{t+s}, s \leq 0; u_{t+s} = 0, s > 0 = u_{t+s}, s \leq 0$ and a_i and b_j are the autoregressive and moving average coefficients respectively. s is the number of steps ahead. For equation (1) above, let $f_{t,s}$ denote the forecast variable and u_{t+s} denote the error term – the AR and the MA process.

Box et. al. (2015) recommended the use of seasonal autoregressive (SAR) and seasonal moving average (SMA) terms for monthly or quarterly data with systematic seasonal movement. Processes with SAR and SMA terms are ARMA models constructed using products of lag polynomials. These products produce higher order ARMA models with nonlinear restrictions on the coefficients. Both the SAR and the SMA are not intended to be used alone.

Since the appropriate reaction of monetary policy to inflationary pressures depends, among other things, on the sources of inflation, it is useful to monitor, analyse, and forecast sub-indices of headline inflation which are defined at the level of product types contained in the CPI (Moser et al. 2007).

The intermediate model will first forecast the CPI of overall inflation (LNCPI_O) and then derive the Year-on-Year Inflation using the equation below:

$$LNINF\_YOY_t = [(CPI\_O_t | CPI\_O_{t-12}) - 1] * 100 \quad (2)$$

Where $LNYOY\_INF$ is the Year-on-Year headline inflation; CPI_O_{t-12} and CPI_O_t are the CPI of overall inflation at time $t - 12$ and t.

The complex model will forecast the CPI of both food and non-food inflation using the SARMA model, and then using their respective weights in the consumer basket, the year-on-year headline inflation will be derived.

$$LNINF\_YOY_t = (1 + r_f) * w_f + (1 + r_{nf}) * w_{nf} \quad (3)$$

Where r_f and r_{nf} are the derived year-on-year food and non-food inflation rates; w_f and w_{nf} are the weights of food and non-food inflation in Ghana's inflation basket, which is the previous month's share of food (43.89959) and non-food (56.10041) CPIs.

The three models will then be compared to each other to see which model better forecasts headline inflation in Ghana. Statistical tests of a model are commonly conducted by splitting a given data set into an in-sample period, used for the initial parameter estimation and model selection, and out-of-sample period, usually used to evaluate forecasting performance. For this study, the in-sample period is from 2012M01 to 2016M02 and the out-of-sample period is from 2016M03 to 2017M02. However, the sample period may be adjusted in order to select the best-forecast model for each variable.

Empirical evidence based on out-of-sample forecast performance is generally considered trust worthier than evidence based on in-sample performance, which can be more sensitive to outliers and data mining. Out-of-sample forecasts also better reflect the information available to the forecaster in real time (Hansen and Timmermann 2012). Forecasters generally agree that forecasting methods should be assessed for accuracy using out-of-sample tests rather than goodness of fit to past data (in-sample tests) (Tashman 2000).

The argument here is that for a given forecasting method, it is possible for one to understate forecasting errors. Method selection and estimation are designed to calibrate a forecasting procedure to its historical data. But the nuances of past history are unlikely to persist into the future, and the nuances of the future may not have revealed themselves in the past. The other variance to this argument is that methods selected by best in-sample fit may not be good at predicting post-sample data. Bartolomei and Sweet (1989) and Pant and Starbuck (1990) provided more convincing evidence on this argument (Tashman 2000).

For this study, the fit period is used to identify and estimate the models while the test period is reserved to assess the model's forecasting accuracy. By withholding all data about events occurring after the end of the fit period, the forecast-accuracy evaluation is structurally identical to the real-world-forecasting environment, in which we stand in the present and forecast the future. However, one was cautious at 'peeking' at the data while selecting the forecasting method since this pollutes the evaluation environment.

As discussed in the literature review, Kernel models have also shown great promise in financial forecasting. However, they typically scale rather unfavourably with the number of training examples, thus pose a degree of freedom problem (Binner et al, 2010). Also based on the argument that the correspondence between the various components of the inflation variable may be both non-linear and time varying (Binner et al. 2010), this study will not use VAR models. This is regardless of the fact that VAR models are major tools for investigating linear relationships between small groups of variables (Duarte–Rua, 2007). As a result, this study will stick to the use of SARMA models to forecast the inflation variables using ARMA models as a control experiment.

4. Data Analysis

Data used for this study was obtained from the website of the Ghana Statistical Service (GSS). The Consumer Price Index (CPI) according to the GSS measures changes over time in the general price level of goods and services that households acquire for the purpose of consumption, with reference to the price level in 2012, the base year, which has an index of 100. The data spans from January 2012 to February 2017 (2012M01 to 2017M02).

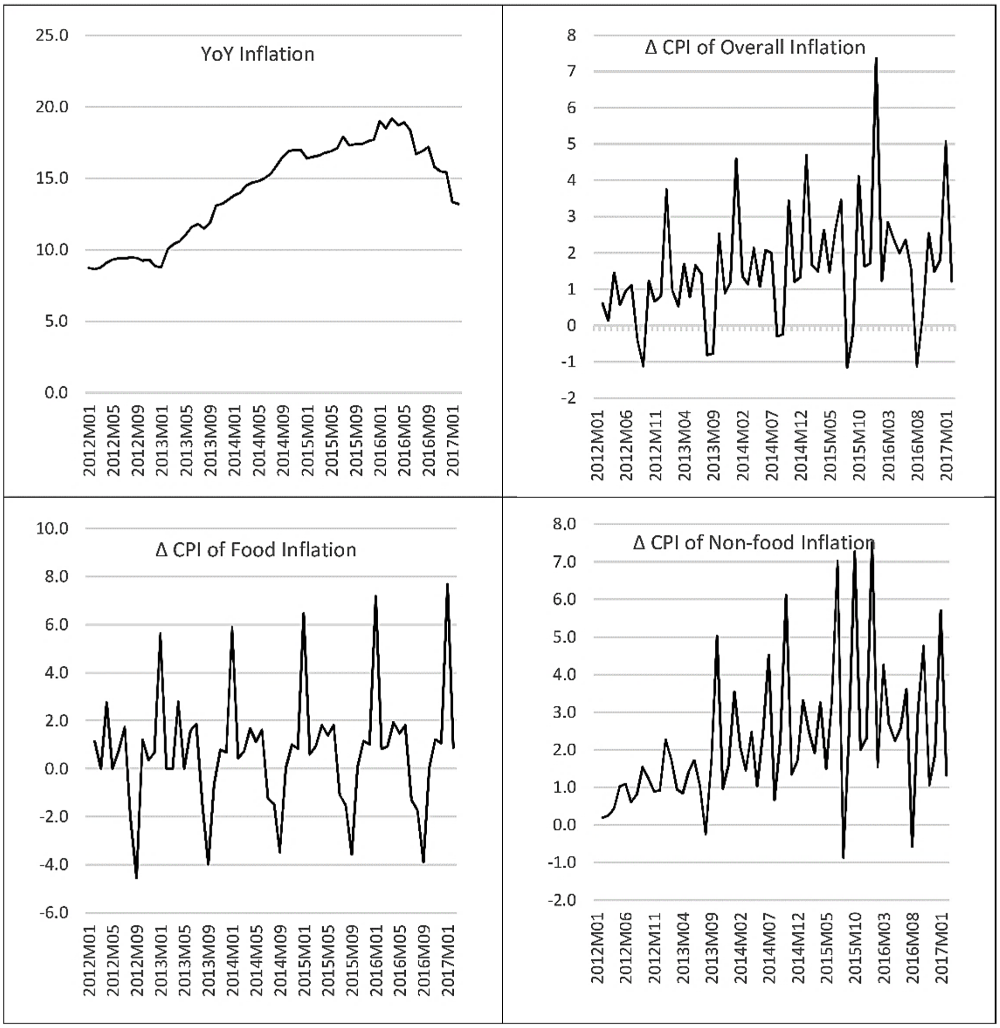
In order to prevent situations where the errors do not have a constant variance (heteroscedasticity) the variables (CPI_F, CPI_NF, CPI_O, and YOY_INF) were transformed into logs (LNCPI_F, LNCPI_NF, LNCPI_O, and LNYOY_INF). The consequences of using OLS in the presence of heteroscedasticity include the possibility that the OLS estimation will still give unbiased coefficient estimates, but that they are no longer BLUE (Best, Linear, Unbiased, Estimator). Also, the standard errors could be inappropriate and hence any inferences we make could be misleading. Also transforming the variables into logs will also prevent specification errors. By this we linearize many previously multiplicative models into additive ones.

There is also the problem of confusing trends in a variable with the presence of seasonality in the variable. For instance, Lütkepohl and Xu (2011), focused explicitly on seasonally unadjusted price series – an interpretation for series with unit roots. They used autoregressive (AR) models which they refer to as seasonal differences (stochastic seasonality models) and models with seasonal dummies for the first differences, called deterministic seasonality models. They found these two models to be more successful in forecasting seasonal time series than models with both first and seasonal differences and models without any differences at all. However, these are normal processes one has to follow when using univariate models for forecasting. In fact, dealing with unit roots cannot be equated to dealing with seasonality. Thus, the dominance of a series by a trend can obscure the presence of seasonal effect in a series, showing it is necessary to de-trend time series data before conducting test for seasonality. The Student t-test and Wilcoxon Signed-Ranks test have been recommended for detection of seasonality but this study uses the graphical method for this purpose as shown in Figure 1 (Nwogu et al. 2016).

A critical examination of the CPIs of the various components of inflation shows an underlying seasonality in Ghana's inflation data. This seasonality is more visible if a first difference is computed and plotted for the CPIs of the various components of inflation. The seasonality is not very pronounced in the CPI of non-food inflation. This is expected as its components involve goods that are very volatile in nature and are mainly imported products that rely heavily on foreign exchange.

As a result, a seasonal autoregressive (SAR) and seasonal moving average (SMA) terms were introduced (products of lag polynomials). Their addition produces ARMA models of a higher order, referred to in this study as SARMA models with non-linear restrictions on coefficients. For this purpose, we will use the overall CPI to demonstrate how a SARMA model can be generated from an ARMA model.

Figure 1 Year-on-year Inflation Rate and Tests for Seasonality



Source: Author's Construction based on Ghana Statistical Service data (2019)

For instance, the SAR and SMA terms can be added to an ARMA (2.2) model of the log of overall inflation (LNCPI_O) to derive the following:

$$LNCPI\_O_t = \phi_1 LNCPI\_O_{t-1} + \phi_2 LNCPI\_O_{t-2} + \theta_1 \mu_{t-1} + \theta_2 \mu_{t-2} + \mu_t \quad (4)$$

Where $\emptyset$ *and* θ are the coefficients of the AR and MA terms denoted by $LNCPI\_O_{t-1,2,3}$ and $\mu_{t-1,2}$.
Using a lag operator L

$$LNCPI\_O_t = (1 - \emptyset_1 L - \emptyset_2 L^2) LNCPI\_O_t + \mu_t (1 + \theta_1 L + \mu_2 L^2) \quad (5)$$

For data on year-on-year inflation, we might wish to add a SAR (12) and a SMA (12) term because we believe that there is a correlation between a year and the previous year.

$$LNCPI\_O_t = (1 - \emptyset_1 L - \emptyset_2 L^2)(1 - \rho_{12} L^{12}) LNCPI\_O_t + \mu_t (1 + \gamma_{12} L^{12})(1 + \theta_1 L + \mu_2 L^2) \quad (6)$$

The parameter $\rho\ and\ \gamma$ are associated with the seasonal part of the process with non-linear restrictions on the coefficients. Thus the current values of the log of the Consumer Price Index (CPI) of the overall component over the study period depends on its own previous two values plus a combination of current and the previous two values of white noise error terms, where $\epsilon(\mu_t) = 0; \epsilon(\mu_t^2) = \sigma^2; \epsilon(\mu_t \mu_s) = 0; t \neq s$

The best SARMA model was estimated and thus selected for the variables (LNCPI_F, LNCPI_NF, LNCPI_O and LNYOY_INF) using the Akaike Information Criterion (AIC). This is meant to determine the type of model that best fits the set of data, and also choose the best model from which to forecast that data. The graphical plots of the autocorrelation function and the partial autocorrelation function were not used to determine the best model because they were difficult to interpret, as there was no clear trend. Information criteria are the most common model selection tool used in econometrics. By using the AIC, the aim is to choose the number of parameters, which minimises the value of the information criterion. It is based on the estimated log-likelihood of the model, the number of parameters in the model and the number of observations.

Since SARMA (a seasonal ARMA) models are non-theoretical (not based on any economic or financial theory) the focus of this study was to determine whether the SARMA model selected describes the log of the Inflation and CPIs in Ghana well and produces accurate forecasts. A year less observations was taken from the full sample to estimate the SARMA model selected in table 1 below. That of the ARMA model is shown in appendix table 1.

Table 1 The Best SARMA Models for Study Variables

CPI	**Adjusted Sample**	**SARMA Model**	**AIC**
LNCPI_F	2014M05 2017M02	(1,0) (12,0)	–6.263680
LNCPI_NF	2014M10 2017M02	(2,1) (12,0)	–6.349343
LNCPI_O	2013M02 2017M02	(2,2) (12,12)	–7.074027
LNINF_YOY	2014M08 2017M02	(4,1) (12,6)	–3.808331

Source: Author’s Construction

The maximum order for both the AR and the MA terms for the alternative ARMA models in appendix table 1 were set at par with that of the SARMA models – with the exception of the model for the log of the overall CPI. That notwithstanding, the SARMA models had a better AIC than the alternative ARMA model in all but one case – the log of non-food CPI.

An out of sample dynamic forecast was then conducted for the period 2016M03 to 2017M02. The forecast function is of the form similar to equation (1) but this time $f_{t,s}$ denote the forecast made using the SARMA (2, 2) (12, 12) model at time t for s steps into the future for Overall CPI in Ghana – the same forecast function was applied to the best models selected for the other variables (Brooks, 2008).

Table 2 Properties of Forecast SARMA Models a One Year Forecast Period

Properties	**LNCPI_F**	**LNCPI_NF**	**LNCPI_O**	**LNINF_YOY**
Root Mean Squared Error (RMSE)	0.00878	0.01351	0.01745	0.15413
Mean Absolute Error (MAE)	0.00698	0.01228	0.01397	0.11846
Mean Absolute Percentage Error	0.14235	0.22769	0.26759	4.38936
Theil Inequality Coefficient	0.00090	0.00126	0.00168	0.02698
Bias Proportion	0.57094	0.27814	0.35378	0.47153
Variance Proportion	0.10794	0.65805	0.61792	0.50700
Covariance Proportion	0.32113	0.06381	0.02830	0.02147

Source: Author's Estimates

Table 2 above shows the output table of the dynamic forecast for the selected SARMA models. The Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) for the models are fairly small. The criterion for these two statistics is that the smaller the error, the better the forecasting ability of the model. The bias proportion indicates that the mean of the forecast is not far from the mean of the actual series. The gap between the variation of the forecast and the variation of the actual series as shown by the Variance Proportion are also fairly small. The Covariance Proportion shows the remaining unsystematic forecasting errors. The forecast for the various SARMA model for the log of the CPI of food inflation can be said to be accurate, followed by the non-food CPI, the overall CPI and then the year-on-year inflation rate. The result for the Theil Inequality Coefficient is also consistent with the results above. The forecast properties for the SARMA models where better than the ARMA forecast properties for both the log of food CPI and the log of non-food CPI. The same cannot be said of the log of overall CPI and the log of the year-on-year inflation (see appendix table 2).

To test for consistency in the result, we varied the out-of-sample period from one year to two months. By doing this the CPI of food lost its forecast accuracy to the CPI of non-food followed by the CPI of overall. This evaluation was based on the forecast properties shown in table 3 below. Just like the forecast properties with the one year out-of-sample period, the forecast properties of the SARMA models appeared to be good in predicting the CPI of food and non-food inflation. The same could not be said for the year-on-year inflation and the CPI of overall inflation (see appendix table 3). In both tables (2 and 3), the average of the CPI forecast for both food and non-food inflation appeared to have better forecast properties than the CPI forecast of the overall inflation and the year-on-year inflation. This confirms the observations of Stock and Watson (2003, 2004) that the best predictive performance is obtained by simply averaging these forecasts. This applied only to the SARMA models – the ARMA models showed mixed results.

Studies have, however, criticised the use of these forecasting properties for deciding the accuracy of a forecasting model. For instance, Makridakis and Hibon (1995) have argued that some of these properties may be influenced by outliers, and as a result has little intuitive meaning. For this reason, the out-of-sample forecasts were compared to the actual data graphically to determine if the findings noted in the forecast properties above would still hold.

Table 3 Properties of Forecast Models for a Two Months Forecast period

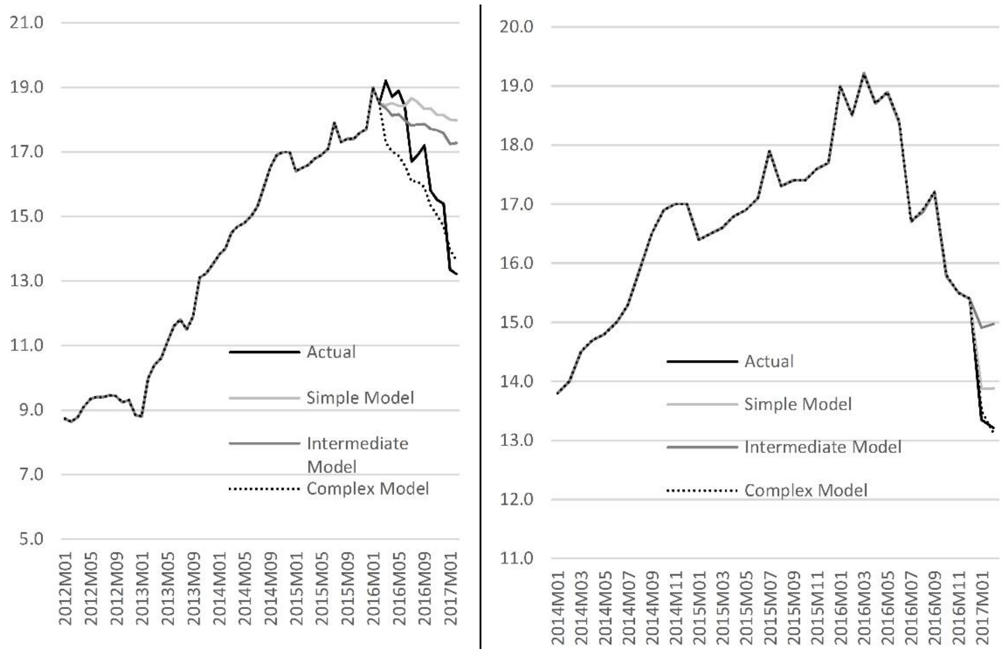
Properties	**LNCPI_F**	**LNCPI_NF**	**LNCPI_O**	**LNINF_YOY**
Root Mean Squared Error (RMSE)	0.02106	0.00183	0.01443	0.04430
Mean Absolute Error (MAE)	0.02103	0.00167	0.01441	0.04393
Mean Absolute Percentage Error	0.42557	0.03059	0.27439	1.69893
Theil Inequality Coefficient	0.00213	0.00017	0.00137	0.00849
Bias Proportion	0.99770	0.17139	0.99749	0.98334
Variance Proportion	0.00231	0.82862	0.00252	0.01121
Covariance Proportion	0.00000	0.00000	0.00000	0.00545

Source: Author's Estimates

Figure 2 shows that the complex model (the model that uses both the forecast values of both food and the non-food CPI) is close to the actual data followed by either the simple model or the intermediate model, depending on the out of sample period. The graph also shows some deviations from the actual inflation values during the forecast period – the variation was worse for the forecast based on just the ARMA models (see appendix figure 1). This observation is not surprising considering the tight monetary stance taken by the government after the signing the IMF program in April-2015. The monetary policy stance tightened over early 2016. The past tightening, together with the fiscal consolidation under the IMF program contributed to the sharp decline in inflation (IMF, 2017).

This change in policy was not fully captured by the models since the out of sample period started at about the time when inflation was peaking and the policy rate was reduced after it had peaked some months into the out of sample period. This observation confirms the observations of Lucas (1977) that backward-looking forecasting models are not good at predicting future events because of changes in policies. To correct for this, we reduced the out of sample period to two months using the forecast outcome from table 3 above. A graphical representation of this showed the forecast abilities of the models to have improved (figure 2 and appendix figure 1). The complex model appeared to be the most accurate, followed by the simple model and the intermediate model. The complex model predicted inflation values of 13.5 and 13.1 for January 2017 and February 2017 respectively. This compares with 13.3 and 13.2 actually realised for the same period. All the models were able to mimic the movement of the actual inflation data but with some variance.

Figure 2 One year and Two Month Out of Sample Forecast Models Versus Actual



Source: Author's Construction based on Ghana Statistical Service data (2019)

5. Conclusion

The adoption of the IT framework by Ghana and other developing countries, implies that monetary aggregates have lost their relevance as explanatory variables to inflation or the price level. It has thus become imperative for monetary policy to target inflation specifically. In other words, for monetary policy to be effective in doing so, a lot of stall will have to be placed on inflation forecasting. A good forecasting model will help the Monetary Policy Committee (MPC) to know when to tighten or loosen their policy stance in order to achieve the desired inflation rate.

As discussed earlier, the components of inflation in Ghana show some seasonality, which most univariate (ARMA) models ignore. By comparing forecast models of both ARMA and SARMA models, this study showed that SARMA model, as demonstrated by its forecast ability, appears to be the best univariate model for inflation forecasting in developing economies (Ghana) due to its seasonal components. The results also show that the complex approach which uses the forecast values of both food and non-food CPI, is very good at forecasting inflation in Ghana. Its predictability however, is most effective in the very short run e.g. two months. This to some extent corresponds to the work of Moser et al. (2007) who favoured the aggregation of sub-indices forecasts over a forecast of headline inflation itself (simple approach).

The results of this study could help monetary policy authorities determine the path of the inflation variable after they have made their policy decisions. This is because the complex model performs very well when policy decisions are captured in the in-sample period. An avenue for further studies is to estimate the role of expectations and other variables in determining the inflation variable in developing economies, allowing for better forecast models. Also, it would be interesting to compare the approach in this study to multivariate models to confirm if the conclusions arrived at still hold.

Acknowledgements

I would like to express my gratitude to the Tempus Foundation through its Stipendium Hangaricum program for granting me the scholarship to undertake this study.

References

Atkeson, A. – Ohanian, L. E. (2001): Are Phillips curves useful for forecasting inflation? Federal Reserve Bank of Minneapolis. *Quarterly Review-Federal Reserve Bank of Minneapolis*, 25, 1, 2.

Balcilar, M. – Gupta, R. – Jooste, C. (2017): Long memory, economic policy uncertainty and forecasting US inflation: a Bayesian VARFIMA approach. *Applied Economics*, 49, 11,1047–1054.

Bartolomei, S. M. – Sweet, A. L. (1989): A note on a comparison of exponential smoothing methods for forecasting seasonal series. *International Journal of Forecasting*, 5, 1, 111–116.

Binner, J. M. – Tino, – P.Tepper, J. – Anderson, R. – Jones, B. – Kendall, G. (2010): Does money matter in inflation forecasting? *Physica A: Statistical Mechanics and its Applications*, 389, 21, 4793–4808.

Box, G. E. – Jenkins, G. M. – Reinsel, G. C. – Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Cecchetti, S. G. – Genberg, H. – Lipsky, J – Wadhwani, S. (2000). *Asset prices and central bank policy*. Centre for Economic Policy Research.

Duarte, C. – Rua, A. (2007): Forecasting inflation through a bottom-up approach: How bottom is bottom? *Economic Modelling*, 24, 6, 941–953.

Fisher, J. D. – Liu, C. T. – Zhou, R. (2002): When can we forecast inflation? *Economic Perspectives-Federal Reserve Bank of Chicago*, 26, 1, 32–44.

Ghana Statistical Service (2019). CPI Coicop Annexes. Retrieved from https://statsghana.gov.gh/nationalaccount_macros.php?currentpage=2&Stats=MjI0MzI3MDczMi4wMDI1&&CPI%20Coicop%20Annexes

Hansen, P. R. – Timmermann, A., (2012):*Choice of Sample Split in Out-of-Sample Forecast Evaluation* (No. 2012–43). Department of Economics and Business Economics, Aarhus University.

Hassani, H. – Coreman, J. – Heravi, S. – Easaw, J. (2018): Forecasting Inflation Rate: Professional Against Academic, Which One is More Accurate. *Journal of Quantitative Economics*, 16, 3, 631–646.

International Monetary Fund (2017) *Ghana: 2017 Article IV Consultation, Fourth Review Under the Extended Credit Facility Arrangement, Request for Waiver for Nonobservance of Performance Criteria, and Request for Extension and Rephasing of the Arrangement-Press Release; Staff Report; Staff Supplement; and Statement by the Executive Director for Ghana.* International Monetary Fund. (Washington).

Lucas Jr R. E. (1977):Some international evidence on output-inflation tradeoffs: reply. *American Economic Review*, 67, 4, 731–731.

Lütkepohl, H. – Xu, F. (2011): Forecasting annual inflation with seasonal monthly data: Using levels versus logs of the underlying price index. *Journal of time series econometrics*, 3, 1.

Hibon, M. – Makridakis, S. G. (1995): *Evaluating accuracy (or error) measures*. *INSEAD View Article PubMed/NCBI*. Fontainebleau.

Mohammed, O. B. – Kolo, A. J. – Solomon, A. A. (2015): Comparative study of inflation rates forecasting using feed-forward artificial neural networks and auto regressive (AR) models. *International Journal of Computer Science Issues (IJCSI)*, 12, 2, 260.

Moser, G. – Rumler, F. – Scharler, J. (2007): Forecasting austrian inflation. *Economic Modelling*, 24, 3, 470–480.

Nwogu, E.C. – Iwueze, I. S. – Nlebedim, V.U. (2016): Some tests for seasonality in time series data. *Journal of Modern Applied Statistical Methods*, 15, 2, 24.

Pant, P. N. – Starbuck, W. H. (1990): Innocents in the forest: Forecasting and research methods. *Journal of Management*, 16, 2, 433–460.

Sbrana, G. – Silvestrini, A. – Venditti, F. (2017): Short-term inflation forecasting: the META approach. *International Journal of Forecasting*, 33, 4, 1065–1081.

Stock, J. H. – Watson, M. W. (1999): Forecasting inflation. *Journal of Monetary Economics*, 44, 2, 293–335.

Stock, J. H. – Watson, M. W. (2001): Forecasting output and inflation: the role of asset prices (No. w8180). National Bureau of Economic Research.

Stock, J.H. – Watson, M.W. (2003): Has the business cycle changed? Evidence and explanations. *Monetary policy and uncertainty: adapting to a changing economy*, 9–56.

Stock, J.H. – Watson, M.W. (2004): Combination forecasts of output growth in a seven-country data set. *Journal of forecasting*, 23, 6, 405–430.

Stock, J. H. – Watson, M. W. (2008): *Phillips curve inflation forecasts* (No. w14322). National Bureau of Economic Research.

Stock, J. H. – Watson, M. W. (2010): *Modeling inflation after the crisis* (No. w16488). National Bureau of Economic Research.

Stock, J.H. – Watson, M.W. (2016): Core inflation and trend inflation. *Review of Economics and Statistics*, 98, 4, 770–784.

Tashman, L. J. (2000): Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, 16, 4, 437–450.

Wright, J. H. (2009): Forecasting US inflation by Bayesian model averaging. *Journal of Forecasting*, 28, 2, 131–144.

Appendix

Table 1 The Best ARMA Models for Study Variables

CPI	Adjusted Sample	ARMA Model	AIC
LNCPI_F	2012M05 2017M02	(3,3)	–5.030210
LNCPI_NF	2012M05 2017M02	(4,4)	–6.465971
LNCPI_O	2012M11 2017M02	(9,5)	–6.396296
LNINF_YOY	2012M05 2017M02	(4,3)	–3.604795

Source: Author's construction

Table 2 Properties of Forecast Models a One Year Forecast Period (ARMA)

Properties	LNCPI_F	LNCPI_NF	LNCPI_O	LNINF_YOY
Root Mean Squared Error (RMSE)	0.06805	0.03104	0.01475	0.08898
Mean Absolute Error (MAE)	0.06169	0.02692	0.01317	0.06910
Mean Absolute Percentage Error	1.25468	0.49854	0.25295	2.54904
Theil Inequality Coefficient	0.00698	0.00288	0.00142	0.01574
Bias Proportion	0.82167	0.61766	0.04900	0.24765
Variance Proportion	0.01261	0.34787	0.71786	0.70753
Covariance Proportion	0.16572	0.03447	0.23314	0.04483

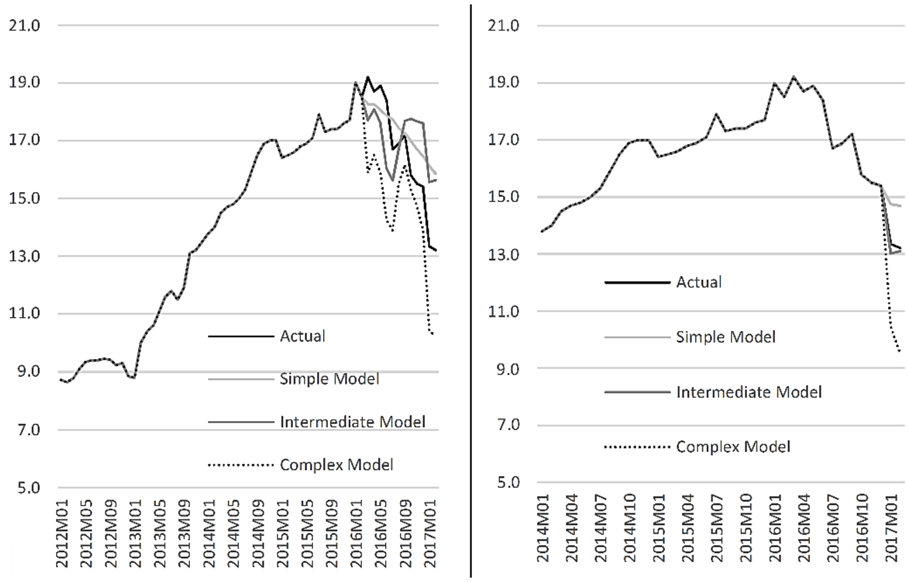
Source: Author's construction

Table 3 Properties of Forecast Models for a Two Months Forecast period (ARMA)

Properties	LNCPI_F	LNCPI_NF	LNCPI_O	LNINF_YOY
Root Mean Squared Error (RMSE)	0.03806	0.01039	0.00226	0.10296
Mean Absolute Error (MAE)	0.03765	0.01000	0.00210	0.10292
Mean Absolute Percentage Error	0.76175	0.18364	0.04005	3.97978
Theil Inequality Coefficient	0.00387	0.00096	0.00022	0.01952
Bias Proportion	0.97867	0.92577	0.86709	0.99920
Variance Proportion	0.00326	0.07423	0.13291	0.00081
Covariance Proportion	0.01808	0.00000	0.00000	0.00000

Source: Author's construction

Figure 1 One year and Two Month Out of Sample Forecast Models Versus Actual (ARMA)



Source: Author's Construction based on Ghana Statistical Service data (2019)

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 303–325.

Measuring fiscal distress at local municipalities

Balázs Tóth

Due to the rise of New Public Management (NPM), accounting systems and organizational best practices have changed greatly in the public sector. The idea promotes the public sector adopting processes and methods from the private sector. This widespread paradigm has a tremendous effect on the importance of measurement of the efficiency, effectiveness, and economic performance of the public sector. However, traditional profitability indicators provide a misleading picture of the financial viability of municipalities. The financial performance of local government can be better evaluated with the level of their fiscal distress.

Besides these new approaches, the financial problems of large American cities called for better monitoring systems several decades ago. Recently, the global crises of the last decade have drawn attention to the importance of this issue. However, there is no universally accepted model for measuring fiscal distress. Due to the special characteristics of the public sector, the measurement of financial distress should be customized for the sector. The goal of this research is to introduce and compare definitions, models, and theories which describe the importance of the evaluation of the fiscal health of the local governments.

Keywords: fiscal distress, public sector, distress prediction

1. Introduction

The aim of this study to compare the definitions and select the most relevant variables of fiscal distress models, on the theoretical level. However, the private sector has several well-known financial distress models (e.g. Altman's multi-discriminant analysis, Ohlson's logit model, Zmijewski's probit model, or the Black-Scholes option-pricing model) (Wu et al. 2010), while in the public sector, there are no universally used models. There are several barriers which have prohibited measuring this kind of financial performance. Besides the changes and the differences in the legal environments, the financial reporting system can cause difficulties, too. Earlier, the cities of the United States used cash or modified accrual accounting. The systems applied here varied greatly. This made the comparability of the public sector entities difficult (Clark 1977). Without an up-to-date standardized accounting system, there is no possibility of creating an efficient predictive model for fiscal distress.

The financial accounting environment of the public sector has changed significantly. From the 1970s, public sector reforms were promoted internationally. These reforms aimed to reduce the inefficiency of the public sector (Christensen et al. 2018). Due to the rise of the NPM, the importance of efficiency, effectiveness, and accountability increased. NPM called for more business-like solutions (Hood 1995). The reform of the accounting system was one of the key areas.

The adoption of accrual accounting in the public sectors started in the 1990s. Firstly, we can mention New-Zealand and Australia as early adopters who were followed by other Anglo-Saxon countries. This process cannot be described as an

English-speaking 'club' phenomenon, as several OECD countries adopted it, or started to adopt accruals by 2000 (Carlin 2005). According to PwC's survey, 80% of OECD countries are planning to introduce accrual accounting by 2020 (PwC 2015). These changes in the financial reporting system have made it more transparent, comparable, and reliable. The improvement of the accounting system enabled the creation of early warning systems. The appearance of fiscal distress modeling is the consequence of both broader public management reforms and the inefficient operation of the public sector. However, fiscal distress modeling is usually not connected to NPM, but the changes which were promoted by the idea provided tools for these kinds of investigations. The fiscal distress models can be a proper indicator of the effectiveness of local governments. The goals of these organizations can be reduced to two points: first they have to meet their financial obligations, and secondly, they have to provide services to their citizens. These two motives can be measured with fiscal distress indicators.

In the 1970s there was increased attention paid to fiscal distress predictions because of the bankruptcies of American cities (Gorina et al. 2018). From this period there were several public sector-specific distress models created, mostly in the Anglo-Saxon countries. This research aimed to recognize financial emergencies (Kloha et al. 2005). However, in the 1990s there was a drop in the amount of research regarding this issue, although interest in fiscal distress measurement recovered in the new millennium (Gorina et al. 2018). In Europe, the measurement of fiscal distress was enabled after the reform of public sector accounting (Cohen et al. 2012). Accrual accounting can provide a more punctual and reliable picture of an organization. This accounting method can improve the valuation of the assets of an organization, and it can show a more actual picture of its revenues and obligations (Simon 2011). The higher validity of the accounting system and the available data provide an opportunity to create and test fiscal distress models. Accrual accounting can capture all of the transactions immediately, full updating accounting reports (European Commission 2013)

The Hungarian public sector (similarly to several European countries) adopted accrual accounting in 2014 (Balog–Jakab 2017), making the measurement of fiscal distress among Hungarian local governments topical and viable. The previous accounting system was not able to create a punctual report on the processes of local governments. Their debt burden was not predictable, and the data provided by accounting reports was not reliable (Simon 2011).

Besides the changes in the financial reporting system enabling proper measurement, the problems of the local government sector highlighted the importance of fiscal distress measurement. During the crisis, the financial health of the Hungarian local governments destabilized (Halmosi 2013). The inappropriate regulation of the local government sector had a key role in indebtedness of the sector (Lentner 2014). These problems called for higher efficiency, more effective control activities, and performance measurement.

The paper introduces nine fiscal distress indicators, which have been applied in previous research. The measures presented can be useful for local managers to examine the conditions and the risk of their municipality, but they are suitable for external users (auditors, creditors, and other stakeholders) as well. The research was based on different datasets. There are examinations from the United States, Australia, and European countries. The comparability of the results and the models are limited due to the different methodical approaches, and the different regulation of the local governments involved. Ziolo (2015) highlighted that it is impossible to adapt to the financial variables that are used in international papers, which could explain the lack of cross-country studies. These measurements are usually customized by one particular country and its regulations. The paper is not trying to create answers and tools for local managers to handle financial problems, the aim is to examine different methods and approaches to measure and predict the level of fiscal distress.

The remainder of the paper organized as follows: In Section 2 the paper discusses the definition and the possible causes of fiscal distress, Section 3 shows researches and variables regarding the measurement of fiscal distress. Section 4 collects the expectations regarding fiscal distress models. Finally, Section 5 summarizes the main findings of the paper.

2. Theoretical background

In this section, different definitions will be introduced, along with major features of the relevant research. Firstly, the different definitions of fiscal distress literature. Most of the research uses similar definitions, however, they are not equivalent. Table 1 presents the commonly used definitions.

Table 1 Commonly Used Definitions Found in The Fiscal Distress Literature

Term	**Definition**	**Author**
Fiscal Health	Underlying or structural ability to deliver public services to its residents, independent of the budgetary decision made by city officers	Ladd–Yinger (1989)
Fiscal strain	An institutional lack of adaption to a changing environment	Clark–Appleton (1989)
Fiscal Stress	The imbalance between the revenue raising capacity and the expenditure needs of a local government	Badu–Li (1994)
Financial distress	Occurs when the entity, municipality or province, is no longer able to perform its essential functions and deliver due services, or when it is no longer able to meet debt within third parties through the ordinary means of restoring fiscal balance or the debt instrument with a balance sheet	Costanz–Rossi–Zito (2012)

Source: Ziolo (2015), p. 16.

The terms and definitions can be dissimilar in different works of research. Clark (1977) examined fiscal strain and aimed to apply indicators which could forecast fiscal strain. He differentiated three approaches to fiscal strain through the data of the bond market. In this research, he defined fiscal strain as a situation when there is a drop in the offering price of a new issue, or it can be indicated by changes in the price of the bond on the secondary market. Both of the indicators related to the probability of the issue meeting regular payments of principal and interest. He examined the third approach. Fiscal strain may then be viewed as deriving from the relationship between resources and expenditures established by municipal leaders. As the socioeconomic features of a municipality or its resources change, municipal leaders differentially adopt new policies to cope with their changing environment. Some local governments adapt rapidly enough to avoid straining themselves fiscally, while others adapt much more slowly, and in the meantime develop considerable fiscal strain. These changes largely influenced by national trends, the leaders of the municipalities have only modest control on any of them (Clark 1977).

Kloha et al. (2005) have a similar definition of fiscal distress. They defined fiscal distress as a failure of the municipalities to meet their standards in the areas of operating position, debt, and community need and resources over successive years. They also examined different definitions, too. "Fiscal distress reflects short-term considerations, such as a local government's ability to meet is payroll and generally make payments in a timely manner" (Kloha et al. 2005). It also could be described as imbalances between the level of the allocated resources and the potentially available resources. However, the definition may also include long-term considerations, including the tax base relative to its expenditures and commitments. Nevertheless, the meeting of financial obligations is not the only task of the municipalities. They also have to meet with needs of the community. The local governments have a wide range of purposes and tasks, and this aspect of distress is particularly difficult to operationalize (Kloha et al. 2005).

According to the study of Kloha et al. (2005), there are four causes of fiscal distress. Shifts in the population and the job market, governmental growth, interest group demands, and poor management. Population and job market shifts represent what is called the migration and tax base erosion model, which is the point that the major cause of the fiscal distresses originate from negative changes of the demographical features. Governmental growth, which is also referred to as bureaucratic growth focuses on the lack of market signals in the public sector. The government's spending increases rapidly. The government has to calculate population growth and inflation. According to the third approach, the cause of the financial vulnerability is the demands of the interest groups. The elected mayors or officials usually must meet with the expectations of different interest groups and this could reduce the effectiveness of the municipality. The fourth reason for fiscal distress is poor management. In this scenario, the accounting and management practices can be described as unfertile practices, the budgeting and estimation procedures are inaccurate (Kloha et al. 2005). Carmelli and Cohen (2001) have a similar explanation for the financial problems. They attributed fiscal distress to a lack of organizational resources and managerial skills, which in turn lead to inappropriate service delivery and inefficient adaptation to changing conditions (Carmelli and Cohen 2001).

Jones–Walker (2007) used a different approach. Distress can be defined as an inability to provide services at pre-existing levels. In order to provide services to the community, municipalities are expected to invest in infrastructure and to maintain legacy infrastructure. In their research, they used the estimates developed by local governments of the cost of restoring infrastructure to a satisfactory condition as a measure of degrees of distress. The binary classification (failed or not failed) is not suitable ord relevant) for the public sector (Jones–Walker 2007).

Trussel–Patrick (2013) defined fiscal distress as the existence of financial and other problems that could cause the municipality to reduce its current level of public services. The situations are caused by the imbalances between the financial resources a municipality has committed to providing services and the potential resources it has available to provide those services (Trussel–Patrick 2013).

Beck–Stone (2017) examined the relationship between the dissolution of the municipalities with fiscal distress. The financial condition of a municipality can be distributed into two components. The first is the probability that a government will be able to meet both its financial obligation to creditors, consumers, employees, taxpayers, suppliers, constituents, and others as they become due. The second one is to meet service obligations to constituents. The municipalities have to fulfill both kinds of obligations. The first component is similar to the concept of financial condition for private entities, but the second can be described as public sector-specific. They emphasized that fiscal distress is difficult to be isolated from other reasons for dissolution. For example, a decrease in the population lowers the tax base, which affects the financial conditions (Beck–Stone 2017).

In the research, six reasons for the dissolutions were differentiated. The dissolution could be motivated by seeking greater efficiency, it could be caused by low participation in the local government. The exodus of the population could also lead to dissolution. The dissolution also could be forced by the state or country. The authors created a group for other reasons too (Beck–Stone 2017). As they highlighted, fiscal distress can be one of the main reasons for a dissolution, but it is difficult to identify as an isolated reason for this process. They summarized the process of measurement of financial distress.

Financial distress is affected by a wide range of factors according to this concept. As a first step, the environmental conditions should be examined. A decrease in the population or, the increase in the average age can also be an early warning sign. The dependence on a declining or struggling industry or the declining price of property (or the greater inventory of homes for sale) could indicate financial distress as well. As a second step, the likelihood of continuation as a going concern should be assessed. In that step, the local conditions, citizens' initiatives, and state initiatives should be considered. As the third step, the likelihood of severe financial distress could be evaluated. The estimation contains short- and long-term indicators. The performance can be assessed through the ability to meet current and future obligation with the available resources. The recent trends in fiscal conditions should also be reviewed. Besides the financial ratios, the behavior of management matters. The competence of and the steps taken by management affects the probability of fiscal distress (Beck–Stone 2017). However, this line of investigation provides a truly unique approach to understanding fiscal distress, for which no model has yet been created.

In consort with the above theories, Skidmore–Scorsone (2011) differentiated external or economic factors and management-type determinants. In the first group, we can find the increasing cost of health insurance or the property market (through the effects the taxation) could be a key driver, too. These elements cannot be influenced by local managers. The second cause is linked to poor financial management. In their research, they measured fiscal distress through the gap between the change of Government Services Cost Index and the change of the Government Revenue Index. The Government Service cost index is divided into Government Employee Cost Index and Capital Cost Index. Both of the Indices are affected by external conditions (Skidmore–Scorsone 2011).

The definitions and approaches introduced above highlight that several financial and social determinants (which are can used by internal or external factors) should be in the scope of further investigations. However, these things should be revised regularly, because there can be different relations in different places, or at different times. For example, in different countries, the local governments can collect different taxes, and have different obligations. In the studies which examine local government in the United States, the role of the property tax is enhanced and as a consequence property, market-related indicators can be involved in the examinations, too. In other countries, different revenues and markets should be investigated.

3. How to measure fiscal distress?

Cohen et al. (2012) highlighted that limitations of bankruptcy prediction. Municipalities in a majority of countries cannot declare bankruptcy, and as a consequence, researchers cannot rely on historical data to identify the characteristics of that (Cohen et al. 2012). Bankruptcy is the last resort of the central government. The procedure of bankruptcy is not advantageous for any of the entities concerned. The avoidance of bankruptcy is usually encouraged by the regulations. If a local government is not able to pay its obligations, they do not declare (or are not allowed to declare) bankruptcy automatically (Halmosi 2018). As a consequence, the application of private sector bankruptcy models is unsuitable for the public sector.

Another important difference is that profitability is not a goal in the public sector, high ROA (return on assets) and ROCE (return on capital employed) are not socially desirable for the non-profit municipalities. Increased profitability may be interpreted as a result of unjustifiably high taxes and not as an indication of efficiency. Similar to this thread, a high degree of debt may not result in default. The central government can help local governments to solve liquidity problems. Moreover, financially distressed public sector organizations do not always declare bankruptcy (Cohen et al. 2012). The binary classification of the municipalities (failed or not failed) cannot be a reliable dependent variable in these models. The process of the bankruptcy and the criteria for the declaration of the bankruptcy differ from country to country (Halmosi 2018). This also generates an additional problem.

Bond default and downgrades of ratings are also an ineffective way of measuring the outcome of fiscal distress, as creditors of municipalities typically do not have the power to force the sale of municipal assets in the event of bond defaults. Moreover, the bond can be covered by insurance, which is increases the ratings of a bond (Trussel–Patrick 2013).

Along with the financial ratios, social-economic measures also need to be considered, as the definitions presented highlight. The changes in the financial ratios usually indicate the problem too late, but a decline in the quality of the services could appear earlier. Moreover, local governments with different social-economical features could have a different level of distress, while their financial ratios are similar.

We can differentiate relative and absolute indicators. In the case of relative indicators, the values of each of the ratios involved has to be calculated, then the municipalities have to be arranged into an order based on the received results. Then the local governments can be differentiated in that order. The orders can be summarized with the help of different pointing systems. A negative consequence is that someone always has to be at the top and at the bottom, irrespective of how distressed local government is. An additional problem with relative models is that model users have to calculate the value of the indices for each municipality, even if they are only interested the level of fiscal distress in the case of one municipality (Kloha et al. 2005). In the absolute models, the results of the local governments are independent of each other's, as the values of the indicators are compared to a theoretical threshold.

In this section, there will be 5 categories of financial distress measurements introduced. The research examined is summarized in Table 2.

Table 2 The examined researches

Authors	Methodology	Rating	Sample	Examined years
Clark (1977)	Financial ratios	Absolute	US cities	1970–1974
Brown (1993)	Point-based rating	Relative	US cities under 100.000 residents	1992
Kloha et al. (2005)	Point-based rating	Absolute	Cities of State of Michigan	1998 (Applied data from 1993–1998 period)
García-Sánchez et al. (2012)	Point-based rating	Relative	Spanish municipalities with a population of over 50.000	1988–2008
Jones–Walker (2007)	Regression	Absolute	161 Australian Councils	2001–2002
Trussel–Patrik (2013)	Regression	Absolute	US municipalities	1995–2008
Gorina et al. (2018)	Regression	Absolute	Municipalities of Pennsylvania, Michigan, and California	2007–2012
Cohen et al. (2012)	Simulation-based	Relative	Greek municipalities	2007–2010
Ziolo (2015)	Cluster analyses	Absolute	Polish municipalities	2008–2013

Source: own elaboration

3.1. Financial ratios

Some researches only collect relevant variables, and do not create fiscal distress models. One of the first investigations was Clark's (1977). He created 4 indicators (Table 3) to measure fiscal strain. In the research, cities of the U. S. were examined. The research highlighted that bond defaults and interest rates on bonds are not proper measures of fiscal distress (Jones–Walker 2007).

Table 3 Indicators of fiscal strain

Indicator	**Description**
Default	Default, or the probability of default, where the default is defined simply as not meeting regular bond payments, is a presumed criterion of fiscal strain for bond investors and the two major bond- rating agencies.
Ratio Measures	Several ratios such as gross debt divided by the tax base or short-term debt over long-term debt are frequently published in municipal fiscal reports.
Social and economic base characteristics	Population size and change, median family income, and taxable property value are among the variables commonly included under this heading.
Funds flow measures	These measures are also often included in some form in financial reports.

Source: Jones–Walker (2007)

Clark (1977) enhanced the importance of funds flow and examined the 29 indicators. These indicators relied on debt statistics (long- and short-term were differentiated), interest-payments, general revenues and retirement funds, and liabilities. The indicators examined the short- and long-term sustainability of the financial structure (Clark 1977).

These kinds of investigations do not provide one single value which can describe the status of the municipality. However, this can be an advantage. The research related to one single indicator is often criticized because the aggregate scores can hide the weaknesses which can be shown by an individual indicator (Gorina et al. 2018). This weakness has a much higher impact among the relative indicators, as there is always a best (or a group of best) municipality which appears to be fiscally healthy, irrespectively of the fact that the whole local government sector can be fiscally distressed at the same time, theoretically. A weakness of this kind of approach is that usually there is no clear guidance for how to prioritize among the ratios.

3.2. Point-based ratings

Brown (1993) created a 10-point scale to assess the financial condition of smaller US cities. Ratios and their clarifications are summarized in Table 4.

Table 4 Ten Key ratios of financial condition

Ratio	**Clarification of Ratio Components**
Total revenues/Population	*Total revenues* are the total revenues for all governmental funds.
Total general fund revenues from own sources/Total general fund revenues	*Total general fund revenues from own sources* are the difference between total intergovernmental revenues.
General fund resources from other funds/ Total general fund sources	*General fund sources from other fund* are general fund operating transfers in. *Total general fund sources* are the total of general fund revenues and operating transfers in.
Operating expenditures/total expenditures	*Operating expenditures* is the total expenditures for the general, special revenues and debt service funds. *Total expenditures* are the total expenditure for all governmental funds.
Total revenues/Total expenditures	*Total revenues* are the total revenues for all governmental funds. *Total expenditures* are the total expenditure for all governmental funds.
Unreserved general fund balance/Total general fund revenues	*Unreserved general fund balance* is the total of both unreserved designated and unreserved undesignated fund balance for the general fund.
Total general fund and cash investments/Total general fund liabilities	(The components are self-explanatory).
Total general fund liabilities/Total general fund revenues	(The components are self-explanatory).
Direct long-term debt/Population	*Direct debt* is general obligation to be repaid from property tax revenues.
Debt services/Total revenues	*Debt service* is the total expenditures in the debt service fund. *Total revenues* of all governmental funds.

Source: Brown 1993, p. 22.

The model is a relative one, after the calculation of the ratios, Brown sequenced local governments and created four quartiles. The organizations of the best quartile received –1 point, the local governments of the second quartile received 0, then the next quartile received 1, and the worst municipalities received –2. Then the points received were summarized and the local governments were categorized into five groups:

- 10 or more points: Among the best
- 5 to 9 points: Better than most
- 1 to 4 points: About average
- 0 to –4 points: Worse than most
- –5 or less: Among the worst (Brown 1993).

Kloha et al. (2005) examined the responsibility of the state regarding the level of fiscal distress of local governments. They analyzed the 10-point scale of the State of Michigan that used 9 indicators (Table 5). Kloha et al. (2005) highlighted that, there is no single indicator which can create a picture of a government's fiscal position. If a local government reached "bad" value on an indicator, they gained one point on the scale. In the case of a consecutive operating deficit, the entity received 2. The indicators represented short- and long-term dimension of the fiscal distress. Besides the financial ratios, several socio-economic indicators (population growth, taxable value related indices) were used in their valuation (Kloha et al. 2005). This method could be labeled as a mixed approach, as it integrates management-type factors and external (economic) determinants, too. This model focusses on the predictions of fiscal distress rather than assessing reactions to the distress (Skidmore–Scorsone 2011).

The authors set standards for particular indices. In some cases, this was straightforward, while in other cases they created the threshold. In the latter case, the standard deviation from average values was used to identify a small percentage that is performing relatively poorly (Kloha et al. 2005).

The authors also introduced an early warning system, based on their valuation (Table 6). If a government scored 4 or fewer points, it could be labeled as fiscally healthy. In their cases, there is no action required by the state. A value above 5 points is considered relatively high, and if a local government reaches that amount, they are informed by the state. If a local government receives more points, it will be placed on a published list, and a review team also could be appointed (above 8 points) (Kloha et al. 2005).

Table 5 Indicators of fiscal distress

Name	**Description**
Population growth	Two-year growth
Real taxable value growth	Two-year growth
Large real taxable value decrease	Looks for large drop over a two-year period
General fund expenditures as a percentage of taxable value	Current general fund expenses divided by current taxable value
General fund operating deficit	Current general expenditures subtracted from current general fund revenues, divided by general fund revenues
Prior general fund operating deficits	Checks "General fund operating deficit" for two previous years
Size of general fund balance	General fund balance as a percentage of general fund revenues
Fund deficits in current or previous year	Current or previous year deficit in major fund
General long-term debt as a percentage of taxable value	Current general long-term debt divided by current taxable value

Source: Kloha et al. (2005), p. 319.

Table 6 Early Warning System

Points from scale	**Category**	**State action**
0–4 points	Fiscally healthy	No action
5 point	Fiscal watch	Local government is notified about relatively high score
6–7 points	Fiscal warning	Local government notified and placed on published list for current and following year
8–10 points	Fiscal emergency	Local government notified, placed on published list for current and following year, automatic consideration of review team

Source: Kloha et al. (2005), p. 321.

García-Sánchez et al (2012) used a similar approach. They created a relative model and they compared their model's results with the evaluations of the model of Kloha et al. (2005). The models were compared using the data of Spanish municipalities for the 1988–2008 period. The Authors used seven indices to predict fiscal distress, the indicators associated with flexibility, independence, and sustainability are the most important ratios to set the limits of municipal fiscal situations. Table 7 introduces the applied indicators.

Table 7 Indicators of fiscal distress model of García-Sánchez et al. (2012)

Indicator	Descriptions
Net Saving Index (NSI)	Difference between the receivables from current budget resources and the budget obligations from non-financial current expenditures, reduced by annual amortization payment-interest and principle-per inhabitant
Current financial independence index (CFII)	Current budgetary payable divided by current budgetary receivables except current grants
Total Finance Independence Index (FII)	Budgetary payables divided by budgetary receivables except grants
Non-financial budgetary result index (NFBRI)	Current budgetary payables, non-financial capital budgetary payables divided by non-financial current budgetary receivables, non-financial capital budgetary receivables.
Financial Charge per inhabitant (FCII)	Annual amortization payment-interest and principal – per inhabitant
Net Debt Index (NDI)	Annual accumulation variation long-term credit operations per habitant
Fiscal revenue Index (FRI)	Fiscal receivables divided by net current budgetary receivables.

Source: García-Sánchez et al. (2012).

The authors calculated the value of the indicators for each of the municipalities examined, then ranked the municipalities according to their values. After that, the municipalities received points based on their rank. The organizations of the best quartile received 0 point, the next quartiles received 0.25, then 0.5 and 1 point. In the case of Net Saving Index and the Fiscal Revenue Index, the higher values were preferred. In contrasts, the rest of the indicators received lower points for lower values. A municipality could score between 0 and 7 points. The authors proposed the following classification:

- 0 to 1.5 points: Excellent
- 1.6 to 2.5 points: Good
- 2.6 points to 3.5 points: Watching
- 3.6 points to 5 points Warning
- 5.1 points to 7 points: Emergency (García-Sánchez et al. 2012).

After the evaluation of the municipalities with their model, they tested which model had more significant explanatory variables. The authors created an overall model of their and the absolute model of Kloha et al. (2005). During the investigations, only six indicators were statistically significant (at the 99 percent confidence level). All of them originated from the model of Kloha et al. (2005). According to this result, the absolute models were proven to be better than the relative indicators, however, the absolute models need to incorporate certain indicators of financial independence as taken into account in other alert systems (García-Sánchez et al. 2012).

3.3. Regressions

Jones–Walker (2007) created two models: a quantitative and a qualitative regression. In the case of the quantitative regression, the dependent variable was a quantitative measure (e.g. physical output levels) of service delivery, while in the second case the quality of the services was the dependent variable (Jones–Walker 2007). The aim of the research was to create an early warning system. According to the authors, the drop in quality of services can do so, meanwhile Table 8 summarizes the explanatory variables of the regressions.

Table 8 Explanatory variables of the fiscal distress

Jones–Walker (2007) quantitative	Cash flow operations to total assets
	Long-term interest bearing debt to total assets
	Cash resources to total assets
	Interest cover
	Gross debt to operating cash flow
	Operating cash flow to total infrastructure assets
	Ordinary revenue (less waste and sewerage charges) to total assets
	Total expenditure by total assets
	Surplus to total assets
Jones–Walker (2007) qualitative	Population within council boundaries
	Local council large or small
	Rates revenue to total ordinary revenue
	Ordinary revenue (less waste and sewerage charges) to total assets
	Road program costs over total assets
	Number of full-time (equivalent) staff
	Carrying value — total infrastructure

Source: Jones and Walker (2007), pp. 409–410

For the qualitative regression, the authors used data and indicators of financial reports, while in the other regression, some social and service-specific factors appeared. The measurement of the dependent variables is an interesting issue. First, the services which have to be provided by the local governments should be collected. This could differ from country to country.

Jones–Walker (2007) examined Australian councils. At this time (2001–2002), the councils had to collect taxes (rates), and they were responsible for the roads and the collection waste for disposal. The councils were usually providing a wide range of other services, but they were less significant and they were funded by Commonwealth and State governments. For waste disposal management, the following variables were used: domestic waste pickups per week, the number of residential properties receiving waste management services, and total kilograms of recyclables collected. The councils were also obligated to provide and maintain infrastructure. The infrastructure variables in the analysis were: the carrying values for buildings, roads, other transport, water, sewerage and drainage infrastructure; estimated cost to bring buildings, roads, other transport, water, sewerage, and drainage infrastructure to a satisfactory condition; and budgeted maintenance expenditure for buildings, roads, water, sewerage, and drainage infrastructure (Jones–Walker 2007).

The financial variables are similar to other financial distress models, but the authors enhanced the importance of cash flow. The Australian councils have had to publish statements, which includes cash flow, since the early 1990s. This made possible examination of the significance of cash flow based indicators. They also tested operating cash flows (e.g., operating cash flows to total assets); cash position (e.g., cash and short term investments to total assets); liquidity and working capital (e.g., current ratio); rate of return (e.g., reported surplus to total assets); financial structure (e.g., total debt to total assets); and debt servicing capacity (e.g., operating cash flow to interest payments) (Jones–Walker 2007).

The quantitative model was not able to provide a statistically significant relationship between the level of fiscal distress and the explanatory variables. According to the results of the qualitative model, the population within council boundaries and the size of the council were positively associated the level of fiscal distress. This means that the councils with a larger population are relatively more distressed than the smaller councils. The councils' distress level is negatively associated with revenue-generating capacity (this variable has the highest statistical impact), while the entities with a smaller number of full-time equivalent employees appeared to be more distressed. The conditions of the infrastructures (carrying value) were also significant, the more written down assets (these assets are older or they are in poor condition) are positively associated with fiscal distress (Jones and Walker 2007). The paper was able to provide the above-mentioned connections that highlight the importance of the monitoring of the condition of assets and revenue-generating capacity. The proper management of the factors can help to reduce the probability of financial problems.

Trussel–Patrick (2013) examined the fiscal distress of US municipalities in the period of 1995-2008. They operationalized the fiscal distress as a 5% decrease in the public service expenses per capita. They highlighted that the proper measurement of the outputs of the public services are complicated, which was the reason they measured efforts (expenditure per capita).

Based on their hypothesis they examined the effects of six indices (Table 9).

Table 9 Financial Indicators of Public Service Reductions

Indicator	Measure	Expected Relationship with Public Service Reductions
Revenue Risk (REVRISK)	Revenues from Other Governments/Total Revenues	positive
Administrative Cost per Capita (ADMIN)	Administrative Expenditures/Population	negative
Capital Outlays (CAPREV)	Capital Expenditures/Total Revenues	negative
Debt per Capita	Total Liabilities/Population	positive
Debt Issued to Revenue (DEBTISSUE)	Debt Issued/Total Revenues	positive
Revenue Growth (GROWTH)	Change in Total Revenues/Total Revenues	negative

Source: Trussel–Patrick (2013)

The model classified correctly 83 percent of the examined municipalities. According to the result, the municipalities who receive more intergovernmental revenue (compared to their own-source revenues), spend less on capital items relative to total liabilities and bond proceeds, and use more debt tend to reduce their public services (Trussel–Patrick 2013).

Gorina et al. (2018) built a regression too, to test their action-based measure of fiscal distress. They labeled a municipality fiscally distressed in a given year if, its financial management was characterized by at least one of the following: a blanket prohibition of overtime, a comprehensive reduction of employee salaries, personnel furloughs or layoffs that affect multiple employees, late payments to vendors and other payees, large across-the-board budget cuts of at least 10 percent of the budget that produce cuts in services, budget enactment later than two months after the beginning of the fiscal year, pension contributions less than 75 percent of annual required contributions, unusual inter-fund transfers of at least 10 percent of general fund, excluding fund balance reclassifications pursuant GASB 54 (Government Accounting Standard Board), unusual tax rate or fee increase that are not related to debt issuance, declaration of fiscal emergency, default on municipal debt, bankruptcy, auditor doubts that the entity may continue to be a "going concern", or a takeover by state or significant state financial assistance (bailout). This evidence-based measurement of fiscal distress is a unique way to conceptualize fiscal distress (Gorina et al 2018).

One of the hypotheses of the authors was that the revenue structure can be an important determinant of fiscal health because of its effects on revenue collections. Governments with diversified revenues have higher revenue volatility in an economic recession. The empirical model used the general fund balance to capture *Cash Solvency*. For *Budgetary Solvency,* they used the operating ratio and total revenues per capita in governmental funds. *Long-term Solvency* is calculated as the level of debt and annual contributions to the pension plans. *Revenue Structure* captures a share of own-source revenues coming from property tax. The models were also controlled for the size and type of the government, local economic factors, the change in the housing prices, and the change in the population in the previous year. The authors ran logistic regression models with state and year fixed effects. The different economic and institutional environment of the three states (Michigan, Pennsylvania, and California) were taken into account (Gorina et al 2018).

According to the results of the research, cash solvency, long-term solvency, and revenue structure can be used for fiscal distress predictions, while budgetary-level solvency, socio-economic indicators, and government type are not as informative. The property tax was negatively associated with fiscal distress. Even though the dramatic effects 2007-2009 recession on the property market, the local governments which were more heavily depended on the property taxes were less likely to be fiscally distressed (Gorina et al 2018).

Gorina et al. (2018) highlighted three applications for practice. First, the local government officials tend to make certain politically and fiscally difficult decisions when they are confronted with strong fiscal pressure. Their study shows the analysis of the government decisions and actions may be used to determine if a government is in fiscal distress. Secondly, their research enhanced that the fund balance as a share of total expenditures and long-term debt as a share of total revenues provide an early warning of fiscal distress. Besides these, the importance of property taxes and the negative effects of the revenue diversifications were demonstrated (Gorina et al. 2018). The action-based method introduced can be useful for external users to evaluate the financial health of an organization.

3.4. Simulation-based approach

Cohen et al. (2012) examined the level of the fiscal distress of Greek municipalities. The authors combined a simulation analysis approach (stochastic multicriteria acceptability analysis) with disaggregation technique. With the help of financial ratios, they distinguished the financially viable municipalities from the distressed ones.

One of the six indices is the total liabilities to total assets ratio (L/A). This ratio shows the municipality's need for third-party financing. The value of the indicator can be described as favorable if its value is under 50%. Another indicator is the ratio of the own revenues to total liabilities (R/L). In Greece, the municipalities had to cover their interest payments with own revenues, in the period examined (2007–2010). As a liquidity indicator, the model uses the ratio of short-term liabilities to own revenues. Both of the exceedingly high or exceedingly low values of the ratio indicates financial operating problems (Cohen et al. 2012).

The comparison of operating expenses and own revenues are also important. The ratio of operating expenses to own revenues shows how the municipality can rely on its own revenues. The higher value of the ratios is preferred, the increase in the value of the indicator could lower their financial risks (STL/R). However, the municipalities are not able to finance their operation from their own revenues. They usually receive sources from the central government too. The reliance on the sources provided by the central government is measured by the ratio of subsidies to population (S/P). The amount of subsidies is affected by several criteria (Cohen et al. 2012). These criteria vary from country to country. The sixth ratio was of own revenues to population (R/P). The indicator expresses the financial autonomy of the municipality. Theoretically, there are some dependencies between the ratios, because four of the six ratios are contained the own revenues. However, the levels of the correlations were found to be moderate. The selected ratios do not include indicators related to profitability, because the results of the operation are heavily influenced by the subsidies received from the central government. The profitability indicators can be misleading in this context. In the case of the Greek municipalities, the analysis of the operating expenses are more informative (Cohen et al. 2012).

The authors defined relative rating, based on their simulations. The overall performance of a municipality was compared to the overall performance of all municipalities. The municipalities could receive a top, good, intermediate, poor, or very poor rating. The best 10% of the entities were labeled as top. The ratings are summarized in Table 10.

Table 10 Rating of the municipalities

Rating	**Percentile**
Top	90–100
Good	67–90
Intermediate	33–67
Poor	10–33
Very poor	0–10

Source: Cohen et al. (2012)

3.5. Cluster analyses

The research of Ziolo (2015) applied a cluster analysis approach. For the classification of the distressed municipalities, the Hellwig aggregate measure was used. As a first step, the inputs in the matrix had to be expressed, then a normalized matrix could be constructed by the means of standardization. As a third step, the stimulants and the destimulants had to be determined. After that distance of each municipality from the pattern and the worst alternative had to be computed. As a last step, synthetic measure (which value has to be between 0 and 1, where is the 1 most preferred value) can be determined. Ziolo applied nine explanatory variables (Table 11).

Table 11 Description of Variables

Description	Character
General Long-Term Debt as a Percentage of Taxable Value	Destimulant
General Fund Balance as a proportion of General Fund Revenues	Stimulant
Operational Surplus as a Percentage of Current Revenues	Stimulant
Population Decrease	Destimulant
General Fund Expenditure as a Percentage of Taxable Value	Destimulant
Budgetary Payables Divided by Budgetary Receivables Except Grants	Stimulant
Administrative Expenditures as a Percentage of Regional GDP	Destimulant
Fiscal Receivables Divided by Net Current Budgetary Receivables	Stimulant
Revenues from Federal and State as a Percentage of Total Revenues	Destimulant

Source: Ziolo (2015)

According to the Hellwig measure results, there were 3 groups created: fiscally distressed municipalities (Hellwig measure value under 0.1), fiscally neutral municipalities (Hellwig measure value between 0.1 and 0.5) and fiscally stable municipalities (Hellwig measure value above 0.5). Then different ratios were compared for each group to characterize the municipalities of the cluster (Ziolo 2015).

The fiscally stable municipalities have a high level of financial autonomy, they have a high level of independence of the income and expenditure. They are typically urban municipalities with significant revenues from property tax. They have the ability to create an operating surplus. However, the fiscally stable municipalities have liquidity risk arising from high investment activity (Ziolo 2015).

The fiscally neutral municipalities form a very heterogeneous group. They are more dependent on transfers than the fiscally stable ones. They have a controlled process of debt. The policymakers try to minimalize the cost of financing and keep the level of long-term debt at a safe level. The investment of the municipalities is significant. These fiscally distressed municipalities heavily depend on state transfers. In the case of rural municipalities, the agricultural tax is significant, too. The revenue from this tax is largely influenced by external conditions (Ziolo 2015).

4. Expectations regarding fiscal distress models

Financial distress models have to meet multiple expectations. Kloha et al. (2005) collected the shortcomings of the previous indicators. These remarks appear to be useful during the examination or the creation of a particular indicator. The problem with several indicators is the number of variables. If there are too many variables, they can provide an unclear result. Usually, it is possible for a local government to score poorly on a few of the indicators. In some cases, there is no clear guidance to evaluate the fiscal position of a local government (Kloha et al. 2005). Moreover, there can be dependencies between the variables. If there is a strong correlation between the variables, the multicollinearity has to be tested and processed.

Another problem can be the exclusion of the key variables. The examination of balance-sheet data is not always able to predict financial problems (Kloha et al. 2005), the signs of fiscal distress usually appearing after the emergence of the fiscal distress itself (Jones–Walker 2007). Besides the balance-sheet data, social and economic variables should be examined.

The indicators usually have an ambiguous expectation, in some cases, the local governments cannot meet with all of the requirements created by the fiscal distress indicators. Moreover, the indicators usually fail to allow divergence in preferences. The demand for services is not constant, it can differ from time to time. Another problem with the indicators can be their relativity. In the case of relative valuation, it has some governments to be at the bottom of the ranking, even if it is scored at an acceptable value. The reverse of that is also possible, someone has to be the best, even if all of the local governments score poorly (Kloha et al. 2005).

The indicators should be able to focus on one locality. Requiring all municipalities to be measured before a single government can be evaluated may not be a reasonable use of resources when alternatives that rely on the objective rather than relative performances. The reliance on relative performance is further complicated by the possibility that various government may not send their reports or audits on time. The replacement of the missing data is difficult because the late reporting from governments does not appear to be random, it is more likely to proceed from distressed ones (Kloha et al. 2005).

The measurement of fiscal distress should meet the following nine criteria:

- Theoretical validity, so that the components operationalize from theories of fiscal distress
- Predictive ability, with the help of the indicator, the financial emergencies could be prevented
- Relevance to the interest of the state
- The indicators use publicly available, standardized frequently collected data
- Historical sense of the progression of difficulty
- Accessible and easily understood by local officials and the public
- Resistant to manipulation or gaming
- Hope for those in distress and forgiveness for governments that are doing well generally
- Differentiate well among the governments evaluated (Kloha et al. 2005).

The universality of fiscal distress models is very limited. The tasks and the responsibilities of the local governments differ to a great degree from country to country. Moreover, we can find enormous differences in the way of working in the public sectors, local governments can use different resources, and the origin of these resources varies, too. The fiscal distress models have to be suitable for the particular public sector. This problem was highlighted by Ziolo (2015), too.

5. Conclusion

The examination of fiscal distress poses a lot of questions. First of all, the definition of fiscal distress should be specified. As the related literature shows, the interpretation of fiscal distress is not clear-cut. The binary classification of defaults (failed or not failed) is not suitable for the public sector. The regulation of bankruptcy in the public sector differs greatly from the private sector. Moreover, local governments have an essential role in the supply of different public services. Operationalization of fiscal distress is a key element in the research. The level of fiscal distress is usually measurable by two components: it can be examined through financial ratios and the parameters of the public services. Such parameters can be the quality of the services, the physical output of the services, or the efforts regarding them (amount spent on specific tasks). The financial ratios usually indicate the financial problems at a late stage, so different socio-economic factors should be involved as well. The local governments of different countries have to provide different services, this makes the creation of a model even harder. Furthermore, fiscal distress is inseparable from another phenomenon of society.

The selection of variables creates additional questions. The relevant financial ratios can alter. As the construction and the regulation of the public sectors are unique, the characteristic of the public sectors is very dissimilar. The local governments have different revenues, they get various rates and taxes. As a consequence, the different local government sectors are affected by different social or economic changes. The tax autonomy of the municipalities can differ too. Due to these determinants, the reliance on the own revenues can be dissimilar as well. As the own revenues of local governments differ, so the subsidies can play a different role too.

Besides revenues, the expenses and obligations can be different, too. The set of the examined variables should be customized for the particular public sector. Moreover, the manner of financing can vary as well. The regulation of the acquisition of loans and the controls related to them also play a key role. During the evaluation of the fiscal distresses this should be noted as well. Other institutional factors have an impact on financial distress. Regulation of budgeting, the transparency of the operation, and the sustainability of the financial plans are relevant in this context.

Along with the relevant variables and the institutional factors which affect the local governments, the availability of data determines the creation of a model. In some cases, the data for the theoretically valid variable is not collected very frequently. As a consequence, the set of the variable should be modified, or the value of the missing data should be calculated through approximations, or with the replacement with other related indicators. Both solutions could lower the validity of the results.

The examined models highlighted the importance of cash-flows and the fund balances. Besides that, the source of the revenues, and the debt-management-related indicators appeared to be significant. There was a significant overlap between the introduced models. However, there is lack of studies which compare the results of different fiscal distress models on the same sample. Moreover, there are not much cross-country studies in the literature of fiscal distress. The comparison of the data based on accrual accounting and cash-based accounting and the comparison of cash-based indicators with accrual-based indicators could be an unresolved problem, too.

The paper has a limitation of its own. There is not a shred of new empirical evidence involved, it remains in the theoretical ground, the assessment of the possible reactions to the fiscal distress are not included. The examined methods could be tested on the Hungarian local government sector, or a country-specific fiscal distress model could be created. This would provide at least some shred of evidence. As a further step, the relations of the fiscal distress can be examined with other characteristics of the local governments. The connection between fiscal distress and the quality of the internal control, or between fiscal distress and the quality of the management, would certainly result in some interesting conclusions.

References

Balog, E. – Jakab, Á. (2017): Az eredményszemléletű államháztartási számvitel bevezetése a nemzetközi tapasztalatok tükrében. In Farkas B. – Pelle A. (szerk): *Várakozások és gazdasági iterakciók.* JATEPress, Szeged, 81–94.

Beck, A. W. – Stone, M. S. (2017): Why municipalities fail: Implications for uncertainty disclosures. *Research in Accounting Regulation*, 29, 1, 1–9.

Brown, K. W. (1993): The 10-Point Test of Financial Condition: Towards an Easy-to-Use Assesment Tool for Smaller Cities. *Government Finance Review*, 9, 21–26.

Carlin, T. M. (2005): Debating the Impact of Accrual Accounting and Reporting in the Public Sector. *Financial Accountability & Management*, 21, 3, 309–336.

Carmelli, A. – Cohen, A. (2001): The financial crises of Local Authorities of Israel: A resource-based Analysis. *European Forum*, 79, 4, 893–913.

Christensen, M. – Newberry, S. – Potter, B. N. (2018): Enabling global accounting change: Epistemistic communities and the creation of a 'more business-like' public sector. *Critical Perspectives on Accounting*, 58, 53–76.

Clark, T. N. (1977): Fiscal Management of American Cities: Funds Flow Indicators. *Journal of Accounting Research*, 15,1, 54–94.

Cohen, S. – Doumpos, M. – Neofytou, E. – Zopounidis, C. (2012): Assessing financial distress where bankruptcy is not an option: An alternative approach for local municipalities. *Europen Journal of Operational Research*, 218, 1, 270–279.

European Comission (2013): *Report from the Commission to the Council and the European Parliament. Towards Implementing Harmonised Public Sector Accounting standards in Member States. The Suitability of IPSAS for Member States,* COM (2013), 114.: http://www.epsas.eu/en/documents/1_EN_ACT_part1_v5.pdf Accessed: 24/11/2018.

García–Sánchez, I. – Cuadrado – Ballesteros, B. – Frías–Aceuituno, J. – Morden, N. (2012): A New Predictor of Local Financial Distress. *International Journal of Public Administration*, 35, 11, pp. 739–748.

Gorina, E.– Maher, C.– Joffe, M. (2018): Local Fiscal Distress: Measurement and Prediction. *Public Budgeting & Finance*, 38, 1, 72–94.

Halmosi, P. (2013): The Effect of the Economic Crisis on Local Governments in OECD Countries. *Public Finance Quarterly*, 58, 3, 293–306.

Halmosi, P. (2018): Detroit város 2013-as csődesetének tanulságai. *Köz-Gazdaság*, 13, 2, 269–281.

Hood, C. (1995): The „new public management" in the 1980s: Variations on a theme. *Accounting, Organisation and Society*, 20, 2–3, 93–109.

Jones, S. – Walker, R. G. (2007): Explanators of Local Government Distress. *Abacus*, 43, 3, 396–418.

Kloha, P. – Weissert, C. S – Kleine, R. (2005): Developing and Testing a Composite Model to Predict Local Fiscal Distress. *Public Administration Review*, 65, 3, 313–323.

Lentner, Cs. (2014): A magyar önkormányzatok adósságkonszolidációja. *Pénzügyi Szemle*, 59, 3, 330–344.

Simon, J. (2011): *A központi és a helyi kormányzat információs rendszerének kérdései*. Ph. D. thesis, Pécsi Tudományegyetem, Pécs.

Skidmore, M. – Scorsone, E. (2011): Causes and consequences of fiscal stress in Michigan Cities. *Regional Science and Urban Economics*, 41, 4, 360–371.

Trussel, J. M. – Patrick, P. A. (2013): The Symptoms and Consequences of Fiscal Distress in Municipalities: An Ivestigation of Reduction in Public Services. *Accounting and Public Interest*, 13, 1, 151–171.

PWC (2015): Towards a new era in government accounting and reporting, 2nd edition. http://www.pwc.com/gx/en/industries/government-public-services/public-sectorresearch- centre/publications/government-accounting-and-reporting-pwc-global-ipsassurvey.html Accessed: 12/01/21019.

Wu, Y. – Gaunt, C. – Gray, S. (2010): A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics*, 6, 1, 34–45.

Ziolo, M. (2015): Diagnosing fiscal distress: Regional evidence from polish municipalities. *Romanian Journal of Fiscal Policy*, 6, 2, 14–33.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 326–336.

The new form of payout policy: the emergence and theoretical background of share repurchases

Dániel Szládek

The way companies return cash to their shareholders has changed considerably in recent decades. After changes in the legislation, share repurchases or buybacks have gained prominence, while dividends have somewhat lost their role as the main payout method. I intend to explore the reasons for this shift in payout policy.

In this paper, I focus on the theoretical background of share repurchases. Besides surveying the main theories related to share repurchases, I introduce the different ways companies can buy their shares back. Data indicating the emergence of share repurchases is also presented, while empirical studies testing the theories are discussed as well. The aim of this paper is to find out why share repurchases have become popular, and to reveal unanswered questions on the topic which point to future research possibilities.

Keywords: corporate finance, payout policy, share repurchases

1. Introduction

Corporate finance deals with three questions that companies face: how to optimally invest their resources (investment policy), what is the best way to fund these investments (financing policy), and how much money should be returned to shareholders (payout policy). In this study, I focus on the third element of corporate finance, highlighting the change in the way companies return cash to their shareholders.

Dividends were the primary means of transferring money back to shareholders for much of the twentieth century. The form of payout policy, however, has changed considerably in recent decades, as share repurchases or buybacks have gained significance. In this paper, my aim is to reveal why share repurchases have become more and more important in the payout decisions of companies. I show that dividends have somewhat lost their status as the main form of payout, while share repurchases are on the rise. I introduce the different ways companies can repurchase shares, also presenting data about the different methods. A substantial part of my paper focuses on the theories related to share repurchases, which intend to explain why companies buy their own shares back.

The structure of the study is as follows: Section 2 presents data showing that share repurchases have become significant in recent decades, Section 3 introduces the different share repurchasing methods available to companies, Section 4 reviews the literature about the theories related to share repurchases, and finally, Section 5 gives the concluding remarks.

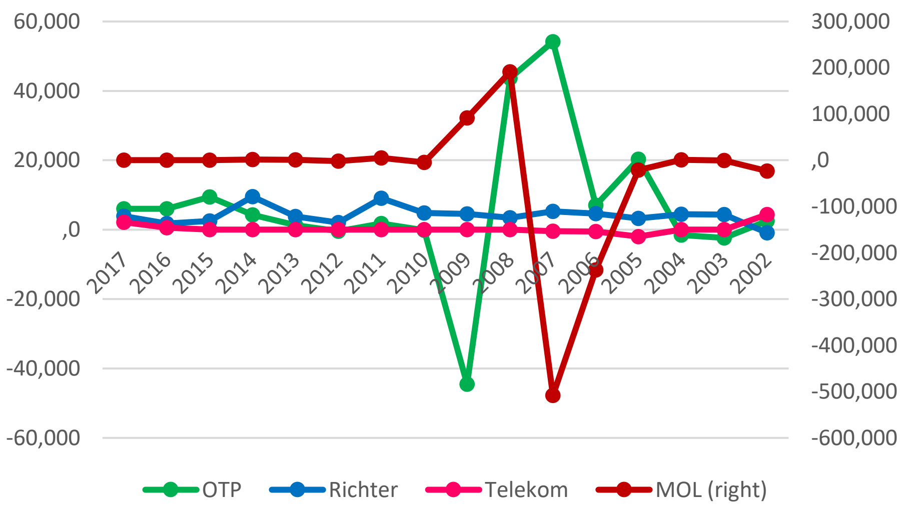
2. The emergence of share repurchases

As already mentioned, dividends were the primary means of transferring money back to shareholders for much of the twentieth century. However, Fama–French (2001) report that in 1973, approximately 52.8% of public companies in the United States (excluding utilities and financial services) paid dividends, the ratio peaked in 1978 (66.5%), then fell sharply: in 1999, only 20.8% of US public companies were dividend payers. This significant decline coincided with the emergence of share repurchases or buybacks. Farre-Mensa et al. (2014) show that the proportion of repurchasing companies has exceeded that of dividend-paying companies in the United States since 1997. A similar pattern is observed in absolute terms: the aggregate dollar amount of share repurchases has overtaken the aggregate dividends paid since 1997, thus buybacks have provided the greater part of the total payout of US public companies in the twenty-first century.

The shift in payout methods is quite apparent in the United States, yet is it a universal phenomenon? In Europe, the proportion of dividend-paying companies has also decreased towards the end of the twentieth century, although not as drastically as in the US (von Eije–Megginson 2008). Dividends remained the major form of payout for European companies at the start of the twenty-first century, although share repurchases have gained prominence in Europe as well.

As we can see, share repurchases have become highly relevant in the world's developed markets. Unfortunately, less developed or developing markets have not drawn the attention of researchers yet, at least concerning share repurchases. In Figure 1, I illustrate the share repurchasing activity of Hungarian blue-chip shares:

Figure 1 Share repurchasing activity of Hungarian blue-chip shares, million forints, 2002–2017



Source: Annual reports of the companies

The hand-collected data from the annual reports in Figure 1 indicate that the share repurchasing activity of Hungarian blue-chip shares has been mixed. For Telekom, buybacks have not been relevant at all. Richter has a steady level of share repurchasing activity, which corresponds to its employee compensation programs. Meanwhile, OTP and MOL have utilized share repurchases the most among the Hungarian blue-chips.

3. Share repurchasing methods

There are several ways in which companies can repurchase their shares. The different methods include fixed-price tender offer or self-tender offer, Dutch auction tender offer, open market repurchase program, privately negotiated share repurchases or targeted stock repurchases, and transferable put-rights distributions.

3.1. Fixed-price tender offer

One of the share repurchasing methods is the fixed-price tender offer. The fixed-price tender offer – as its name suggests – means that the company may buy shares back for a price fixed in advance. Apart from the price, the ratio of shares outstanding which can be repurchased is also set, as well as the duration of the program. This period may be extended, although usually the repurchase is conducted within five or six weeks (Hsieh–Wang 2009).

The repurchase price determined in advance is generally higher than the actual market price because shareholders can be motivated with this premium to sell their shares to the company. If there are too many shareholders willing to sell their shares, then the company may choose to buy more shares back, while in case of lower participation than expected, the repurchase program may be extended or the offer withdrawn. Usually, the management of the company does not participate in the program to preserve the credibility of the repurchase announcement (Hsieh–Wang 2009, Vafeas 1997).

3.2. Dutch auction tender offer

The Dutch auction tender offer is very similar to the fixed-price tender offer. The number of shares to be repurchased and the duration of the program are fixed in advance in this method as well. The difference is regarding the determination of the repurchase price: during a dutch Dutch auction tender offer, there is no fixed price, but rather, a range for the price. Shareholders participating in the repurchase program submit the price and the number of shares they are willing to sell to the company. The company sorts these offers in order, then sets the minimum price, per which it can buy back enough shares to reach its particular goal. Shares offered at or below the minimum price are bought back, while the other shares remain at the shareholders (Hsieh–Wang 2009, Vafeas 1997).

3.3. Open market repurchase program

Among share repurchase methods, open market repurchase programs are the most popular. This method has several distinctions from the above discussed tender offers.

First, the number of shares to be brought back is not determined beforehand, rather, the value of the program is announced. Second, open market repurchase programs are generally conducted through longer time periods, a program can last as much as several years. Third, one of the most significant differences from fixed-price and Dutch auction tender offers is that during an open market repurchase program, the company is not obliged to repurchase shares for the value announced earlier (Hsieh–Wang 2009, Vafeas 1997). Stephens–Weisbach (1998) show that in the United States sample an average of 74% to 82% of the announced value is repurchased in the three years after the announcement is made.

3.4. Privately negotiated share repurchases

The fourth method of repurchasing is the privately negotiated share repurchase. The essence of this method is that the company targets a shareholder or a group of owners who hold a great number of shares. Generally, the price includes a premium above the current market price, thus the term “greenmail’ is widespread. If the premium is big enough, then it is worth selling the shares back to the company, thus the “greenmail” reference. This method can be an effective measure to hinder a hostile takeover attempt, although it may also be used when there is no such event (Hsieh–Wang 2009).

3.5. Transferable put-rights distributions

The last share repurchasing technique is the transferable put-rights distributions. As the name suggests, this method means that the company distributes put options among the shareholders. This option grants the right to shareholders to sell their shares for a given price and maturity. The option can also be sold to other shareholders of the company (Hsieh–Wang 2009).

The distribution of transferable put rights has several advantages compared to other share repurchasing methods. Those shareholders will sell the option and retain their shares who have a higher reservation price concerning the shares. Consequently, shareholders with a lower reservation price will sell their shares back to the company, thus the ratio of shareholders with a higher reservation price will increase, which makes a potential hostile acquisition harder to pull off. Furthermore, investors face different tax rates, and the ability to sell the options means that they can decide to keep the shares or strike the options while optimizing tax payment as well (Hsieh–Wang 2009).

3.6. Statistics of share repurchase methods

Banyi et al. (2008) provide data for four of the above discussed methods in the United States: fixed-price tender offers, Dutch auction tender offers, open market repurchase programs, and privately negotiated share repurchases are analyzed in their work. Open market repurchase programs are the most popular among the share repurchase methods. Between 1985 and 2004, 84% of repurchases (12,931 events) were open market repurchase programs. In dollar amounts, about 90% (approximately 1,749 billion dollars) were paid through open market repurchase programs. Data for the other methods: fixed-price tender offer: 747 events and about 71 billion dollars, Dutch auction tender offer: 253 events and about 41 billion dollars, and privately negotiated repurchases: 1,369 events and about 94 billion dollars. Thus, we can conclude that open market repurchase programs dominate other repurchase methods.

4. Theories related to share repurchases

Research in corporate finance theory gained significance in the middle of the twentieth century. In their influential paper, Miller–Modigliani (1961) state that payout policy is irrelevant regarding the value of the enterprise. The theory is valid when crucial assumptions are made, including no taxes and transaction costs, competitive markets, informational symmetry, and rational investors. Naturally, these assumptions are not fulfilled in real markets, and most of the theories related to payout policy, and to share repurchases as well, originate from the violation of one or more of these assumptions (Farre-Mensa et al. 2014). These considerations and other relevant theories related to share repurchases are discussed below.

4.1. Taxes and regulation

One of the assumptions of Miller–Modigliani (1961) that definitely does not hold in real markets is that there are no taxes. Clearly, individuals and companies have to pay taxes on their income. Taxes are highly relevant for payout policy because dividends and share repurchases have different tax implications in many countries. Usually, income from dividends is taxed according to the personal or ordinary income tax rates, while earnings from buybacks are deemed as capital gains, to which different tax rates apply. Grullon–Michaely (2002) point out that even though dividends were at a tax disadvantage in the United States for much of the twentieth century, payout policy was still overly dominated by dividends and share repurchases played only a marginal role.

Why were dividends preferred instead of share buybacks despite the tax disadvantage? The answer can be found in regulatory measures. The Securities Exchange Act of 1934 regulated share repurchases in the United States. The Securities Exchange Commission (SEC) wanted to prevent companies from manipulating the price of their own shares, but the regulation turned out to be too rigorous, thus companies avoided buybacks and paid dividends instead. However, in 1982, the SEC adopted Rule 10b-18, which eased the regulation (Grullon–Michaely 2002). Since then, share repurchases have gained prominence, as I have shown above.

Regulation is a key factor in payout decisions in other parts of the world as well. For example, Lee et al. (2010) report that share repurchases were prohibited in many European countries in the nineties, which may be one of the reasons why buybacks have not been as dominant lately in the continent as in the United States. Andriosopoulos–Lasfer (2015) assess three European countries, the United Kindom, France, and Germany, their results indicating that local institutional environment and regulatory measures significantly influence the amount of and approach to share repurchases.

In a unique empirical study, Brav et al. (2005) surveyed 384 financial chiefs and interviewed a further 28 company officials to uncover the motivations of management regarding dividends and buybacks. Concerning the tax implications of payout policy for their investors, managers replied that tax considerations are only of secondary importance when payout decisions are made.

4.2. Signaling and undervaluation theory

The informational asymmetry observed in capital markets is a violation of the Miller–Modigliani (1961) assumptions. This asymmetry arises from the fact that the managers of the company have more information about the business they run than the shareholders who own the company. The management may utilize dividends or share repurchases as a costly signal to the market that the company is operating appropriately, and even foreshadow improving future performance (Miller–Rock 1985). Share repurchases are especially suitable signals according to the undervaluation theory: managers of the company may initiate buyback programs when they think the shares of the company are undervalued in the market (Ofer–Thakor 1987).

Ofer–Thakor (1987) summarize the stylized facts about the effects of the announcement of dividends and share repurchases: (1) announcement of dividends and buybacks have a significant positive effect on share price, (2) announcements of repurchases generate a greater response in share price than those of dividends do, (3) companies repurchase shares with a premium, thus the buyback price is higher than the earlier market price, (4) in many cases, share price drops after the repurchase is finished, (5) despite this drop, the share price after the repurchase is still greater than it was before the buyback.

Ofer–Thakor (1987) provide a theoretical framework, in which they prove that the signaling theory holds for dividends and share repurchases alike. They also show that the announcement of buybacks conveys more information than that of dividends does. Howe et al. (1992) utilized fixed-price repurchase offers and special dividends to test – among other things – the signaling theory, and their evidence suggests that the theory holds. Vafeas (1997) reveals a subtle difference between different repurchase methods: he argues that fixed-price tender offers do provide a signal, while open market repurchase programs can be employed to exploit short-term undervaluation. Dittmar (2000) analyzes a sample ranging from 1977 to 1996 to simultaneously test several hypotheses. Her Tobit model results indicate that the undervaluation theory is valid for every year in her sample.

Contradicting these theoretical and empirical results, managers do not believe they use share repurchases as a costly signal to the market. However, they gladly repurchase shares, when the shares are deemed undervalued (Brav et al. 2005).

4.3. Agency costs and the theory of free cash flows

In most cases concerning publicly traded companies, the management and the owners of the company are not the same group of people, which can lead to conflicting interests between these sides. Agency costs arise when the management of the company does not operate the business in the best interest of the shareholders but seeks to maximize its own wealth or influence (Damodaran 2014).

Easterbrook (1984) and Jensen (1986) uncover the relationship between agency costs and payout policy. At the time, dividends were the main instruments of returning cash back to the shareholders, thus the researchers focused solely on dividends. They argue that increasing dividends reduce the cash balance available to managers, thus leaving them with limited resources.

Free cash flows that are not returned to the owners can damage shareholder value in two ways (Hsieh–Wang 2009). First, the available cash reserves could be used to invest in projects which have a negative net present value. This overinvesting behavior reduces the value of the enterprise, thus shareholder value is decreased as well. Second, cash balances could be utilized to directly serve the interests of the management, for example increasing managers' compensation, which reduces free cash flow available to investors (Hsieh–Wang 2009).

Howe et al. (1992) analyze the Tobin Q measure of companies, which is the ratio of the company's market value and the replacement cost of its assets. Companies with lower Q values are the overinvesting companies, while a high measure indicates successful value maximization. Howe et al. (1992) find that the two groups of companies react approximately the same to announcements of share repurchases, thus they reject the theory of free cash flows.

Perfect et al. (1995), however, oppose the results of Howe et al. (1992). They believe that the rejection or acceptance of the free cash flow hypothesis depends greatly on the calculation method of the Tobin Q measure, and criticize that Howe et al. (1992) used three years of data for the calculation of the measure, instead of considering the latest year, which is the standard procedure in the literature. Perfect et al. (1995) thus argue that the theory of free cash flows cannot be rejected.

Vafeas–Joy (1995) evaluated 162 share repurchase announcements between 1985 and 1991. They also utilized the Tobin Q measure, and their results indicate that Jensen's (1986) theory holds, thus the free cash flows hypothesis is valid. Dittmar (2000) reports evidence that the theory holds for several years in her sample as well, further proving the free cash flows theory.

Perhaps unsurprisingly, the Brav et al. (2005) paper contradicts most of the above-mentioned empirical results. The surveyed and interviewed managers, perhaps unsurprisingly, do not believe that the agency theory is adequate. Naturally, it is not expected that managers would willingly admit running the business in a way that is not entirely in conjunction with the interests of the shareholders.

4.4. Substitution hypothesis

The way companies return cash to their shareholders does not matter in the Miller–Modigiliani (1961) framework, as because of the no taxes and transaction costs assumptions, the value of the company is indifferent to payout decisions The substitution hypothesis originates from this classic proposition, which suggests that dividends and share repurchases are perfect substitutes.

As discussed above, these assumptions do not hold in real markets, thus the substitution hypothesis has been tested on several occasions. Grullon–Michaely (2002) inquire whether the emergence of buybacks at the end of the twentieth century was caused by companies replacing dividends with share repurchases. Applying Lintner's (1956) dividend forecasting model to test the relationship between dividends and share repurchases, they prove that the substitution theory holds.

Jiang et al. (2013) also report evidence justifying the substitution theory. They utilize the catering theory of Baker–Wurgler (2004), which states that managers cater to the investors' demand for dividends when dividend-paying shares trade with a higher price than other companies' shares. Jiang et al. (2013) find that the premium related to share repurchases (dividends) has a negative effect on the willingness to pay dividends (repurchase shares), which supports the substitution theory.

Most of the empirical studies related to share repurchases have been conducted using data from the United States. Andriosopoulos–Hoque (2013), however, test the substitution hypothesis in three European countries, the United Kingdom, France, and Germany, considering open market repurchase programs and the relevant legal system in these countries. The researchers report that the substitution hypothesis can be rejected in the UK and Germany, but the evidence backs the theory in France (Andriosopoulos–Hoque 2013).

Opposing the above results, Dittmar (2000) finds that the substitution theory does not hold in most of the years in her sample. Brav et al. (2005) claim that managers do not think of dividends and share repurchases as perfect substitutes for each other. They explained that dividends are sticky and hard to reduce without negative consequences to share price, while buybacks provide flexibility in determining the amount of cash returned to shareholders.

4.5. Other motivations for share repurchases

The previous sections discussed the theories related to share repurchases which can be derived from the violation of one or more of the Miller–Modigliani (1961) assumptions. There are, however, other factors which may motivate managers to repurchase shares. The motivations considered below are takeover deterrence and the avoidance of earnings per share dilution.

Share repurchases may hinder hostile takeover attempts through various channels. First, according to the signaling and undervaluation theory, the announcement of a repurchase program increases the price of the shares, thus the potential acquisitor would have to offer a higher price for them. Second, owners of the company who sell their shares in a repurchase program have a lower reservation price regarding the share price, leading to an ownership structure where shareholders with a higher reservation price remain in place. The remaining owners would only accept a higher potential bid from the acquisitor, thus a buyback program may reduce the possibility of a successful hostile takeover (Hsieh–Wang 2009).

Denis (1990) analyzed the payout policy of targeted companies, which faced hostile acquisitors. He reports that share repurchases and special dividends are effective tools against hostile acquisition attempts. Lee et al. (2010) show that takeover deterrence is a relevant motive for share repurchases in the European countries evaluated. Dittmar (2000), however, finds that hindering hostile takeovers is a significant factor only in some years of her sample.

Companies sometimes compensate their workforce with stock options. If these options are exercised, then the company has to issue new shares, and the increased number of total shares causes the earnings per share (EPS) to decline. Many managers have EPS as one of their performance indicators, thus they are motivated to reduce the number of shares outstanding, which can be achieved by buying shares back (Farre-Mensa et al. 2014).

The surveys and interviews conducted by Brav et al. (2005) reveal that avoiding earnings per share dilution is indeed one of the major motivations of managers initiating share repurchase programs.

5. Conclusion

The payout decision is one of the major questions of corporate finance. The way companies return cash to their shareholders has changed significantly over the last few decades. Share repurchases or buybacks have gradually replaced dividends as the main form of payout in the United States, and have gained prominence in other capital markets of the world as well.

The importance of share repurchases has motivated researchers and professionals alike to discover why payout policy is shifting from dividends to buybacks. Several theories originate from the violation of the Miller–Modigiliani (1961) assumptions, such as taxes and regulatory considerations, signaling and undervaluation theory, the agency costs and theory of free cash flows, and the substitution hypothesis. These theories are discussed above, as well as other motivating factors, for example, takeover deterrence and the avoidance of earnings per share dilution.

Overall, empirical studies testing the hypotheses often report contradictory results, and thus no dominant theory relating to share repurchases has arisen. Further research is needed to better understand the motivations behind buybacks and reduce this ambiguity. Moreover, existing research focuses mainly on the United States and some other developed countries, while less developed and developing markets have not been in the center of attention, an area which is open for new research. Classical valuation theory and methods concentrate on shareholder cash flows of dividends, but the emergence of share repurchases may prompt the reassessment of such models, highlighting the theoretical and practical significance of share repurchases or buybacks.

Acknowledgments

This research was partially supported by the Pallas Athene Domus Educationis Foundation. The views expressed are those of the author's and do not necessarily reflect the official opinion of the Pallas Athene Domus Educationis Foundation.

References

Andriosopoulos, D. – Hoque, H. (2013): The determinants of share repurchases in Europe. *International Review of Financial Analysis*, 27, C, 65–76.

Andriosopoulos, D. – Lasfer, M. (2015): The market valuation of share repurchases in Europe. *Journal of Banking & Finance*, 55, C, 327–339.

Banyi, M. L. – Dyl, E. A. – Kahle, K. M. (2008): Errors in estimating share repurchases, *Journal of Corporate Finance*, 14, 4, 460–474.

Baker, M. – Wurgler, J. (2004): A Catering Theory of Dividends. *Journal of Finance*, 59, 3, 1125-1165.

Brav, A. – Graham, J. R. – Harvey, C. – Michaely, R. (2005): Payout policy in the 21st century. *Journal of Financial Economics*, 77, 3, 483–527.

Damodaran, A. (2014): *Applied Corporate Finance*, *4th Edition.* Hoboken, NJ: Wiley.

Denis, D. J. (1990): Defensive Changes in Corporate Payout Policy: Share Repurchases and Special Dividends. *Journal of Finance*, 45, 5, 1433–1456.

Dittmar, A. K. (2000): Why Do Companies Repurchase Stock? *The Journal of Business*, 73, 3, 331–355.

Easterbrook, F. H. (1984): Two Agency-Cost Explanations of Dividends. *The American Economic Review*, 74, 4, 650–659.

Fama, E. – French, K. (2001): Disappearing dividends: changing company characteristics or lower propensity to pay? *Journal of Financial Economics*, 60, 1, 3–43.

Farre-Mensa, J. – Michaely, R. – Schmalz, M. (2014): Payout Policy. *Annual Review of Financial Economics*, 6, 1, 75–134.

Grullon, G. – Michaely, R. (2002): Dividends, Share Repurchases, and the Substitution Hypothesis. *Journal of Finance*, 57, 4, 1649–1684.

Howe, K. M. – He, J. – Kao, G. W. (1992): One-Time Cash Flow Announcements and Free Cash-Flow Theory: Share Repurchases and Special Dividends. *Journal of Finance*, 47, 5, 1963–1975.

Hsieh, J. – Wang, Q. (2009): *Stock Repurchases: Theory and Evidence.* http://dx.doi.org/10.2139/ssrn.1395943 Accessed: 2018.11.05.

Jensen, M. (1986): Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *The American Economic Review*, 76, 2, 323–329.

Jiang, Z. – Kim, K. – Lie, E. – Yang, S. (2013): Share repurchases, catering, and dividend substitution. *Journal of Corporate Finance*, 21, C, 36–50.

Lee, C. I. – Ejara, D. D. – Gleason, K. C. (2010): An empirical analysis of European stock repurchases. *Journal of Multinational Financial Management*, 20, 2–3, 114–125.

Lintner, J. (1956): Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes. *The American Economic Review*, 46, 2, 97–113.

Miller, M. – Modigliani, F. (1961): Dividend Policy, Growth, and the Valuation of Shares. *The Journal of Business*, 34, 4, 411–433.

Miller, M. – Rock, K. (1985): Dividend Policy under Asymmetric Information. *Journal of Finance*, 40, 4, 1031–1051.

Ofer, A. R. – Thakor, A. (1987): A Theory of Stock Price Responses to Alternative Corporate Cash Disbursement Methods: Stock Repurchases and Dividends. *Journal of Finance*, 42, 2, 365–394.

Perfect, S. B. – Peterson, D. R. – Peterson, P. P. (1995): Self-tender offers: The effects of free cash flow, cash flow signalling, and the measurement of Tobin's q. *Journal of Banking & Finance*, 19, 6, 1005–1023.

Stephens, C. P. – Weisbach, M. (1998): Actual Share Reacquisitions in Open-Market Repurchase Programs. *Journal of Finance*, 53, 1, 313–333.

Vafeas, N. (1997): Determinants of the Choice between Alternative Share Repurchase Methods. *Journal of Accounting, Auditing & Finance*, 12, 2, 101–124.

Vafeas, N. – Joy, M. (1995): Open market repurchases and the free cash flow hypothesis. *Economic Letters*, 48, 3–4, 405–410.

Von Eije, H. – Megginson, W. L. (2008): Dividends and share repurchases in the European Union. *Journal of Financial Economics*, 89, 2, 347–374.

Chapter IV

International relations

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 338–356.

Income Inequality and Economic Growth:An Empirical Analysis of Kenya

John Kibara Manyeki – Balázs Kotosz

Despite the well-known potential benefits of equality of income in economic growth, the statistics for Kenya show that income is heavily skewed in favour of the rich and against the poor with the country's top 10% households controlling 42% of the total income while the bottom 10% controls less than 1%. The existing empirical evidence for the Kenyan economy does not shed light on whether there is a robust association between income inequality and growth over time. This paper provides an empirical investigation on the relationship between income inequality and economic growth and the hypothesis addressed was: inequality is harmful for growth. We contribute to the literature by employing an autoregressive distributed lag model using a time series date spanning from 1990–2015. The study found a significant positive but weak long run relationship between income inequality and growth. The short run was a strong positive relationship, which was significant at 1% level. This income inequality favours the rich, and therefore, to ensure fair distribution of wealth and a balanced growth, a policy goal should be equity in income distribution to reduce excessive income disparities. More research should be carried out on all other measures of inequality, to bring to light, which among them is more influential to GDP growth.

Keywords: GINI coefficient, ARDL estimation, income inequality, GDP growth, Kenya

1. Introduction

Why does a country like Kenya in different periods grow at such fluctuating different rates (Figure 1)? Can the growth trend be associated with the level of income inequality (Figure 2)? Or in short, is there a relationship between income inequality and economic growth? This question of income inequality and economic growth has been a major concern for social scientists and frustrated our political energies in Kenya for decades. Ever since Kenyan independency, the guise of equality of wealth and income has been the guiding motive of successive Kenyan governments. This has seen the introduction of such measures as free primary education, constituency development funds, and other sector-level reform initiatives to reduce income inequality, hence stimulate growth (Constitution of Kenya 2010, National Treasury 2015). However, statistics for Kenya show that income inequality is ever increasing and is heavily skewed in favour of the rich and against the poor, with the country's top 10% households controlling 42% of the total income while the bottom 10% controls less than 1% (SID 2004).

Figure 2 Trend of GDP growth in percentage

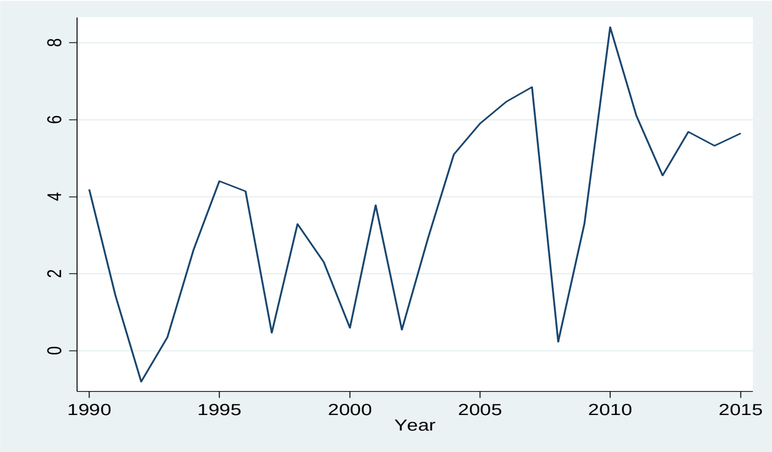
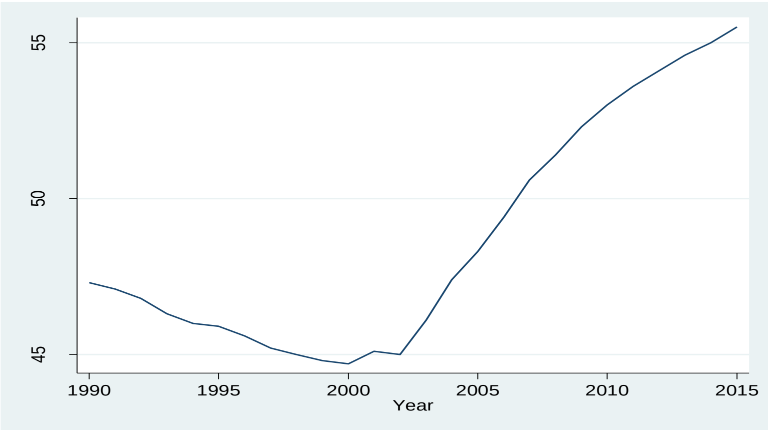


Figure 3 Trend in GINI Coefficient in percentage



Source: Own construction based on United Nations Development Programme (UNDP) and World Bank World Development Indicators (WDI) databases

On the effect of inequality on growth, a contradictory view has been gaining currency. Whereas the convectional textbook approach is that inequality is a good incentive for growth (Kuznets 1955), essentially insofar as it generates an incentive to work and invest more, or can trigger more investment, given that high-income groups tend to save and invest more. On the other hand, development economies have long expressed counter-arguments although not in a formalized way (Todaro 1992, Galor 2009). One of the main arguments by Cingano (2014) and Campos (2017) was that greater inequality can reduce the professional opportunities available to the most disadvantaged groups in society and therefore decrease social mobility, limiting the economy's growth potential. Similarly, greater inequality can also negatively affect growth if, for example, it encourages populist policies (Eisler 2016, Garcia and Arenas 2017) or leads to an excessive rise in credit, which ends up acting as a brake on growth (Morron 2017). While classical and neoclassical approaches have underlined a growth-promoting effect, modern perspectives highlight a potential growth-dampening impact of inequality on growth, the question of which of these effects are predominant depends strongly on the degree of income inequality already reached (Galor 2009). In this paper, as our empirical aim is to identify differentiated negative and/or positive effects income inequality has on Kenyan economic growth, we were, therefore, tempted to follow the position of the modern economists and summarize our tentative hypothesis in a simple statement: inequality is harmful for growth.

Before subjecting this hypothesis into empirical investigation, it would be prudent to provide some insight into various forms of inequalities associated with economic growth. Inequality is observed not only in incomes, but also in terms of social exclusion and the inability to access social services and socio-political rights by different population groups, genders and even races. According to Mount (2008) and Rugaber and Boak (2014), inequality can, therefore, be classified as, 1) Income inequality which measures the gap between the rich and the poor or people with similar background, status, qualifications but with different incomes, 2) Gender inequality as manifested in wages, discrimination, domination of positions of power and responsibility – it limits the extent to which women or men can make it to the top, 3) Opportunity inequality, which is measured in terms of ease of access to education, work, and housing, markets on the basis of race, ethnicity or gender, even across countries, and 4) Asset/wealth inequality, which measures the disparity not just in quantity but also in quality of natural resources, infrastructure, raw materials, and amount of human capital and assets. The focus of this article will be income inequality and its contribution to economic growth[1].

Why the focus on income inequality rather than some other measurable quantity? The reason is twofold: firstly, income as a proxy for economic welfare, and secondly, income as command over resources (Cowell 2007). In this article, income inequality is, therefore, defined as the degree to which distribution of economic welfare generated in an economy differs from that of equal shares among its inhabitants which means that one segment of the population has a disproportionately large share of income compared to other segments of that population (SID 2004). It may also entail comparison of certain attributes or well-being between two persons or a group of people and the differences in share of these attributes. However, the existing empirical evidence for the Kenyan economy does not shed light on whether there is a robust association between income inequality and growth over time. Therefore, this paper attempt to establish whether any relationship between income inequality and GDP growth really exists. In particular, our contribution is the use of autoregressive distributed lag to empirically assess whether and/or how inequality affects economic

[1]Note economic growth will be used interchangeably with GDP growth

growth. These results are crucially important for policy makers, as their challenge is to find out how, and not just if, inequality is affecting the process of economic growth. The remainder of the paper is organized as follows. Section II briefly reviews the theoretical foundation of income inequality on economic growth. This section also reviews some empirical literature on the relationship of income inequality and economic growth for Kenya and other parts of the world. Section III sets out the methodology, including the empirical model and database used in this study. Section IV displays the main results, which include tests result and empirical findings. Finally, section V concludes the paper.

2. Theoretical foundation of income inequality analysis

The most influential contribution in modern economic literature addressing explicitly the issue of economic inequality is grounded on the theoretical foundation developed by Kuznets (1955). Basing on empirical evidence, Kuznets maintains that inequality tends to rise in the early stages of economic development, as a consequence of industrialization, and then it declines in later stages, as capitalism matures. In this way income inequality presents the classical inverted-U shaped trend in time. Kuznets describes a positive relationship between income inequality and economic growth in the early phases of growth and a negative relationship in the later phases.

The research by Kuznets laid a significant foundation for studying the relationship between economic growth and income inequality. Beyond the theoretical sphere, many authors have attempted to provide empirical evidence of inequality's effects on economic growth although with contradictory findings. While some scholars found income inequality as negatively associated with growth in the long run (e.g. Alesina–Rodrik 1994, Perotti 1996, Panizza 2002, Nel 2003), other scholars found mixed results comparing income inequality and growth for developed and developing countries. For instance, Barro (2000) and Voitchovsky (2005) found a negative relationship between inequality and growth for poorer countries, but a positive relationship in the case of richer countries. However, this positive impact relies on panel data analysis and is either associated with short-term economic growth (Forbes 2000) or is dependent on national income (Barro 2000), on the initial income distribution itself (Chen 2003), on the profile of inequality (Voitchovsky 2005), or on the process of urbanization (Castells-Quintana–Royuela 2014).

In the stream of income inequality and GDP growth analysis, the critics of Kuznets' hypothesis question the nature of causality between the two variables, especially on the basis of empirical economic literature. The most relevant finding was that of Fields (2002) who observed that it is not economic growth per se which gives rise to economic inequality but it is the nature of economic growth which determines the development of inequality. In particular, Fields claimed that the effect of growth on inequality depends on the size and structure of the economy (which can be classified as a developed or developing economy).

In Kenya, there is scanty information in the literature on the relationship between income inequalities and economic growth. Most of the studies focus on explaining the contribution of income inequality on poverty levels experienced in the country. By way of example, Njuguna (2005) investigated the extent of poverty and the level of inequality in the Kenyan economy. The author compared changes in poverty and inequality between regions and their robustness using stochastic dominance analysis. The study used Welfare Monitoring Survey data spanning from 1994 and 1997 to shed some light on the intertemporal patterns of changes in welfare levels and distribution in Kenya across geographical and socio-economic groupings of policy interest. The author found that for a wide range of poverty lines, poverty and inequality increased in Kenya over the period. Along the same lines, a study by Suri et al. (2008) using time series data that ranged from 1997–2007 found that income inequality has been declining. The salient recommendation of these two studies was that poverty reduction requires economies to address inequality and economic structures – in addition to sustaining high levels of economic growth.

A more recent study by Gakuru and Mathenge (2012) followed the same line and seeks to highlight the levels of income inequality in Kenya and its implications on various policy options targeted at reducing poverty. The study applied the 2003 Kenyan SAM to develop a multiplier simulation model which tracks the linkages among the demand-driven shocks on economic growth, income generation, and consequently income distribution implications on different economic groups. The empirical results from the multiplier analyses show that due to high inequality in Kenya, stimulation of growth mainly benefit the richest urban household deciles, who own most of the factors of production. The authors recommended that Kenya will need to focus not only on economic growth, but also on inequality in order to effectively tackle poverty in the country.

The study of the impact of income inequality on the country's economic growth is highly relevant today. However, we were only able to discover one paper that tried to include income inequality proxies in the growth regression. This study was carried by Wanyagathi (2006) who used an ordinarily least squares estimation procedure to investigate the relationship between income inequality and economic growth using time series data spanning from 1950–2008. The author used control variables such as total expenditure on education, health, and population growth. Although this study did not consider the problems associated with time series data (the presence of multicollinearity, heteroscedasticy, autocorrelation etc.) the study provides some insight concerning the relationship between income inequality and growth. The author found that GINI coefficient, which is the measure of income inequality, to be negatively related to growth, which contradicts the Kuznets hypothesis. The author associated this to the social problems linked to inequality for the developing economy such as Kenya's, include corruption, civil wars, political instability and many issues.

Broadly speaking, as indicated in the reviewed literature, there is no single, universal mechanism behind the relationship between inequality and growth. Nevertheless, the reviewed literature has provided mechanisms supporting both possibilities, and the empirical literature attempting to discriminate between these mechanisms in the Kenyan situation has been largely inconclusive. Further, the econometric method employed in most studies was ordinary regression using OLS estimation technique, which does not take in account the problem associated with multicollinearity, heteroscedasticy, autocorrelation, or model specification among other things. Even if the results of the previous studies provide an important step in understanding the impact income inequality has on economic growth, given these problems, the findings of these studies can be considered preliminary in nature. Establishing a relationship between inequality and economic growth can be severely obstructed by inadequate econometric techniques. We tried to overcome this problem by applying autoregressive distributed lag (Nkoro–Uko 2016) model, specifically system ARDL, which improves the ability to handle endogeneity and avoids problems resulting from non-stationary time series data typically found in economic growth regressions. (Laurenceson–Chai 2003, Manyeki–Kotosz 2017). Further, the ARDL model was considered because it is flexible and combines both short run and long run effect into a single equation, which was the aim of the paper.

3. Methodological Framework and Data

The main objective of the research is to find out the relationship between inequality in income and growth of the economy in Kenya. Ultimately, in order to test our hypothesis that inequality is harmful to growth, we needed to optimize the model to capture both the short-run and long-run effects, and therefore this section continues with discussion of the empirical model and data applied.

3.1. Empirical model: ARDL Approach

Before specifying the model for use in this analysis we had to investigate whether the variables are stationary or not. The stationarity of the variables was examined to avoid the existence of spurious estimation results. Stationarity can be done in two ways; 1) KPSS test for stationarity that consider the null hypothesis H_0 that the series is stationary, and 2) unit root tests, such as the Dickey-Fuller test and its augmented version, the augmented Dickey–Fuller test (ADF) (Dickey and Fuller, 1979, 1981), or the Phillips–Perron test (PP) (*Phillips and Perron 1988)*, for which the null hypothesis H_0 is the opposite, that the series possesses a unit root and hence is not stationary. In this study, unit root test was adopted and both ADF and PP test conducted. Both ADF and PP tests the null hypothesis of a unit root in a time series sample. If a series is stationary without any differencing, it is said to be $I(0)$ or integrated of order 0. On the other hand, if a series is stationary after d-difference, it is said to be $I(d)$ or integrated of order d.

The second step involved is cointegration test. Several methods are available for conducting the cointegration test and the most commonly and widely used methods include the residual based Engle-Granger (1987) test, the maximum likelihood test based on Johansen (1991, 1998) and Johansen and Juselius (1990) tests. Due to the low power and other problems associated with these test methods, the OLS based autoregressive distributed lag (ARDL) approach to cointegration was adopted. The main advantage of ARDL modelling lies in its flexibility, in particular that it can be applied when the variables are of different order of integration (Pesaran–Pesaran 1997). Compared to other cointegration test approaches that require order of integration of the variables to be determined first which may lead to misclassification of variables as $I(0)$ or $I(1)$, a ARDL uses a bounds testing procedure to draw conclusive inference without knowing whether the variables are integrated of order zero ($I(0)$) or one ($I(1)$) (Pesaran et al. 2001). Another advantage of this approach is that the model takes sufficient numbers of lags to capture the data generating process in a general-to-specific modelling framework (Laurenceson–Chai 2003), and this accounts for the autocorrelation issue. Its popularity also stems from the fact that cointegration of nonstationary variables is equivalent to an error-correction (EC) process[2], and the ARDL model has a reparameterization in EC form (Engle and Granger 1987, Hassler and Wolters 2006). The EC integrates the short-run dynamics with the long-run equilibrium without losing long-run information, and the existence of a long-run cointegrating relationship can be tested based on the EC representation. In addition, it is also argued that using the ARDL approach avoids problems resulting from non-stationary time series data (Laurenceson–Chai 2003).

The existence of the long-run relation between the variables under investigation is tested by computing the Bound F or t statistic (bound test for cointegration) in order to establish a long run relationship among the variables. This bound for t-statistic is carried out on each of the variables as they stand as endogenous variables while others are assumed as exogenous variables. This approach is illustrated by using an ARDL (p,q) regression with an I(d) regressor as follows

$$\begin{aligned} GDPG_t = C_0 + \beta_1 GDPG_{t-1} + \cdots + \beta_p GDPG_{t-p} + \alpha_0 GINICoeff_t \\ + \alpha_1 GINICoeff_{t_{t-1}} + \cdots + \alpha_q GINICoeff_{t_{t-q}} + \mu_t \end{aligned}$$

or

$$\begin{aligned} GINICoeff_t = C_0 + \beta_1 GINICoeff_{t-1} + \cdots + \beta_p GINICoeff_{t-p} + \alpha_0 GDPG_t + \\ \alpha_1 GDPG_{t-1} + \cdots + \alpha_q GDPG_{t-q} + \mu_t \qquad (1) \end{aligned}$$

Where I =1,2,…T and $\mu_t \sim iid(0, \sigma^2)$, C_0 is the drift and $GINICoeff_t$ and $GDPG_t$ are the gross real GINI coefficient in percent and gross domestic product growth in percentage, respectively.

[2]A dynamic error correction model (ECM) can be derived from ARDL through a simple linear transformation (Banerjee et al. 1993)

The GINI coefficient is the measurement of the income distribution in a country. It is based on a scale from 0 to 100% with 0 being perfect equality and 100% representing perfect inequality. The $I(d)$ process can be generated by

$$GDPG_t = GDPG_{t-1} + \varepsilon_t \text{ Or } GINICoeff_t = GINICoeff_{t-1} + \varepsilon_t \tag{2}$$

Note u_t and ε_t are uncorrelated for all lags such that $GDPG_t$ or $GINICoeff_t$ is strictly exogenous with respect to u_t.ε_t is a general linear stationary process. In practice the ARDL ($p, q_1, q_2 \ldots\ldots q_k$) model for cointegration testing is expressed as;

$$\Delta GDPG_t = C_0 + \sum_{i=1}^{k} \beta_i \Delta GDPG_{t-i} + \sum_{j=0}^{k} \alpha_j \,\Delta GINICoeff_{t-j} + \delta_1 GDPG_{t-1} + \delta_2 GINICoeff_{t-1} + v_{1t}$$

or

$$\Delta GINICoeff_t = C_0 + \sum_{i=1}^{p} \beta_i \Delta GINICoeff_{t-i} + \sum_{j=0}^{q} \alpha_j \,\Delta GDPG_{t-j} + \delta_1 GINICoeff_{t-1} + \delta_2 GDPG_{t-1} + v_{1t} \tag{3}$$

Here, k is the ARDL model maximum lag order and chosen by the user. The F-statistic is carried out on the joint null hypothesis that the coefficients of the lagged variables (δ_1, δ_2) are zero. The null of non-existence of the long-run relationship is defined by; H_0: $\delta_1 = \delta_2 = 0$ (null, i.e. the long run relationship does not exist) H_1: $\delta_1 \neq \delta_2 \neq 0$ (Alternative, i.e. the long run relationship exists). The model is "autoregressive", in the sense that $GDPG_t$ or $GINICoeff_t$ is "explained (in part) by lagged values of itself. Pesaran et al. (2001) provide lower and upper bounds for the asymptotic critical values depending on the number of regressors, their order of integration, and the deterministic model components based in F-test or t-test. Based on Pesaran et al. (2001), you fail to reject the null H_0^F or H_0^t respectively if the test statistic is closer to zero than the lower bound of the critical values, and reject the null H_0^F or H_0^t respectively if the test statistic is more than the extreme upper bound of the critical values. The existence of a (conditional) long-run relationship is confirmed if both null H_0^F or H_0^t are rejected. If a long run relationship exists between the underlying variables, while the hypothesis of no long run relations between the variables in the other equations cannot be rejected, then ARDL approach to cointegration can be applied. The optimal lag orders *p* and *q* (possibly different across regressors) can be obtained with proper model order selection criterion, e.g. the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) or Hannan–Quinn Criterion (HQC). For this case, we adopted AIC criteria

Having confirmed that there exists a long-run relationship among variables, then the ARDL model can be reparameterization in conditional ECM form as follows;

$$\Delta GDPG_t = C_0 - \gamma(GDPG_{t-1} - \vartheta GINICoeff_{t-1}) + \sum_{i=1}^{p-1} \varphi_{GDGG_i} \Delta GDPG_{t-i} + \sum_{j=0}^{q-1} \varphi_{GINICoeff_i} \Delta LnGINICoeff_{t-j} + \mu_t \quad (4)$$

Where μ_t is a random "disturbance" term (white noise error term) and C_0, $GDPG_t$ and $GINICoeff$ are as defined earlier, with the speed-of-adjustment coefficient $\gamma = \sum_{j=1}^{p} \varphi_i$; and the long-run coefficients $\vartheta = \frac{\sum_{j=0}^{q} \beta_i}{\gamma}$.
Where φ_{LGDPG_i} and $\varphi_{GINICoeff_i}$ are the short-run dynamic coefficients of the model convergence to equilibrium. If the value of speed of adjustment is zero, it means that there is no long-run relationship. If it is between −1 and 0, there exists partial adjustment; a value smaller than −1 indicates that the model over-adjusts in the current period; a positive value implies that the system moves away from equilibrium in the long run (Oktayer–Oktayer 2013).

The model was further subjected to diagnostic and the stability tests to ascertain the appropriateness of the ARDL model. The diagnostic tests include a check for normality (Shapiro-Wilk W test for normal data: H_o: Normal), serial correlation (LM Test – Breusch–Godfrey LM test for autocorrelation: H_0: no serial correlation), the autoregressive conditional heteroscedasticity (ARCH Test - Breusch–Pagan / Cook–Weisberg test for heteroskedasticity: H_o: Constant variance), and finally the functional form of the model (Ramsey RESET test 1978) using powers of the fitted values: H_o: model has no omitted variables). In addition, the stability tests of ARDL model for long-run and short-run parameters was conducted by using the cumulative sum of recursive residuals (CUSUM) and the cumulative sum of squares (CUSUM square) of recursive residuals.

3.2. Data

The study relied entirely on secondary data sources. Income inequality is measured by the GINI coefficient. The GINI coefficients data was obtained from the World Income Inequality Database (WIID) put together by the United Nations Development Programme (UNDP) and GDP growth data was sourced from World Bank World Development Indicators Database (WDID). The researcher collected annual time series data of the GINI coefficients and GDP growth for Kenya spanning from 1990 to 2015. We intended to use data from 1963 to 2017 but the main constraint encountered while collecting data for the model was with the amount of information available for the GINI coefficient used to measure income inequality and GDP growth. The GDP data between 1963 up to 2015 was available but beyond 2015 was not recorded while the GINI coefficients had a lot of gaps forcing us to use the shorter data series of between 1990–2015.

The paths and patterns of income inequality represented by GDP growth (Figure 1) and GINI coefficient (Figure 2) differ over time period. Figure 1 shows GDP growth fluctuating over time with a big drop recorded around 1992 (coinciding with the beginning of multiparty in Kenya) and 2007 (the post-election violence). Income inequality in percentage shows a declining trend up to 2000. From early 2002 onwards, the increase in income inequality became gradual (Figure 2) widening gap between rich and poor. Figure 2 also shows a great depression around 2000–2001 which can be associated with the transitions between two government regimes. The former Moi regime,was more of one party system of government compared to president Kibaki's regime, which was more inclusive and development oriented in that it created a favorable environment for investment. In this regard, the rich people with high propensity to save, expended their investment, thereby accumulating wealth that widened the gap between rich and poor to about 55%.

4. Results

The main objective of this article is to establish whether the current upsurge of income inequality in Kenya is growth- promoting or growth-dampening. Following the influential paper by Nelson and Plosser (1982) that provided statistical evidence of the presence of stochastic trends in many macroeconomic time series (like GNP, GINI coefficient, etc.), the first part of the empirical analysis focuses on testing the unit root and cointegration of the GINI coefficients and GDP series. The notion of unit/cointegration arose out of the concern about spurious or nonsensical regressions in time series. The second part involve the estimation of the ARLD model and the last part contain model diagnostic and stability tests.

4.1. Unit root test

This section investigates whether the income inequality measured using GINI coefficients and growth in GDP data has a unit root. The test was carried out in order to eliminate any possibility of spurious regressions and erroneous inferences. This involved determining the order of integration of the time series through unit root test. Accordingly, ADF and PP test were conducted at level and at different levels of difference and the results of the two are reported in Table 1 below. As indicated in the table, both tests, the ADF and PP test failed to reject the null hypothesis of unit root at level for GINI coefficients implying that the variables are non-stationary in level. But at second difference, null hypothesis is rejected implying that the variables become stationary at second difference. For GDP growth, the two tests confirm the present of unit root at level at 5% level of significance. However, since first and second order differencing in all cases eliminates the unit root of most of the variable under consideration, the maximum order of integration can be concluded to be $I(2)$.

Table 1 Summary result of Unit Root Test

Test	Variable	Level (I(0)) T-statistics	Level (I(0)) Lags	1st Difference (I(1)) T-statistics	1st Difference (I(1)) Lags	2nd Difference (I(2)) T-statistics	2nd Difference (I(2)) Lags
ADF	GDP Growth	–2.998** (0.0351)	0	–5.922*** (0.0000)	0	–	–
	GINI Coeff	2.055 (0.9987)	0	–1.580 (0.4938)	0	–6.333*** (0.0000)	0
Phillips-Perron	GDP Growth	–2.972** (0.0376)	3	–6.685*** (0.0000)	3	–	–
	GINI Coeff	0.969 (0.9939)	3	–1.407 (0.5791)	3	–6.473*** (0.0000)	3

Note: P-value at the parenthesis
Source: Own computation based on analysis

4.2. ARDL modelling approach to cointegration analysis

The next step was to examine the existence of cointegration. Since the variables are of different order of integration, we used ARDL modelling due to the fact that it can be applied when the variables are of different order of integration (Pesaran and Pesaran 1997).

The bounds test approach on the two variables has been used to examine long-run relationship between the variables. The maximum lag length of the variables in ARDL model, were selected using the AIC. Based on the result there is strong evidence of cointegration between GDP growth and GINI coefficients because the calculated F-statistic is 11.986, which is greater than the critical values of upper bound at 1% level of significance. The causality is a unidirectional relationship with GINI coefficient Granger cause GDP growth.

Table 2 Result of the cointegration test using ARDL Approach

Dependent variable	Independent variable	F-test Statistic	Cointegration
GDP Growth	GINI Coeff	11.986****	YES
GINI Coeff	GDP Growth	0.868	NO

	*10% Sign. Level		**5% Sign. level		***2.5% Sign. level		****1% Sign. level	
K	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
F_c	4.04	4.78	4.94	5.73	5.77	6.68	6.84	7.84

Source: Own computation based on analysis

4.3. Estimation of the model

Presented in table 3, are the result for the long run and short run coefficient. The result shows a weak positive long run relationship between GDP growth and GINI coefficients at 1% level of significant and a strong positive short run relationship between the two variables at 5% significant level which is against the hypothesis. Basing therefore on these results, a one percent point change in the level of inequality will result to an increase in GDP growth by 0.019percentage point in short run and by 0.003 percentage point in long run if appropriate policies that check on redistribution of income are put in place. Although the results confirm the Kuznets hypothesis that describes a positive relationship between income inequality and economic growth in the early phases of growth, which can be associated in this case with the short run, a negative relationship in the later phases means Kuznets hypothesis was contradicted. In addition, the result also contradicts the prediction of some of the theories about how inequality might impact growth in a non-linear way (Kuznets 1955). For our case, the short run and long run relationship is positive and linearly related. One of the possible causes for the differences in the result may be attributed to the data points used for GINI coefficients and GDP growth, model specification and the control variables being included in the estimation. While the author of this study used percent values and the ARDL model which uses lagged difference as instrument variables, most other scholar used absolute values and OLS regression, and include variables such as education, employment, health expenditures and population growth as control variables. This behavior can similarly be explained using some of the political economy and socio-political instability theories (e.g. Benhabib 2003), which suggest that while some inequality is unlikely to cause unrest and provides growth-enhancing incentives in short run, inequality can disrupt economic relations after it reaches some ‘tipping point’ by inviting political interference through rent-seeking behavior and appropriation. This paradigm was experienced recently in Kenya as a result of changes from the presidential system to a parliamentary system rather than the recommended establishment of a devolved governmental system.

ARDL model select instrument variables (lagged levels for first differences). The p-values indicate that the instrument variables are significant 1% or 5% level of significant, which is clear evidence of the absence of “instrument proliferation”, which has been shown to lead to severe biases and weakened tests of instrument validity in the GMM model (Roodman 2009). The adjustment variable (ADJ_GDP Growth_ L1) provides the feedback and/or speed of adjustment from short-run to long-run equilibrium. There are two important things about this adjusted variable. Firstly, the coefficient should be significant, and secondly it must be negative, so that it provides further proof of stable long-run relationship (Shahbaz and Rahman 2010). The results of the short run model show that adjustment is very strong and negative as well as significant, thus we can rely on adjustment for short-run to converge to long run equilibrium. The R^2 value for two variable cases (GDP growth and GINI coefficient) is 0.74923981. The value implies that 74.9% variation in the GDP growth can be explained by the GINI coefficient while 25.1% is unexplained. The adjusted R^2 of 0.64893573 indicate good fit and correctness of the model specification.

Table 3 Estimated Long Run Statistic Using ARDL Approach

Models	Coefficient	Std. Error	t	P>|t|
Long run				
LR_GINI Coeff_L1	0.2990922***	0.0511981	5.84	0.000
Short run				
GDP Growth_LD	0.991758**	0.3416773	2.90	0.011
GDP Growth_L2D	0.440398*	0.2523159	1.75	0.101
GDP Growth_L3D	0.3736959*	0.1888455	1.98	0.066
GINI Coeff_D1	1.908839**	0.7052704	2.71	0.016
ADJ_GDP Growth_L1	–2.038949***	0.4310692	–4.73	0.000
_Cons	–22.83591***	7.7809530	–2.93	0.010
R-squared	0.74923981			
Adj R-squared	0.64893573			

* Statistically significant at 10% level; ** Statistically significant at 5% level;
*** Statistically significant at 1% level

Source: Own contribution based on analysis

4.4. Diagnostic tests

The model was further subjected to diagnostics to ascertain the appropriateness of the ARDL model in estimating the effect of GINI coefficients on GDP growth. The diagnostic tests involved checking for normality (Shapiro–Wilk W test for normal data), serial correlation (Breusch–Godfrey LM test), the autoregressive conditional heteroscedasticity (ARCH Test–Breusch–Pagan / Cook–Weisberg test), the functional form of the model for omitted variables (Ramsey RESET test), and Durbin–Watson d–statistic. Based on the results of different diagnostic tests, the statistics reported the ARDL model was fit to be used for the estimation purpose since the tests show that there is absence of autocorrelation, functional form misspecification, heteroscedasticity in the models, and the errors follow the normal distribution since the p-values in all test are greater than 0.05. For Durbin-Watson d-statistic test, a rule of thumb is that test statistic values in the range of 1.8 to 2.2 are relatively normal. Values outside of this range could be cause for concern. Field and Miles (2010) suggests that values under 1 or more than 3 are a definite cause for concern.

Table 5 Results of Diagnostic Tests at constant prices

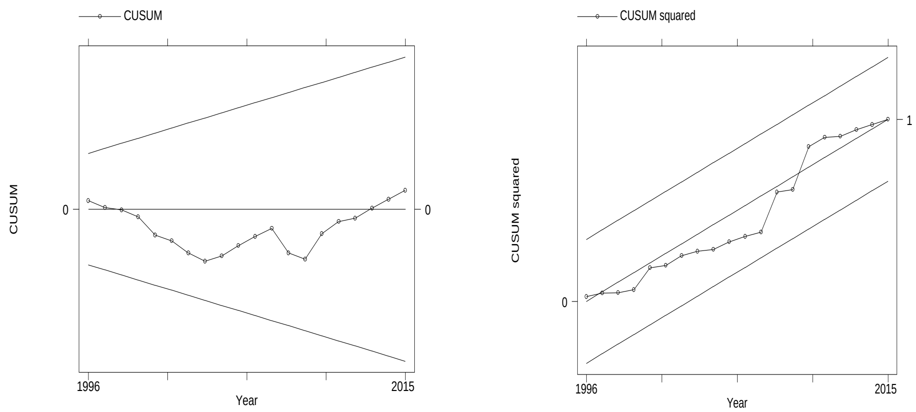
Test	**Statistic**	**Prob> chi2**
Shapiro–Wilk W test for normal data	0.98340	0.96045
LM test for autoregressive conditional heteroskedasticity (ARCH)	0.186	0.6661
RESET test	2.31	0.1283
Breusch–Pagan / Cook-Weisberg test for heteroskedasticity	2.24	0.1341
Breusch–Godfrey LM test for autocorrelation	0.559	0.4546
Durbin–Watson d–statistic	2.159444	

Source: Own computation

4.5.Model stability test

Assessment of model stability was done by plotting the CUSUM and CUSUM squares. The CUSUM test is based on the residuals from the recursive estimates, and based on this test, the null hypothesis implies that the statistic is drawn from a distribution called the CUSUM distribution developed by Page in 1954 (Grigg et al. 2003). If the calculated CUSUM statistics appear to be too large to have been drawn from the CUSUM distribution, we reject the null hypothesis (of model stability). The output will be a graph of the CUSUM statistics and bands representative of the bounds of the critical region for a test at the 5% significance level. In the figure below, the straight lines represent critical bounds at 5% significance level and since the plots of these two tests do not cross the critical value line, it implies that there is a stable long-run relationship between GDP growth and GINI coefficients.

Figure 3 Parameter Stability Test



Source: Own construction based on the analysis

5. Conclusion and policy recommendations

From the literature search, it is clear that the topic of income inequality has not been a major topic of discussion in Kenya, yet it is an important element in the process of economic growth. This paper focus on analyzing the relation between income inequality and GDP growth for the Kenyan economy. More precisely, the study focuses on investigation whether there is a relationship between income inequality and GDP growth. A GINI coefficient was used as a measure for income inequality. An Autoregressive-Distributed Lag (ARDL) model that combines long and short run into a single equation was applied. The ARDL approach was selected because of its flexibility and that it can be applied when the variables are of different order of integration. The ARDL model for GDP growth and GINI coefficients was fitted and a weak positive but significant long run relationship was found between the two variables. There was a strong positive short run relationship between GINI coefficient and GDP growth, and very strong and significant at 1% adjustment rate. The diagnosis and stability tests accepted the model as stable for predicting GDP growth.

Results emphasize the complexity of the relationships between income inequality and economic growth in Kenya data. In this matter, what is interesting is not whether inequality is harmful or beneficial for growth, since the finding are contradictory, but rather the magnitude of the relationship. Although the study found a positive effect of inequality on growth, for a balanced welfare, one policy goal should be equity in income distribution to reduce excessive income disparities. As suggested by Todaro (1997), a more equitable distribution of income can stimulate healthy economic expansion by acting as a powerful material and psychological incentive to widespread public participation in the development process. This endeavor as currently being realized in Kenyan economy through implementation of the Constitutional Development Fund (CDF), the Economic Stimulus Project, free primary education and free tuition in secondary schools, and recently, universal free medical care, should be encouraged and if possible enhanced, although the impact of these remain unknown.

This recent research has focused attention on the impact of income inequality on economic growth. More research should be carried out on all other measures of inequality, to bring it into light which among them is more influential on GDP growth.

References

Alesina, A. – Rodrik, D. (1994): Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109, 2, 469–490.

Banerjee, A. – Dolado, J. J. – Galbraith, J. W. – Hendry, D. (1993): Co-integration, error correction, and the econometric analysis of non-stationary data. *OUP Catalogue*. https://ideas.repec.org/s/oxp/obooks.html Accessed: October 27, 2019.

Barro, R. J. (2000): Inequality and growth in a panel of countries. *Journal of Economic Growth*, 5, 1, 5–32.

Benhabib, J. (2003): The trade-off between inequality and growth. *Annals of Economics and Finance*, 4, 2, 491–507.

Campos, A. (2017): How does inequality affect economic growth? http://www.caixabankresearch.com/sites/default/files/documents/34-35_dossiers_2_ing_eng.pdf Accessed: October 27, 2019.

Castells-Quintana, D. – Royuela, V. (2014): Agglomeration, inequality and economic growth. *The Annals of Regional Science*, 52, 2, 343–366.

Chen, B. (2003): An inverted-U relationship between inequality and long-run growth. *Economics Letters*, 78, 205– 212.

Cingano, F. (2014): Trends in Income Inequality and its Impact on Economic Growth. *OECD Social, Employment and Migration Working Papers*, No. 163, OECD Publishing, Paris.

Constitution of Kenya (2010): The Constitution of Kenya, 2010. National Council for Law, Government Printer, Nairobi, Kenya. https://www.kenyalaw.org Accessed: October 27, 2019.

Cowell, F. A. (2007): Income Distribution and Inequality. STICERD, London School of Economics, DARP 94. http://eprints.lse.ac.uk/3780/1/Income_Distribution_and_Inequality.pdf Accessed: October 27, 2019.

Dickey, D. – Fuller, W. A. (1979): Distribution of the estimators for autoregressive time series with a Unit Root. *Journal of the American Statistical Association*, 75, 2, 427– 431.

Dickey, D. A. – Fuller, W. A. (1981): Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, 49, 1057–1072.

Eisler, D. (2016): Income inequality and the rise of U.S. populism: A cautionary tale for Canada. https://www.schoolofpublicpolicy.sk.ca/documents/research/policy-briefs/JSGS-policybriefs-Income-Inequality-FINAL-web.pdf Accessed: October 27, 2019.

Engle, R. F. – Granger, C. W. (1987): Co-integration and error correction: Representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 55, 2, 251–276.

Field, A. – Miles, J. (2010): *Discovering statistics using SAS: (and sex and drugs and rock'n'roll)*. Sage. http://sro.sussex.ac.uk/id/eprint/13842 Accessed: October 27, 2019.

Fields, G. S. (2002): Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States. Cornell University, ILR school site. http://digitalcommons.ilr.cornell.edu/articles/265/ Accessed: October 27, 2019.

Forbes, K. J. (2000): A reassessment of the relationship between inequality and growth. *American Economic Review*, 90, 4, 869–887.

Gakuru, R. – Mathenge, N. (2012): Poverty, Growth, and Income Distribution in Kenya: A SAM Perspective. AGRODEP Working Paper 0001.

Galor, O. (2009): *Inequality and Economic Development: The Modern Perspective*. Edward Elgar Publishing Ltd.

Garcia-Arenas, J. (2017): Inequality and populism: myths and truths. http://www.caixabankresearch.com/en/inequality-and-populism-myths-and-truths Accessed: October 27, 2019.

Grigg, A.O. – Farewell, V. T. – Spiegelhalter, D. J. (2003): The Use of Risk-Adjusted CUSUM and RSPRT Charts for Monitoring in Medical Contexts". *Statistical Methods in Medical Research*. 12, 2, 147–170.

Hassler, U. – Wolters J. (2006): Autoregressive distributed lag models and cointegration. *Allgemeines Statistisches Archiv*, 90, 1, 59–74.

Johansen, S. – Juselius, K. (1990): Maximum Likelihood Estimation and Inference on Cointegration-With Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics*, 52, 2, 169–210.

Johansen, S. (1991): Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, 59, 1551–1580.

Johansen, S. (1998): Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. *Econometric Theory*, 14, 517–524

Kuznets, S. (1955): Economic growth and income inequality. *The American Economic Review*, 45, 1, 1–28.

Laurenceson, J. – Chai, J. (2003): Financial Reform and Economic Development in China. Cheltenham, UK, Edward Elgar.

Manyeki, K. J. – Kotosz, B. (2017): Empirical Analysis of the Wagner Hypothesis of Government Expenditure Growth in Kenya: ARDL Modelling Approach. *Theory methodology practice: club of economics in miskolc*, 13, 02, 45–57.

Morron, S. A. (2017): Can inequality cause a financial crisis? http://www.caixabankresearch.com/sites/default/files/documents/38-39_dossiers_4_ing_eng.pdf Accessed: October 27, 2019.

Mount, F. (2008): Five types of inequality. http://equality-ne.co.uk/downloads/337_FivetypesofInequality.pdf Accessed: October 27, 2019.

National Treasury (2015): Quarterly Economic and Budget Review (2014/15 for the period ending 30th June 2015, Government Printer, Nairobi, Kenya. http://www.treasury.go.ke Accessed: October 27, 2019.

Nel, P. (2003): Income Inequality, Economic Growth, and Political Instability in Sub-Saharan Africa, Journal of Modern African Studies, 41, 4, 611– 639.

Nelson, C. R. – Plosser, C. R. (1982). Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of monetary economics*, 10, 2, 139–162.

Njuguna, F. (2005): Poverty and income inequality; A stochastic dominance approach, University of Nairobi. http://erepository.uonbi.ac.ke:8080/xmlui/handle/123456789/19845 Accessed: October 27, 2019.

Nkoro, E. – Uko, A. K. (2016): Autoregressive Distributed Lag (ARDL) cointegration technique: application and interpretation. *Journal of Statistical and Econometric Methods*, *5*, 4, 63–91.

Oktayer, A. – Oktayer, N. (2013): Testing Wagner's law for Turkey: Evidence from a trivariate cuasality analysis. *Prague Economic Papers*, 2, 284–301.

Page, E. S. (1954). Continuous Inspection Scheme. *Biometrika*. 41, 1/2, 100–115.

Panizza, U. (2002): Income Inequality and Economic Growth: Evidence from American Data. *Journal of Economic Growth*, 7, 1, 25– 41.

Perotti, R. (1996): Growth, Income Distribution and Democracy. *Journal of Economic Growth*, 1, 2, 149–87.

Pesaran, M. H. – Pesaran, B. (1997): Working with Microfit 4.0: Interactive Econometric Analysis. Oxford, Oxford University Press. https://trove.nla.gov.au/work/7618564 Accessed: October 27, 2019.

Pesaran, M. H. – Shin, Y. – Smith, R. J. (2001): Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16, 3, 289–326.

Phillips, P. C. B. – Perron, P. (1988): Testing for a Unit Root in Time Series Regression. *Biometrika*, 75, 2, 335–346.

Ramsey, J. B. (1978): Tests for specification errors in classical linear least squares regression analysis. *Royal statistical society*, 31, 350−371.

Roodman, D. (2009): A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71, 1, 135–158.

Rugaber, C. S. – Boak, J. (2014): Wealth gap: A guide to what it is, why it matters. AP News https://www.businessinsider.com/wealth-gap-a-guide-to-what-it-is-and-why-it-matters-2014-1 Accessed: October 27, 2019.

Shahbaz, M. – Rahman, M. M. (2010): Foreign capital inflows growth nexus and role of domestic financial sector: An ARDL co-integration approach for Pakistan. *Journal of Economic Research*, 15, 3, 207–231.

SID (Society for International Development) (2004): Pulling Apart: Facts and Figures on Inequality in Kenya, Nairobi: SID. https://www.sidint.net/sites/www.sidint.net/files/docs/pulling-apart.pdf Accessed: October 27, 2019.

Suri, T. – Tschirley, D. – Irungu, C. – Gitau R. – Kariuki, D. (2008): Rural incomes, inequality and poverty dynamics in Kenya, Tegemeo in institute of Agricultural policy and Development, Egerton University. https://ageconsearch.umn.edu/record/202613/files/Tegemeo-WP30-Rural-incomes-inequality-poverty-dynamics-Kenya.pdf Accessed: October 27, 2019.

Todaro, M. P. (1992): The meaning of development, *Economics for a developing world: An introduction to principles, problems and policies*, Longman: London.

Todaro, M. P. (1997): *Economic Development*. London: Longman.

United Nations Development Programme (UNDP). (2016): Data on inequality – Income inequality in percentage. Human Development Data (1990–2015). http://hdr.undp.org/en/data#

Voitchovsky, S. (2005): Does the Profile of Income Inequality Matter for Economic Growth? *Journal of Economic Growth*, 10, 3, 273–296.

Wanyagathi, A. M. (2006): Income inequality and economic growth in Kenya. College of Humanities and Social Sciences (CHSS). University of Nairobi. http://hdl.handle.net/11295/61803 Accessed: October 27, 2019.

World Bank – World Development Indicators (WDI). (2016): GDP per capita growth (annual percentage). World Bank national accounts data. https://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 357–396.

The De-industrialization Process In Azerbaijan: Dutch Disease Syndrome Revisited

Ibrahim Niftiyev

This paper focuses on the de-industrialization processes of Azerbaijan adopting the Dutch disease syndrome as the theoretical framework. After the emergence of Dutch disease hypotheses, resource-rich countries have become its main object of research. The consequences of Dutch disease syndrome are chronically appreciating national currency, a shrinking manufacturing sector compared to the booming sector, and the services sector. In order to shed light on this aspect of the Azerbaijan economy, important literature examples regarded de-industrialization and Dutch disease were examined and descriptive statistics applied to visualize the economy's recent timeline. This research mainly brings back the actuality of the Dutch disease phenomena to Azerbaijan's economy, connecting it to the de-industrialization process on employment, output, and trade level. The main intention is to depict and to compare policy responses of the national government during and after such crisis periods like 2008–2009 and 2014–2015 in a systematic detailed manner.

Keywords: Dutch disease, de-industrialization, oil tradable sector, non-oil tradable sector, non-tradable sector

1. Introduction

The de-industrialization process is a decline in the value-added manufacturing sector as a component of employment, output, and trade of the economy. Various reasons may lead to the de-industrialization process and one of them is Dutch disease syndrome, as it was modeled in the work of Corden and Neary (1982) in its direct and indirect form. In fact, Palma (2008) mentioned Dutch disease as the reason for possible de-industrialization in developing countries. Mainly, Dutch disease is the consequence of a booming sector which increases domestic income and demand for goods (Badeeb et al. 2017). Discovery or dependency on natural resource extraction and exports are the most popular forms of Dutch disease. As Bresser-Pereira (2013) indicated, Dutch disease results in chronic exchange rate appreciation and inflation. This leads to an increase in the price of the non-booming sector commodities, leading to the lower competitiveness of these, as well as lowering investments they might otherwise attract. This negative effect on economic growth is called spending effect with indirect de-industrialization outcome. The second effect is the resource movement effect which happens when labor and capital move out from the rest of the economy to the booming sector resulting in direct de-industrialization. Resource movement effect might not happen if the booming sector is relatively non-labor intensive or they can occur at the same time. Therefore, structural changes in the economy at the expense of the manufacturing sector leads to de-industrialization process.

Many studies support the argument that industrialization, particularly manufacturing, is the engine for the economic growth (Lewis 1954, Kaldor 1966, 1967, Szirmai 2013, Szirmai et al. 2013). Accordingly, any lagging in the industrialization process may lead to economic slowdowns with structural disequilibrium. Several studies examined Dutch disease symptoms in the Azerbaijan economy (Hasanov 2010, 2013); however, the results were not unanimous on the existence of the phenomena and specifically regarded de-industrialization tendencies (Egert 2009, 2012, Gahramanov–Fan 2002). In fact, exchange rate appreciation, changes in real wages, and a shrinking manufacturing sector and other industrial sub-sectors on the background of the 2008–2009 – global financial crisis and in 2014–2015 – sharp commodity price downturns showed consistent symptoms of Dutch disease literature.

After the collapse of the Soviet Union, many former member states faced extreme economic adaptation hardship adjusting to the new challenges of independence. Previously established common markets, industrial structures, and trade relations were lost, which led to a troublesome shift towards the market economy. Political instability, short-lived governments, and military conflicts did not allow for recovery from the economic downturns which started in the early 90s. All of this required new structural transformations, the creation of effective reforms in the economy. The Republic of Azerbaijan was no exception; however, having rich oil and gas resources brought new challenges to the transition process, pushing the government to use the easily available mineral resources, crowding out manufacturing and even the agriculture sector. Although the country's proven total oil and gas reserves are only 0.4% and 0.7% of total world reserves respectively (British Petroleum Company 2018), natural resources played a major role in attracting revenue to the economy during and after the transition period. It also influenced the non-oil sector of the economy. In fact, in 2010 the share of manufacturing in industrial production was 74% less than in 1990 at the current prices, while GDP derived from agriculture decreased from 32.5% in 1991 to 12.4% and 5.6% in 2003 and 2017 years respectively (SSCRA 2019a). Furthermore, manufacturing contributed just 5% in 2007, while the mining industry had a 53.7% share. Also, the mining sector dominated in attracting both local and foreign investments between 2005–2016 at 73.8% of the annual average, while manufacturing got only 8.7% (SSCRA 2019b).

The first boom in oil revenues occurred during the beginning of the last century and oil and gas extraction started to become the main industry of Azerbaijan during Soviet administration. In the early 90s the main trade objective of international economic relations and foreign policy was oil and oil products, as it used to be during the USSR era. with the main difference being that compared to Soviet years, the country's economy polarized into extractive industry and crude oil and chemical products exports. Subsequently, the overwhelming dominance of the oil and gas sector in the economy brought various risks to the country's economy and one of them was the de-industrialization process.

This study found overlapping trends among the various sectors in Azerbaijan related to resource movement effect and spending effect of Dutch disease, which causes direct and indirect de-industrialization. Therefore, during 2008–2011 which was a period of the huge foreign revenue inflow into the country, sharp decline in manufacturing value-added employment, output and trade data point to the de-industrialization process as the result of resource movement effect. Also, the progress in the share of the tertiary sector indicates the spending effect of Dutch disease as a result of high government expenditure.

This work illustrates the possible descriptive paths of the de-industrialization processes as the main extension of Dutch disease hypotheses in Azerbaijan. The present paper uses descriptive statistics to compare the oil tradable sector, the non-oil tradable sector, and the non-tradable sector in the context of employment, output, and trade to track particular decline trends and de-industrialization patterns.[1] It is important to note that the factor endowments of the country identify the optimal structure of the industry also determining the country's competitiveness during a particular period of time (Lin 2015). Azerbaijan's industrial heritage can be observed from the 90s and early 2000s even those years are accompanied by economic stagnation and crisis. In this case, observing the increasing dominance of extractive industry and decreasing role of non-oil and non-tradable sector will provide us a general impression of the presence of the Dutch disease effects. Therefore, the most important aim of the work is to identify de-industrialization trends, clarifying certain connections to the previous Dutch disease studies, at the same time, contrasting policy responses of the national government according to the critical years of 2008–2009 (global financial crisis) and 2014–2015 (sharp commodity price downturns). So far, there has not been any special emphasis among the international publications regarded the de-industrialization of Azerbaijan's economy in terms of employment, output and trade level. Subsequently, this paper intends to explore some aspects of lagging industrialization in Azerbaijan.

The evaluation of the general government programs, goals and certain policy tools indicate that from the economic policy perspective the government tried to apply some policies, however, no strong emphasis on Dutch disease phenomenon or the de-industrialization process can be observed. Mainly, the government aimed to incentivize non-oil sector agents via certain programs and to preserve vanishing sectors. Certain institutional reforms were completed, however all government goals and programs still heavily consider oil revenue and its management techniques. In terms of more concrete policy measures and tools, national currency devaluation and fiscal framework – "New Budget Policy" are first attempts to transfer the economy from the pro-cyclical policy to the countercyclical policy. Hence, there are many challenges in front of those government programs, goals, and policy measures to overcome, and since the acknowledgment of Dutch disease or de-industrialization is still lacking, there is long way to achieve a diversified and sustainable economy.

[1] More detailed information regarded this sectoral division has been presented in the third section of the paper.

The paper is organized as follows. Section 2 provides a literature review of the theoretical background related to de-industrialization, the importance of the manufacturing sector, and Dutch disease syndrome. In Section 3, data sources, key indicators, the definitions of industry and sectors, and methodology are depicted to explain how the analysis of the de-industrialization process of Azerbaijan was carried out. Section 4 demonstrates the descriptive characteristics of de-industrialization patterns through Dutch disease effects and sectoral disproportions on three levels: employment, output, and trade. In Section 5, policy responses, their accordance with the industrialization process during the last crucial years and recommendations on the future policy implications are mentioned. Section 6 draws a conclusion.

2. Literature Review

Starting from the eighteenth century, several technological advancements and innovations, alongside the expanding markets, led to rapid industrialization. Later on, the Industrial Revolution term was coined to describe those radical changes related to textile production and the application of steam power in Britain. Consequently, the industrialization process was a function of technological advance leading to the transformation process of production (factory system, specialization-based division of labor, etc.) and a remarkable increase in income per capita. The economic growth resulting from industrialization and specifically from manufacturing raised labor productivity and concentrated the production process in large-scale enterprises (Kemp 2013).

Several studies have confirmed that industrialization contributed to the economic growth of developed and developing countries over time, arguing in particular that the manufacturing sector is the engine of the development. For instance, Fagerberg and Versbagen (1999, 2002) regressed real GDP growth rates and growth rates of the manufacturing sector in 76 countries concluding that industrialization produced higher growth rates in Latin America and in the East Asia region, but the effect was lesser in advanced economies excluding the period of 1950–1973. Some other recent research has focused on low and middle-income countries like South Africa and African countries (Olamade–Oni 2016, Moholwa 2017), Turkey (Ozturk–Altinoz 2018), and on some Central and Eastern European countries (Ulbrych 2017) defending the thesis that the output of the manufacturing sector plays a key role in the economic growth of that country. Furthermore, industrialization has also played a big role in the catching-up process of developing countries since 1950 (Szirmai et al. 2013). That is why policy-makers should pay attention to it and take measures if manufacturing underperforms. However, the second half of the twentieth century brought new challenges to the industrialization process for both developed and developing countries. Significant declines in the share of manufacturing in GDP and manufacturing employment among upper and middle-income countries brought attention to the de-industrialization phenomenon which appeared as a result of the post-industrial development stage or structural changes.

Research on the de-industrialization process emerged in the 1960s and 1970s in the UK as a result of concern with economic growth slowdowns. Singh (1977) depicted de-industrialization as a structural disequilibrium in the UK where the manufacturing sector was losing competitive ability despite rising productivity and price competitiveness. Later on, positive and negative de-industrialization concepts were brought up by Rowthorn and Wells (1987). This meant that if the services sector can absorb the labor resulting from manufacturing, this is positive de-industrialization, while the opposite meant negative de-industrialization. Also, the same authors proposed a third type of de-industrialization – trade de-industrialization as a change in the structure of net exports away from manufacturing goods towards other goods and services. Koritz (1991) described de-industrialization as a reversal of the industrialization process while pointing out the differences between de-industrialization and destructuring, and de-industrialization and restructuring. Kollmeyer (2009) indicated de-industrialization was behind the declining trend in employment in the processing industry. Then, as cited in Hegyi-Kéri (2016), Gregory et al. (2009) identified de-industrialization as responsible for the continued downturn of industrial activity and capacity, especially in the processing industry.

Rodrik (2016) mentions employment de-industrialization (as a decline of manufacturing employment in the share of the total employment) and output de-industrialization (as a drop in the value-added manufacturing share of GDP at current and constant prices). In his research, Rodrik used the term premature de-industrialization for developing countries, initially introduced by Dasgupta and Singh (2006), instead of de-industrialization. Sampling mainly developing and late industrialized countries, Rodrik (2016) concluded that countries in Latin America and in sub-Saharan Africa without having experienced sufficient industrialization had turned into services economies and faced drops in their manufacturing employment and share in GDP.

De-industrialization is a quite commonly observed phenomenon among the early industrialized, as well as, so-called “latecomer” countries. It was investigated in countries like the UK (Forsyth and Kay 1980,1981), the USA (Koritz 1991), and India (Amirapu and Subramanian 2015, Simmons 1985). However, it remains an ongoing process in Latin America (Palma 2008), East Asia (Koo et al. 2016), and sub-Saharan African (Noorbakhsh–Paloni 1999) countries. Analysis of the historical data shown in the work of Rowthorn and Ramaswamy (1997) also determined a de-industrialization trend among European Union countries, the United States, Japan, and some other countries between 1960–1994. The catastrophic decline of employment in manufacturing and basic industries triggered the study of de-industrialization with extensive consideration of the phenomenon in developing countries (High 2013). Thus, the growing body of literature regarded the importance of the manufacturing sector and industrialization as an engine of growth leads to acknowledging the de-industrialization issue as the main policy measurement for economic development.

The most popular hypotheses of de-industrialization for a long time was an “Inverted-U” relationship between manufacturing employment and income per capita. It was accepted as a natural process in mature economies because they start to provide specialized services based on their development. However, de-industrialization differs from country to country based on historical development and the forces behind it. Van Neuss (2018) summarized existing explanations of de-industrialization as follows: non-homothetic preferences – changes in demand resulting from changes in income, technology – movement of labor to stagnant or non-progressive sectors of the economy due to improved technology or productivity growth in the particular sectors, input-output linkages, and outsourcing – the changing nature of the production system via intermediates, and international trade – comparative advantage may crowd out other non-competitive sectors due to the relative-price effects. Naturally, various countries or regions experience de-industrialization differently. So, early tendencies are a mixture of the first and the second explanations by Van Neuss (2018) and according to Robert-Nicoud (2008), despite technological advancements leading to industrialization, it also changed the fact that the goods need to be produced near the point of final consumption due to the rapidly falling trade costs since 1960. Moreover, as cited in Robert-Nicoud (2008), Autor et al. (2003) reported that between 1960 and 2000 the decomposition of the US labor force showed a lower mean share of routine tasks, pointing to the role of offshoring in de-industrialization process. Falling trade costs because of globalization is a popular hypothesis among the explanations of de-industrialization. For instance, according to Bogliaccini (2013), the reason for de-industrialization in the middle-income Latin American countries[2] between 1980 and 2000 was a causal link between trade liberalization and income inequality. Furthermore, Brady and Denniston (2006) indicated different effects of globalization in the process of industrialization and de-industrialization process in liberal market economies, European countries and command market economies. The main finding is that there is a curvilinear relationship between globalization and manufacturing. At low levels, globalization supports manufacturing via differentiation and specialization, while high levels of globalization lead to saturation lowering manufacturing's share in the labor force. Moreover, on the subsample level of the study, globalization had significant effects among liberal market economies and European countries but not in command market economies. To sum up, having all these differences from country to country based on the particular force of de-industrialization, certain specific structural changes like having a booming sector is also highly related to de-industrialization as discussed on the following paragraphs.

Palma (2008) emphasized the other source of de-industrialization which is Dutch disease. In fact, the de-industrialization process as a result of having a booming sector originates from the original theory of Corden and Neary (1982). Despite the fact that the reasons why a country may experience Dutch disease can differ (commodity boom, tourism, oil, etc.), the main reason why it is linked with de-industrialization relies on crowding out effects from the booming sector. In fact, while having a booming sector, countries like the Netherlands, United Kingdom, Brazil, etc. witnessed rapid falls in manufacturing employment and output. Furthermore, rising

[2] Argentina, Brazil, Chile, Costa Rica, Mexico, Uruguay, and Venezuela

levels of expenditure on services and globalization were considered as possible reasons of de-industrialization (Williamson 2006), despite Iversen (2001) indicating that globalization or trade openness has a little or insignificant role in these processes.

The intellectual basis for the importance of the manufacturing sector for economic growth can be traced back to Kaldor (1966, 1967) and Lewis (1954). Kaldor's model encompassed both supply and demand side, emphasizing dynamic economies of scale like a positive correlation between manufacturing output and manufacturing productivity. Accordingly, he proposed three laws: higher manufacturing output leads to higher economic growth; the productivity of the manufacturing sector positively influences manufacturing growth (also known as Verdoorn's law); the growth of the manufacturing sector is positively connected to the productivity of non-manufacturing sector. Lewis argued that industrial development presents a clear path to capital accumulation and economic growth because the capitalist system expands continuously via reinvesting profits if there is unlimited labor supply at subsistence wage levels. Rodrik (2016) mentioned that manufacturing activities are crucial for growth due to features like technological dynamism, the absorption capacity of high quantity unskilled labor and being tradable. Moreover, there are other popular arguments especially in the case of manufacturing: (i) the empirical connection between industrialization and per capita levels point to the reason why rich countries are rich, (ii) the manufacturing sector has higher rates of productivity growth than the other sectors of the economy (iii) favoring manufacturing is a good way of overcoming structural change burden (iv) manufacturing contributes to capital accumulation more than agriculture or services (v) the availability of economies of scale is higher than in other sectors (vi) to make technological advancements is easier in manufacturing than services (vii) linkage and spillover effects are stronger in manufacturing (viii) due to Engel's law, primary and agricultural product exporters will not benefit from the expanding world markets for manufactured goods (Szirmai 2013). From this perspective of the importance of the manufacturing sector for economic growth, we can argue that the development of the manufacturing sector should be a crucial policy choice for national governments. Nevertheless, a discovery or abundance of either natural resources or agricultural products lead to the so-called Dutch disease syndrome and Palma (2008) proposed Dutch disease as a new trigger of de-industrialization in contrast to conventional approaches to de-industrialization.

Coming to Dutch disease hypotheses, in 1977 "The Economist" introduced the term to illustrate how the Dutch manufacturing sector declined after the discovery of the large natural gas field in the North Sea in 1959 (The Economist 1977). The phenomena explain how the manufacturing or agriculture sector shrinks because of the great amount of foreign currency inflow into the economy's "booming sector." The first definitive model of Dutch disease was coined by Corden and Neary (1982) to describe the effects and structural changes of the boom-generated economic growth. The classic model of the Dutch disease theory identifies two sectors: tradable - the booming sector and manufacturing (lagging) or non-tradable – including services. Generally, the extraction of natural resources such as oil, natural gas, nonferrous metals, or the production of coffee or cocoa are the main sources of the booming sector. However; usually, manufacturing and agriculture are the lagging sectors. If a resource boom happens, this influences the economy in two ways: the resource movement effect and the spending effect. The resource movement effect is the shift in production away from the manufacturing sector through increased labor demand. This effect can be negligible if the booming sector uses relatively little labor and capital compared to the other sectors. On the other hand, the spending effect is the flow of revenue towards the non-tradable sector via additional revenue that the booming sector generates.

Several studies have tried to evaluate countries like Canada (Stanford 2012), Australia (Corden 1996), the United Kingdom (Ross 1986), Nigeria (Ezeala-Harrison 1993), Norway (Larsen 2006) in terms of Dutch disease syndrome. Consistently with the model, the common claims were dependence on primary exports, poor productivity growth, real exchange rate appreciation, and declines in agriculture, value-added manufacturing employment, and value-added manufacturing output. Moreover, Gurbanov and Merkel (2010) gave some country examples like Norway, Nigeria, Indonesia, and Botswana where Dutch disease has been successfully overcome.

At the end of the 90s and beginning of the 2000s, early attempts were made to investigate the Azerbaijan economy within the Dutch disease framework. Singh and Laurila (1999) drew attention to Dutch disease phenomenon in Azerbaijan in the middle and in the long term because of the risk of the overshooting in real exchange rates. The authors mentioned that Dutch disease might be a policy issue if the inherited industrial heritage were not used as the basis for restructuring. In other words, the real sector of the economy might experience pressures.

Other research related to early Dutch disease diagnostics in Azerbaijan came from Gahramanov and Fan (2002). The authors used the extended version of the Balassa (1964) and Samuelson (1964) model (included variables TOT, i.e. prices of export over prices of imports) to determine if Dutch disease could be found in terms of increases in the purchasing power of the Azerbaijani Manat in real terms. Briefly, Balassa (1964) and Samuelson (1964) model predicts that real appreciation in the national currency will occur as the result of faster productivity growth in tradable sector compared to non-tradable goods and services. Therefore, the price of tertiary sector services will rise in response to fast productivity growth in the tradable sector leading to high probability of inflation. The study revealed that a10% rise in the price level caused the real exchange rate to fall (appreciate) by 8.7%, real wages to increase due to the lower increase in the consumer price index, while external oil price shocks hardly influenced Azerbaijani Manat between 1994 and 2001. Thus, the authors concluded that the monetary side of Azerbaijan economy did not show any vulnerability to Dutch disease syndrome.

Kaser (2003) described possible threats of Dutch disease to the Azerbaijani economy in line with the diversification and political risks based on the macroeconomic comparison of transition countries. A similar study was done by Auty (2001) on Caspian basin region countries concluding that there is evidence of Dutch disease syndrome among the resource-rich transition countries. However, for the later years, Egert (2009, 2012) argued that due to the declining relative oil prices and lack of data, it is not possible to track Dutch disease effects among post-soviet countries.

Considering these examples from the literature, the research questions for this paper are as follows:

1. Does the employment, output, and trade data show the direct and indirect de-industrialization patterns as the result of the resource movement and spending effect of Dutch disease in Azerbaijan?
2. What have the policy responses of the national government been as a reaction to critical years like 2008–2009 and 2014–2015, to shelter the primary export-based economy and particularly value-added manufacturing sector from a potential de-industrialization process?

3. Data and Methodology

Conducting descriptive analysis of Dutch disease related de-industrialization is complex research and requires sophisticated approaches from different angles. In order to do so, the economy of Azerbaijan was analyzed through three aggregated sectors – oil tradable sector, non-oil tradable sector, and non-tradable sector, similar to Hasanov (2013) who conducted a Dutch disease related study on Azerbaijan. The sectors are the summation of related sub-sectors of the economy. For instance, oil tradable sector contains oil and gas extraction, chemical industry, electricity, gas and steam production, distribution and supply services. Despite such sectors like chemical industry or electricity production being considered as manufacturing sector in the statistical classifications, they are based on natural resources and have comparative advantages compared to the non-oil sector, so they were included in oil tradable sector to have a clearer non-oil tradable sector. Moreover, the role of the distribution and supply aspect, meaning the influence of the services related to oil and gas extraction or related industries, is very small, so it does not distort the big picture related to oil tradable sector. In turn, mainly non-oil tradable sector is the value-added manufacturing goods such as textiles, machinery, production of electrical and electronic equipment, weaving and cloth production, food and beverage production, metallurgy and etc. Also, it should be noted that, besides specific sectoral data related to non-oil tradable sector on the aggregated scope, some data sets combine agriculture, fishery, and forestry data to draw a comparative format, even though the decline in agriculture, forestry or fishery is not de-industrialization. The non-tradable sector consists of service sectors such as the construction, transportation, communication, and other tertiary sectors. Hence, the relevant sub-chapters will compare changing trends in the aggregated sectors (namely oil tradable sector, non-oil tradable sector and non-tradable sector) and based on the knowledge from the literature review, de-industrialization process will be analyzed as a shrinkage of the value-added manufacturing sector as a part of non-oil tradable.

Due to availability issues, certain data types can be found either in the form of mining or in a form of oil tradable sector, and the data related to value-added manufacturing also represented in a form of different kinds of simple and complex industrial products in various parts of the paper. Generally, when we say mining sector in Azerbaijan, it is extraction of crude oil and natural gas, mining of metal ores, mining of stone, sand, gravel, salt and other mining industry production, as well as mining support service activities. However, the main purpose of this paper is to descriptively compare oil tradable sector and non-oil and non-tradable sectors with the sub-sectoral levels like textiles, food and beverage production, etc. In fact, the data presented as mining does not distort our analysis because the average percentage share of oil sector (extraction of crude oil and natural gas) as a share of mining used to be 89.66% between 1990–2017.[3] Furthermore, the main difference between non-oil tradable sector and value-added manufacturing is that non-oil tradable sector may contain agriculture data but value-added manufacturing strongly sticks to the such sectors as production of television receivers, manufacture of food and beverages, manufacture of furniture, manufacture of machinery and equipment, etc.

On the sub-sectoral data level, particular sectors were emphasized in the relevant sub-chapters to compare pre-boom and post-boom trends. Moreover, various literature examples have been connected to each de-industrialization dimension to explore the drivers behind the trends. Consequently, the fourth chapter of this study frames de-industrialization on three levels: employment, output, and trade. Again, it should be noted that the term "de-industrialization" has not been used for non-oil tradable sector but only for value-added manufacturing because non-oil tradable sector also may contain such data as agriculture, forestry or fishery.

Main data sources are from the State Statistical Committee of the Republic of Azerbaijan, the Central Bank of Azerbaijan and the World Bank national accounts data, and OECD National Accounts data files. In the study, descriptive statistics were applied to compare the annual and periodic percentage changes of given variables. Due to the availability issues, some chapters and sub-chapters incorporate data from the 90s while the others illustrate only the 2005–2017 period. Considering the fact that the country's huge oil revenue started to flow into the economy in 2008, it still can present the necessary information for an overall impression.

Also, certain time periods on the graphs were set at 3 or 5 year ranges to depict a more understandable overview of the economy while certain graphs present consecutive time series data. It does not bias the visualization and interpretation of the constructed data because there are not any big deviations from the general trends. The ranges were organized in a way to encapsulate important breaking points.

The visual patterns of both three-sector and sub-sector de-industrialization data provide an initial understanding of the Azerbaijan economy in the context of

[3] This can be easily found on the data set included in the Reference List named SSCRA (2019c): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/industry/en/007en.xls Accessed: January 18, 2019.

Dutch disease hypotheses. Notwithstanding, our brief evaluation of the national policy decisions of the government allows us to see the underlying reasoning from the economic policy-making point of view. Policies included are the results of presidential decrees published on the electronic database of the Ministry of Justice of the Republic of Azerbaijan[4].

The findings will contribute to further research regarded Dutch disease effect in Azerbaijan economy and its possible relationship with this particular type of de-industrialization, for example, employment de-industrialization to understand how labor market structures can be influenced by the booming sector. Moreover, precisely this de-industrialization phenomenon in the Azerbaijan economy has never been a focal point among interested researchers, so this study could well serve as a general analytical framework connecting the de-industrialization experience of resource-rich countries to more mainstream de-industrialization studies.

4. De-industrialization Patterns in Azerbaijan

The Azerbaijan economy experienced severe industrial downsizing like other former member countries of the USSR during the 90s. In fact, Azerbaijan's GDP had already shrunk by 17% between 1985 and 1991 (Cornell 2015). However, losing old economic relations and common markets after the collapse of the Soviet Union were not the only reasons why the economy faced these difficulties. Indeed, the loss of 20% of territory containing fertile land as the result of the devastating war with Armenia, having one million internally displaced refugees, very negatively influenced the socio-economic environment of the country (Ibrahimov 2016). This chapter will analyze how on the background of the booming sector other sectors of the economy experienced structural changes, while simultaneously making comparisons with previous Dutch disease-related studies. In order to do so, first it is important to get a glimpse of the general macroeconomic and fiscal data, then employment, output, and trade data will show the main trends of de-industrialization.

During Soviet times, the central government used central planning tools to shape the industrial structures of the member countries. Industrial production and the division of labor and trade were regulated by Moscow. Azerbaijan's main exports were oil products, machinery, weaving, and wine products, and the main trade partners were countries like Russia and Ukraine (Nuri Aras et al. 2016). Most of all, the oil and gas based extractive industry played a huge role in being a priority in the central government's economic policy. Naturally, the current industrial structure is a legacy of the Soviet period and several development stages of the oil and gas industry. According to Karpov (2010) there were five stages in the growth of the oil and gas industry in the post-soviet area: the first – from the mid-1860s to 1920, the second from the beginning of 1920 until 1932, the third 1933–1950, the fourth from the middle of 1950 until the end of 1980, and the fifth 1980–1990.

[4] - http://www.e-qanun.az/

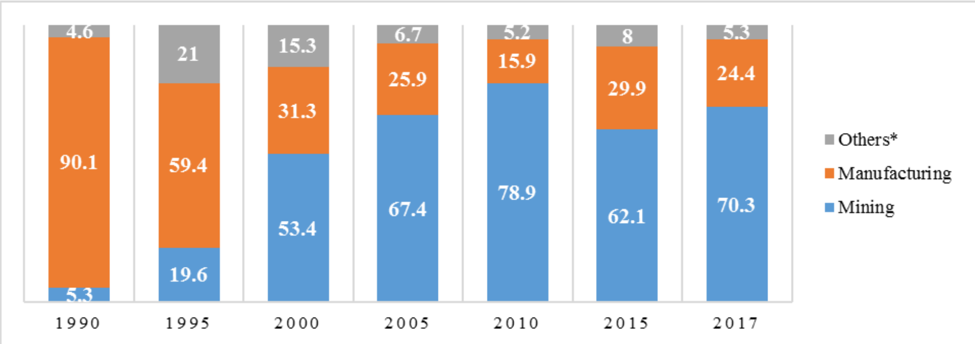
The oil sector's role in the economy was very prominent during both the pre-Soviet and Soviet period. In fact, the world's first oil boom happened in 1874 in Azerbaijan thanks to the efforts of the Swedish Nobel brothers, culminating in 1901 (over 10 million tons a year). So, during the first stage of oil industry growth in the post-soviet period, western investment dominated. Azerbaijan had in fact accounted for 70% of crude oil production in the whole USSR area in 1941. Then, starting from 1960 oil production moved to Siberia and Kazakhstan leading to decreases in the oil industry's output and the necessary investment towards oil and gas (Laurila 1999). Consequently, the fifth stage of the development of the oil and gas industry which encompassed the 1980–1990 period was not actual for Azerbaijan. This resulted in an indicator of 3% of the Soviet Union's oil production and labor force migration (Cornell 2015). Then, the first years of independence brought enormous stress to economic activities and relations because of the transition process.

The first years of independence were characterized by a severe fall in industrial production, structural changes and refurbishments of the old oil and gas industrial infrastructure. More precisely, it was a 38% decline in industrial production in 1994 compared to 1990. At the same time, real average monthly wages were 17% less than in 1989 with 25% spent on household consumption (Cornell 2015). These early trends were followed by a booming sector – oil tradable sector and decreasing manufacturing share in the industrial production. After the recovery from the 1998 oil price downturns, the oil and gas industry started to operate remarkably successfully through Azerbaijan International Operating Company and State Oil Company. Baku–Supsa pipeline started to deliver the oil to Black Sea ports and Baku-Tbilisi-Ceyhan was completed in 2006 leading to large-scale oil production and huge oil revenue. Over time, the oil and gas extraction or mining sector took over the lion's share of industrial production.

Figure 1 shows how the structure of industrial production in Azerbaijan changed over the period of 27 years, from 1990 to 2017 (at current prices as a share of overall industrial production). At the risk of oversimplification, this major trend depicts a dramatic decline in manufacturing while notable increases in the mining industry up to 2010. However, it does not mean that crude oil and gas fields were discovered in 1990 leading to an exact repeat of Dutch disease as happened in 1959 in the Netherlands. This structural transformation frames how mining, mainly the extractive industry became a so-called "booming sector" whilst manufacturing's share dropped from 90.1% to 15.9% in 2010 and contributed 29.9% and 24.4% in 2015 and 2017 respectively.
Another dimension of the topic lies in the answer to the following question: did the transition effect play a predominant role in the de-industrialization process of the manufacturing value added sector? Despite a definitive answer being very difficult, Figure 2 shows that the economy of Azerbaijan recovered in the year 2005to a level of 24.8 billion USD compared to 22.2 billion USD (constant 2010 USD) in 1991. Between these dates, we observe a gradual decrease in real GDP until 1995, then moderate increase until 2005. The trend is more or less the same with GDP in current prices, it shows the recovery of the economy in 2004 (8.7 billion USD) compared to the year 1991's indicator of 8.2 billion USD. So, because of exceptional circumstances, we observe a shrinking manufacturing value added sector during the 90s, however, the de-

industrialization trend after 2005 and during the oil revenue boom period – roughly between 2008–2014 years should be treated independently of any so-called “transition effect”.

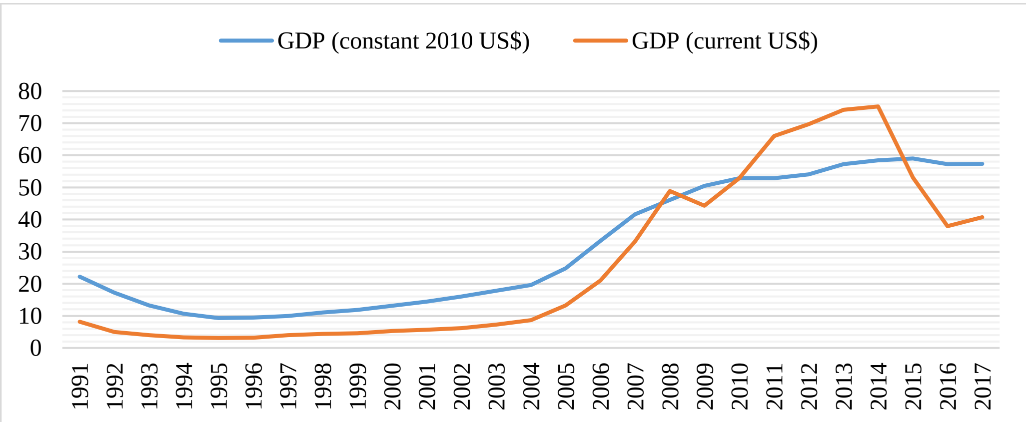
Figure 1 Sectoral Structure of the Industrial Production in Azerbaijan (in % of gross total)



*Others include electricity, gas and water related production, recycling and logistics
Source: Own construction based on SSCRA (2019c)

Moreover, further data and analysis to this research will show that a sharp rise in GDP after 2005 was because of the oil tradable sector, and there were particular fluctuations, including sharp downturns in the value-added manufacturing sector (non-oil tradable sector). Consequently, we can accept the prevalence of the transition effect during 1991 and in the years 2004/2005. Given this, however, successive time periods should at least have contributed to an improvement for the manufacturing sector and we must somehow assume that the deterioration of value-added manufacturing sector or the other sectors were not the sole responsibility of the transition process or any effect with an expectation of upward trends in the non-oil tradable sector.

Figure 2 GDP of Azerbaijan in constant and current USD (billions USD)



Source: World Bank national accounts data, and OECD National Accounts data files.

It is also worth noting the distorted character of Figure 1 regarding early 90s mining data. Crude oil and gas production experienced a gradual fall between 1990 until 1994, from 9.9 million tons to 6.4 million tons (Table 1). However, we observe the increasing share of the mining sector in both industrial production and in the total exports because of the inflated volume of its value in terms of price. So, it is important to bear in mind the fact that the drops in the non-mining sector exaggerated the role of mining in industrial production in the early 90s. Starting from 1996, the oil and gas sector became a stable, growing industry gaining more and more share in total industrial production.

Table 1 The production and export data of mining and oil and gas sectors during early 90s

	1990	**1991**	**1992**	**1993**	**1994**	**1995**
Crude oil and gas production, in millions of tons	9.9	8.6	7.9	6.8	6.4	9.9
Volume of mining sector at factual price, million Manat	12.9	36.8	404.7	1888.6	19.0	346.5
Share of oil and gas in total exports, in %	12.2	10.9	18.4	16.4	22.3	12.2

Source: World Bank (1995), SSCRA

At the same time, Table 2 outlines the changes regarded gross value added in GDP by economic activity between 1991 and 2017. The table shows that agriculture played a very major role during the early 90s, however after 2004 its share hovered around 5–7 %. Manufacturing value added followed the same trend bottoming out below 10% after 1996. The contribution of services fluctuated over time achieving high shares in recent years like 2015, 2016, and 2017 at 41.4%, 40.1%, and 38.8% respectively. Meanwhile, an increasing trend in construction and services point out the spending effect of Dutch disease.

Table 2 Gross value added in GDP by economic activity in Azerbaijan, 2005–2017 (%)

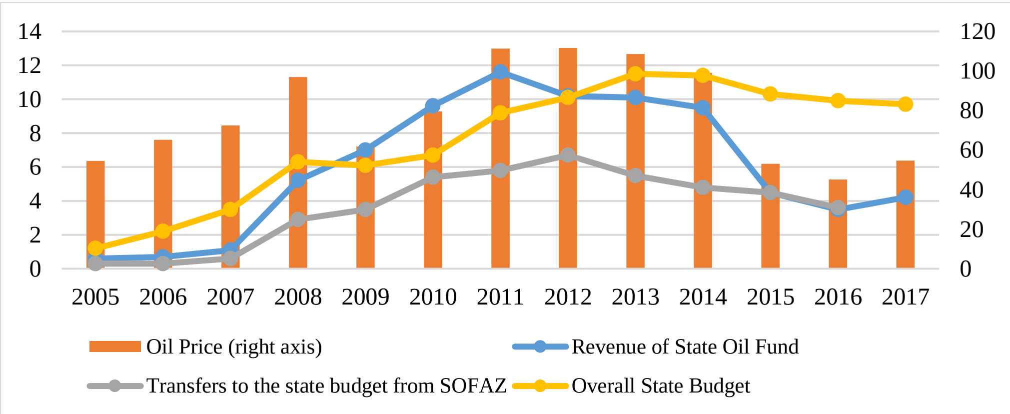
Years	Agriculture*	Industry			Construction	Services
		Overall***	Mining**	Manufacturing value added		
1991	32.5	31.6	-	-	-	36.6
1992	26.1	36.6	-	21.9	-	28.9
1993	27.1	32.1	-	12.9	-	35.9
1994	32.3	27.6	-	12.6	-	38.0
1995	25.3	31.0	-	11.5	-	37.9
1996	24.9	35.1	-	10.4	-	30.8
1997	20.1	37.0	-	8.3	-	35.8
1998	18.1	35.0	-	8.3	-	43.3
1999	18.4	39.1	-	6.0	-	39.0
2000	16.0	42.5	-	5.2	-	35.7
2001	14.8	43.4	-	6.1	-	34.4
2002	13.9	46.1	-	7.4	-	32.4
2003	12.4	48.5	-	8.6	-	31.9
2004	10.9	50.7	-	8.2	-	31.7
2005	9.1	48.7	42.2	6.5	9.0	26.2
2006	7.1	56.7	50.9	5.8	7.7	24.0
2007	6.7	58.7	53.7	5.0	6.4	22.7
2008	5.6	57.4	52.7	4.7	7.0	24.7
2009	6.1	47.9	42.4	5.5	7.2	32.5
2010	5.5	50.6	45.9	4.7	8.1	29.0
2011	5.1	52.0	48.0	4.0	8.0	29.0
2012	5.1	47.3	43.1	4.2	10.1	31.3
2013	5.4	43.4	39.2	4.2	11.6	32.9
2014	5.3	39.0	34.3	4.7	12.6	35.6
2015	6.2	31.4	26.4	5.0	12.0	41.4
2016	5.6	35.6	30.7	4.9	10.5	40.1
2017	5.6	38.8	34.1	4.7	9.5	38.8

* The data also includes the forestry and the fishery data
** Starting from 2005 year, the mining sector was mainly based on extractive industry, meaning oil and gas extraction. The share of oil and gas in the extractive sector ranged between 90–96%, only dropping to 87% in 2017.
*** The data includes construction data between 1991–2005.

Source: Own Construction based on SSCRA and World Bank data

Figure 3 depicts the main trends in the Azerbaijan economy related to the oil price and revenue. According to the Figure 3, there are three main boom periods in oil: the first, until 2008 when the oil prices kept growing while revenue of the State Oil Fund and Transfers from it to the state budget were low; the second, when huge amounts of foreign currency flowed into the economy increasing the fund revenues and transfers to the state budget (2008–2011); the third, gradual (2011–2014) and sharp (2014–2015) declines in both oil revenue and transfers. These trends are compatible with the changes in the oil and non-oil tradable, as well as, the non-tradable sectors of the economy, as will be discussed later on. The following sub-chapters examine these trends against the background of sectoral changes and Dutch disease research devoted to Azerbaijan.

Figure 3 Fiscal aspects of oil sector and oil price (Brant trademark), in USD (right axis) and revenue of the State Oil Fund and State Budget (billion USD); 2005–2017



Source: State Oil Fund of the Republic of Azerbaijan (2017), Annual Report; Macrotrends Brent Crude Oil Prices

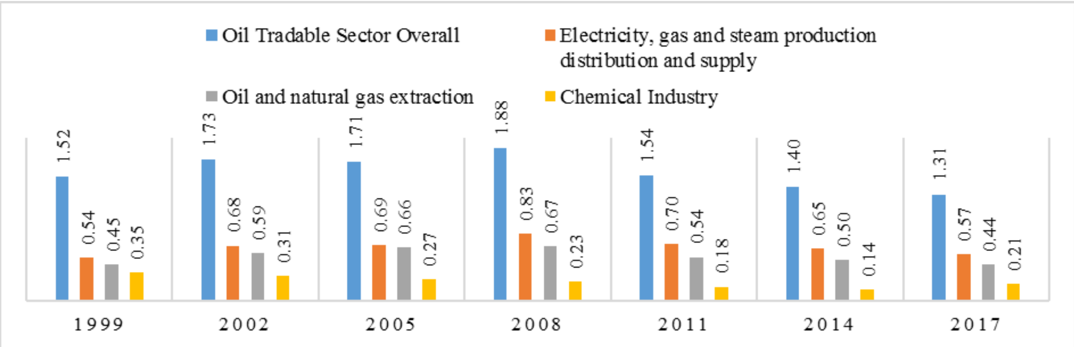
4.1. Employment De-industrialization

According to the Dutch disease model, resource movement effect causes direct de-industrialization in the manufacturing and service sectors, reducing the output of non-tradable sector because of the rise in demand and profitability for labor in the booming sector. According to Corden and Neary (1982), resource movement consists of two parts: movement of labor from the lagging sector to the booming sector and from the non-tradable sector to the booming sector. In contrast, the spending effect raises the output of non-tradables because government expenditures seek to provide extra demand for the non-tradable sector through the accumulated foreign currency. In other words, based on the magnitude of a certain effect, we can track Dutch disease related de-industrialization patterns through employment data.

Figure 4 illustrates employment trends in the oil sector as share in total employment in Azerbaijan between 1999 and 2017, a period of 18 years. Consistent with the Dutch disease theory, when a booming sector is labor-intensive and labor is mobile among the sectors, it should attract labor force from the other sectors. Overall, the share of the oil tradable sector in total employment slightly increased until 2008 and rose to 1.88% when it experienced decline and bottomed out at around 1.31% in 2017. So, it is interesting that, despite having “booming sector” status, the oil sector also faced a slight employment decline starting from 2008 until 2017. Following the same development trend, the other two biggest sub-sectors of oil tradable sector – electricity, gas and steam production, and oil and gas extraction showed gradual growth until 2008 while the chemical industry shrank to almost half, from 0.4% to 0.2% over the eighteen years’ timescale.

The employment trends in the oil sector are related to the execution and completion phases of the upstream and downstream projects in the oil and gas production. For example, the largest oil field - Azeri Chirag Guneshli (ACG) was prepared to provide the necessary crude oil to the Baku–Tbilisi–Ceyhan pipeline and the other oil pipelines in 2008. Consequently, until the year 2008, there were three development stages: 1. development of Central Azeri started in 2005; 2. Development of East Azeri started in 2005, and in 2006 West Azeri platforms were established; 3. The launch of Deepwater Gunashli platform in 2008. In the gas production, the largest gas field is Shah Deniz and the estimated total investment is 10 billion USD. Subsequently, the relevant exploitation projects were finished around 2006 (Ciarreta and Nasirov 2012). Meanwhile, because of the transition process and decreases in energy consumption, modernization issues caused employment decline in other sub-sectors of oil tradable sector such as electricity, gas, and steam production and chemical industries.

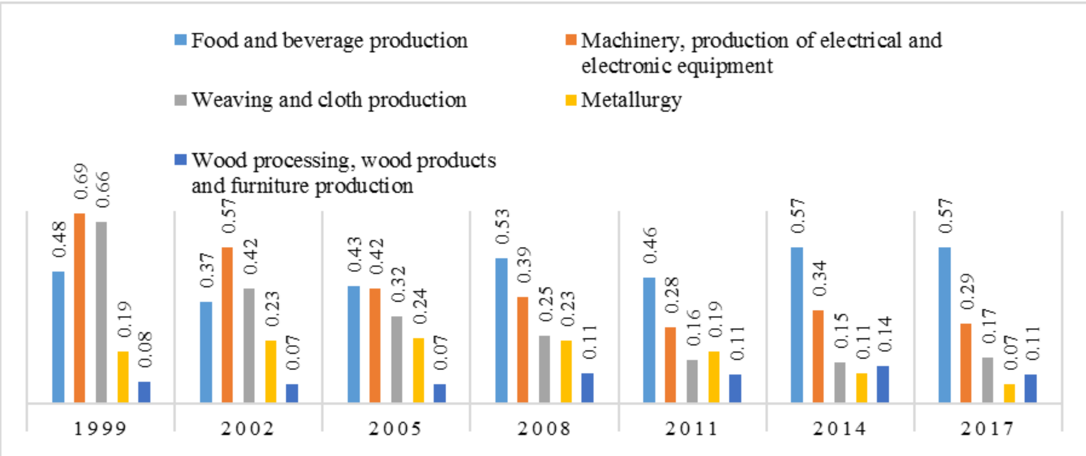
Figure 4 Employment data of the oil tradable sector, 1999–2017
(% of the total employment in the economy)



Source: Own construction based on SSCRA data (2019d; 2019e)

As the biggest representative of non-oil tradable sector, agriculture also experienced employment decline over the 1999–2017 period while other sub-sectors of non-oil tradable fluctuated (Figure 5). Agricultures sector employed 41.49% of the labor force in 1999, however during and after oil booming, it had dropped to 36.35% by 2017. In terms of de-industrialization, particular manufacturing sectors showed a downward trend. For instance, weaving and cloth production accounted for a very low share in total employment at 0.15% in 2014 compared to the late 90s when it was 0.66%. Machinery, production of electric and electronic equipment showed constant decreases starting from 1999 until 2011, from 0.66% to 0.16%, but slightly enlarged its share around 2017. In fact, there has been an overall relative employment de-industrialization in the biggest representatives of the non-oil value added manufacturing tradable sectors of the economy. After the oil boom period as well as in the years of crisis, metallurgy sector downsized noticeably to 0.07% in 2017 from the pre-crisis period of 2005 – 0.25%. The general trend in wood processing, wood products and furniture production is more or less stable – between 0.08 % and 0.11%. The last sector mainly addresses domestic and regional markets of the country, however, still it is part of non-oil tradable sector.

Figure 5 Structure of employment of non-oil tradable sector, 1999–2017 (% of the total employment in the economy)

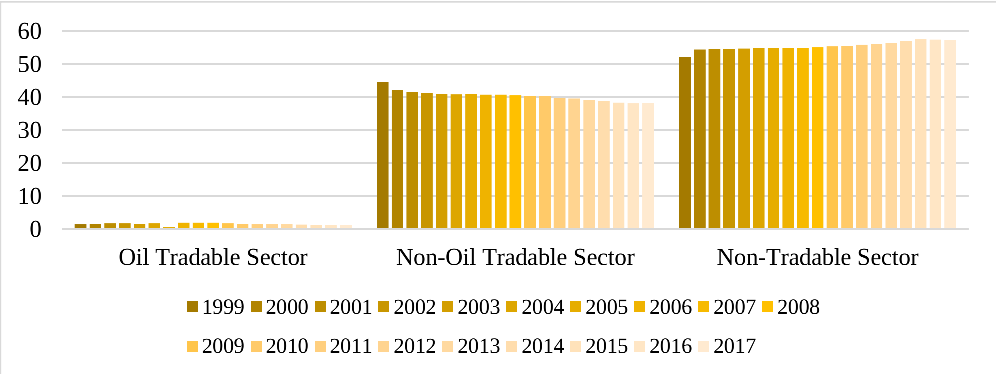


Source: Own construction based on SSCRA data (2019d; 2019e)

Figure 6 depicts the general comparison of oil tradable, non-oil tradable and non-tradable sectors. The overall impression is that oil tradable sector experienced employment fall and non-oil tradable sectors witnessed employment de-industrialization between 1999 and 2017 while non-tradable sector increased its share in total employment from 52.1% to 57.3%. In terms of resource movement and the spending effect of Dutch disease, both effects are slightly apparent in a descriptive manner.

Hence, Sadik-Zada et. al. (2019) conducting Input-Output analysis for 2006, 2008 and 2009 data in Azerbaijan, concluded that the oil sector did not get integrated into the local economic structure despite having a wealth of qualified human capital, local content policies, and infrastructure. This highlights the currency of Hartwick's rule (Hartwick 1977) for sustainable development in a natural resource abundant country. It is a well-known fact that natural resources like oil and gas are non-renewable exhaustible resources. Society can reach a point when dependence on natural resource extraction and exports cannot be fulfilled with the low stock levels of depleted natural resources. So, Hartwick's rule in resource economics states the amount of spending towards produced capital (buildings, roads, knowledge stocks, etc.) that is needed to counterbalance of declining natural resource stocks.

Figure 6 Sectoral structure of employment, 1999–2017
(% of the total employment in the economy)



Source: Own construction based on SSCRA data (2019d; 2019e)

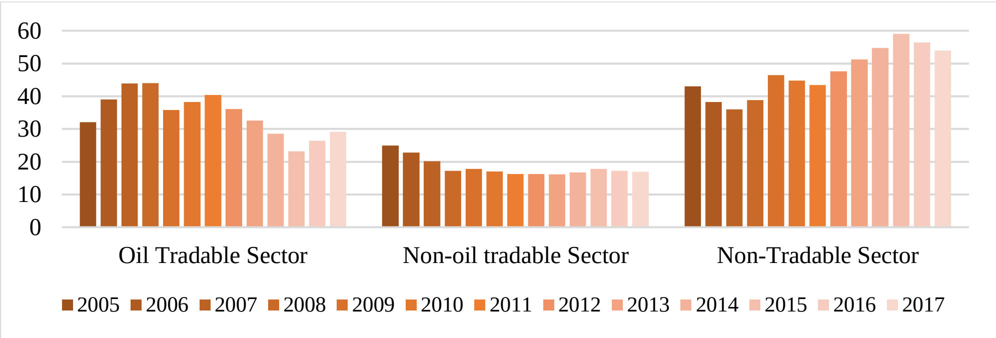
4.2. Output De-industrialization

As a result of certain labor force movements, the output levels of sectors change commensurately. The Dutch disease model describes direct de-industrialization (resource movement effect) as the output fall in manufacturing, rise in wage rate in the booming sector at a constant real exchange rate, and reduced service sector output. On the other hand, indirect de-industrialization (spending effect) is an increase in the non-tradable sector's output as a response to providing excess demand against a background of fluctuating exchange rate towards services, sourcing from the extra revenue which a booming sector generates. Moreover, both effects might occur and lead to a real appreciation. To picture the general trend, two perspectives, total output in the economy and industrial production were analyzed.

The peak of the revenue of the State Oil Fund of Azerbaijan was 11.6 billion USD in 2011 (Figure 3). Starting from the same year (2011) the share of oil tradable sector in the total output production started to decline – from 40.4% in 2011 to 23.2% in 2015 and 29.1% in 2017 – giving away its growth status to the non-tradable sector (Figure 7).

Furthermore, as can be observed in Figure 7, consistent with the Dutch disease hypotheses, the non-oil tradable sector did not achieve solid growth between 2005 and 2017, while non-tradable sector showed a sharply increasing trend after the oil revenue boom. During the oil revenue boom period, 2007–2011, transfers from the windfall fund to the state budget also increased, and non-tradable output grew to 59% in 2015. This aspect supports the statement related to spending effect which is the result of government expenditures sourcing from the boom.

Figure 7 Share of the sectors in total production output, 1999–2017 (% of total output)

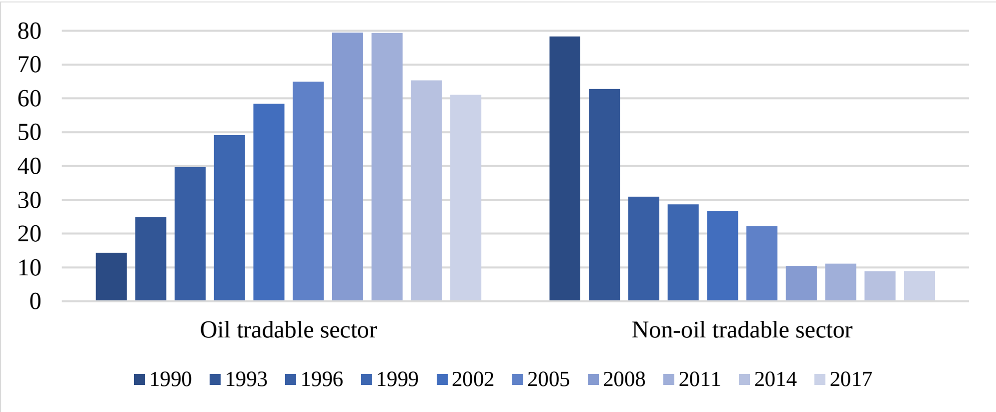


Note: The data for non-oil tradable sector also incorporates agriculture, forestry, and fishery data.
Source: Own construction based on SSCRA (2019f) data

The picture gets clearer when we consider oil tradable sector and non-oil tradable sector against the background of industrial production over a longer time range. From Figure 8 we can observe a stable increase of oil tradable sector starting from 1990 until 2008–2011 period, which is almost 80% of all industrial products, while non-oil tradable sector after the oil boom bottomed out at 9% in 2017. Considering the fact that the data related to non-oil tradable sector encompasses manufacturing value added, excluding such activities like the manufacture of refined petroleum products, chemical industry, and electricity, gas and steam production, distribution of supply, the argument of output de-industrialization of non-oil tradable sector does not seem so far-fetched.

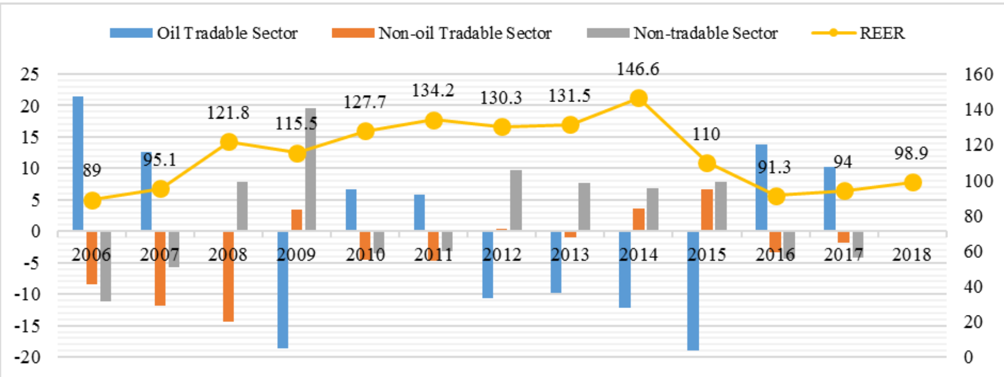
Figure 9 shows the output growth rates and the real exchange rate appreciation starting from 2006. Dutch disease is a phenomenon of a country’s chronic exchange rate overvaluation resulting from the discovery of cheap and abundant resources (Bresser-Pereira 2013). We saw that in Azerbaijan’s case it was not exactly “the discovery” but a preference for oil tradable sector during and after the painful transition process. Fast transferal of the natural resources to the international markets attracts huge revenue and capital to the economy. Surveyed literature from Magud and Sosa (2013) showed forty-one studies which supported that Dutch Disease led to the appreciation of the national currency while only eight research studies contradicted this. At the same time, the authors indicated that thirty-one investigations defended the thesis that natural resource revenue and capital inflow booms cause appreciation while opposite views were found seven times less. It is consistent with the case of Azerbaijan because, until the "low oil price" era of 2014–2015, overvalued national currency reinforced economic growth led to the highest GDP growth rates since independence from the USSR (34.5% in 2006). However, following the commodity price downturns, the economy showed national currency devaluations, banking bankruptcies, and declines in growth.

Figure 8 Share of the sectors in industrial products, 1990–2017
(% of total industrial products)



Source: Own construction based on SSCRA data (2019g)

Figure 9 Output growth (total output) in %, compared to the previous year, Real Effective Exchange Rate (right axis) (base year: 2000), 2006–2018



Sources: Own construction based on SSCRA (2019f) and Central Bank of Azerbaijan data

Figure 9 illustrates three main periods (2006–2008, 2009–2014 and 2015–2018) related to the annual output growth compared to the previous year among the economic sectors.[5] During the first period between 2006 and 2008, only oil tradable sector performed positive growth rates while both non-oil tradable sector and non-tradable sector indicated negative output performance. From Figure 3 we can observe that in 2008 year the revenue of the State Oil Fund increased 372.72 % compared to the previous year, and during the same year the real exchange rate overshoot was by 28%.

In the second period from 2009 until 2014, the most vulnerable sector in terms of output growth declines was oil tradable sector, dropping to –18.6% in 2009. Generally, between 2004–2011 years production and exports continuously rose in the oil and gas sector. However, coinciding with the global financial crisis, crucial oil and gas extraction projects were finalized in 2009. So, starting from 2011 oil production and exports declined due to the reductions in the outcome of the main oil and gas extraction project "Azeri-Chirag-Deepwater Gunashli." In the same period, non-oil sector did not show notable growth rates while non-tradable sector performed positively during 2012–2014. The appreciated position of Azerbaijani Manat did not create an export-friendly environment for economic agents, on top of which the spending effect of Dutch disease showed itself in increased share of non-tradable sector's outcome (Figures 7 and 9).

The third period draws attention to the sharp devaluation of the national currency and decreases in oil price. First, the oil price dropped to approximately 53$ in 2015 from almost 99$ in the previous year. Subsequently, the price effect influenced oil tradable sector's output pushing it down to –18.9% in 2015, however, non-oil tradable and non-tradable sectors performed solid output growth by 6.6% and 7.9% respectively. Second, for the years 2016 and 2017, the output of oil tradable sector was 13.8% and 10.2% respectively. This can be interpreted as additional stimulus of the main players of oil and gas industry because of the extension of the "The Contract of the Century"[6] which determined the main production and export patterns of the industry. Accordingly, the work from Hasanov and Samadova (2010) introduced the relationship between the real exchange rate and non-oil exports. The research indicates that a 1% appreciation of real exchange rate leads to a 1.63% decrease in non-oil exports in real terms and 1.46% decline in non-oil GDP in the long-term. However, in the short-term, a 1% appreciation of the real exchange rate leads to a 4.18% and 3.89% decline respectively among the aforementioned indicators.

Figures 10 and 11 describe the most vulnerable sub-sectors of the non-oil tradable sector in output to de-industrialization during the oil boom period. For example, Figure 10 is a good contrast between the old and new manufacturing value-

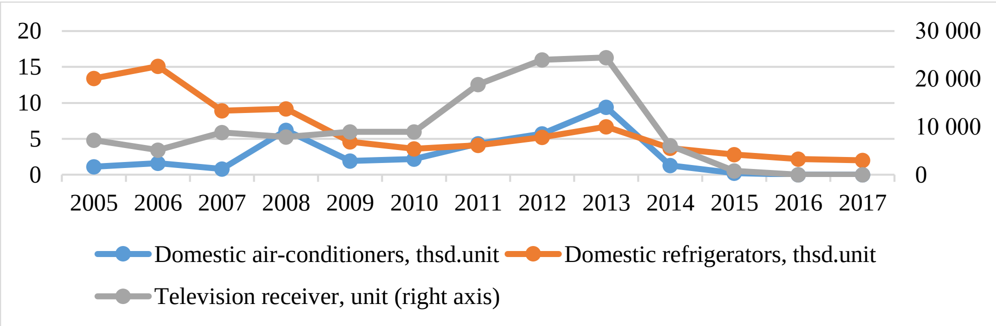
---

[5] As there is not a precise trend on the graph, the author divided the timescale into three conditional periods to shed light on the adopted research questions.

[6] "The Contract of the Century" drawn up in 1994 between Azerbaijan and eleven of the biggest oil companies from eight countries, was the first project among the post-soviet countries attracting a huge amount of western multinational investments.

added sectors. Still in Soviet years, Azerbaijan was the first union country where domestic air-conditioners started to be produced. The factory was opened in 1975 in Baku, under license from Toshiba. The production capacity was 425–430 thousand units (Bulanova 2019). According to the available data, domestic air conditioners witnessed slight increases during 2008–2013, however, in 2014 production bottomed out at 1.3 thousand units, which are just 0.3% of the possible capacity. The main reason why the revival of this production line has not been organized is related to the Soviet union's production traditions. In other words, during its period of operations, the air conditioners were constructed from imported finished components from various other union countries. Consequently, the collapse of USSR and not having appropriate government policies supporting manufacturing during the 90s and the oil boom period led to a sad ending. A worrying drop happened in the domestic refrigerators output too. Output fell from 15.1 in 2006 thousand units to 2 thousand units in 2017. During the aforementioned years, the appreciated national currency highly hindered domestic production. In contrast to the indicated sectors, production of the television receivers was a new industry in Azerbaijan, however, the output data clearly shows that despite remaining stable until 2013, the trend did not continue for long, leading to noticeable decreases. Also, there is no official data yet for the years like 2016 and 2017.

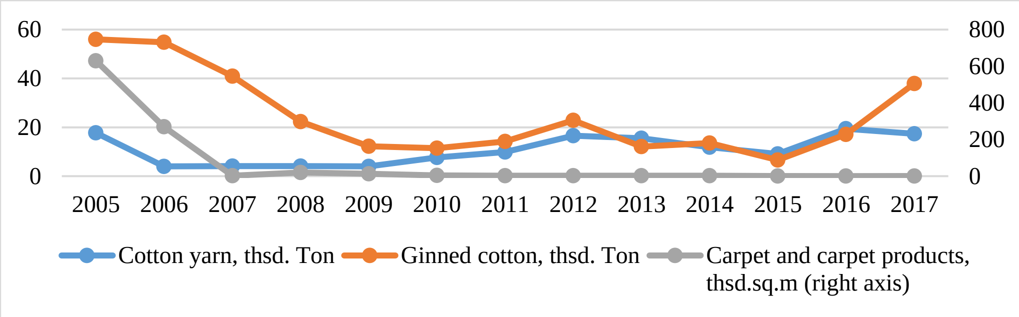
Figure 10 Television receiver, domestic air-conditioner, and domestic refrigerator productions, 2005–2017



Source: SSCRA (2019h)

From Figure 11 the overall impression reveals that output de-industrialization happened in the traditional manufacturing fields like cotton yarn, ginned cotton, and carpet production. Ginned cotton experienced a more detectable reduction in output terms, falling from 56 thousand tons in 2005 to 6.6 thousand tons in 2015. The only rise was during 2015–2017 surging to 37.9 thousands of tons in 2017, which is related to changes in the agriculture policy of the state. During the same time period, the output of carpets and carpet products fell by 99.9% in 2016 compared to 2005. Thus, it was an insignificant increase in the output of carpet and carpet products in 2017, by 0.4 thousand square meters. The main reason of de-industrialized carpet production can be viewed as a consequence of the withdrawal of state investments in the industry.

Figure 11 Production of cotton yarn, ginned cotton and carpet and carpet products, 2005–2017



Source: SSCRA (2019h)

4.3. Trade De-industrialization

Figure 12 illustrates export de-industrialization of particular value-added manufacturing sectors and the trends in agriculture and livestock export patterns from 1994 until 2017, while Figure 13 depicts import patterns of the same sectors of Azerbaijan. Figure 12 shows that some of the sectors underwent considerable export de-industrialization. For instance, textiles sector suffered the highest percentages during 1994 to 1998 (min. 9.2% – max. 19.5%) which started to fall from 1999 (2.9%) and bottomed out at around 0.1% in 2008.[7] Between the 2008–2017 period, a gradual increase was observed and the highest share was 0.6% in 2017.

The same fate was shared by machinery and mechanical appliances, electrical equipment and apparatus category as well, even though it fell significantly right after 1994 from a 13.7% share to 3.8% in 1999. After 2000 machinery and mechanical appliances, electrical equipment and apparatus category showed fluctuations but downsized to a historically low indicator – 0.1% in 2008. In 2017 the share of the category was 0.4%.

The overall downward trend occurred in prepared foodstuffs, beverages, spirits and vinegar, and tobacco (hereby food production). While exports declined from 7.5% in 1994 by 6.7% in 2017, the imports experienced gradual increases and fluctuations after 2004 (Figure 12). Agricultural exports shared a 2.5% portion of exports in the late 90s but after declines, they reached 3.9% in 2017. As can be seen from the “employment de-industrialization” section of this paper, employment in the agriculture sector did not witness a dramatic collapse, however, a large proportion of

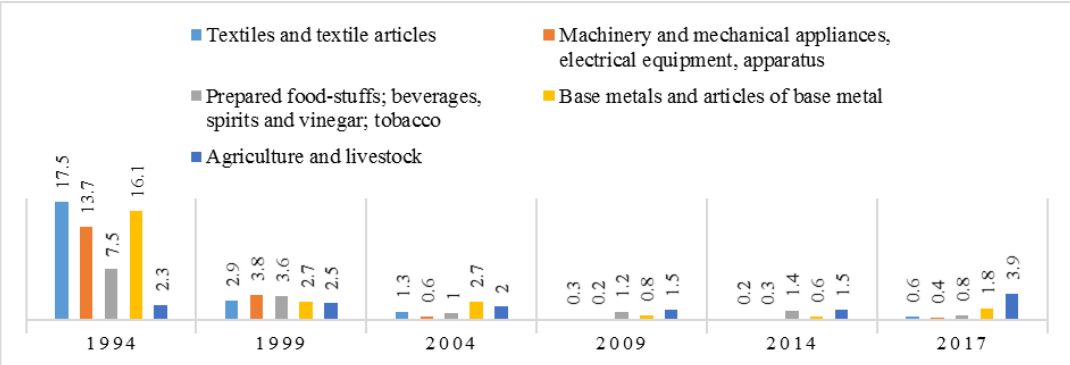
---

[7] Because of the long time series data, some of the years mentioned like 1993, 1994, 1995, 1996, 1997 and 2008 were not depicted in order to gain cleaner data visualization, however, indicated verbally to show the important signals. This aspect also was mentioned on the data source and methodology section of the paper.

the population is employed in this sector – 36.35% in 2017. Naturally, the increasing value added should be expected in exports.

Against the background of the aforementioned trade (export) de-industrialization, the extension of the Agreement of the Century - New Contract of the Century, until 2050 is the solid example of how the oil and gas sector plays an influential role in the economic policy of the Republic of Azerbaijan, and will continue its supremacy during in the near future. Thus, besides the oil and gas sector, other industries and especially, the manufacturing sector did not exhibit industrialization patterns after independence as can be be appreciated in Figures 7 and 8. In contrast to this, since 2000, the "mineral fuel, lubricants, and similar materials" category in the exports have dominated at 80%; peaking in 2008 at 97.1%, and continuing its dominance until 2017 by 89.5% (SSCRA 2019i).

Figure 12 Trade de-industrialization of Azerbaijan, 1999–2017
(% of the total exports)

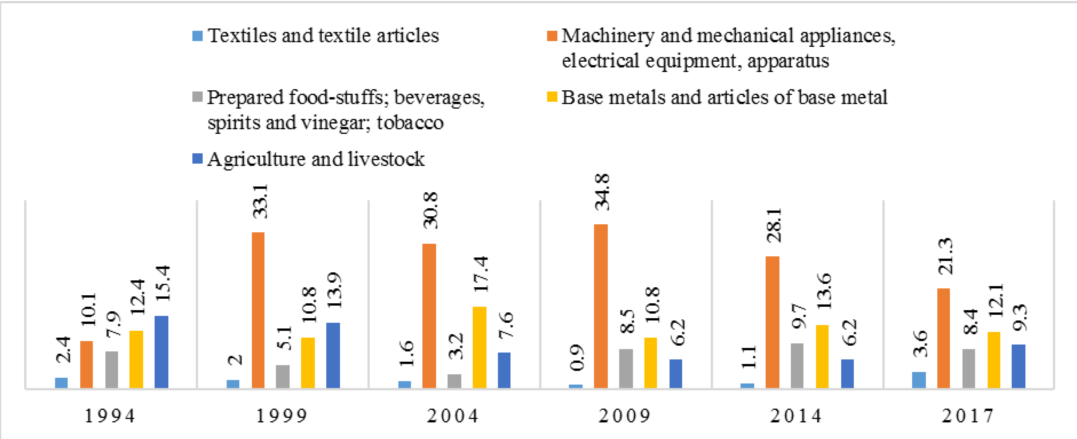


Source: SSCRA (2019j)

Consistent with the de-industrialization literature, some of the abovementioned categories showed import growth rates according to Figure 13. For instance, disease, an increase in the non-tradable sector's share in GDP. A similar conclusion was drawn from the investigation of Gurbanov et al. (2017). The authors argued that despite fiscal regulations and oil revenue management practices of SOFAZ, the production of industrial products had declined. Moreover, despite massive government expenditures and because of high volatility, those expenditures were not to serve as a fuel in non-oil production.the biggest category was machinery and mechanical appliances, electrical equipment, and apparatus, which had a high export share between 1994–1999 but due to the turndown started to be imported heavily from 1999 – 33.1%. The other categories depicted in Figure 13 had higher import shares compared to export shares during the given period.

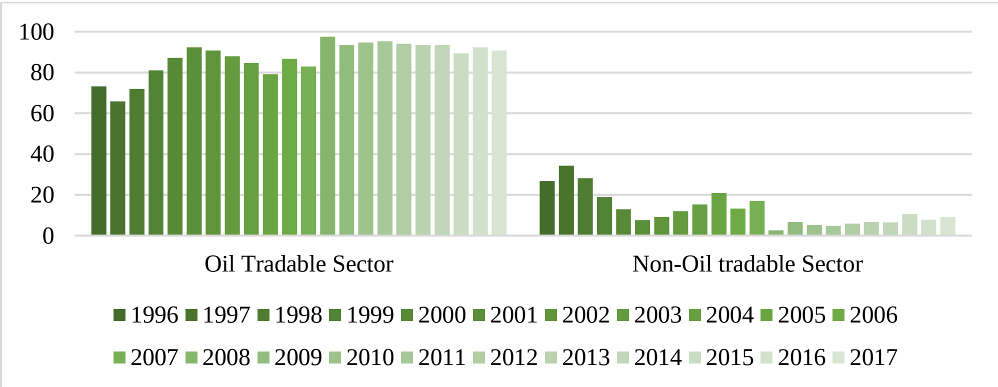
One could blame so-called "transition effect" of the transition period among the post-soviet economies starting from the early 90s, however, Hasanov (2013) indicated that transition effect did not play as statistically significant a role in the process of relative de-industrialization as did oil price as a proxy of the oil sector. He defended the thesis that the Dutch disease theory is an actuality for the Azerbaijan economy because of relative de-industrialization observations in the non-oil tradable sector, and quite naturally for Dutch

Figure 13 Trade de-industrialization of Azerbaijan, 1994–2017 (% of total imports)



Source: SSCRA (2019k)

Figure 14 Trade de-industrialization of Azerbaijan, 1996–2017 (% of total exports)



Source: Own construction SSCRA (2019k)

To sum up the trade de-industrialization aspect of this paper, Figure 14 illustrates how the share of oil tradable sector and non-oil tradable sector changed over 1994–2017 period considering the export data of Azerbaijan. The share of oil tradable in total exports achieved continuous growth until 2001, from 73.2% to 92.4%, however, starting from 2002 it gradually declined to bottom out at around 79.1% in 2005. After 2008 oil tradable mostly remained steady with slight downturns around

2015. This is because of the stabilization of the biggest oil and gas projects. We start to see a striking difference between oil tradable sector and non-oil tradable sector after 2007. For example, 97.5% for oil tradable sector compared to 2.5% of non-oil tradable in 2008. Two big periods explain non-oil trades: the first period, 1996–2007, when the highest share was 34% and the lowest share was 7.6%. The second period which started with the historically low indicator – 2.5% against 97.5% oil tradable sector in 2008 and moderate rises to 10.6% and 9.2% in 2015 and 2017 respectively. To conclude, trade data clearly indicates the de-industrialization pattern of non-oil tradable sector, especially overlapping with the oil revenue boom year period like 2008–2014.

5. Examination of Government Goals and Programs and Policy Measures

The national government always considered effective revenue management and diversification actions; however, most of them were mainly political tools on the eve of elections. It became clear especially in 2014–2015 that dependence on oil prices and revenue raised many challenges for the Azerbaijani economy. The Global Financial Crisis in the 2008–2009 period and recent downturns in oil prices influenced GDP growth, inflation, investments, and sectoral outputs. In fact, the non-oil tradable sector experienced negative growth rates between 2006–2009. Sharp drops in oil prices in 2014-2015 created severe external shocks for the Azerbaijan economy. Weak macroeconomic foundations and oil dependence of the economy influenced the country’s financial sector immediately, leading to double-digit inflation, less oil revenue, and a decline in industrial output (Ibadoghlu et al. 2013). In 2017, GDP fell 45.8% compared to 2014, whereas GDP per capita in PPP declined from $17,926.7 USD to $17,453 USD (2016). Furthermore, the budget deficit was –1.6% in 2017; however, in 2014 the indicator was –0.5% (State Oil Fund of the Republic of Azerbaijan 2017). While the real effective exchange rate was decreasing, the output of oil tradable fell by 12.2% and 18.9 % in 2014 and 2015 respectively. Drops in non-oil sector and non-tradable sector followed this trend during 2016 and 2017 (Figure 9). For these reasons, we can observe some comprehensive and multidimensional strategic approaches of the national government level starting in 2016.

5.1. Examination of Government Goals and Programs

During the years of oil revenue boom, we do not observe any noticeable non-oil tradable sector associated policy mechanisms. A “Long-Term Strategy for the Management of the Oil Revenue” was adopted to identify the main principles of oil revenue management and the middle-range cost policy in 2004. Only in 2011 was “Azerbaijan 2020: Outlook Into The Future” Concept of Development provided to shape a strategic approach based on the current opportunities and resources and attain a stage characterized by sustainable economic growth and high social welfare. The other pre-crisis policy was “State Program on the Development of Industry in the Republic of Azerbaijan in 2015–2020” in 2014. The program aims to increase the competitiveness of the overall industry by upgrading its structure and increasing non-oil GDP. Involvement of science, innovation, and human capital development are the main objectives of the state program. However, following the years 2016 and 2017, the non-oil tradable sector did not exceed the 2014 results in terms of almost all three industrialization levels (employment, output, and trade). Nuri Aras et al. (2016) and Bulut and Suleymanov (2012), indicated that government expenditures and investments are measures against Dutch disease in Azerbaijan, however instead of improving the performance of lagging sectors and increasing the observable outcome of the "2015–2020 State Program on Industrial Development", government spending and investments, raised the relative price of non-tradable sector, leading to indirect de-industrialization.

After the 2014–2015 period, the most important policy response was the Strategic Roadmap for the National Economic Outlook of the Republic of Azerbaijan in 2016. The document considers twelve strategic roadmaps in eleven sectors of the economy and makes responsible, appropriate policies for the identified targets. Establishment of the Legal Entity of Public Law Center for Analysis of Economic Reforms and Communication was another crucial step to build a holistic framework for the application and monitoring of the policies. The establishment aims to ensure transparency and agility by creating and supporting additional regulations for securities markets, investment funds, banking, and insurance, as well as payment systems, to respond to the recent developments in the world economy. Furthermore, the creation of the Financial Stability Committee aims to minimize the effects of sharp decreases in oil and gas revenue considering the volatile nature of commodity exports. It takes advance steps to protect the negative influences of the global crisis in the mid and long-term periods. Meanwhile, various other policy actions like “Additional Actions to Investment Promotions”, “Additional Measures on Promotion of the Non-oil Products” and “The Cancellation of the Inspections in Entrepreneurship” have been made in order to minimize the effects of future commodity price shocks mainly via non-oil tradable sector.

Additionally, among the policies implemented, the Regional State Development Program of the Azerbaijan Republic on Economic and Social Development stands out. It included the years 2004–2008 and 2009–2013, 2014–2018. The new program was approved recently by the president of the country and has included the period of 2019–2023. The main focus of these programs is regional development. It aims to improve such objectives as increasing local production, enhancing entrepreneurship potential, increasing employment, and stimulating exports-oriented production. On the last decree, it was mentioned that more than 2 million new workplaces were provided, of which 1.5 million are permanent positions. Nonetheless, according to the official state statistics, both oil and non-oil tradable sector on employment, output and trade level witnessed downsizing. This is especially inconsistent with the targets and objectives of the indicated policy responses before and after the low oil and gas price era of 2014–2015. After such comprehensive policy mechanisms and non-oil development initiatives, the rural regions of Azerbaijan are still underperforming. According to the year’s data from 2017, the three largest sources of household revenue are, free farming – 29.3%, patronage – 18.1%, and pensions – 15.5%, while the private sector is 4.2% and agriculture is only 0.4% of household revenue (SSCRA 2019l).

We have seen certain state development programs targeting specific sub-sectors of manufacturing in order to respond to such severe de-industrialization process. For instance, the "State program on carpet weaving art and development in the Republic of Azerbaijan for 2018–2022 years" was adopted to improve the export potential of carpet products by upgrading human and physical capital of the sector. The other policy is related to cotton production – "State program on development of cotton growing in the Republic of Azerbaijan for 2017–2022 years" which targets export-oriented production with innovative infrastructure. It also demands the necessary institutional and legal regulations in this sphere. This is the first time after independence that specific manufacturing sectors like the production of the carpet and carpet products and cotton have become an object for the development program. More recent policy and institutional activities involve the establishment of the Social Research Center and Ensurement of Innovative Development (which aims to systematize innovation infrastructure and coordinate the relationship among the important government institutions like National Academy of Sciences, Ministry of Finance, Ministry of Economy and etc.).

5.2. Examination of Policy Tools and Measures

This sub-chapter analyzes two important and concrete policy tools related to oil revenue management aside from the abovementioned government goals and programs. In a chronological order, the first policy tool is the devaluation of the national currency in 2015, and the second policy measure is the adoption of the fiscal framework which covers the stabilization mechanism of transfers from the State Oil Fund, regulation of middle-term expenditures and a strategy for government debt management in the Azerbaijan Republic in the middle and long-term in 2018.

Because of its heavy reliance on oil tradable exports, the economy of Azerbaijan is utterly dependent on the oil revenue. As Bayramov and Abbas (2017) noted, Azerbaijan's GDP performance dropped by 5.8% to 2.8% in 2014 and then to 1.1% in 2015 as the latest petroleum crisis led to a substantial recession. In addition, earnings in the 2014 budget decreased from US$ 23.6 billion in 2014 to US$ 16.9 billion. In 2015, 60% of the entire budget was sourced from the State Oil Fund of the Republic of Azerbaijan amounting to US$ 10.2 billion. Therefore, the painful oil shock of the sharp commodity price downturns in 2014 and 2015 led to the adoption of certain policy measures. The Central Bank of Azerbaijan identified the new fixed exchange rate of US$ to AZN as 1.05 AZN on 21 February 2015 (Statement of the Central Bank 2015a). The Central Bank of Azerbaijan declared "a floating exchange rate" and a second devaluation on 31 December 2015 as 1 US$=1.5610 AZN (Statement of the Central Bank 2015b). According to The Central Bank of Azerbaijan, this action created more maneuvering capabilities for the government to adapt the value of the national currency according to oil prices. Furthermore, such a policy tool is commonly seen to overcome the competitiveness issue of the tradable sector. In Azerbaijan's case, there was a rational explanation why the government used this policy measure, however, following that year (2015) 15 banks closed because they could not meet the necessary capital requirements and still the floating exchange rate regime is not adequate for the economy. Hence, it is extremely important to evaluate this policy tool in a brief way of pros and cons from the international literature perspective.

The role of exchange rate manipulation or currency devaluation is a controversial topic. The orthodox opinion has defended devaluation as serving a beneficial and significant purpose in stabilizing the equilibrium of transactions (through its cost-switching impacts and enhanced tradeable outputs), while the New Structuralist School has noted that exchange-rate adaptation has contractionary impacts (Agenor 1991). For example, Corsetti et al. (2000) emphasized the positive impact of competitive devaluations. In Azerbaijan's case, the devaluation of 2015 was not solely caused by the decision related to increasing competitiveness, however, the general welfare mechanism can be considered the same. Moreover, considering Bhalla (2008) work, investment activity and economic growth can be explained by the currency undervaluation, because devaluation promotes economic growth while the opposite harms the growth. In contrast, there are several nuances to the national currency devaluation which makes it highly uncertain policy tool. Mironov (2015) indicated that the outcomes of the undervalued or balanced exchange rate can be different on output and certain macroeconomic indicators and overall economic growth. It is a very important policy implication for the decision-makers in the government. Economists have always reserved a certain level of skepticism for the effect of national currency devaluation, supporting their claims by empirical analysis. An early indication came from Krugman and Taylor (1976) emphasizing the risks of devaluation for developing countries because of the high probability of structural reforms failing. Similarly, the contractionary effects of devaluation on output were identified by Edwards (1985) analyzing the effect of the real exchange rates on real output growth in a group of twelve countries between 1965–1980. The research for the given years provided the insight that in the short-run devaluation was contractionary and after one year it may have an expansionary effect, however, in the long-run it has no effect on output. Also, as Mironov (2015) summarizes, devaluation is a country-specific phenomenon always highly sensitive to the tools used against the background of degrading credit ratings, slowing investment activity and bearing a high risk of creating inflation. Thus, this policy tool requires a broader time range to observe its consequences.

The second policy tool is the adoption of the "New Budget Policy"[8] to regulate government spending to overcome the high-level dependency on the oil

[8] This is an umbrella term used among experts in Azerbaijan. There is no official document which names these policy measures as described here. Due to the reason that there are several parts and stages (some

revenue. It means putting certain qualitative limitations on the different parameters of the state budget to overcome the issue, sourcing from short-term based interests of different decision-makers and institutions and prioritizing long-term strategic interests of the national economy. In fact, well-developed budget rules can cushion the blow of the shocks for the macroeconomic institutions, stabilize exchange rate, neutralize inflation and minimize the non-resource revenue deficit of the budget, and it is being used among 80 countries (Eurasia Extractive Industries Knowledge Hub). Accordingly, fiscal rules against the pro-cyclical budget policy were included in Strategic Roadmap for the National Economic Outlook of the Republic of Azerbaijan after the oil price shocks of 2015. In fact, "Strategic Target 1: Empowering Fiscal Sustainability and Adoption of the Sustainable Monetary Policy" under the "Strategic Targets" chapter, strongly emphasized the dependency of the Azerbaijan national economy on oil revenue and formed conceptual foundations of the new multidimensional policy tool. To do so, the adoption of the "Golden Rule" principle was specified to regulate specific parts of the planned fiscal framework. The four main components of this rule are the following: regulation of the transfers of the oil revenue to the state budget, creation of the mechanism of the expenditure and investment discipline, emboldening mid-term expenditure framework and execution of the result-oriented budget mechanism (Strategic Roadmap for the National Economic Outlook of the Republic of Azerbaijan).

As is well known, macroeconomic stability can be threatened considerably by the pro-cyclical budget policy. Unlimited or unregulated government spending may create huge government debt risks for the economy consistent with the Dutch disease theory. To reduce the dependency level of non-oil sector from the resource revenue and to boost its revenue-creation potential, new fiscal rules seem to be important institutional response. So, the general mechanism of the possible fiscal rules covers the necessary qualitative restrictions on the size of public debt, budget deficit, level of spending and volume of income. As the first step towards the "New Budget Policy", Parliament made the necessary amendment to the "Law on Budget System" on June 29, 2018 (Eurasia Extractive Industries Knowledge Hub). The changes cover the following: firstly, the upper limit of each following financial year's consolidated budget spending may not exceed 103% of the previous year's approved spending. Secondly, so-called the "Golden Rule" for the spendable oil revenue has to be calculated. The mechanism is the following:

- First, 30% of the difference between the net financial assets and forecast oil revenue for the predicted budget year must be derived;
- Second, 20% of the obtained difference must be calculated;
- Third, the number found on the second step has to be summed with the least indicator among the abovementioned two.

---

still ongoing) of this policy measure, evaluating it is a complex process. For practical purposes, occasionally this term will be used in the necessary places.

Secondly, sounding with fiscal rules, certain regulations will address budget discipline. Briefly, budget discipline simply tries to enhance productivity and sustainability of expenditure policy based on the "Golden Rule". More effective expenditure policy will facilitate only those projects and expenditures that aim for rational implementation based on profitability and return on investment. So, more concrete and accurate rationale will be required to finance expenditures instead of traditional resource distribution method implemented since independence.

The third element of this conceptual approach is a mid-term expenditure framework which aims to achieve macroeconomic stability, better sustainable economic growth based on fiscal discipline, to prioritize expenditure directions among the sectors sourcing from resource revenue, and to create the preconditions for the measurable budget mechanism (Rules of the Preparation of Middle-term Expenditure Framework 2018). However, this is still an ongoing process compared to the other parts of the fiscal framework and will take several years before the complete application. The lack of the data and insight does not allow us to evaluate its role and perspectives yet.

The last piece of the conceptual approach covers the policy measures that aim to focus firmly on strategic goals and programs instead of conventional expenditure items, to clearly state the expectations of state entities, to design the state budget based on thee identified expectations from these state entities, to create supervision mechanism related to the spending patterns of these state entities, and to form the budget based on the evaluation of the achievement rates and efficiency of the state entities. This last measure has a very high potential to boost the effectiveness of state budget expenditures on the institutional level if applied transparently.

Some associated opportunities and risks related to the new fiscal framework should also be noted at this point. Surely, limiting the transfers and the spending sourced directly from oil revenue might create a better transition to countercyclical policies, while reducing state debt might lead to better macroeconomic stability. However, this policy tool still does not guarantee lower dependency on oil revenue, and it does not provide the necessary insurance for the leverage of savings. The formula is still dependent on the exchange rate and the legal conditions for budgetary law in the fiscal year are rather weak, and the effect of fiscal frameworks can be reduced substantially by periodic disruptions during the implementation of these laws (Eurasia Extractive Industries Knowledge Hub).

6. Concluding Remarks

This study sought to investigate the patterns of the de-industrialization process that happened through the resource movement effect and spending effect of Dutch disease hypotheses in Azerbaijan after independence. In order to do so, two research questions were constructed: Does the employment, output and trade data show the direct and indirect de-industrialization patterns as the result of the resource movement and the spending effect of Dutch disease in Azerbaijan? What are the policy responses of the national government in response to critical years like 2008–2009 and 2014–2015 to alleviate the primary export-based economy and particularly value-added manufacturing sector from the de-industrialization process?

The main findings were based on the descriptive statistics on three levels: employment, output and trade de-industrialization, as well as three sector level, namely, oil tradable sector, non-oil tradable sector and non-tradable sector. Consequently, despite oil tradable sector is the booming sector of the economy, it showed slight employment decline between 2008 and 2017 (from 1.88% to 1.31%) but considerable output falls starting from 2011 (from 40.4% to 23.2% in 2015). However, it greatly dominated in exports starting from 2008 till 2017 by 93% average annual percentage. Consistent with Dutch disease syndrome, the most vulnerable to the direct and indirect de-industrialization were non-booming sectors, in Azerbaijan's case non-oil tradable sector. The employment data shows that agriculture witnessed a 5% decrease while manufacturing value-added sectors like machinery, production of electrical and electron equipment, weaving, and cloth production experienced severe de-industrialization. The non-oil tradable sector did not escape output and trade decrease either. High government expenditures sourcing from the huge foreign currency led to high shares of non-tradable in employment and output.

The national government created more sophisticated policy responses after the "low oil price" era; however, prior considerations concerning diversification and regional developments are not consistent with the statistical data. The economy of Azerbaijan is still dependent on the revenue from crude oil and gas exports and experienced de-industrialization during the 90s and oil revenue boom.

The government of Azerbaijan should consider exactly the certain de-industrialization process as a consequence of Dutch disease hypotheses, so the implementation of solutions for this issue may find its way to the table of the decisionmakers. As a policy response, protection of the traditional value-added sectors like carpet production and cotton production can be observed, however other policy measures aim to regulate oil revenue transfers from the sovereign oil fund and to regulate expenditures in short and middle term. Surely, it will lead to better environment for the non-state investments but carry associated risks alongside specific approaches which have to be considered. Finally, exchange rate policies should be developed to protect foreign trade structures and regulate competition-related issues.

Value-added manufacturing and in the case of Azerbaijan, the non-oil tradable sector is vitally important to stable economic growth during fluctuations in the international commodity markets. Starting from the beginning of the twentieth century, many studies have supported manufacturing led growth hypotheses. Even in today's world, under the dominance of the services sector, the manufacturing sector is crucial for the evolutionary path of a country's economy. Subsequently, the de-industrialization process, namely shrinkage of employment, output, and trade in the value-added manufacturing or in other sectors lead to the structural changes. The phenomenon gets even more important in the case of post-soviet and transition countries. Dutch disease hypothesis is one of the theoretical models which explain de-industrialization through the booming and lagging sectors. Thus, the symptoms of Dutch disease can be observed in the Azerbaijan economy through main trends in the booming sector, non-oil tradable sector, and non-tradable sector. Appropriate policy measures also must follow economic policy decisions in a direct and clear way addressing the institutions and economic agents of the national economy.

References

Agenor, P. R. (1991): Output, devaluation and the real exchange rate in developing countries. *Review of World Economics*, 127, 1, 18–41.

Amirapu, A. – Subramanian, A. (2015): Manufacturing or services? An Indian illustration of a development dilemma. Washington, DC: *Center for Global Development and Peterson Institute*, 408.

Autor, D. H. – Levy, F. – Murnane, R. J. (2003): The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118, 4, 1279–1333.

Auty, R. (2001), Reforming resource-abundant transition economies: Kazakhstan and Uzbekistan. In Auty R. (ed.), *Resource Abundance and Economic Development*, New York: Oxford University Press, 260–276.

Badeeb, R. A. – Lean, H. H. – Clark, J. (2017): The evolution of the natural resource curse thesis: A critical literature survey. *Resources Policy*, 51, 123–134.

Balassa, B. (1964): The purchasing-power parity doctrine: a reappraisal. *Journal of Political Economy*, 72, 6, 584–596.

Bayramov, V. – Abbas, G. (2017): Oil shock in the Caspian Basin: Diversification policy and subsidized economies. *Resources Policy*, 54, 149–156.

Bhalla, S. (2008): Economic development and the role of currency undervaluation. *The Cato Journal*, 28, 2, 313–340.

Bogliaccini, J. A. (2013): Trade liberalization, deindustrialization, and inequality: Evidence from middle-income Latin American countries. *Latin American Research Review*, 48, 2, 79–105.

Brady, D. – Denniston, R. (2006): Economic globalization, industrialization and deindustrialization in affluent democracies. *Social Forces*, 85, 1, 297–329.

Bresser-Pereira, L. C. (2013): The value of the exchange rate and the Dutch disease. *Brazilian Journal of Political Economy*, 33, 3, 371–387.

British Petroleum Company (2018): *BP statistical review of world energy*. https://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review/bp-stats-review-2018-full-report.pdf Accessed: December 13, 2018.

Bulanova, O. (2019): Vpervie v SSSR: Bakinskiy Zavod Bitovix Kondicionerov. *Azer History*, https://azerhistory.com/?p=12833, Accessed: February 2, 2019.

Bulut, C. – Suleymanov, E. (2012): Importance of economic collaboration between Azerbaijan and Turkey in preventing Dutch Disease in Azerbaijan. *International Symposium on Regional Cooperation and Development*, 17–20.

Ciarreta, A. – Nasirov, S. (2012): Development trends in the Azerbaijan oil and gas sector: Achievements and challenges. *Energy Policy*, 40, 282–292.

Corden, W. M. – Neary, J. P. (1982): Booming sector and de-industrialisation in a small open economy. *The economic journal*, 92, 368, 825–848.

Corden, W. M. (1996): James Meade 1907–1995. *Economic Record*, 72, 217, 172–174.

Cornell, S. E. (2015): *Azerbaijan since independence*. London and New York: Routledge.

Corsetti, G. – Pesenti, P. – Roubini, N. – Tille, C. (2000): Competitive devaluations: toward a welfare-based approach. *Journal of International Economics*, 51, 1, 217–241.

Dasgupta, S. – Singh, A. (2006): Manufacturing, services and premature deindustrialization in developing countries: A Kaldorian analysis. WIDER Research Papers, *United Nations University (UNU)*, 49.

Edwards, S. (1985): Are devaluations contractionary? Working Paper No. 1676. *National Bureau of Economic Research*

Egert, B. (2009): Dutch disease in former Soviet Union: Witch-hunting? BOFIT Discussion Paper No, 4/2009, *CASE Network Studies and Analyses* No. 380.

Egert, B. (2012): Dutch Disease in the post-soviet countries of Central and South-West Asia: How contagious is it? *Journal of Asian Economics*, 23, 5, 571–584.

Eurasia Extractive Industries Knowledge Hub. http://eurasiahub.khazar.org/pdf/Summary_EN_Fiscal_Rules.pdf Accessed: 25 June, 2019.

Ezeala-Harrison, F. (1993): Structural Re-Adjustment in Nigeria: Diagnosis of a severe Dutch Disease syndrome. *American Journal of Economics and Sociology*, 52, 2, 193–208.

Fagerberg, J. – Verspagen, B. (1999): 'Modern Capitalism' in the 1970s and 1980s. In Setterfield M. (ed): *Growth, employment and inflation*. London: Palgrave Macmillan, 113–126.

Fagerberg, J. – Verspagen, B. (2002): Technology-gaps, innovation-diffusion and transformation: an evolutionary interpretation. *Research Policy*, 31, 8–9, 1291–1304.

Forsyth, P. J. – Kay, J. A. (1980): The economic implications of North Sea oil revenues. *Fiscal Studies*, 1, 3, 1–28.

Forsyth, P. J. – Kay, J. A. (1981): Oil revenues and manufacturing output. *Fiscal Studies*, 2, 2, 9–17.

Gahramanov, E. – Fan, L. S. (2002): The "Dutch Disease" in Caspian Region: The Case of Azerbaijan Republic. *Economic Studies*, 5, 10, 9–30.

Gregory, D. – Johnston, R. – Pratt, G. – Watts, M. – Whatmore, S. (eds.) (2009): *The Dictionary of Human Geography*. Malden and Oxford: John Wiley & Sons, 380–382.

Gurbanov, S. – Merkel, E. T. (2010): Avoiding the Dutch Disease: a comparative study of three successful countries. *Journal of Qafqaz University*, 29, 21–27.

Gurbanov, S. – Nugent, J. B. – Mikayilov, J. (2017): Prudent management of oil revenues: impacts of government capital expenditures in Azerbaijan. *In Energy Economics Emerging from the Caspian Region: Challenges and Opportunities*, 1st IAEE Eurasian Conference, Aug 28–31, 2016. International Association for Energy Economics.

Hasanov, F. (2010): The impact of real oil price on real effective exchange rate: The case of Azerbaijan. *German Institute for Economic Research Discussion Paper* 1041.

Hasanov, F. (2013): Dutch disease and the Azerbaijan economy. *Communist and Post-Communist Studies*, 46, 4, 463–480.

Hasanov, F. – Samadova, I. (2010): The impact of real exchange rate on non-oil exports: The case of Azerbaijan. *MPRA Paper*, 29556, 1–19.

Hartwick, J. (1977): Population growth, technical progress, intergenerational equity and the investment of resource rents. *American Economic Review*, 67, 5, 972–974.

Hegyi-Kéri, Á. (2016): A Study of the process and types of deindustrialization in the Visegrád countries. *Theory Methodology Practice: Club of Economics in Miskolc*, 12, 2, 5–13.

High, S. (2013): “The Wounds of Class”: A historiographical reflection on the study of deindustrialization, 1973–2013. *History Compass*, 11, 11, 994–1007.

Ibadoghlu, G. – Alasgarov, K. – Bayramov, G. (2013): Oil and gas revenue management in Azerbaijan. Policy Paper on Revenue Management in Azerbaijan. Baku: *Economic Research Center*.

Ibrahimov, A. (2016): Impact of post-soviet transition on the economy of Azerbaijan, *Geopolitica.info*. https://www.geopolitica.info/impact-of-post-soviet-transition-the-economy-of-azerbaijan/ Accessed: February 1, 2019.

Iversen, T. (2001): The dynamics of welfare state expansion: trade openness, de-industrialization, and partisan politics. Pierson, P. (ed.): *The New Politics of the Welfare State*. Oxford: Oxford University Press.

Kaldor, N. (1966): *Causes of the Slow Rate of Economic Growth of the United Kingdom*. Cambridge: Cambridge University Press.

Kaldor, N. (1967): *Strategic factors in economic development*. New York State School of Industrial and Labor Relations. Cornell University

Karpov V. P. (2010): Neft i gaz v promishlennoy politike SSSR (Rosii). *Vestnik Nijnevartovskogo Gosudarstvennogo Universiteta*, 4, 75–88.

Kaser, M. (2003): The economic and social impact of systemic transition in central Asia and Azerbaijan. *Perspectives on Global development and Technology*, 2, 3, 459–473.

Kemp, T. (2013): *Historical patterns of industrialization*. London and New York: Routledge.

Koritz, D. (1991): Restructuring or destructuring? Deindustrialization in two industrial heartland cities. *Urban Affairs Quarterly*, 26, 4, 497–511.

Kollmeyer, C. (2009): Explaining deindustrialization: How affluence, productivity growth, and globalization diminish manufacturing employment. *American Journal of Sociology*, 114, 6, 1644–1674.

Koo, H. – Hayashi, Y. – Weng, D. – Bi, J. (2016): De-industrialization and the changes in occupational structure in three east Asian cities. *Development and Society*, 45, 3, 439–465.

Krugman, P. R. – Taylor, L. (1976): Contractionary effects of devaluation. *Journal of International Economics*, 8, 445–456.

Larsen, E. R. (2006): Escaping the resource curse and the Dutch disease? When and why Norway caught up with and forged ahead of its neighbors. *American Journal of Economics and Sociology*, 65, 3, 605–640.

Laurila, J. (1999): Power politics and oil as determinants of transition: the case of Azerbaijan. *Russian & East European Finance and Trade*, 37, 3, 25–76.

Lewis, W. A. (1954): Economic development with unlimited supplies of labour. *The manchester school*, 22, 2, 139–191.

Lin, J. Y. (2015): The Washington Consensus revisited: a new structural economics perspective. *Journal of Economic Policy Reform*, 18, 2, 96–113.

Macrotrends, Brent Crude Oil Prices: https://www.macrotrends.net/2480/brent-crude-oil-prices-10-year-daily-chart Accessed: January 10, 2019.

Magud, N. – Sosa, S. (2013): When and why worry about real exchange rate appreciation? The missing link between Dutch disease and growth. *IMF Working Papers 10/271*. International Monetary Fund.

Mironov, V. (2015): Russian devaluation in 2014–2015: Falling into the abyss or a window of opportunity? *Russian Journal of Economics*, 1, 3, 217–239.

Moholwa, M. (2017): The impact of re-industrialization on economic growth and employment in South Africa. Doctoral dissertation. https://repository.up.ac.za/bitstream/handle/2263/64885/Moholwa_Impact_2018.pdf?sequence=3&isAllowed=y (Accession No. 02606356). Accessed: 13 February, 2019.

Noorbakhsh, F. – Paloni, A. (1999): Structural adjustment programs and industry in Sub-Saharan Africa: restructuring or de-industrialization? *The Journal of Developing Areas*, 33, 4, 549–580.

Nuri Aras, O. – Suleymanov, E. – Mammadov, K. (2016): Economy of Azerbaijan. Baku: Sharg-Garb Publishing House.

Olamade, O. – Oni, O. (2016): Manufacturing and economic growth in Africa: A panel test of kaldor's first growth law. *Journal of Economics and Sustainable Development*, 7, 22, 126–140.

Ozturk, S. – Altingoz, B. (2018): Imalat Sektoru Firma Karliliginin Ekonomik Buyume Uzerine Etkisi: Turkiye Icin Bir Uygulama. *Omer Halisdemir Universitesi Iktisadi ve Idari Bilimler Fakultesi Dergisi*, 11, 3, 18–25.

Palma, J. G. (2008): De-industrialization, 'premature' de-industrialization and the Dutch Disease. *The New Palgrave Dictionary of Economics*, 1–8, 1297–1306.

Robert-Nicoud, F. (2008): Offshoring of routine tasks and (de) industrialisation: Threat or opportunity – And for whom? *Journal of urban Economics*, 6, 2, 517–535.

Rodrik, D. (2016): Premature deindustrialization. *Journal of Economic Growth*, 21, 1, 1–33.

Rowthorn, B. – Wells, J. R. (1987): *De-industrialization and foreign trade*. Cambridge: Cambridge University Press.

Rowthorn, R. – R. Ramaswamy (1997): Deindustrialization: Causes and Implications. *IMF Working Paper 97/42*. Washington, DC: IMF

Ross, A. C. (1986): The United Kingdom's Experience with North Sea Oil and Gas. *IDS Bulletin*, 17, 4, 1–6.

Rules of the Preparation of Middle-term Expenditure Framework (2018). Available: https://president.az/articles/29735, Accessed: July 18, 2019.

Sadik-Zada, E.R. – Loewenstein, W. – Hasanli – Y. Miner Econ (2019): Production linkages and dynamic fiscal employment effects of the extractive industries: input-output and nonlinear ARDL analyses of Azerbaijani economy. *Mineral Economics*, 1–16.

Samuelson, P. A. (1964): Theoretical notes on trade problems. *The Review of Economics and Statistics*, 145–154.

Singh, A. (1977): UK industry and the world economy: a case of deindustrialization. *Cambridge Journal of Economics*, 1, 2, 113–136.

Singh, R. – Laurila, J. (1999): Azerbaijan: recent economic developments and policy issues in sustainability of growth. *Bank of Finland Institute for Economies in Transition*, Discussion Papers No. 5.

SSCRA (2019a): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/system_nat_accounts/en/021en.xls Accessed: January 15, 2019.

SSCRA (2019b): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/industry/en/016en.xls Accessed: January 15, 2019.

SSCRA (2019c): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/industry/en/007en.xls Accessed: January 18, 2019.

SSCRA (2019d): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/labour/en/002_1-2en.xls Accessed: January 18, 2019.

SSCRA (2019e): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/industry/en/014_2en.xls Accessed: January 19, 2019.

SSCRA (2019f): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/system_nat_accounts/en/018en.xls Accessed: January 19, 2019.

SSCRA (2019g): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/industry/en/006en.xls Accessed: January 20, 2019.

SSCRA (2019h): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/industry/en/008en.xls Accessed: January 20, 2019.

SSCRA (2019i): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/industry/en/016en.xls Accessed: January 20, 2019.

SSCRA (2019j): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/trade/en/f_trade/xt008_4en.xls Accessed: January 25, 2019.

SSCRA (2019k): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/trade/en/f_trade/xt008_2en.xls Accessed: January 26, 2019.

SSCRA (2019l): State Statistical Committee of the Republic of Azerbaijan. Available: https://www.stat.gov.az/source/budget_households/az/174-175.xls Accessed: January 26, 2019.

Simmons, C. (1985): 'De-industrialization', Industrialization and the Indian Economy, c. 1850–1947. *Modern Asian Studies*, 19, 3, 593–622.

Stanford, J. (2012): A cure for Dutch disease: Active sector strategies for Canada's economy. *Ottawa, ON: Canadian Centre for Policy Alternatives*.

State Oil Fund of the Republic of Azerbaijan (2017): Annual Report. Available: https://www.oilfund.az/report-and-statistics/get-download-file/7_2017_tam_en.pdf Accessed: February 12, 2019.

Statement of the Central Bank of the Republic of Azerbaijan (2015a). Available: https://www.cbar.az/press-release-919/statement-of-the-central-bank-of-the-republic-of-azerbaijan?language=en Accessed: June 21, 2019

Statement of the Central Bank of the Republic of Azerbaijan (2015b). Available: https://www.cbar.az/press-release-863/statement-of-the-central-bank-of-the-republic-of-azerbaijan Accessed: June 21, 2019

Strategic Roadmap for the National Economic Outlook of the Republic of Azerbaijan https://static.president.az/pdf/38542.pdf Accessed: July 17, 2019

Szirmai, A. (2013): Manufacturing and economic development. In Szirmai, A., Naudé, W., Alcorta, L. (Eds.): *Pathways to industrialization in the twenty-first century. New challenges and emerging paradigms*. Oxford: Oxford University Press.

Szirmai, A. – Naude, W. – Alcorta, L. (2013): Introduction and overview: the past, present and future of industrialization. In Szirmai, A. – Naudé, W. – Alcorta, L. (Eds.): *Pathways to Industrialization in the Twenty-First Century. New Challenges and Emerging Paradigms*. Oxford: Oxford University Press.

Ulbrych, M. (2017): The role of manufacturing in Central and Eastern EU member states. In The Proceedings of the 17th International Joint Conference, *Central and Eastern Europe in the Changing Business Environment*. Bratislava: Vydavateľstvo Ekonóm, 245–260.

The Economist (1977): The Dutch disease. *The Economist*, 26, 82–83.

Van Neuss, L. (2018): Globalization and deindustrialization in advanced countries. *Structural Change and Economic Dynamics*, 45, 49–63.

Williamson, J. G. (2006): Globalization, de-industrialization and underdevelopment in the third world before the modern era. *Revista de Historia Economica-Journal of Iberian and Latin American Economic History*, 24, 1, 9–36.

World Bank (1995): *Statistical Handbook 1995, States of the Former USSR*. Washington: The World Bank.

Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 397–414.

Is the Ricardian Equivalence Hypothesis Valid? An Empirical Study for Ethiopia

Sisay Demissew Beyene – Balázs Kotosz

There are two main dimensions of dealing with the topic of Ricardian Equivalence Hypothesis (REH) – the Keynesian proposition and the REH itself. According to the REH, today's borrowing to stimulate the economy or tax reduction – a substitution of debt for taxes -does not affect demand and consumption levels. However, the reverse is true for the Keynesian proposition. The objective of this paper is to test the validity of the REH in Ethiopia using annual data running from 1990 to 2011 by employing the bound testing – ARDL approach. The study included three main variables (the budget deficit, government consumption expenditure, and government debt) which contribute to the REH along with another variables. The result shows that government debt failed to fulfil REH. This implies that, in this study, we found limited evidence of the validity of the REH in Ethiopia.

Keywords: REH, ARDL, Ethiopia

1. Introduction

The Ricardian equivalence theorem was formulated, as the name suggests, by the British classical economist David Ricardo, who went on to immediately dismiss it as being irrelevant. However, the neoclassical economist Robert Barro forcefully argued that REH is worthy of professional attention and yields important policy prescription (Heijdra 2002).

There are two main points of view when dealing with the REH. These are the Keynesian proposition and the Ricardian equivalence hypothesis itself. According to Okpanachi and Abimiku (2007) cited in Daylop (2010), Keynesians argue that an increase in government spending by running budget deficits enhances domestic output and this stimulates the economy in the short run by making households feel wealthier, thus raising total private and public consumption expenditure. Besides, according to the Keynesian argument, consumers treat government debt as net wealth. Therefore, substitution of debt for taxes has a positive impact on private consumption and aggregate demand, even though it increases the real interest rate and leads to crowd out private investment and the economy slowing as well (Marinheiro 2001). However, in a sticky price model within a large economy, a fiscal expansion increases the real interest rate. This increase, in turn, leads to a fall in private consumption (Kim–Roubini 2008).

Nevertheless, the Ricardians argue that since a tax cut now is a tax increase in the future, the substitution of debt for taxes has no effect on aggregate demand or on the interest rate. Besides, potential consumers assume that today's borrowings are the postponed taxes of the future, the consumption level of consumers remaining unchanged due to their savings today. Further, under the REH, consumers respond to a reduction in tax by increasing their savings by buying securities, which helps to pay the increased future taxes and to repay future debt. Hence, when private savings increase by the same amount as the budget deficit, both the national savings and the interest rate will be unchanged. Hence, for a given expenditure path, financing the public expenditure either by debt or taxation does not affect private consumption (Marinheiro 2001). In addition to taxation, issuing bonds is one of the ways of financing government expenditure. Since these bonds are considered loans, it will be paid back in increasing tax revenues. This is the choice between periods of "tax now or later.". If the government chooses the tax later principle to finance its deficit, according to REH, the taxpayers expect higher taxes in the future. Hence, fearing future taxes, economic agents will increase their savings by reducing their current consumption level. Similarly, if the government had chosen to tax now principle, the effect on aggregate demand would be the same.

Also, there has been contradictory empirical evidence on the existence of the REH. For example, studies that support the REH are (Tanner 1979, Kormendi 1983, Evans 1988, Leiderman–Razin 1988, Kormendi–Meguire 1990, Evans 1993, Issler–Lima 2000, Giorgioni–Holden 2003, Olasunkanmi–Akanni 2013, Mosikari–Eita 2017). In contrast, some studies (Yawitz–Meyer 1976, Buiter–Tobin 1979, Modigliani–Sterlling 1986, Bernheim 1987, Kazmi 1994, Graham–Himarios 1996, Drakos 2001, Marinheiro 2001, Onafowora–Owoye 2006, Vamvoukas–Gargalas 2008, Fang et al. 2010, Waqas–Awan 2011, Saeed–Khan 2012, Onyeiwu 2012, Odianye–Ebi 2013, Aderemi 2014) do not confirm the REH. Further, some studies found inconclusive results (Gupta 1992, Kaadu–Uuskula 2004). Further, the issues surrounding the effect of fiscal policies (the variables in the REH) are on today's global agenda, but there is a lack of country-relevant empirical studies in the case of Ethiopia (even Pickson–Ofori-Abebrese (2018) did not included Ethiopia to test the REH for SSA countries). This has resulted in a knowledge gap in the literature, thus necessitating the need for a systematic examination.

Hence, based on the above contradicting theories, inconclusive empirical findings, and lack of country-relevant empirical studies in the case of Ethiopia, we tested the REH to fill the literature gap by employing the ARDL cointegration approach. The main objective of our study is to provide an empirical test of the REH in the case of Ethiopia using time series data extending from 1990 to 2011.

2. Literature Review

This section presents the literature which is relevant to the topic. Specifically, it has theoretical literature about the theoretical precondition for the existence of REH and the theoretical framework of this study. Besides that, it has empirical literature, some of which supports the existence of REH and some that does not or provides inconclusive (mixed) results.

2.1.Theoretical Requirements for the Existence of the REH

The infinite time horizon of individuals is one of the requirements for the existence of the REH. That means, the time horizon of individuals should be at least the same as the lifetime of the government. This is because if the individual's lifespan is limited and shorter than that of the government, borrowing will increase the net wealth of the individual if that person dies. Here individuals are linked with future generation by altruistic gifts. Since individuals care about their children's well-being, they do not decide for a tax cut by increasing their consumption. Instead, they will buy securities and other fixed assets and transfer these assets to their families (Marinheiro 2001).

The existence of a perfect capital market (liquidity unconstraint) is an essential element to maintain REH. According to Hayashi (1987), if consumers face quantity constraint (due to the high-interest rate) on their borrowing, they face liquidity constraint. Therefore, they are not able to smooth out their consumption over an entire lifetime, and they will lack an opportunity to select the tax burden, and they will become indifferent to the issue.

The other prerequisite for the existence of REH is the presence of lump-sum taxes. Lump-sum taxation requires that a tax now be precisely equivalent to a tax next year, and by assumption raises the same present value of revenue. Debt and taxes must be equivalent. Moreover, failure to allow fully for the future by virtue either of finite horizons or fiscal targets are inconsistent with the lump-sum assumption. Any lump-sum tax must be intertemporally neutral, both in the sense that it does not distort between the present and future consumption when used in all periods at a constant rate, and in the sense that a tax differential between periods does not induce any taxpayer response (Brennan and Buchanan, 1980). However, in reality, taxes are not lump-sum. The reality is that tax liability is substantial if future income is high, and it is low if the income is small. Hence, with household lifetime resources becoming uncertain, this may lead to an increase in current consumption (Romer 1996, Marinheiro 2001). According to Romer (1996), if individuals do not optimise their consumption fully over the long term, the Ricardian equivalence will not hold. Further, the perfect foresight assumption is one of a basic assumptions for the occurrence of REH, even though it is difficult to achieve in an uncertain world (De Grauwe 1996, Marinheiro 2001).

2.2. The Theoretical Framework

The two main methods of testing the REH are the consumption function and the interest rate approaches. The consumption function approach tries to assess whether increases in government debtare considered net wealth by individuals and create increases in private consumption. However, the interest rate approach evaluates whether deficits lead to an increase in interest rates (Marinheiro 2001, Aderemi 2014). However, discriminating between the Ricardian equivalence and the perfect capital markets hypothesis for an open economy is the main problem of the interest rate approach. In an open economy, even when the consumers are not Ricardian, the interest rate may remain unchanged even though their consumption increases in response to a budget deficit, and when it is assumed the interest rate across countries are equalised by international capital flows. Given the domestic interest rate, the budget deficit may be financed by an inflow of capital. However, in this case, it would create a deficit in the current account, which leads to the so-called twin-deficit phenomena when the Ricardian equivalence does not hold (Marinheiro 2001).

Further, the consumption function approach can be examined by the reduced-form (the structural consumption functions) and Euler equation approach. The approaches of Kormendi (1983), Modigliani and Sterlling (1986), Bernheim (1987), Perelman and Pestieau (1993), Cardia (1997), and Leachman (1996) are the popular approach from the reduced-form (the structural) consumption functions categories. Among these approaches, our study follows that of Bernheim (1987), which is the reduced (the structural consumption), approach to testing the existence of REH in the case of Ethiopia. The reason for using the structural consumption approach relative to Euler is: the Euler equation approach needs several restrictions in order to obtain an observable consumption function, such as the imposition of a constant real rate of return, the specification of a specific form of utility function, like the quadratic utility function in order to aggregate the Euler equation across individuals (Adji 2007). Besides that, the Euler approach requires incorporating rational expectations optimising framework (Aschauer 1985, Gupta 1992). However, the structural consumption function is less restrictive compared to the Euler approach (Kormendi 1983, Bernheim 1987). Further, we chose the Bernheim (1987) among other types of reduced-form (the structural) consumption functions due to different reasons (see data sources, model specification, and methodology of the study section). Hence, his standard model of private consumption is:

$$C_t = B_0 + \beta_1 Y_1 + \beta_2 DEF_t + \beta_3 G_t + \beta_4 D_t + \beta_5 W_t + \beta_t X + u_t \quad (1)$$

Where *C* is private consumption, *Y* is GDP, *DEF* is a budget deficit, *G* is government expenditure, *D* is government debt, *W* is wealth and *X* represents a vector of variables capturing the socio-economic conditions of the countries.

2.3. Empirical Literature

In this section, the empirical studies are presented on the REH. Here we clustered the studies based on their findings – those who support the REH, those who do not support the REH, and those who found mixed (inconclusive) results in the course of their research.

Table 1 Empirical Literature

Author and Year	**Model type**	**The scope of the study**	**Results**
Tanner (1979)	Yawitz-Meyer and the Life Cycle Model.	From 1947–1974, USA	Supports the REH
Buiter–Tobin (1979)	Kochin regression approach	From 1949–1976, USA	REH does not hold
Kormendi (1983)	OLS	From 1930–1976, USA	Supports the REH
Modigliani–Sterling (1986)	OLS	From 1952–1976, USA	REH does not hold
Evans (1988)	GMM	Quarterly data from 1947: II–1985: IV, USA	Supports the REH
Leiderman–Razin (1988)	non-linear least squares (from the TSP program)	Monthly data from1980:9–1985:12, Israel	Supports the REH
Kormendi–Meguire (1990)	Engle and Granger approach	From 1931–1985, USA	Supports the REH
Gupta (1992)	Aschauer (1985) model	From 1963–1986, developing countries	Inconclusive result. The study supports the REH only for the case of South Korea, Singapore, Pakistan, and Thailand.
Evans (1993)	Hansen (1982) approach	From 1960–1988, for 19 OECD countries	Supports the REH
Kazmi (1994)	OLS	From 1960–1988, Pakistan	REH does not hold
Himarios (1995)	based on the Euler condition	From 1953–1986,	Inconclusive
Issler–Lima (2000)	Johansen cointegration	From 1947–1992, Brazil	Supports the REH
Drakos (2001)	VECM	Quarterly data from Q1, 1981 to Q3, 1996, Greece	REH does not hold
Marinheiro (2001)	Both the Structural and Euler consumption functions approaches are adopted. Besides that, he used Kormendi (1983) consumption function, along with the Error Correction Method.	From 1954 to 1997, Portugal	REH does not hold
Giorgioni–Holden (2003)	OLS, Fixed Effect and Random Effect	From 1976–1998, for Ten developing economies	Supports the REH
Kaadu–Uuskula (2004)	Instrumental variable technique and full information maximum likelihood method	Quarterly data from 1997Q1–2002Q4, Estonia	Inconclusive
Onafowora–Owoye (2006)	Granger causality test and Vector Error Correction Method (VECM)	From 1970 to 2001, Nigeria	REH does not hold
Vamvoukas–Gargalas (2008)	Cointegration analysis, Granger causality tests and impulse response	From 1948 to 2001, Greece	REH does not hold
Fang *et al.* (2010)	Structural Vector Autoregressive (SVAR) estimation technique	Monthly data from January 1992 - June 2009, China	REH does not hold
Waqas–Awan (2011)	Johansen Cointegration	From 1973–2009, Pakistan	REH does not hold
Saeed–Khan (2012)	Johansen cointegration.	From 1972–2008, Pakistan	REH does not hold
Onyeiwu (2012)	Ordinary Least Squares (OLS) and Error Correction Method (ECM).	Quarterly time-series data from 1994–2008, Nigeria	REH does not hold
Odianye–Ebi (2013)	VECM	Quarterly time series data from Q1 1970– Q4 2010, Nigeria	REH does not hold
Olasunkanmi–Akanni (2013)	Johansen Cointegration and the Error Correction Mechanism	From 1981–2011, Nigeria	Supports the REH
Aderemi (2014)	Ordinary Least Squares (OLS)	From 1981 to 2012, Nigeria	REH does not hold
Mosikari–Eita (2017)	ARDL	Two sample periods, 1980–2014 and 1988–2014, Lesotho	Supports the REH
Pickson–Ofori-Abebrese (2018)	ARDL	From 1981–2014, for sub-Saharan countries (Botswana, Ghana, Gambia, Nigeria, and Kenya)	REH does not hold

Source: Authors' construction

From the above empirical literature, some of the studies support the REH and the others do not. Besides, there are few studies whose results are inconclusive. This is because of differences in the variables included in the model, the methodology, the time scope, and the case studies. Generally, relative to developed countries, REH does not exist in the case of developing countries. This implies in most developed nations that the main requirements of REH are mostly fulfilled relative to developing nations

3. Data Sources, Model Specification, and Methodology of the Study

In this section, the data type, sources, and data analysis of the study are presented. Furthermore, using Bernheim (1987) approach as a theoretical framework, it offers a way to specify the model. Finally, the ARDL estimation technique along with the estimation procedures are also presented.

3.1. Data Type, Source, Data Analysis, and Model Specification

This study used secondary time series data from 1990 to 2011. The sources of data were WDI, IMF, and countryeconomy.com (see Appendix 1). Further, in the study, we used only econometrics to test the REH for the case of Ethiopia.

This study used the reduced-form (the structural) consumption functions. In addition, it followed the Bernheim (1987) approach to test the existence of REH in the case of Ethiopia. Hence, his standard model of private consumption is:

$$CON_t = B_0 + \beta_1 GDP_t + \beta_2 DEF_t + \beta_3 GOVCE_t + \beta_4 GOVD_t + \beta_5 W_t + \beta_t X + \varepsilon_t \qquad (2)$$

where CON is private consumption, GDP is the gross domestic product, DEF is a budget deficit, GOVCE is government consumption, GOVD is government debt, W is wealth, and X represents a vector of variables capturing the socio-economic conditions of the countries. However, estimating equation (2) for the case of Ethiopia will have various problems such as the unavailability of data on wealth. To solve the above problems, Bernheim (1987) suggested dropping the variable wealth because of its unavailability. Furthermore, Bernheim (1987) used growth in GDP and growth of population as the socio-economic factors. However, for our case, adding these variables leads to the regression result “singular matrix”, so as a result, we dropped them. Finally, we modified the original model of Bernheim as follows:

$$CON_t = \beta_0 + \beta_1 GDP_t + \beta_2 DEF_t + \beta_3 GOVCE_t + \beta_4 GOVD_t + \varepsilon_t \qquad (3)$$

We used the natural logarithm to measure the elasticity for all variables of the model. More specifically, the final model we used is:

$$lnCON_t = \beta_0 + \beta_1 lnGDP_t + \beta_2 lnDEF_t + \beta_3 lnGOVCE_t + \beta_4 lnGOVD_t + \varepsilon_t \qquad (4)$$

Where, β_0 is an intercept term, and β_1, β_2, β_3, and β_4 are the long run coefficients that will be estimated. Further, the REH holds when $\beta_2 = \beta_3 = \beta_4 = 0$. If government consumption substitutes private consumption, then $\beta_3 < 0$, while if it complements it, then $\beta_3 > 0$.

3.2. ARDL Model Specification

To empirically analyse the long run relationships and dynamic interactions among the variables of interest, the model has been estimated by using the bounds testing (or Autoregressive Distributed Lag (ARDL)) cointegration procedure, which was initially presented by Pesaran and Shin (1999) and further extended by Pesaran and Smith (2001). The procedure is adopted because of its several advantages over the conventional type of cointegration techniques. Firstly, relative to other multivariate cointegration techniques such as Johansen and Juselius, the bounds test procedure is simple since it allows the cointegration relationship to be estimated by Ordinary Least Squares once the lag order of the model is known. Secondly, the bounds testing procedure does not require the order of integrations of all the variables included in the model to be the same like other techniques such as the Johansen approach. It is appropriate irrespective of whether the regressors in the model are purely I(0), purely I(1) or mutually integrated (Fosu–Magnus 2006, Bakry–Almohamad 2018). The results of Augmented Dicky Fuller unit root test in Table 1 indicate that the time series variables under examination are integrated of different orders (mixed). Hence, the choice of the ARDL cointegration approach enables us to test the long-run relationships among these variables. Thirdly, the ARDL approach is considered to be a statistically significant approach and more valid than other cointegration techniques for small sample size (Fosu–Magnus 2006, Bakry–Almohamad 2018). This study uses yearly time series data from 1990 to 2011: which is considered a small sample (Narayan et al. 2004). The procedure will, however, fail in the presence of I(2) series (Fosu–Magnus 2006). Fourth, unalike Johansen and Juselius residual-based cointegration tests, this method is efficient and cannot lead to contradictory results, especially when there are more than two I (1) variables under consideration. When we see our variable, four out of five variables are I(1), so ARDL approach eliminates contradictory findings. Fifth, this method (ARDL) includes information on the structural break in time series data and does not suffer from low predicting power. In Ethiopia, there were three structural breaks from 1990 to 2011 (during 1992, 1993, and 2003). Therefore, the choice of the ARDL cointegration approach enables us to consider structural breaks in our study. The sixth advantage of this approach is that the model takes a sufficient number of lags to capture the data generating process in a general to specific modelling framework (Muhammad 2009). Seventh, it estimates the short and long-run components of the model simultaneously, removing the problems associated with omitted variables and autocorrelation. Eighth, this technique generally provides unbiased estimates of the long-run model and valid t-statistic even when some of the regressors are endogenous (Srinivasan et al. 2011). Having the above advantages and following the Bernheim (1987) approach as a framework, the ARDL general model we used in this study is:

$$\Delta lnCON_t = \alpha_0 + \sum_{i=1}^{p} \beta \Delta lnCON_{t-i} + \sum_{i=0}^{p} \delta \Delta lnGDP_{t-i} + \sum_{i=0}^{p} \gamma \Delta lnDEF_{t-i} + \sum_{i=0}^{p} \varphi \Delta lnGOVCE_{t-i} + \sum_{i=0}^{p} \sigma \Delta lnGOVD_{t-i} + b_0 lnCON_{t-1} + b_1 lnGDP_{t-1} + b_2 lnDEF_{t-1} + b_3 lnGOVCE_{t-1} + b_4 lnGOVD_{t-1} + v_t \quad (5)$$

Where, $b_1, b_2, b_3,$ and b_4 are long-run multipliers, α_0 is drift (constant term), and the coefficients of lagged values of difference of the variables show the short-run dynamic structure. Further, Δ is the first difference operator, and p is the optimal lag length.

3.3. ARDL Cointegration Procedures

Testing for the stationarity status of all variables to determine their order of integration is the initial step in ARDL, since unit root tests could be undertaken following the general formula of Augmented Dickey-Fuller (ADF) test. Accordingly, to verify the stationarity of variables, ADF is undertaken:

$$\Delta y_t = \delta + \beta t + \alpha y_{t-1} + \sum_{t+2}^{n} \gamma_i \Delta y_{t-i} + \varepsilon_t \quad (6)$$

The hypothesis to be tested $H_0: \alpha = 0$
$H_1: \alpha < 0$

Reject H_0 if $t_{\alpha=0}$ is less than critical values

Where, y_t represents variables subject to ADF test of stationary condition., in which all variables of the model are tested following the above formula of the unit root test. Here all variables should be either integrated order zero or one, or mixed. To avoid spurious results, it is necessary to confirm that none of the variables is integrated of order 2 or beyond (Fosu–Magnus 2006). Following the unit root test, the second step of ARDL approach is the selection of the maximum lag length for general and optimal lag length for the long run and short run equations using different information criteria before we estimate the model. The most common information criteria for the selection of lag length are Akaike Information Criteria (AIC) and Schwarz Bayesian Criteria (SBC). However, Pesaran and Shin (1999) and Narayan (2004) suggested choosing two as the maximum order of lags if the observations are annual. Once the maximum lag length is determined, the third step is an estimation of the general equation (5) and then testing the existence of a long-run relationship among the variables by conducting F-test for the joint significance of the coefficients of the lagged levels of the variables. That means the null hypothesis (H_0) for no cointegration among variables in equation 7 against the alternative hypothesis (H_1) is

$$H_0 : b_0 = b_1 = b_2 = b_3 = b_4 = 0$$
$$H_1 : b_0 \neq b_1 \neq b_2 \neq b_3 \neq b_4 \neq 0 \quad (7)$$

The F test has a non-standard distribution which depends on (i) whether variables included in the model are I (0) or I (1), (ii) the number of regressors, and (iii) whether the model contains an intercept and/or a trend. The test involves asymptotic critical value bounds, depending on whether the variables are I (0), I (1) or a mixture of both. Two sets of critical values are generated which one set refers to I (1) series and the other for I (0) series. Critical values for I (1) series are referred to as upper bound critical value whilst the critical values for I (0) series are referred to as the lower bound critical values. If the F-test statistic exceeds their respective upper critical values, we can conclude that there is evidence of the long run relationship between the variables regardless of the order of integration of the variables. If the test statistics are below the upper critical value, we cannot reject the null hypothesis of no cointegration (Duasa, 2007). The fourth step is an estimation of the long-run and short-run relationship simultaneously. Once cointegration is established, the conditional ARDL long-run model is:

$$lnCON_t = \beta_0 + \sum_{i=1}^{p} \beta lnCON_{t-i} + \sum_{i=0}^{q} \beta_1 lnGDP_{t-i} + \sum_{i=0}^{r} \beta_2 lnDEF_{t-i} + \sum_{i=0}^{s} \beta_3 lnGOVCE_{t-i} + \sum_{i=0}^{t} \beta_4 lnGOVD_{t-i} + \varepsilon_t \quad (8)$$

This involves selecting the orders of the ARDL *(p, q, r, s, t*) for the model using AIC or SBC. The ARDL specification of the short-run dynamics derived by constructing an Error Correction Model (ECM) in the following form:

$$\Delta lnCON_t = \mu_{0+} \sum_{i=1}^{p} \mu \Delta lnCON_{t-i} + \sum_{i=0}^{q} \mu_1 \Delta lnGDP_{t-i} + \sum_{i=0}^{r} \mu_2 \Delta lnDEF_{t-i} + \sum_{i=0}^{s} \mu_3 \Delta lnGOVCE_{t-i} + \sum_{i=0}^{t} \mu_4 \Delta lnGOVD_{t-i} + \lambda ecm_{t-1} + \varepsilon_t \quad (9)$$

Where all coefficients of the short run equation are coefficients relating to the short-run dynamics of the model convergence to equilibrium, λ is the speed of adjustment parameter, and ecmt-1 is the one period lagged error correction term. Finally, it is necessary to run diagnostic tests such as serial correlation using Breusch-Godfrey serial correlation LM test, heteroskedasticity test using Breusch-Pagan-Godfrey test, normality using Jarque-Bera test, and stability tests using CUSUM and CUSUM of squares.

4. Econometric Estimation Results and Discussion

This section contains the empirical results and their interpretations along with the theoretical and empirical support. More specifically, the unit root test using Augmented Dickey-Fuller with intercept and trend, cointegration test, long-run and short-run dynamics, and diagnostic (normality, heteroscedasticity, autocorrelation, and stability) tests of the model are presented.

4.1. Unit Root & Cointegration Tests

The result of the unit root test shows that all variables, except government consumption expenditure (it is I(0)), included in the model are I (1) at one percent level of significance (see Table 2). Hence, having this mixed order of integration, we can proceed with the ARDL cointegration technique.

Table 2 Unit root test

Variables	ADF test statistics (with intercept and trend)		Order of integration
	Level	First difference	
LNCON	–0.966575	–5.370663***	I(1)
LNGDP	–2.261874	–5.019909***	I(1)
LNDEF	–3.134442	–5.595000***	I(1)
LNGOVCE	–4.901242***	–2.648994	I(0)
LNGOVD	–0.972786	–4.653030***	I(1)

Note *** Significant at 1% level, All the values in the table are t-statistics,

Source: Authors construction from using EViews 9 result, 2019.

Comparing the calculated F statistics with the upper bound critical values at one percent critical level of significance is the way to check the existence of cointegration among the variables. Our result implies that the null hypothesis of no cointegration rejected at one percent level. As a result, in the model, there is cointegration relationship between the variables (see Table 3).

Table 3 Cointegration test

Test statistics	Value	No. of independent variables	Significance level	Bound critical values	
	REH model				
F-statistics	8.06	4		I(0)	I(1)
			10%	2.45	3.52
			5%	2.86	4.01
			2.5%	3.25	4.49
			1%	3.75	5.06

Source: Authors construction from using EViews 9 result, 2019

4.2. The Long-run & Short-run Estimations

In the model, the explanatory variables included together explain around 99 percent of the systematic variation in consumption during the period being studied. The F-statistics are highly significant at the one percent level. Since the range of Durbin-Watson is between 0 to 4 and near to 2, the D-W result (1.71) of our model shows the absence of serial correlation of the residuals in the system. However, the long run equilibrium coefficients and their asymptotic standard error, t-values, and p-values are presented in Table 4.

Table 4 Estimated Long run Coefficients ARDL (1, 0, 2, 0, 0) selected based on AIC
The dependent variable is LNCON

Variables	Coefficients	Std. Error	T-statistics [Prob]
LNGDP	1.346	0.106	12.68 [0.000]***
LNDEF	–0.025	0.054	–0.475[0.643]
LNGOVCE	0.016	0.085	0.194 [0.848]
LNGOVD	–0.102	0.055	–1.829[0.092]*
Constant	–4.480	1.486	–3.013[0.010]**
NB: The following values are from the estimation of the general model			
R-squared	0.994	F-statistic	317.08
Durbin-Watson stat	1.714	Prob(F-statistic)	0.0000

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level
Source: Authors construction from using EViews 9 result, 2019

As we discussed earlier, the REH holds when $\beta_2 = \beta_3 = \beta_4 = 0$. Hence, the result shows that only government debt affects private consumption negatively and significantly. The coefficient of government debt is –0.1, which indicates that, while other things were constant, a one percent increment was responsible for a 0.1 percent reduction in private consumption during the period under study. Since $\beta_4 \neq 0$, our result does not support the REH in Ethiopia. In our result the government debt is a negative and significant effect on private consumption, and it is in line with the Keynesian crowding-out effect. Generally, these findings are largely in line with conventional Keynesian economics; hence we can conclude that Ethiopia is a non-Ricardian economy.

The result of the Error Correction Model (ECM) is presented in Table 5. In the short-run, private consumption expenditure is positively and significantly affected by the level difference of GDP. However, it is negatively and significantly affected by the level difference of government debt and lag difference of deficit. In the short run, the rise in the level difference of GDP by one percent results in a rise in the private consumption level of Ethiopia by 1.53%. However, a one percent increment in the level difference of government debt and lag difference of deficit reduces the private consumption by 0.11 and 0.05%, respectively. In both the long-run and short-run $\beta_2 = \beta_3 = \beta_4 = 0$ did not hold. Hence, Ethiopia is a non-Ricardian economy.

The error correction term indicates the speed of adjustment to restore equilibrium in the dynamic model. The ECM coefficient shows how slowly variables converge to equilibrium and theoretically, it should have a statistically significant coefficient with a negative sign. This condition occurs in our model. Besides that, the highly significant error correction term confirms the existence of a stable long-run relationship between variables. The coefficient of ECM (–1) = –1.139, implying that about 113% of the deviation of the actual private consumption from its equilibrium value is eliminated every year; hence, in this study, full adjustment to reach equilibrium would require less than a year.

Table 5 Error Correction Representation ARDL (1, 0, 2, 0, 0) selected based on AIC
The dependent variable is D(LNCON)

Variables	Coefficients	Std. Error	T-statistics [Prob]
D(LNGDP)	1.535	0.239	6.413 [0.0000]***
D(LNDEF)	–0.055	0.039	–1.407 [0.184]
D(LNDEF(–1))	–0.052	0.027	–1.876 [0.085]*
D(GOVCE)	0.018	0.097	0.194 [0.849]
D(GOVD)	–0.116	0.064	–1.814 [0.094]*
CointEq(–1)	–1.139	0.164	–6.909 [0.0000]***

*** Significant at 1% level * Significant at 10% level

Source: Authors construction from using EViews 9 result, 2019

Finally, the diagnostic tests of the model such as normality test of Jarque-Bera, serial-correlation of Breusch-Godfrey LM, heteroskedasticity test of Breusch-Pagan-Godfrey, stability test of recursive residual (CUSUM) and CUSUM of square (CUSUMSQ) tests have been conducted. Hence, the estimated residuals did not provide any significant evidence of non-normality, serial-correlation, non-stability, or heteroskedasticity effect in the error term (see Appendix 2).

5. Conclusion and Future Studies

The Keynesian proposition and the REH are the two main competing views when dealing with the REH. According to the REH, today's borrowing to stimulate the economy or tax reduction – a substitution of debt for taxes -does not affect demand or consumption level. However, Keynesians argue that an increase in government spending by running budget deficit and substitution of debt for taxes can stimulate the economy (aggregate demand), thus raise total private and public consumption expenditure.

The primary objective of this study is to test the existence of REH empirically in the case of Ethiopia using annual time series data from 1990 to 2011 by employing the ARDL estimation approach. To do so, we conducted the unit root test using ADF and hence all variables except natural logarithm of government consumption expenditure are I(1). Further, we conducted the cointegration test to confirm whether there is longrun relationship among the variables in the model. Following the unit root and cointegration tests, we estimated both the long-run and short-run equilibirium relationships and the results provide substantial evidence against the prevalence of

REH in Ethiopia and support for Keynesian debt non-neutrality. This is because the REH holds when all budget deficit, government consumption expenditure, and government debt does not affect the private consumption level. Even though our result passed the first two requirements (deficit and government consumption expenditure), it did not fulfil the third requirement (government debt). Theoretically, the REH will be valid if there is the same discount rate for both public and private sectors, perfect capital market, no liquidity constraint, consumers are rational, certainty in the future incomes and taxes, and non-distortionary tax. However, all of the above assumptions are not found in Ethiopia. Therefore, it is an expected result for a developing country like Ethiopia.

Finally, this study has its own limitations even though it tried to fill the literature gap. We dropped some variables due to unavailability of data (wealth) and singularity of the regressed variable because of the short time series data relative to the variables included in the model. Hence, in future researches could extend similar investigations by taking these factors into account.

References

Aderemi, T. (2014): An Empirical Application of Debt Neutrality to the Nigerian Economy: Further Evidence. *International Journal of Economic Practices and Theories*, 4, 4, 493–503.

Adji, A. (2007): Essays on Ricardian Equivalence: Economic dissertations. Department of Economics Georgia State University.

Aschauer, D. (1985): Fiscal policy and aggregate demand. American Economic Review, 75, 117–127.

Bakry, W. – Almohamad, S. (2018): Further evidence on Middle East and North Africa financial markets integration*, International Journal of Economics and Business Research*, 15, 1, 1–22.

Bernheim, B. D. (1987): Ricardian Equivalence: an evaluation of theory and evidence with three 'Comments', in Stanley Fisher (ed.). NBER Macroeconomics Annual, 263–315.

Brennan, G. – Buchanan, J. M. (1980): The Logic of the Ricardian Equivalence Theorem, FinanzArchiv / Public Finance Analysis, New Series, 4–16.

Buiter, W. – Tobin, J. (1979): Debt Neutrality: A Brief Review of Doctrine and Evidence, in von Furstenberg, G. (ed.). *Social Security versus Private Saving*, Cambridge, Mass., Ballinger, 39–63.

Cardia, E. (1997): Replicating Ricardian Equivalence Tests with Simulated Series, *The American Economic Review*, 87,1, 65–79.

Daylop, G. T. (2010): Fiscal Deficits and the Growth of Domestic Output in Nigeria. *Jos Journal of Economics*, 4, 1, 153–173.

De Grauwe, P. (1996): International Money, 2nd ed., Oxford, Oxford University Press.

Drakos, K. (2001) Testing the Ricardian Equivalence Theorem: Time Series Evidence from Greece. *Journal of Economic Development*, 26, 1, 149–160.

Duasa, J. (2007): Determinants of Trade Balance: A case of Malaysia, *Journal of Economic Cooperation*, 28, 3, 21–40.

Evans, P. (1988): Are Consumers Ricardian? Evidence for the United States. *Journal of Political Economy,* 96, 5, 983–1004.

Evans, P. (1993): Consumers are not Ricardian: evidence from nineteen countries. *Economic Inquiry,* XXXI, 534–548.

Fang, W. – Ma, J. – Sheng, Y. (2010): An Empirical Test on the validity of China's Fiscal Policy based on the Ricardian Equivalence Proposition. *International Journal of Business and Management,* 5, 8, 209–214.

Fosu, O. – Magnus, F. (2006): An Examination of Foreign Direct Investment, Trade and Growth Relationships: A Case of Ghana. *American Journal of Applied Science*. 3, 11, 2079–2085.

Giorgioni, G. – Holden, K. (2003): Does the Ricardian Equivalence hold in Developing Countries? *International Review of Applied Economics*, 17, 2, 209–221.

Graham, F. – Himarios, D. (1996): Consumption, Wealth, and Finite Horizons: Tests of Ricardian Equivalence. *Economic Inquiry,* XXXIV, 527–544.

Gupta, K. (1992): Ricardian Equivalence and Crowding out in Asia. *Applied Economics*, 24, 19–25.

Hayashi, F. (1987): Tests for Liquidity Constraints: a critical survey and some new observations, in Truman F. Bewley (ed.), Advances in Econometrics Fifth World Congress, vol. 2, Cambridge University Press.

Heijdra, B. J. (2002): The foundation of modern macroeconomics: Oxford university press Inc., New York.

Himarios, D. (1995): Euler Equation Tests of Ricardian Equivalence. *Economics Letters*, 48, 165–171.

Issler, J. V. – Lima, L. R.(2000): Public Debt Sustainability and Endogenous Seigniorage in Brazil. *Journal of development Economics,* 62, 1, 131–147.

Kaadu, H. – Uuskula, L. (2004): Liquidity Constraints and Ricardian Equivalence in Estonia, *Working Papers* of Eesti Bank No. 7.

Kazmi, A. (1994): Private Consumption, Government Spending, Debt Neutrality: Resolving Kormendi-Feldstein-Modigliani Controversy. *The Pakistan Development Review,* 33, 4, 1055–1071.

Kim, S. – Roubini, N. (2008): Twin deficit or twin divergence? Fiscal policy, current account, and real exchange rate in the U.S. *Journal of international economics*, 74, 362–383.

Kormendi, R. (1983): Government Debt, Government Spending, and Private Sector Behavior. *American Economic Review,* LXXIII, 5, 994–1010.

Kormendi, R. – Meguire, P. (1990): Government Debt, Government Spending, and Private Sector Behavior: Reply and Update, *American Economic Review*, 80, 3, 604–617.

Leachman, L. L. (1996): New Evidence on the Ricardian Equivalence Theorem: A Multico-integration Approach. *Applied Economics*, 28, 6, 695–704.

Leiderman, L. – Razin, A. (1988): Testing Ricardian Neutrality with an Intertemporal Stochastic Model. *Journal of Money, Credit and Banking,* 20, 1, 1–21.

Marinheiro, C. F. (2001): Ricardian Equivalence: An Empirical Application to the Portuguese Economy. Faculty of Economics of the University of Coimbra and Katholieke Universiteit Leuven, Discussions Paper Series 01.12, 1–41.

Modigliani, F. – Sterling, A. G. (1986): Government debt, government spending and private sector behaviour: a further comment. *The American Economic Review,* 80, 3, 600–603.

Mosikari, T. J. – Eita, J. H. (2017): Empirical test of the Ricardian Equivalence in the Kingdom of Lesotho, Cogent Economics & Finance, 5, 1–11.

Muhammad, S. (2009): Nominal and Real Devaluations Relations: A Case of Pakistan, *International Journal of applied Econometrics and Quantitative Studies Cooperation*, 9, 1, 85–108.

Narayan, P. – Smith, R. – Nandha, M. (2004): Interdependence and dynamic linkages between the emerging stock markets of South Asia, Accounting and Finance, 44, 3, 419–439.

Odianye, J. C. – Ebi U. K. (2013): The Relationship between Budget Deficit and Interest Rate: Evidence from Nigeria. *European Journal of Business and Social Sciences*, 2, 1, 158–167.

Okpanachi, U.M. – Abimiku, C.A. (2007): Fiscal Deficit and Macroeconomic Performance: Survey of theory and Empirical Evidence, In Ogiji P. (ed.) *The Nigerian Economy: Challenge and Directors for Growth in the Next 25 Years*, Makurdi, Aboki Publisher.

Olasunkanmi, O. I. – Akanni, O.P. (2013): Testing the Validity of Ricardian Equivalence in Nigeria. *International Centre of Science Education and Academic Research for Scholars*, 1, 1, 164–174.

Onafowora, A. O. – Owoye, O. (2006): An Empirical Investigation of Budget Deficit and Trade Deficit: The Case of Nigeria. *Journal of Developing Areas,* 39, 2, 153–174.

Onyeiwu, C. (2012): Domestic Debt and the Growth of Nigerian Economy. *Research Journal of Finance and Accounting*, 3,5, 45–56.

Perelman, S. – Pestieau, P. (1993): The Determinants of the Ricardian Equivalence in the OCDE Countries. in Verbon and Winden (eds.), 181–194.

Pesaran, M. H.. – Shin, Y. (1999): An Autoregressive Distributed Lag Modeling Approach to Cointegration Analysis, S. (Ed), The Ragnar Frish Centennial Symposium, Cambridge University Press.

Pesaran, M. H. – Smith, R. J. (2001): Bound Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics*, 16, 3, 289–326.

Pickson, R. B. – Ofori-Abebrese, G. (2018): Ricardian Equivalence Hypothesis in the Sub-Sahara African Countries. *Journal of Economic Integration*, 33, 3, 466–487.

Romer, D. (1996): Advanced Macro-Economics: McGraw Hill, Library of Congress Cataloging-in-Publication Data.

Saeed, S. – Khan A. M. (2012): Ricardian Equivalence Hypothesis and Budgetary Deficits: The Case of Pakistan. International Islamic University, Islamabad Pakistan.

Srinivasan, P. – Kumar, S. – Ganesh, L (2012), "Tourism and Economic Growth: A case of Sri Lanka," *The Romania Economic Journal*, XV, 45, 211–225.

Tanner, J. (1979): An Empirical Investigation of the Tax Discounting – A Comment. *Journal of Monetary Economics*, 31, 207–227.

Vamvoukas, A. G. – Gargalas, N. V. (2008): Testing Keynesian Proposition and Ricardian Equivalence: More Evidence on the Debate. *Journal of Business and Economics Research*, 6, 67–76.

Waqas, M. – Awan, M.S. (2011): Are Pakistan Consumers Ricardian?" *Economics and Business Review* 13, 3, 161–177.

Yawitz, J. – Meyer, L . (1976): An Empirical Test of the Extent of tax Discounting. *Journal of Money, Credit and Banking*, 8, 247–254.

Appendix

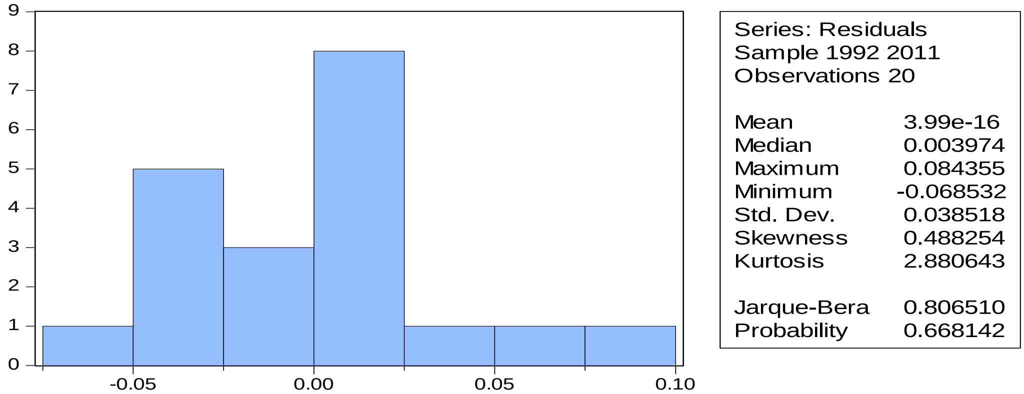
Appendix 1 Definitions, Measurement and Data Sources

Variables	Definition and Measurement	Source
LNCON	Natural logarithm of private consumption expenditure measured as US$	WDI
LNGDP	Natural logarithm of Gross Domestic Product measured as US$	WDI
LNDEF	Natural logarithm of government fiscal deficit measured as US$	Countryeconomy.com
LNGOVCE	Natural logarithm General government final consumption expenditure measured as US$	WDI
LNGOVD	Natural logarithm government debt measured as US$	IMF

Source: Authors construction

Appendix 2 Diagnostic tests

A. Normality test



B. Autocorrelation test

Breusch-Godfrey Serial Correlation LM Test:

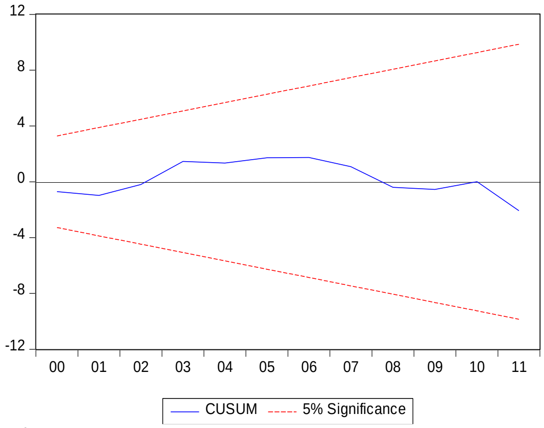
F-statistic	0.522148	Prob. F(2,10)	0.6086
Obs*R-squared	1.891106	Prob. Chi-Square(2)	0.3885

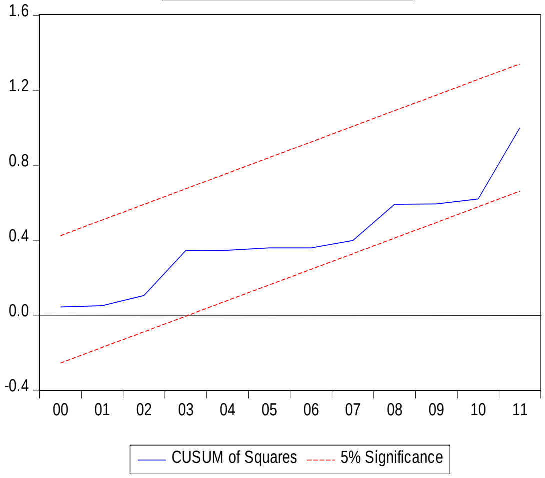
C. Test of Heteroskedastisity

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	1.933448	Prob. F(7,12)	0.1507
Obs*R-squared	10.60082	Prob. Chi-Square(7)	0.1570
Scaled explained SS	3.588544	Prob. Chi-Square(7)	0.8258

D. Stability Test





Udvari B. (ed) 2020: Proceedings of the 4th Central European PhD Workshop on Technological Change and Development. University of Szeged, Doctoral School in Economics, Szeged, pp 415–439.

The magnitude of trade misinvoicing in Ghana and Hungary: Commodity and trading partner level analysis

Isaac Kwesi Ampah

The substantial and the persistent nature of trade misinvoicing in developing countries trade with advanced nations has gained considerable attention in academia and in policy cycles, especially due to its linkages with corruption and tax evasion, and its impact on global trade and domestic resources mobilization. Using the Harmonised System (HS) revision 2 commodity codes of the United Nations Commodity Trade Statistics (UN-COMTRADE) in 2017, this paper examines the magnitude and the nature of trade misinvoicing in Ghana and Hungary with a specific focus on the commodities, as well as the trading partners that are heavily involved in these misinvoicing practices. The evidence at both commodity and trading partner level indicate that Ghana and Hungary lose billions of dollars from trade due to misinvoicing practices in their economies. The results also highlight the need for both governments to increase their access to data, especially at their border sites, and possibly track custom valuations declared at their border stations to that of their trading partners to detect any possible trade corruption and institute sanctions against the companies and individuals involved to deter others from engaging in it.

Keywords: Ghana; Hungary; Export and Import misinvoicing; trade; UN-COMTRADE

1. Introduction

International trade is seen as a vital ingredient in the socio-economic development of any nation, especially in developing and emerging countries. Not only does it enhance their competitiveness by helping them reduce the cost of inputs and increase their value-added, but it also encourages innovation by facilitating the exchange of technology and technical know-how. Trade among countries also promotes export diversification by allowing countries involved to access new markets and new ideas which otherwise would not be available to them. However, for the benefit of trade to be sustainable and more inclusive, it is essential that the countries involved are able to amass their legitimate revenue and gains from it. Unfortunately, for many developing and emerging countries, the institutional framework coupled with imperfect monitoring and weak enforcement has created incentives for trade corruption by agents seeking to maximise their private profits and other gains.

Ideally, Ghana's total exports to Japan in a year should be equal to Japan's total imports from Ghana within the same year, after adjustment is made for the cost of transport, insurance, and duties. However, in practice, one can expect differences to occur as a result of arithmetical or statistical errors. GFI (2015) and Ndikumana et al. (2015) note that if such errors are genuine, one may expect it to be relatively small problem, as the capacity, experience, and training among customs agencies and statistical compilers in various countries has improved. Indeed, UNCTAD (2016) and Ndikumana and Boyce (2010) added that such errors will not persist or increase over a relatively long period of time and may be rotating around a mean of zero or diminishing with time. But unfortunately, for many developing and emerging countries, these bilateral trade discrepancies are quite substantial, indicating the presence of either excessive normal or perverse discrepancies (UNCTAD 2016). The "perverse discrepancies" occur when the importers' value of a consignment of goods is significantly lower than the value reported by the exporters of the same goods, plus the cost of transport, insurance, and duties, signifying either over-invoicing on the part of the exporters or under-invoicing on that of importers, or both (UNCTAD 2016). These "excessive normal discrepancies" also occur when the value reported by the importers for the same good is greater than the value reported by the exporters by an amount that is considerably larger than the reasonable value acknowledged to be the costs of transport, insurance, and freight, indicating either under-invoicing in the case of the exporters or over-invoicing in the case of the importers or both (UNCTAD 2016).

Trade misinvoicing, which refers to either perverse discrepancies or excessive normal discrepancies, has received substantial attention in academia and policy circles recently due to its impacts on global trade and revenue mobilisation, especially among developing nations. The latest estimates by Global Financial Integrity (GFI) (2019) using the UN-COMTRADE indicates that, between 2006 and 2015, trade misinvoicing alone constituted about 1.128 trillion USD annually on average, representing about 22 per cent of their total trade. In terms of the dollar value of over-invoicing of trade, European nations such as Hungary, Romania and Bulgaria, as well as Latin American countries like Mexico, Brazil, Argentina and Peru were all among the top 30 countries found culpable of this erroneous act. Asian nations such as Malaysia and Thailand as well as African nations such as South Africa and Tunisia were all among top countries exhibiting cases of such trade misinvoicing act. Although many reasons can be attributed to this huge amount of trade misinvoicing, an emerging fear, is that some of the differences in trade transactions are as a result of deliberate actions by traders to circumvent capital controls, and avoid taxes and non-tariff measures, among other fraudulent motivations. Studies by GFI (2018) and Baker et al. 2014 note that trade misinvoicing in developing and emerging countries not only weaken their objective of reducing poverty and inequality, and enhancing growth in living standards, it also depresses government revenues and allows wealthy individuals and corporations to hide stolen money, evade taxes, and avoid the adverse impacts of currency depreciation.

In this paper, we disentangle the data that is currently available in the UN-COMTRADE to examine the nature of trade misinvoicing in Ghana and Hungary with a specific focus on the commodities as well as the trading partners that are heavily prone to these misinvoicing practices. By this level of estimation, this paper hopes to contribute to the literature on how trade misinvoicing disturbs both low and high-income countries trade performance and revenue mobilization, and also shed some light on how to move forward by offering possible solutions in dealing with the issues associated with it. The next section presents a brief review of the relationship between bilateral trade discrepancies and trade misinvoicing. Section three is devoted to the methodology used for the computation of trade misinvoicing and describes the method of compiling the data. The results by country, commodities, and trading partner countries are presented in section four, and finally section five sums up the paper with conclusions drawn from the results and some policy recommendations.

2. Literature review

The concept of bilateral trade discrepancies and trade misinvoicing has generated various responses and even serious debate among scholars and policy analysts since the seminal work of Morgenstern (1963) and Bhagwati (1964) in the 1960s. There is one school of thought that believes that large and systematic trade discrepancies are essentially motivated by the intrinsic desire of both exporters or importers to evade tariffs or taxes, or hide payment to an associate by declaring a value that does not reflect the real value of the goods. Meanwhile, there is another school of thought that explicitly admits that trade discrepancies do exist but may not necessarily relate to misinvoicing. This section briefly reviews this literature.

2.1. Trade discrepancies correlate with and are caused by misinvoicing

Studies directing to the positive relationship between trade discrepancies and trade misinvoicing essentially derive their evidence from the notion that trade discrepancies occur as a result of trader’s desire to avoid bureaucratic and lengthy administrative procedures from customs authorities, or desire to maximize profit by dodging tariffs or taking advantage of tax incentives aimed at promoting exports. Furthermore, it is claimed that trade discrepancies are intentionally created by importers and exporters to take advantage of the premiums in the exchange rate system (UNCTAD 2016). Along these lines, Bhagwati (1964) found that products facing high tariffs experienced substantial import under-invoicing relative to products facing low tariffs in Turkey. Epaphra (2015) also found that trade misinvoicing is highly correlated with tax rates with import misinvoicing being greater or higher for commodities facing higher tax rates than commodities facing low tax, authenticating the results of Bhagwati (1964). Again, a recent study by Kellenberg and Levinson (2016) also found evidence that tariff evasion is one of the ways that several lead firms ‘intentionally misreport’ trade data.

Fisman and Wei (2004) also studied the effect of tariff rates on trade misinvoicing practise between China and Hong Kong. The evidence shows that firms that engage in cultural property and antique trade in these countries mis-invoice their trade transactions to take advantage of the differences in tax rates across the products. Likewise, Yeats (1990) found that smuggling is widespread in trade among African countries because importers intentionally under-invoice to avoid high tariffs or quotas. Berger and Nitsch (2012) also found that trade discrepancies are highly positively correlated with corruption. In addition, empirical evidence from Bahmani-Oskooee and Goswami (2003), Barnett (2003), and Biswas and Marjit (2005) also suggest that traders engage in import over-invoicing and export under-invoicing to generate additional foreign exchange currencies to trade in goods and services with premiums on the black market.

2.2. Trade discrepancies are not necessarily the product of misinvoicing

Several authors have also provided evidence that trade discrepancies may not necessarily indicate trade misinvoicing. Östensson 2018 cited accidental errors in the classification of goods, recorded export destinations being different from actual ones as in the case of transit points, price changes while goods are in transit and discrepancies between estimated and actual freight cost as reasons behind trade discrepancies. Empirically, Hangzhou (2009) finds the attribution of imports to the country of origin, attribution of exports to the country of last known destination, and different valuations, as key reasons for the unusually large and growing statistical discrepancies in bilateral trade between China and the United States. Ferrantino and Zhi (2008) also found valuation issues, U.S. tariffs, and re-exporting through the United States itself as the cause of the robust discrepancies. Ajayi (1998) also mentioned diversion of goods en route to the final destination, re-exports, reporting lags, currency conversions, and exchange rate variations as potential reasons beyond misinvoicing. He also added that "in Sub-Saharan Africa, one of the basic causes of trade discrepancy stems from the fact that most imported or exported goods are routed through several countries before the final destination is reached. Martin (2016) confirmed that discrepancies also arise because of different definitions of exports and imports, different definitions of territory, timing, declarations of the country of origin, exchange rates, and intermediation, in addition to under-invoicing.

2.3. Literature adopted for the current study

While it is possible that discrepancies may exist in trade transactions between exporters and importers due to statistical, measurement or other related errors as shown in section 2.2, GFI (2015) noted that such errors should be relatively small as the capacity, experience, and training among customs agencies and statistical compilers in various countries has improved. Moreover, UNCTAD (2016) and Ndikumana and Boyce (2010) added that such errors will also not persist or increase over a relatively long period of time. Östensson (2018) writes that it is very doubtful that developing countries would accidentally omit over US$1 trillion from their economies from trade year after year, and never put in any mechanism to correct it.

Empirically, recent studies such as GFI (2019), UNCTAD (2016), Ndikumana et al. (2015), and Baker et al. (2014) using the standard International Monetary Fund (IMF) Direction of Trade Statistics (DOTS) have also noted that when trade discrepancies are greater than 10 per cent of the export value of the traded goods, then such discrepancies are in the region of trade misinvoicing. Other studies such as Ndikumana and Boyce (2018) examining used the same analytical framework when estimating trade misinvoicing in African sub-Saharan countries. Similarly, Jha and Truong, 2014, Beja, 2007 and Kar, 2010 also used the same framework analysing trade misinvoicing for India and other Asian countries. This paper is therefore premised on the notion that trade discrepancies greater than 10 per cent of the export value of the traded commodity are due to misinvoicing and this assumption is based on the aforementioned empirical studies.

3. Methodology

In achieving the objectives of this study, three empirical exercises are performed. The first is an investigation of the commodities or product in the study countries that are heavily involved in trade misinvoicing practices. The second empirical exercise also involves computation of trade misinvoicing for trading partners and the last part involves the computation of trade misinvoicing by commodity at the trading partner level.

3.1. Trade Misinvoicing by commodities Level Computation

In examining the commodities that are heavily involved in the trade misinvoicing practices in the study countries, two computation approaches are adopted.

- The first approach is the identification of the most traded commodities of each of the study countries. This is done by extracting the exports or imports of all commodities to the world (as the trading partner) using the Harmonised System (HS) revision 2 commodity codes and datasets from the UN-COMTRADE database. With the imports, since "the world" as a reporter does not appear in the UN Comtrade database, this paper estimates the imports as the sum of imports by all individual partners. This computation method has been used by UNCTAD 2016, the GFI (2015) and Baker et al. (2014). The major commodities are commodities with a relatively large share of the total export or imports in the study year.
- Once, the key traded commodities are determined, the second approach involves the computation of the trade misinvoicing. This is accomplished by comparing the exports or import data of the major commodities exported or imported in the world and the world's imports or export and interpreting the difference after the cost of freight and insurance are made as evidence of misinvoicing.
- For clarity, the computation of export misinvoicing by country *A* and product *i* at any time ***t*** is given as:

  $$EMA_{it} = PMA_{it} - [XA_{it} + CIF(XA_{it})] \quad (1)$$

  Where EMA_{it} is the export misinvoicing of country A for commodity *i* at time *t*. PMA_{it} is the value of the trading partners' imports from country A of product *i* at time *t*, XA_{it} is country A's exports to all the trading partners' as reported by country A, and the *CIF* is the factor which represent the costs of freight and insurance.

  Positive values of EMA_{it} are evidence for export under-invoicing whereas negative values of the difference are evidence for export over-invoicing.

- Similarly, the import misinvoicing of major commodities involved in the trade misinvoicing of country *A* and product *i* at time ***t*** is given as:

  $$IMA_{it} = MA_{it} - [PXA_{it} + CIF(PXA_{it})] \quad (2)$$

  Where IMA_{it} is the import misinvoicing of country A of product *i* at time *t*. PXA_{it} is the value of all trading partners export of product *i* to country A at time *t*, MA_{it} is country A's imports from all the trading partner and *CIF* is the factor, representing the costs of freight and insurance.

  In this case, positive values of IMA_{it} are import over-invoicing whereas negative values of the difference are evidence for import under-invoicing.

3.2. Trade Misinvoicing by Partner Level Computation

- Trade misinvoicing at the trading partner level also follows the same process as the product level computation, however, with this computation, the centre for the computation is the trading countries and not specific commodities as in the previous calculation. Once again, the main trading partners based on the relative shares in cumulative exports or imports in the study period are the first to be investigated. Once these leading trading partners have been identified, a similar computation follows to calculate trade misinvoicing by trading partner level.
- For export misinvoicing for country A at time t, the computation is given as

  $$EMA_t = PMA_t - [XA_t + CIF(XA_t)] \quad (3)$$

  Where PMA_t is the value of the trading country's imports from country A as reported by the trading partners, XA_t is country A's exports to the trading countries as reported by country A, and CIF is the factor, representing the costs of freight and insurance.

  Again, positive values are evidence for export under-invoicing whereas negative values of the difference are evidence for export over-invoicing.

- For import misinvoicing, for the country, A at time t is given as

  $$IMA_t = MA_t - [PXA_t + CIF(PXA_t)] \quad (4)$$

  Where MA is country A's imports from its trading partners, and PXA_{it} is the trading partners' exports to country A, and CIF is the factor representing the costs of freight and insurance.

  Similarly, positive values of the difference are evidence for import over-invoicing whereas negative values of the difference are evidence for import under-invoicing.

3.3. Trade Misinvoicing by commodity and trading partner level

- Export misinvoicing for the main commodities at trading partner level is given as

  $$EMA_{jit} = PMA_{jit} - \left(XA_{jit} + CIF(XA_{jit})\right) \qquad (5)$$

  Where EMA_{jit} is the export misinvoicing of country A from its trading partner *j* of product *i* at time *t*. PMA_{jit} is the value of the trading partner *j* imports from country A of product *i* at time *t*, XA_{jit} is country A's exports to its trading partner j as reported by country A, and *CIF* is the factor representing the costs of freight and insurance.

  Positive values of EMA_{jit} are evidence for export under-invoicing whereas negative values of the difference are evidence for export over-invoicing.

- Import misinvoicing of major product at trading partner level is also given as

  $$IMA_{jit} = MA_{jit} - [PXA_{jit} * CIF\left(PXA_{jit}\right)] \qquad (6)$$

  Where IMA_{jit} is the import misinvoicing of country A from its trading partner *j* of product *i* at time *t*. PXA_{jit} is the value of the trading partner *j* export to country A of product *i* at time *t* as reported by *j*, MA_{jit} is country A's imports from trading partner *j* as reported by country A, and *CIF* is the factor, representing the costs of freight and insurance.

  Similarly, positive values of IMA_{jit} are import over-invoicing whereas negative values of the difference are evidence for import under-invoicing.

4. Result and Discussion

Tables 1 and 2 and Figures 1 to 4 discuss the main results of the trade misinvoicing estimates for the major commodities and the main trading partners of Ghana. It also provides trade misinvoicing estimates for specific major commodity against its major trading partner level.

4.1.1. Ghana's Trade Misinvoicing by commodities group level

In Table 1, the major commodities and its associated misinvoicing are reported. The results in Table 1 shows that Ghana's exports are dominated by a few primary commodities such as precious pearls, metal and stones, mineral fuel and oil as well as cocoa and cocoa preparations. These three commodities alone contribute about US$11.94 billion out of the US$14.35 billion total value of Ghana's export. Per this value, these three commodities contribute about 83 per cent of Ghana's total export. In the case of imports, the shares of vehicles other than railway or tramway; Nuclear reactors, boilers, machinery and mechanical appliances, and salt, sulphur and stone contribute about 33 per cent of total imports representing about US$4.25 billion out of the total imports of about US$12.72 billion.

Table 1 Ghana's trade misinvoicing by commodity level

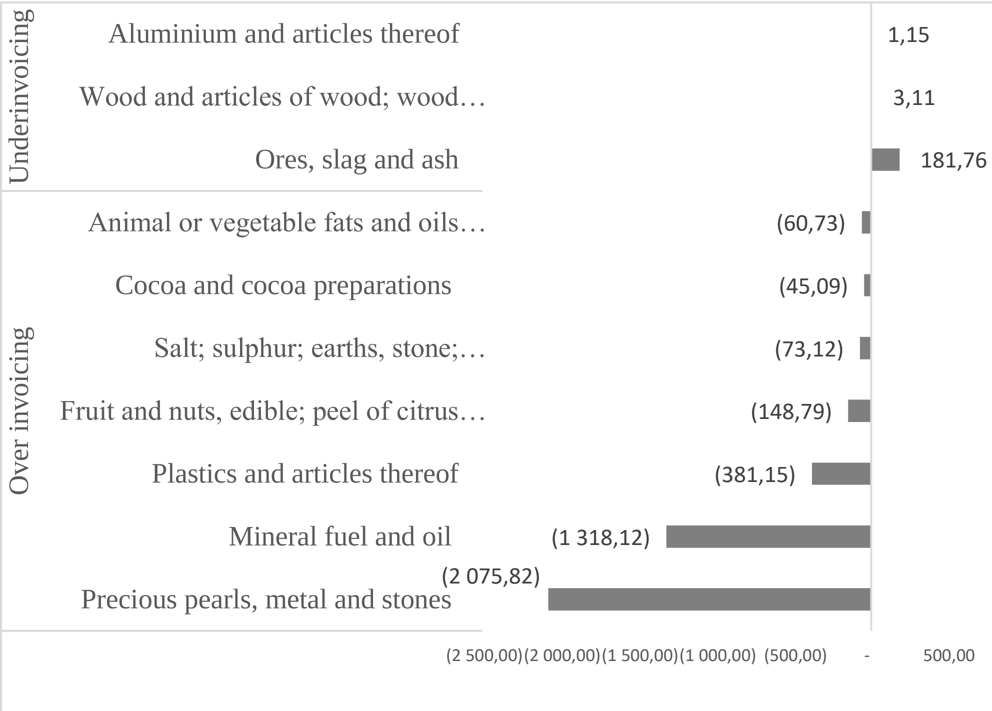
Co. Codes	Commodities	Export/ Import in million US$	Misinvoicing (% of export /import)
	Export		
71	Precious pearls, metal and stones	5,861.49	35.41
27	Mineral fuel and oil	3,639.25	36.22
18	Cocoa and cocoa preparations	2,433.74	1.85
8	Fruit and nuts	409.42	36.34
39	Plastics and articles thereof	370.32	102.93
44	Wood and articles of wood	189.03	1.65
15	Animal or vegetable fats and oils	187.97	32.31
26	Ores, slag and ash	186.09	97.67
25	Salt, Sulphur and stone	92.03	79.45
	TOTAL	*14,358.51*	
	Import		
87	Vehicles; other than railway or tramway	1,872.47	41.40
84	Nuclear reactors, boilers, machinery and mechanical appliances	1,392.84	21.45
25	Salt, Sulphur and stone	980.18	80.45
85	Electrical machinery and equipment	785.49	64.08
10	Cereals	716.15	68.82
48	Paper and paperboard	524.20	63.45
73	Iron or steel articles	517.46	9.78
39	Plastics and articles thereof	468.53	52.17
87	Iron and steel	412.37	4.87
	TOTAL	*12,718.14*	

Source: Author's computation using UN Comtrade data

Comparing Ghana's total export or import values of the identified commodities to her trading partners' data reveals excessive discrepancies for all the major commodities traded by Ghana. In the case of export, the evidence of export over-invoicing is common for all the trading commodities with the exception of ores, slag and ash, wood and articles of wood, and aluminium and articles thereof as shown in Figure 1. Precious pearls, metal and stones and mineral fuel and oil are the largest and most significant contributors to over-invoicing in the country, with the total misinvoicing amounting to US$2,075.82 million and US$ 1,318.12 million respectively. The misinvoicing of these two exported commodities cost Ghana about 24 per cent of her total export. Also, over 100 per cent of Ghana's export value of plastics and other articles of plastics exported by companies in Ghana and recorded in

Ghana's export data actually did not reach the exported partners countries as shown in Table 1. In the same way, about 97.67 per cent of imports of ores, slag and ash by trading partners of Ghana supposedly to be recorded as an export from Ghana were not recorded in Ghana's export data.

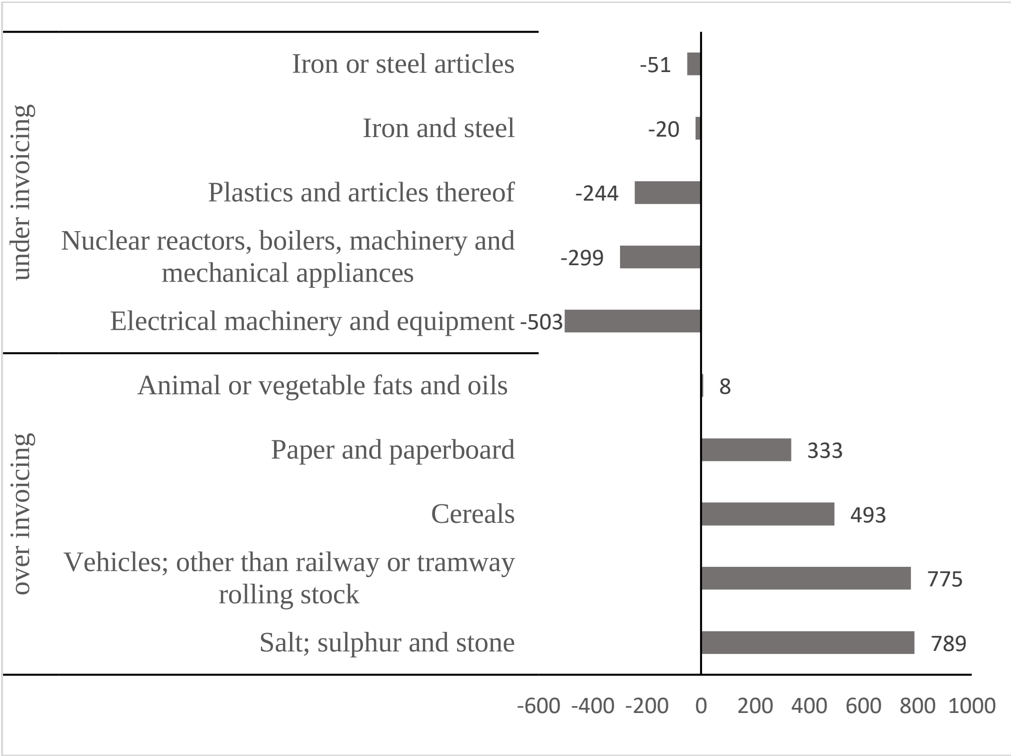
Figure 1 Ghana's export misinvoicing by commodity level



Source: Author's computation using UN Comtrade data

For imported commodities by Ghana, the trend is the same. over-invoicing is recorded for commodities such as vehicles other than railway or tramway, salt, sulphur and stones, cereals and finally, paper and paperboard. For commodities such as nuclear reactors, boilers, machinery and mechanical appliances, electrical machinery and equipment, iron or steel articles and plastics and articles, the study found evidence of under-invoicing. As shown in Figure 2, the evidence of misinvoicing is especially large for imported commodities such as salt, sulphur and stone, vehicles other than railway or tramway as well as electrical machinery and equipment. Table 1 also shows that about 80 and 69 per cent of the value of salt, sulphur and stone and cereals respectively imported by Ghana, were not actually recorded in the exporter's account. Likewise, the result in Table 1 also revealed that 52 per cent of the value of plastics and articles thereof imported by Ghana's trading partners were not recorded in Ghana's import.

Figure 2 Ghana's export misinvoicing by commodity level



Source: Author's computation using UN Comtrade data

4.1.2. Ghana's Trade Misinvoicing by trading partner's level

Table 2 shows the results of trade misinvoicing for the main trading partners in the case of Ghana.

The results show evidence of both under and over misinvoicing for imported and exported trading partners. Ghana's export to China, Canada and India exhibits very high levels of export over-invoicing, with a total of US$767 million, US$264 million, and US$195 million respectively. This practice is also observed in Ghana's exports to the United Kingdom, Germany, Spain, Netherlands, and Switzerland, but the proportion is relatively small compared to that of China, Canada, and India. Evidence of under-invoicing is also seen in Ghana's exports to USA and France, with the USA having the total amount of approximately US$332 million as misinvoicing out of total export of US$407 million. Ghana's trade with Canada needs further investigation since almost ninety-two per cent of Ghana's total export to Canada were not reported in Canada's import data even though they are reported in Ghana's export. Another case is the USA. About 82 per cent of imports recorded by the USA as goods from Ghana were not recorded in Ghana as export to the country.

Table 2 Ghana's trade misinvoicing by trading partner level

Country	**Export/ Import in million US$**	**Misinvoicing**	**Misinvoicing (% Export or import)**
	Export		
India	2,689.42	(195.17)	7.26
China	2,381.36	(766.56)	32.19
Switzerland	1,660.10	(1.30)	0.08
Netherlands	884.61	(37.74)	4.27
USA	*407.97*	*332.40*	*81.48*
United Kingdom	329.47	(123.83)	37.58
Canada	*287.37*	*(264.34)*	*91.99*
France	261.80	72.79	27.81
Germany	233.07	(52.02)	22.32
Spain	213.18	(69.98)	32.83
	Import		
China	*2,134.18*	*(3,173.10)*	*148.68*
USA	1,200.06	254.09	21.17
United Kingdom	*1,099.10*	*561.24*	*51.06*
Spain	*754.55*	*475.41*	*63.01*
Belgium	718.81	287.68	40.02
South Africa	410.86	22.99	5.60
Canada	*363.51*	*156.08*	*42.94*
Germany	341.51	8.46	2.48
Turkey	325.15	78.14	24.03
Rep. of Korea	307.17	38.91	12.67
Malaysia	303.47	22.06	7.27
Italy	289.76	2.51	0.87
Netherlands	*237.00*	*(598.45)*	*78.80*
Japan	200.59	56.30	28.07
France	192.90	(89.41)	34.84

Source: Author's computation using UN Comtrade data

Probing into the case of Ghana's imports, the results in Table 2 also reveal large discrepancies between the values reported in Ghana and those recorded in her trading partners' records. In Table 2, the results show that United Kingdom, Spain, USA, and Canada recorded the largest amount of over-invoicing with a total amount of US$ 561.24 million, US$ 475.41 million, US$ 254.09 million, and US$ 156.08 million respectively. There was also import under-invoicing in trade with other major trading partners. Along with France, the top two trading partners that accounted for the largest share of under-invoicing are China and Netherlands, with China amounting to US$ $3 billion and the Netherlands amounting to US$ 598.45 million. France also accumulated a total amount of US$ 89.41 million under-invoicing.

Likewise, more than 100 per cent of China's export to Ghana (specifically *148.68 per cent)*, reported in China's export was not reported in Ghana's import data. A similar situation is also recorded in the Netherlands, where, about 79 per cent of its export to Ghana, though recorded in Netherlands export data, was never recorded in Ghana's imports. On the other hand, about 63%, 51% and 40% respectively of goods imported by Ghana from Spain, United Kingdom and Belgium were actually not recorded by these countries as export to Ghana.

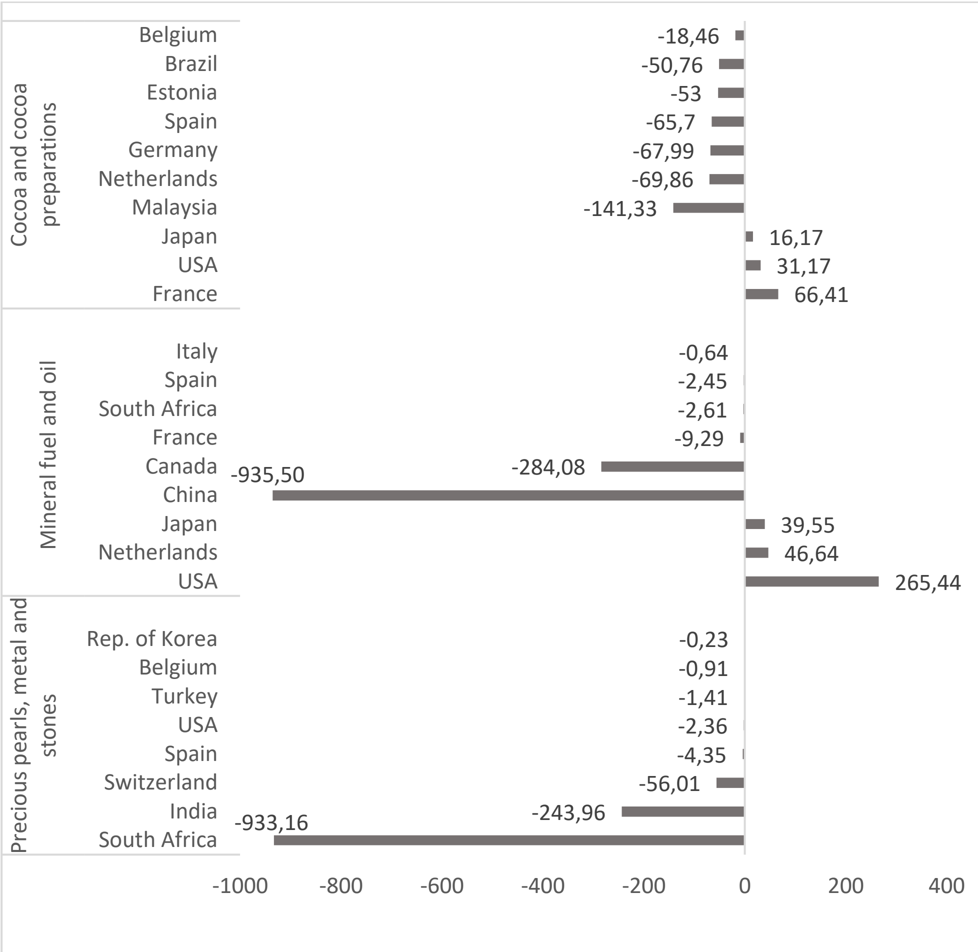
4.1.3. Ghana's Trade Misinvoicing by commodities and trading partner level

In this estimation, we focused on the major traded commodities and the partner countries that are heavily involved in misinvoicing for each of these commodities. The results in Table 1 in the Appendix shows that the export of precious pearls, metal and stones by Ghana exhibit heavy concentration among two trading countries, India and Switzerland. These two countries together accounted for about over 71% of the country's total exports: India with 43.27% and Switzerland with 27.82%.

The results in relation to export misinvoicing reported in Figure 3 show consistent under-invoicing occurring in all trading partners trade as far as the export of precious pearls, metal and stones is concerned. South Africa, India and Switzerland are the largest partner country destinations of export under-invoicing with the total amounting to about US$ 1.2 billion. Again, the results reported in Table 1 in the Appendix indicate that 99%, 95% and 91% of precious pearls, metal and stones imported by Spain, China and South Africa from Ghana respectively were not recorded in Ghana's export data. This needs further investigation. In the case of mineral fuel and oil, the results in Figure 3 shows large-scale under-invoicing in exports to China and Canada, accounting for the lion's share at US$ 935.50 million and US$ 284.08 million respectively. Over-invoicing was also recorded in Ghana's export of mineral fuel and oil to the USA. The amount involved is US$ 284.08 million representing about 90% of Ghana's export to the USA.

With respect to Ghana's export of cocoa and cocoa preparations, the result revealed systematic under-invoicing with Malaysia (US$141.33 million), the Netherlands (US$69.86 million), Germany (US$67.99 million), Spain (US$65.7 million), Estonia (US$53 million), Brazil (US$50.76 million), and Belgium (US$18.46 million) whiles countries like France (US$ 66.41 million), USA(US$31.37 million), and Japan (US$16.17 million) showing cases of trade over-invoicing. In fact, about 50% of Malaysia's imports from Ghana were not reported in Ghana's export data. Similarly, about 50% and 54% of imports recorded by Spain and France respectively as goods from Ghana were not recorded in Ghana's export to these countries. This is reported in Table 1 in the Appendix.

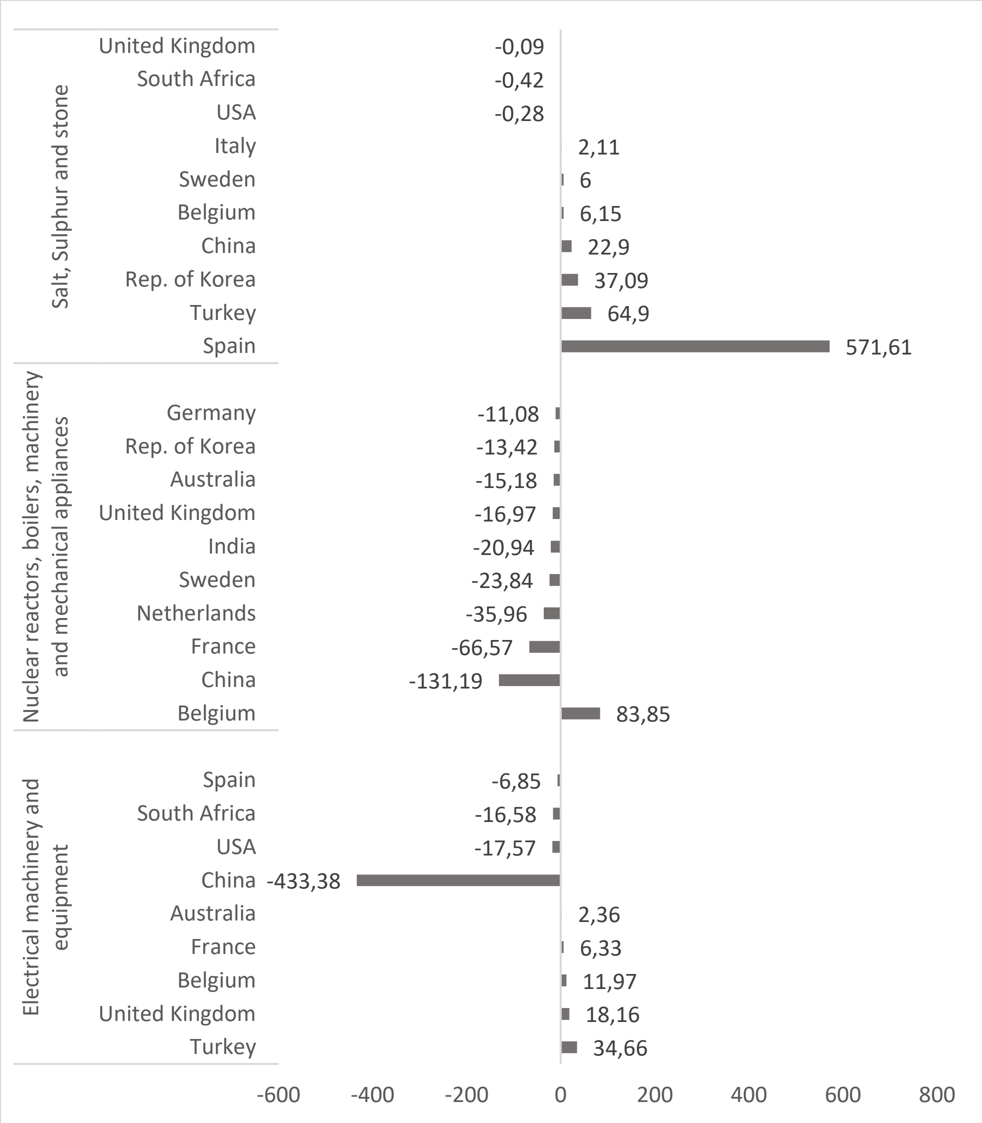
Figure 3 Ghana's Export misinvoicing by commodity and partner trading level



Source: Author's computation using UN Comtrade data

Figure 4 also reports Ghanaian import misinvoicing of the three main imported commodities and their associated bilateral trading partners. In the case of electrical machinery and equipment, import under-invoicing cases were seen in Ghana's imports from China, USA, South Africa, and Spain, with China recording the largest amount of under-invoicing totalling US$ 433.38 million. Also, the import of electrical machinery and equipment from Turkey, the United Kingdom, France, Belgium, and Australia reveal cases of over-invoicing. The results show that Turkey is the major trading partner of Ghana that is prone more import over-invoicing practices as far as the import of electrical machinery and equipment is concerned with a total of about US$ 34.66 million. Also, about half of China's export of electrical machinery and equipment to Ghana was not recorded in import data. This is shown in Table 2 of the Appendix.

Figure 4 Ghana's imports misinvoicing by commodity and partner trading level



Source: Author's computation using UN Comtrade data

Concerning Ghana's import of nuclear reactors, boilers, machinery and mechanical appliances, the results in Figure 4 also reveals large discrepancies between the values reported in Ghana's records and that of her trading partners' data. The analysis of the data in the UN-Comtrade revealed large discrepancies with import under-invoicing being predominant. China, France, and the Netherlands recorded the largest amount of under-invoicing with a total amount of US$ 131.19 million, US$

66.57 million, and US$ 35.96 million respectively. There was also import over-invoicing in an US$ 83 million trade deal with Belgium. Again, about 86.42 per cent of Netherland's export of nuclear reactors, boilers, machinery and mechanical appliances to Ghana, was not reported in Ghana's import data. Also, about 81.19% of export of nuclear reactors, boilers, machinery and mechanical appliances from the Republic of Korea was never recorded in Ghana's official imports. On the other hand, about half of Ghana's imports from Belgium was actually not recorded by these countries as exports to Ghana. This is shown in Table 2 in the Appendix. Unlike the case of imports of nuclear reactors, boilers, machinery and mechanical appliances, which was consistently afflicted with under-invoicing, the imports of salt, sulphur and stone were rather stricken with over-invoicing. Spain, Turkey, Rep. of Korea, China, Belgium, Sweden, and Italy are countries with the largest amount of import misinvoicing as far as salt, sulphur and stone are concerned.

4.2. Trade Misinvoicing in Hungary

Tables 3 and 4, as well as Figures 5 and 6, discuss the main results of the trade misinvoicing estimates for the major commodities and the main trading partners of Hungary. It also provides trade misinvoicing estimates for specific major commodity against its major trading partner level.

4.2.1. Hungary's Trade Misinvoicing by commodities group level

Unlike Ghana's exports, which show a heavy concentration on a few primary commodities such as precious pearls, metal and stones, mineral fuel and oil, as well as cocoa and cocoa preparations, that of Hungary in Table 3 shows an export concentration on a few industrial commodities. The share of electrical machinery and equipment, nuclear reactors, boilers, machinery and mechanical appliances and vehicles other than railway or tramway contribute about 56% of Hungary's total export. These three commodities contribute about US$63.07 billion out of the US$113.40 billion total value of Hungary's export. In the case of imports, these same commodities contribute the largest shares of Hungary's imports. Together, they contribute about 48% of total imports, representing about US$49.81billion out of the total imports of about US$104.28 billion.

Just like the case of Ghana, analysis of trade misinvoicing at the commodity level in Hungary reveals excessive discrepancies for both imported and exported commodities. In the case of exports, the results reveal systematic export over-invoicing for all the major trading commodities. The largest amount of export over-invoicing is recorded in commodities such as electrical machinery and equipment (US$4.24 billion); vehicles other than railway or tramway (US$ 2.31 billion), nuclear reactors, boilers, machinery and mechanical appliances (US$ 1.71 billion), and pharmaceutical products (US$ 1.26 billion). This same practice is also observed in Hungary's export of mineral fuels, oils and products; rubber and its articles; optical, medical or surgical instruments and apparatus; organic chemicals, and furniture and fittings.

Table 3 Hungary's trade misinvoicing by commodity level

Co. Codes	Commodities	Export/ Import in million US$	Mis-invoicing	Misinvoicing (% of Export/ Import)
		Export		
85	Electrical machinery and equipment	23,062.39	(4,236.94)	18.37
84	Nuclear reactors, boilers, machinery and mechanical app	20,709.38	(1,713.47)	8.27
87	Vehicles other than railway or tramway rolling stock	19,292.91	(2,291.59)	11.88
30	Pharmaceutical products	5,209.10	(1,259.83)	24.19
39	Plastics and articles thereof	4,365.18	(670.10)	15.35
90	Optical, medical or surgical instruments and apparatus	4,171.31	(487.49)	11.69
27	Mineral fuels, and oil	2,834.45	(751.70)	26.52
40	Rubber and articles thereof	2,554.53	(671.22)	26.28
29	Organic chemicals	1,865.16	(177.08)	9.49
Total		***113,382.08***		
		Import		
85	Electrical machinery and equipment	21,318.63	(594.99)	2.79
84	Nuclear reactors, boilers, machinery and mechanical app	17,214.52	(1,252.23)	7.27
87	Vehicles; other than railway or tramway	11,275.47	(1,763.99)	15.64
27	Mineral fuels and oils	8,090.65	3,441.16	42.53
39	Plastics and articles thereof	4,874.20	(143.28)	2.94
30	Pharmaceutical products	4,219.15	79.27	1.88
73	Iron or steel articles	2,441.79	(248.77)	10.19
72	Iron and steel	2,433.53	(217.81)	8.95
90	Optical, photographic, medical or surgical instruments	2,320.11	(1,171.11)	50.48
76	Aluminium and articles thereof	2,079.16	(170.15)	8.18
Total		***104,283.80***		

Source: Author's computation using UN Comtrade data

For imported commodities by Hungary, the trend is the same, with the only exception being Hungary's imports of mineral fuels and oils which recorded under-invoicing of US$ 3.44 billion. The import of mineral fuels and oils in Hungary calls for further investigation since about 50% of imports were not recorded in foreign exports accounts. As shown in Table 3, the evidence of over-invoicing is especially large for imported commodities such as electrical machinery and equipment; nuclear reactors, boilers, machinery and mechanical appliances; and vehicles other than railway or tramway. This significant over-invoicing problem in Hungary's exports and imports can be related to the refundable and non-refundable incentives available from the Hungarian government and the EU Funds, which seek to encourage investors keen to enlarge their trade base.

4.2.2. Hungary's Trade Misinvoicing by trading partner's level

Trade in Hungary appears to exhibit a heavy concentration on the EU market, with Germany alone accounting for about 27% of its total export value of US$113,4 billion, and import value of US104.4 billion. Trade here also shows substantial misinvoicing.

In terms of export, the analysis in Table 4 shows substantial over-invoicing, occurring with nine (9) out of the eleven (11) trading partners selected for this study. However, the aggregate results are heavily influenced by Germany. Export to Germany generates a cumulative amount of $4.3 billion in export over-invoicing which represent 14% of the total export of Hungary to Germany. The USA and the Chinese are the other trading partners in the sample for whom trade with Hungary exhibits export under-invoicing accounting for US$ 1.67 and US$ 1.15 billion respectively. Also, about 46% of Hungary's export to the Netherlands were not reported in the Netherlands's import data. Similarly, about 53% and 43% of imports recorded by USA and China respectively as goods from Hungary were not recorded in Hungary's official exports to those countries.

As is the case for Ghana, Hungary's imports are also flawed with large sums of discrepancies between itself and its trading partners as recorded in Table 4. Both over and under-invoicing could be identified in this case. Imports from countries like China, the Netherlands, Austria, Russia, USA, and the United Kingdom exhibited various forms of over-invoicing, with China and the Netherlands being the most or the largest contributors at US$ 3.3 billion and US$ 1 billion respectively. Also, countries like Germany, Italy, Czechia along with other countries exhibited large sums of import under-invoicing. Again, about 63% of goods recorded in Hungary's data as an import from China is not recorded in Chinese accounts as exported to Hungary. The same situation can be said for the Netherlands where about 20% of its exports to Hungary were not recorded in Hungary's official imports.

Table 4 Hungary's trade misinvoicing by trading partner level

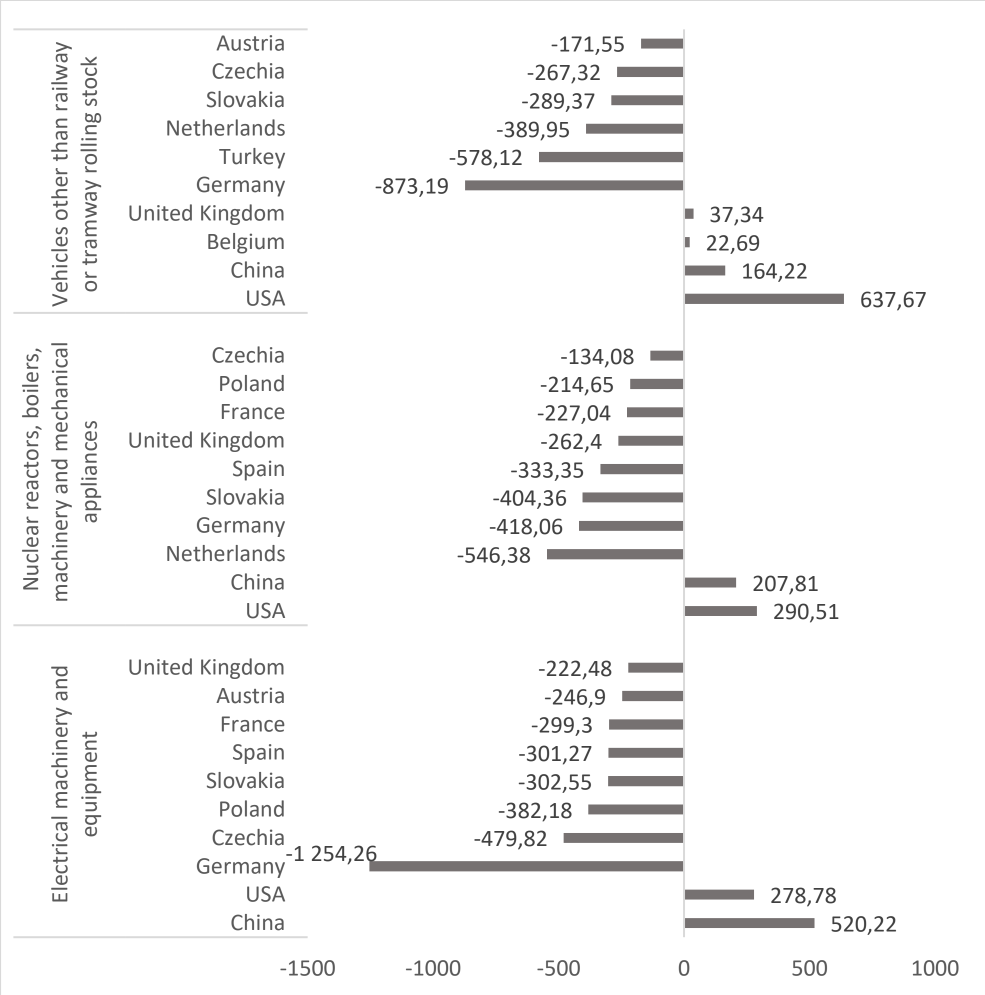
Country	Export/ Imports in million US$	Export misinvoicing	Misinvoicing (% of Export or Import)
	Export		
Germany	31,009.86	(4,392.63)	14.17
Austria	5,488.25	(1,570.75)	28.62
Slovakia	5,374.15	(1,891.77)	35.20
France	4,977.77	(747.24)	15.01
Czechia	4,879.21	(1,495.52)	30.65
Poland	4,673.16	(1,392.51)	29.80
United Kingdom	3,946.51	(763.22)	19.34
Netherlands	*3,898.95*	*(1,783.50)*	*45.74*
USA	*3,188.60*	*1,672.63*	*52.46*
Spain	3,170.26	(714.67)	22.54
China	*2,663.85*	*1,146.99*	*43.06*
	Import		
Germany	27,675.68	(3,234.66)	11.69
Austria	6,438.02	602.32	9.36
Poland	5,837.58	(543.30)	9.31
Slovakia	5,672.03	73.94	1.30
China	*5,291.35*	*3,336.61*	*63.06*
Netherlands	5,262.20	1,045.71	19.87
Czechia	5,140.03	(842.55)	16.39
Italy	4,973.52	(849.29)	17.08
France	4,198.13	(33.88)	0.81
Russian Federation	3,574.32	920.62	25.76
Romania	3,032.62	(620.42)	20.46
Belgium	2,313.50	(426.83)	18.45
USA	2,219.10	142.05	6.40

Source: Author's computation using UN Comtrade data

4.1.3. Hungary's Trade Misinvoicing by commodities and trading partner

The result of Hungary's export misinvoicing by commodity and trading partner in Figure 5 shows excessive negative discrepancies, suggesting export over-invoicing for all trading partners in the three main commodities considered. Whereas trade with Germany exhibits substantial export over-invoicing worth US $2.5 billion for the three main exported goods, that with China shows excessive under-invoicing for all the major commodities amounting to US$890 million.

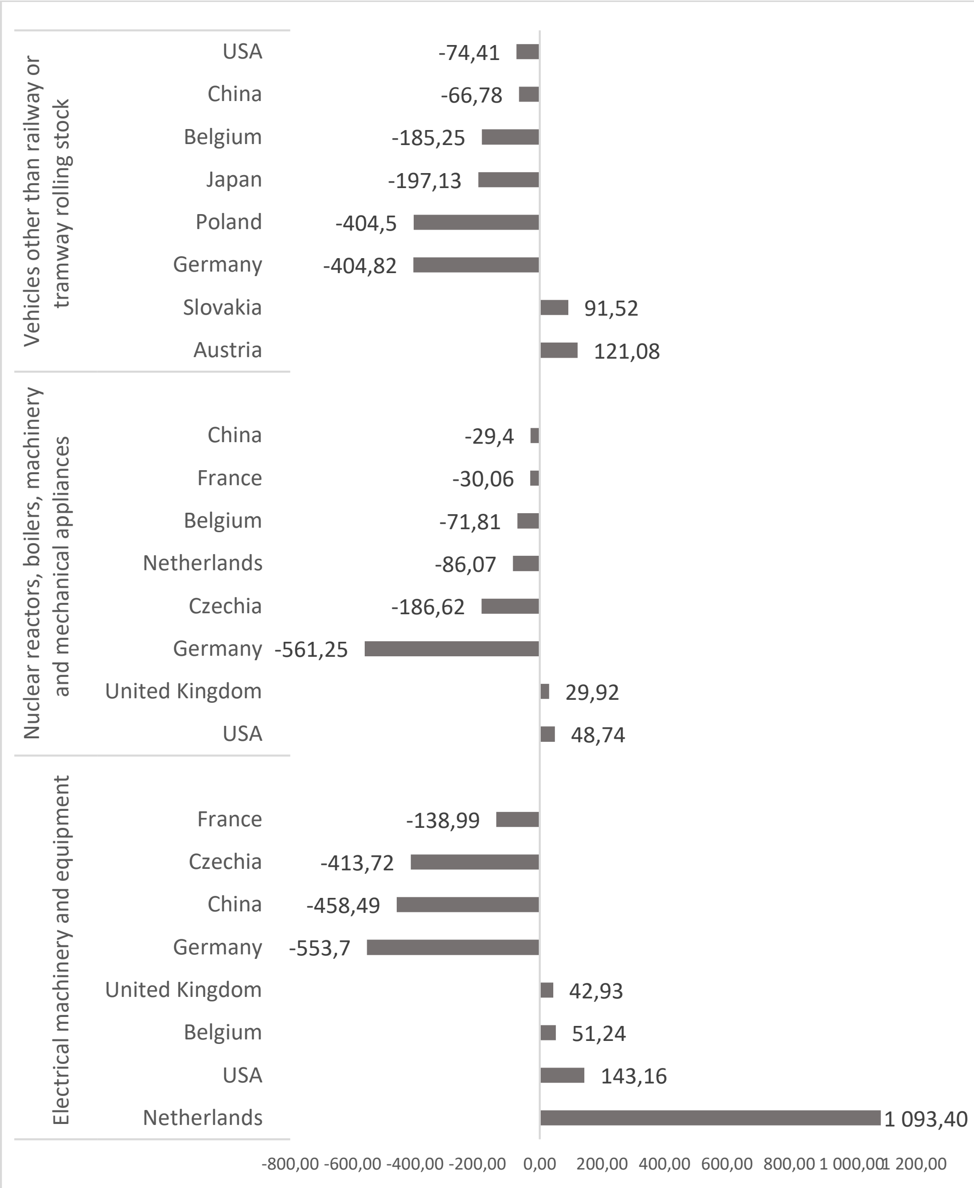
Figure 5 Hungary's Export misinvoicing by commodity and trading partner



Source: Author's computation using UN Comtrade data

Import misinvoicing by commodities and trading partner level in Hungary is also reported in Figure 6. The result shows evidence of both over and under-invoicing across all the three main traded goods. The results show that Hungarian imports from Germany, China, and Belgium across all the three main trading commodities witnessed under-invoicing. Other countries like the Netherlands and the United States witnessed both over and under import misinvoicing. In the case of electrical machinery and equipment, 60 per cent of Hungary's import is not recorded in its import data. Again, import misinvoicing of vehicles other than railway or tramway rolling stock account for the same value to the total import of Hungary from Belgium.

Figure 6 Hungary's import misinvoicing by commodity and trading partner



Source: Author's computation using UN Comtrade data

5. Conclusion and policy recommendations

This paper examines the magnitude of trade misinvoicing in Ghana and Hungary with particular emphasis on key commodity groups and trading partners that are heavily involved in the misinvoicing practice. The significant nature of the estimates highlights the seriousness of trade misinvoicing as far as revenue mobilization, and total trade is concerned. The results from a handful of commodity groups used for the analysis shows that Ghana loses about US$3 billion in the export of precious pearls, metal and stones and mineral fuel and oil alone. In imports, Ghana loses about US 2.3 billion to import over-invoicing in vehicles other than railway or tramway rolling stock; salt, sulphur and stone; cereals and paper and paperboard. Similarly, the Hungarian government also lost about US$9.5 billion in over-invoicing of commodities such as electrical machinery and equipment, nuclear reactors, boilers, machinery and mechanical appliances, vehicles other than railway or tramway rolling stock and pharmaceutical products. An additional US$3.6 billion is also lost to import under-invoicing to the former three commodities groups.

Ghana's trading partners, which were significantly prone to export over-invoicing include China, Canada, and India, whereas countries such as the USA and France contribute significantly to import under-invoicing. In terms of imports, Ghana's trade with the United Kingdom, Spain, the USA, and Canada recorded a large amount of over-invoicing while Ghana's trade with China, Netherlands, and France showed the opposite. For Hungary, exports to countries like Germany, the Netherlands, the United Kingdom, Belgium, Poland, Czechia, and France exhibit substantial over-invoicing. Moreover, imports from countries like China, the Netherlands, Austria, Russia, USA and the United Kingdom exhibited various forms of over-invoicing, with China and the Netherlands being the most significant or largest contributors. In contrast, countries like Germany, Italy, Czechia along with other countries exhibited large sums of import under-invoicing.

These results highlight the need for both governments to increase their access to data especially at their border sites and possibly track custom valuations declared at their border stations to that of their trading partners to detect any possible trade corruption, and institute punitive sanctions against companies found misinvoicing to deter others from engaging in it. Also, since trade incentives, high tariffs and non-tariff barriers encourage misinvoicing, a more pragmatic trade liberalisation policy should rather be pursued, especially by developing countries. However, if tax incentives are to be used, then it is imperative that it is targeted at nontraditional products to facilitate its trade.

References

Ajayi, S. (1998): Adjusting for international trade-faking. In Iqbal, Z. and Khan, M. S. (Eds.). *Trade reform and regional integration in Africa.* International Monetary Fund, (Vol 763).

Bahmani-Oskooee, M. ‒ Goswami, G. G. (2003): Smuggling as another cause of failure of the PPP. *Journal of Economic Development*, 28, 2, 23–38.

Baker, R. ‒ Clough, C. ‒ Kar, D. ‒ LeBlanc, B. ‒ Simmons, J. (2014): *Hiding in Plain Sight: Trade Misinvoicing and the Impact of Revenue Loss in Ghana, Kenya, Mozambique, Tanzania, and Uganda: 2002–2011*. Washington DC: Global Financial Integrity.

Barnett, R. C. (2003): Smuggling, non-fundamental uncertainty, and parallel market exchange rate volatility. *Canadian Journal of Economics*, 36, 3, 701–727.

Beja, E. L. (2007): Brothers in Distress: Revolving Capital Flows of Indonesia, Malaysia, and Thailand. *Journal of Asian Economics*, 18, 6, 904–914.

Berger, H. ‒ Nitsch, V. (2012): Gotcha! A profile of smuggling in international trade. In Storti, C. C. and De Grauwe, P. (Eds.): *Illicit Trade and the Global Economy*: CESifo Seminar Series. Cambridge and London: MIT Press, 49–72.

Bhagwati, J. (1964): On the under-invoicing of imports [with application to recent Turkish experience]. *Bulletin of the Institute of Economics and Statistics (Oxford University)*, 26, 389–397.

Biswas, A. K. ‒ Marjit, S. (2005): Misinvoicing and trade policy. *The Journal of Policy Reform*, 8, 3, 189–205.

Epaphra, M. (2015): Tax rates and tax evasion: evidence from missing imports in Tanzania. *International Journal of Economics and Finance*, 7, 2, 122–137.

Ferrantino, M. J., – Zhi, W. A. N. G. (2008): Accounting for discrepancies in bilateral trade: The case of China, Hong Kong, and the United States. *China Economic Review*, 19, 3, 502–520.

Fisman, R., – Wei, S. J. (2004): Tax rates and tax evasion: evidence from "missing imports" in China. *Journal of Political Economy*, 112, 2, 471–496.

Global Financial Integrity (2017): *Illicit financial flows to and from developing countries, 2005–2014*, Washington, DC: Global Financial Integrity (April).

Global Financial Integrity (2019): *Illicit Financial Flows to and from 148 Developing Countries: 2006–2015*. Washington DC: Global Financial Integrity (January).

Global Financial Integrity (2015): Illicit financial flows to and from developing countries, 2005–2014, Washington, DC: April.

Hangzhou, P. (2009): *Report on the statistical discrepancy of merchandise trade between the United States and China.*

Jha, R. – Truong, D. N. (2014): *Trade misinvoicing and macroeconomic outcomes in India*. The Australian National University, Australia South Asia Research Centre, ASARC Working Papers.

Kar, D. (2010): The drivers and dynamics of illicit financial flows from India: 1948–2008. Washington DC: Global Financial Integrity (November).

Kellenberg, D. – Levinson, A. (2016): M*isreporting trade: tariff evasion, corruption, and auditing standards:* National Bureau of Economic Research, Inc.

Martin, M. (2016): What's the difference? comparing U.S. and Chinese trade data. Congressional Research Service. Retrieved from https://www.fas.org/sgp/crs /row/ RS22640.pdf.

Morgenstern, O. (1963): *On the accuracy of economic observations*. Princeton University Press.

Ndikumana, L. – Boyce, J. K (2018): *Capital flight from Africa: updated methodology and estimates*. PERI Working Paper. Retrieved from https://www.peri.umass.edu/publication/item/1083-capital-flight-from-africa-updated-methodology-and-new-estimates.

Ndikumana, L. – Boyce, J. K. – Ndiaye, A. S. (2015): Capital flight from Africa: Measurement and drivers. In Ajayi, S. I. and Ndikumana, L. (Eds.), *Capital Flight from Africa: Causes, Effects and Policy Issues* (pp. 15–54). Oxford: Oxford University Press.

Ndikumana, L. – Boyce, J. K. (2010): Measurement of capital flight: Methodology and results for sub-Saharan African countries. *African Development Review,* 22, 4, 471–481.

Östensson, O. (2018). Misinvoicing in mineral trade: what do we really know? *Mineral Economics*, 31, 1–2, 77–86.

UN-COMTRADE (2018): https://comtrade.un.org/. Accessed on December 20, 2018.

UNCTAD (2016) *Trade misinvoicing in primary commodities in Developing Countries: the cases of Chile, Côte d'Ivoire, Nigeria, South Africa and Zambia,* UNCTAD/SUC/2016/2, 2nd version.

Yeats, A. J. (1990). On the accuracy of economic observations: do sub-Saharan Africa trade statistics mean anything? *World Bank Economic Review,* 4, 2, 135–156.

Appendix

Table 1 Ghana's export misinvoicing by commodity and partner trading level

Country	Export in a million US$	Misinvoicing	Misinvoicing (% of export)
	Precious pearls, metal and stones		
India	2,536.16	(243.96)	9.62
Switzerland	1,630.36	(56.01)	3.44
South Africa	851.75	(933.16)	91.28
China, Hong Kong	51.28	(54.13)	94.73
Turkey	33.81	(1.41)	4.18
Lebanon	23.27	0.55	2.35
USA	9.16	(2.36)	25.75
Spain	4.39	(4.35)	99.17
Belgium	2.60	(0.91)	35.18
Rep. of Korea	2.20	(0.23)	10.36
TOTAL	5,861.4		
	Mineral fuel and oil		
China	2179.96	(935.50)	42.91
Canada	258.25	(284.08)	90.91
Netherlands	202.23	46.64	23.06
USA	160.45	265.44	60.45
France	91.79	(9.29)	10.12
Togo	62.66	(59.54)	95.02
Italy	52.87	(0.64)	1.21
Spain	52.50	(2.45)	4.66
South Africa	49.18	(2.61)	5.30
Japan	48.47	39.55	21.41
	Cocoa and cocoa preparations		
Netherlands	584.15	(69.86)	11.96
Malaysia	273.79	(141.33)	51.62
Brazil	191.76	(50.76)	26.47
Germany	185.68	(67.99)	36.61
USA	181.11	31.17	17.21
Estonia	130.90	(53.00)	40.49
Spain	130.15	(65.70)	50.48
France	122.70	66.41	54.13
Japan	106.54	16.17	15.18
Belgium	106.17	(18.46)	17.39

Source: Author's computation using UN Comtrade data

Table 2 Ghana's import misinvoicing by commodity and partner trading level

Country	Import in million US$	Misinvoicing	Misinvoicing (% of Import)
Electrical machinery and equipment			
China	214.71	(433.38)	49.54
USA	120.84	(17.57)	14.54
Turkey	65.10	34.66	53.24
United Kingdom	54.18	18.16	33.52
South Africa	31.03	(16.58)	53.44
France	21.32	6.33	29.67
Belgium	21.06	11.97	56.83
Australia	9.85	2.36	23.93
Spain	5.86	(6.85)	116.82
TOTAL			
Nuclear reactors, boilers, machinery and mechanical appliances			
China	313.58	(131.19)	41.84
Belgium	165.47	83.85	50.68
United Kingdom	112.84	(16.97)	15.03
Germany	75.85	(11.08)	14.61
India	62.74	(20.94)	33.38
Netherlands	41.62	(35.96)	86.42
Australia	32.49	(15.18)	46.73
France	23.54	(66.57)	35.37
Rep. of Korea	16.53	(13.42)	81.19
Sweden	12.60	(23.84)	52.85
TOTAL			
Salt, Sulphur and stone			
Spain	602.94	571.61	94.80
Turkey	101.29	64.90	64.08
Rep. of Korea	39.61	37.09	93.65
China	34.82	22.90	65.77
Belgium	11.54	6.15	53.32
Sweden	9.22	6.00	65.11
Italy	4.30	2.11	48.95
USA	0.45	(0.28)	62.45
South Africa	0.23	(0.42)	55.68
United Kingdom	0.14	(0.09)	62.89
TOTAL			

Source: Author's computation using UN Comtrade data