Possible approaches of the Quarterly GDP Estimation

Klára Anwar[1]

Nowadays, quarterly national accounts - in particular quarterly GDP - represent an essential instrument and source of information for both economic and monetary policy as continuously underlined by the major users such as the European Central Bank (ECB), government authorities, economic actors and forecasters.
At the end of 2006, Hungarian Central Statistical Office (HCSO) has introduced full set of current price quarterly calculations based on statistical and administrative data sources. According to this approach, on production side quarterly GDP is not obtained directly but as a difference of output and intermediate consumption plus taxes less subsidies on production. For this compilation, intermediate consumption has been estimated quarterly though there is no direct data source available for this figure. Under continuous methodological work, I would like to explore and analyse the possible approaches to estimate intermediate consumption quarterly.

Keywords: quarterly national accounts, GDP, intermediate consumption, methodology

1. Introduction

The HCSO has started to establish the methodological work on quarterly national accounts in 1993. Its actuality and importance is well defined in the European System of Accounts as "they are the only coherent set of indicators, available with a short time-lag, able to provide a short term overall picture of both non-financial and financial economic activity." (Eurostat 1996)

At the beginning, quarterly GDP was estimated by volume projection method on production side, i.e. available proxy indicators were applied to extrapolate the value added series at two-digit level of NACE[2], the available nomenclature for economic activities at that time. Volume indices were aggregated by using the shares of industries in the total value added of the base year. Thus GDP from the production side was not available at current prices. (KSH 2002) While from the expenditure side the estimation was based on current and constant price estimation from the beginning. This fact evolved a methodological gap between the two

[1] Klára Anwar, statistical counsellor, Hungarian Central Statistical Office, Hungary
[2] NACE is the European classification system of the economic activities by industries (abreviation of Nomenclature générale des Activités économiques dans les Communautés Européennes).

approaches. (Anwar 2008) However, some European recommendations also expected the use of current price approach on both sides. The first results of this approach were published by HCSO in December 2006. (HCSO 2009)

Nevertheless, there are continuously growing demands against the quarterly GDP, like it should be more and more timely and also more and more accurate, comprehensive, and also reasonably detailed, generates continuous research work and improvements.

2. Data sources and methods

Due to obtain value added at current and constant prices both output and intermediate consumption have to be estimated. Its reason is provided by one of the well known National Accounts principle called double deflation. According to it, value added at constant prices can be obtained by deducting the separately deflated intermediate consumption from output. The ESA 1995 therefore defines value added at constant prices "as the difference between output and intermediate consumption at constant prices" (Eurostat 1996). This can be expressed in the following equation:

$$VA = \sum P_O \cdot Q_O - \sum P_{IC} \cdot Q_{IC},$$

where P_O and Q_O are prices and quantities for output and P_{IC} and Q_{IC} are prices and quantities for intermediate consumption.

Similarly to most of the EU countries, no short term data are available for the calculation of intermediate consumption, thus an alternative estimation method should be applied (OECD 1996). According to international methodologies, the application of the intermediate consumption to output ratio from the last available annual accounts is a generally applied method. My purpose is to improve or confirm the quarterly intermediate consumption estimation.

Therefore I analyzed the relationship between the annual intermediate consumption and the annual gross output for non-financial corporations sector at the A6 ESA classification level (Eurostat 1996) at current prices for the period of 1995-2008. I prepared my analysis on current price data due to the chain-linking methodological concept, i.e. the current price of the previous year provide the comparable price for the actual year using annual weighting (Anwar–Szőkéné Boros 2008). As a first step, I chose A6 aggregation level and if a good relationship can be proved at this level than a detailed analysis may be followed.

The A6 classification means 6 groups of branches in the following grouping: agriculture and fishing; industry (mining; quarrying; manufacturing and electricity); construction; trade, repair, hotels, restaurants; transport, storage and communication; financial intermediation and real estate activities; public administration, education,

health, social work and other community, social and personal service activities. The first group of A6 was deducted from the observation due to the fact that the intermediate consumption of agriculture practically remains stable within the year in spite of any changed in its output. While I added the branches, total to the observed series.

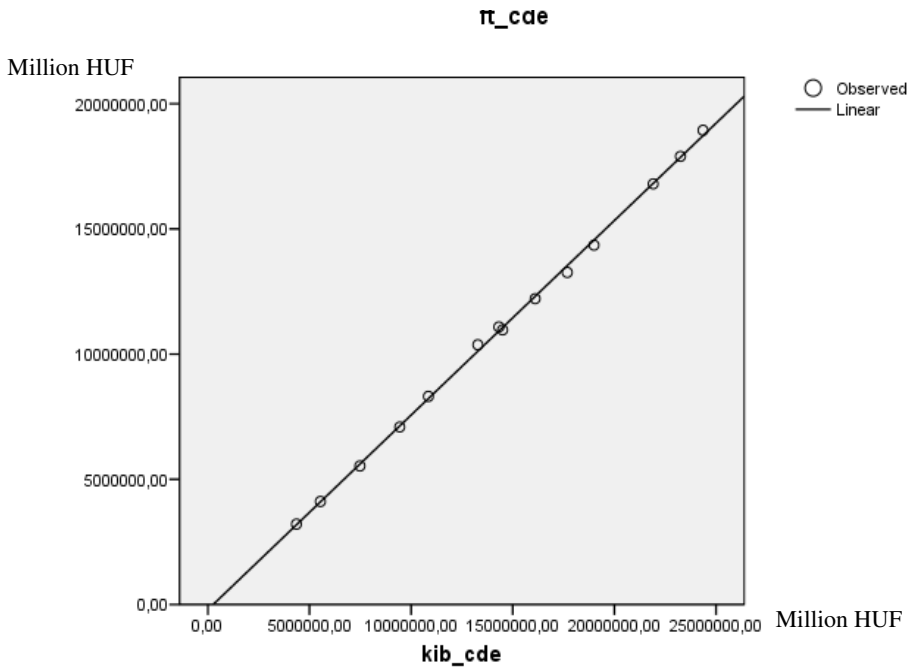
For this study I applied SPSS 15.0 software. I used linear and log linear regressions as well to describe the relationship but similar results were reached. Than, I have performed some time series analysis.

3. Results

Each observed series have 14 elements as HCSO compiled consistent national accounts time series back to 1995. Though more elements would be suitable for any examination, the consistent time series are only available for this period. Each series were examined separately to find out the relationship between the two types of the figures.

In case of the **industry – mining, manufacturing and electricity, gas and water supply** (CDE) R^2 value showed a very strong relationship i.e. variance of the intermediate consumption (IC) can be explained 99.9% by the output, and the rest 0.1% is random error. (See figure 1)

Figure 1. The relationship between output *(kib_cde)* and intermediate consumption *(ft_cde)* for industry



Source: SPPS software results.

The *F* statistic was highly significant, therefore the null hypothesis was rejected according to the coefficients' test to zero value. The results of *t* statistics showed also the same, it was highly significant as well. It was so nice to be true. I checked the Durbin-Watson test. Its value (see table 1) explained everything.

Table 1. **Model Summary[b]**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	1,000[a]	,999	,999	155308,470	,815

a. Predictors: (Constant), kib_cde

b. Dependent Variable: ft_cde

Source: SPPS software results.

Though there is a thumb rule that claims if the result of Durbin-Watson test is under the R^2 value means that there is false regression between the examined

figures, if the number of the observed population is low, the results should be tested by other methods (Darvas 2001).

A possible method to avoid false regression is to set up regression for the differences. This resulted appropriate statistics. Beside the high R^2 value (99.6%) and the significant *F* and *t* statistics, the Durbin-Watson test showed also an acceptable (Hunyadi 2001) result (1.523) beside 1.0097 and 1.3404 critical values, i.e. residuals are linearly independent. The significance level of Kolmogorov - Smirnov test (0.990) showed that the residual distribution is normal.

In case of **construction** (F) the deterministic coefficient was similar to that of the industry, and the *F* and *t* statistics also explained high significance between the 2 figures. The result of the Durbin-Watson test was 1.192 here that fell in the indifferent zone, between the critical values of 1.04495 and 1.35027, i.e. according to this test; no decision could be taken. So I tested the regression for the differences and an acceptable Durbin-Watson figure (1.928) was reached. Furthermore the significance level of Kolmogorov - Smirnov test (0.640) described normal distribution of residuals.

The third observed group of industries was **trade, repair, hotels, restaurants, transport, storage and communication** (GHI). In this case correlation statistics showed that output may explain the intermediate consumption variance by 99.7%. This was justified by significant *F* and *t* statistics. Even the result of Durbin-Watson test (1.812) and the significance of Kolmogorov -Smirnov test (0,832) confirmed this result. Thus the relationship between the two figures of these industries might be expressed by regression.

In case of **financial intermediation and real estate activities** (JK), the deterministic coefficient was significant as well by its extremely high, 99.9%, and it was confirmed by 19 621.69 *F* statistic and by significant *t* statistic value.

The Durbin-Watson test resulted 2.536, i.e. the null hypothesis that residuals do not correlate to each other could be accepted. The significance of Kolmogorov - Smirnov test (0.532) showed that the normal distribution of residuals is granted. Therefore, the relationship between the two figures might be expressed by regression in this case.

The forth case was the group of **public administration, education, health, social work and other community, social and personal service activities** (LMNO) branches. This showed similar results to those of the construction. Though the high deterministic coefficient (99.4%) was justified by the results of *F* test (2 091), that of *t* statistic (see Table 2) and by the significance of Kolmogorov - Smirnov test (0.957). The Durbin-Watson test result fell in the indifferent zone. Thus, the regression for the differences was tested and the new Durbin-Watson test explained (1.967) linearly independent residuals.

The last case was **the industries, total,** i.e. total output and total intermediate consumption. This showed similar facts. R^2 explained significant relationship between intermediate consumption and output that was supported by significant *F*

and *t* statistics values, and also by the significance of Kolmogorov - Smirnov test (0.434), while the Durbin-Watson test value (1.183) fell in the indifferent zone. Therefore the regression for the differences has been tested and acceptable Durbin-Watson figure (2.42) was obtained as according to the critical values this figure should be lower that 2.64973 and higher than 1.35027 i.e. there is no autocorrelation in the model. (Anwar-Ugrósdy 2009)

The results that 3 of the 6 studies were finished by undecisionable Durbin-Watson test induced me for further examination; however regression set up for differences provided acceptable results in these cases.

It can be also stated that the number of the observed population was relatively low; however it was the highest number of population that could be observed at that time as the quarterly GDP series start at 1995.

Even if the regression set up for differences gave acceptable results, this weakens the success of the model. This means that new indicators should be introduced into the analysis, or further examination is needed.

Therefore I applied time series analysis approach to investigate the relationship between the series. (See table 2. for the case of industry.)

Table 2. **Parameter Estimates**

		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	AR1	,711	,280	2,545	,026
Regression Coefficients	kib_cde	,771	,007	117,296	,000

Melard's algorithm was used for estimation.

Source: SPPS software results.

Due to the fact that the level of the autocorrelation was very high in each case, differences should be calculated for each series to reach stationarity. These transformed series had acceptable autocorrelation level i.e. the application of one difference was enough, and even the significance of Kolmogorov - Smirnov test showed normal distribution of residuals. In spite of the expectation, the parameter of the autoregressive model (at lag one) was not significant in any of the observed series; however other statistics confirmed the significance of these models.

3. Conclusion

According to the above mentioned results, in case if there is not enough direct information available for intermediate consumption, it could be advisable to proceed GDP estimation on the following way: to obtain the intermediate consumption to output average ratio of the last period for a given branch and apply it in the estimation of the current period data.

Though according to the above mentioned results, further examination is also needed from another aspect. It seems a possible approach to filter out the trend factor from both output and intermediate consumption series. Then the relationship between the two types of the obtained series should be analysed again for each group of branches. This may lead to better coverage between the two variables and smaller relative error.

References

Anwar, K. 2008: Current Price Approach of Quarterly GDP Estimations from P roduction Side in Hungary, *Hungarian Statistical Review*, Vol. 86. Special No. 12., 45-57 pp.

Anwar, K. – Szőkéné Boros, Zs. 2008: A láncindexek alkalmazása a nemzeti számlákban, *Statisztikai Szemle*, Budapest, 86. Évf. 7-8., 713-731 pp.

Anwar, K. – Ugrósdy, Gy. 2009: New Aspect of the Methodological Improvement of Quarterly GDP Estimation, *Bulletin of the Szent István University*, Gödöllő (under publication)

Darvas, Zs. 2001: *Beezetés az idősorelemzés fogalmaiba*, jegyzet, Corvinus Egyetem, Budapest, 59-70 pp.

EUROSTAT 1996: *European System of Accounts*, ESA 1995, Luxembourg, 337-342, 426 pp.

EUROSTAT 2002: *Quarterly chaining of National Accounts*, Eurostat, B1/CN 516 e number manuscript, Luxemburg

Hunyadi, L. 2001: *Statisztikai következtetéselmélet közgazdászoknak*, KSH, Budapest, 407-419 pp.

HCSO 2009: *First release, Gross Domestic Product 2009 2nd quarter*, Budapest

KSH 2002: *A negyedéves bruttó hazai termék (GDP) számítási módszere* Magyarországon, Budapest

OECD 1996: *Quarterly National Accounts, Sources and Methods* Used by OECD Member Countries, Paris