How can statistical analysis prove the existence of an independent foreign language skill?

Ágnes Dévény[1]

The aim of my research was to prove that foreign language mediation is an independent language skill that can be measured by specific criteria and as part of a foreign language examination it contributes to a more complex assessment of the test-takers' language knowledge.

Besides other methods that are beyond the scope of this paper, the most important part of the research was the descriptive and mathematical statistical analysis of the results of the 18 examination periods covering more than 27000 test results from Budapest Business School Examination Centre. I wanted to present in this paper how I used statistical analyses to prove the existence of an independent foreign language skill.

My research verified my hypothesis that written mediation examination task measures a segment of language knowledge independently; skills (subskills) appearing in it are not present in other examination tasks. Foreign language mediation is an independent language skill; therefore bilingual language examinations can measure skills that are not measureable by monolingual examination systems.

Keywords: assessing foreign language competence, language testing, pedagogy

1. Introduction

Is foreign language mediation an independent language skill? Is it part of our conception of language proficiency and a segment of the language learners'/users' language competence? There have been heated debates in professional circles on the role and function of mediation and the debates have not been settled yet (Szabari 2001; Bárdos 1997, 2005; Heltai 2001). Experts interpret and explain the concept of mediation in many different ways and their opinions vary, reflecting the insecurity of specialists in the field of bilingualism/multilingualism. Debates become even more heated when experts have to decide whether a mediation task (e.g. written mediation test) can be considered as legitimate language examination subtest (task) (Klaudy 1984, 1986a, 1986b, 1990; Fekete 2001, 2002). The scientific investigation of bilingualism/multilingualism in education and testing, as well as the interpretation

[1] Ágnes Dévény, PhD, associate professor, Budapest Business School, Institute of Foreign Languages and Communication, Faculty of Commerce, Catering and Tourism, Hungary (Budapest)

Phillips, D. 1990: *Longman practice tests for the TOEFL*. Longman, New York.
Popham, W. J. 1990: *Modern educational measurement*. Allyn and Bacon, Boston.
Shohamy, E. 1984: Does the testing method make a difference? The case of reading comprehension. *Language Testing,* 1, pp. 146-169.
Urquhart, A.-Weir, C. J. 1998: *Reading in a second language: Process, product and practice.* Longman. London.
Wolf, D. F. 1993: A Comparison of assessment tasks used to measure FL reading comprehension. *Modern Language Journal*, 77, pp. 473-489.

gained about test format effects in this study do not provide a comprehensive conclusion, further research is needed in the area, especially in two directions: how task types influence performance at different levels of proficiency, and how task type effect is related to the conceptual and linguistic difficulty of a reading text.

References

Alderson, J. C. 2000: *Assessing reading.* Cambridge University Press, Cambridge.

Bachman, L. F. 1991: *Fundamental considerations in language testing*. Oxford University Press, Oxford.

Bachman, L. F. 2004: *Statistical analyses for language assessment.* Cambridge University Press, Cambridge.

Baker, R. 1997: *Classical test theory and item response theory. Theory in test analysis* [Special report 2, Language testing up-date]. Department of Linguistics and Modern English Language, Lancaster University, Lancaster.

Bárdos, J. 2002: *Az idegen nyelvi mérés és értékelés elmélete és gyakorlata.* Nemzeti Tankönyvkiadó, Budapest.

Csapó, B. 1993: Tudásszintmérő tesztek. In Falus, I. (ed.): *Bevezetés a pedagógiai kutatás módszereibe.* Keraban Kiadó, Budapest, pp. 277-317.

Dancey, C. P. - Reidy, J. 2004: *Statistics without maths for psychology*. Pearson/Prentice Hall, Harlow, UK.

Gordon, C. M. - Hanauer, D. 1995: The interaction between task and meaning construction in EFL reading comprehension tests. *TESOL Quarterly,* 29, pp. 299-322.

Henning, G. 1987: *A guide to language testing*: *Development, evaluation and research*. Newbury House, Cambridge, MA.

Horváth, Gy. 1997: *A modern tesztmodellek alkalmazása.* Akadémiai Kiadó, Budapest.

Liu, F. 2009: The effect of three test methods on reading comprehension: An experiment. *Asian Social Science, 5,* No. 6. pp. 147-153. Retrieved on 01/05/2010 from http://ccsenet.org/journal/index.php/ass/article/view/2491.

Loch, A. 2007: Task types in assessing reading comprehension. *BGF Tudományos Évkönyv.* BGF, Budapest, pp. 358-362.

Loch, A. 2009a: Native language vs. target language use in language testing. *Acta Linguistica* 8, Vol. 2, University of Matej Bela, Banska Bystrica, Slovakia, pp. 64-70.

Loch, A. 2009b: Minták összehasonlíthatóságának bizonyítása statisztikai eljárásokkal. Paper presented at *New Methods in Applied Linguistics Research Conference*. 26-27. October 2009, Budapest.

McNamara, T. 2000: *Language testing.* Oxford University Press, Oxford.

It was also important whether task type and language use exercised any joint effect on the participants' performance. In the case of *Test 1*, the two variables jointly did not show a significant relationship with the test scores although had a significant effect on means separately (Table 7). This fact as well as the findings that the MC format affected performance on *Test 1* negatively and on *Test 2* positively indicate that task type and other factors may interfere.

Table 7. Tests of significance in relation to the Analysis of Variance for the joint effect of task type and language use

Variable	df	F	Significance	Partial Eta squared
Test 1				
Corrected Model	3	6.925	.000	.082
Task type	1	13.204	.000	.053
Language	1	5.183	.024	.022
Task type*language	1	3.382	.067	.014
Total	238			
Test 2				
Corrected Model	3	1.384	.248	.017
Task type	1	1.595	.208	.007
Language	1	2.504	.115	.011
Task type*language	1	.243	.623	.001
Total	238			

Source: own creation

Statistics confirmed the expectations that the participants' reading scores would significantly correlate at the .01 level with their scores on the TOEFL papers on receptive skills (r = .584). However, the overlap between scores was not particularly large, which indicates the distinctiveness of reading skills

5. Summary of findings

In the light of the results, it is obvious that the research hypotheses gained partial verification only. The statistics showed that in most cases no significant difference was found. However, there were exceptions, both in the task type and the language use comparisons, when the mean differences reached the statistically significant level. Due to these mixed results, no general conclusions can be drawn.

Although results were not consistent, some of the findings strongly suggest that test method variables may exert a joint effect with other factors such as text difficulty or test takers' characteristics. As the mixed results

.080) showed that eight percent of the variation in test scores could be explained by the different languages. In the case of *Test 2* the non-parametric Mann-Whiney U test also indicated significant difference (z = -2,054, Asymp. Sig.= 0,40) (Table 5). This indicates that, at least in some cases, the language of the task had a decisive influence on the response.

Table 5. Test of significance in relation to the Mann-Whitney U test statistics of test scores on the L1 and L2 versions of the MC tests

	Test 1	Test 2
Mann-Whitney U	1223.500	.1451.500
Wilcoxon W	2819.500	3662.500
Z	-3.230	-2.054
Asymp. Sig. (2-tailed)	.001	.040

p<.05

Source: own creation

The third comparison focused on investigating the effect of task type. As the English multiple choice version of *Test 2* was slightly positively skewed, in this case the Mann-Whitney U test was applied. Significant difference was found in one case (Table 6): between the means of test scores on the SAQ and MC versions of *Test 1* in English ($t_{113} = 3.800$, $p < .001$), with an effect size $d = .72$. Performance on the SAQ test highly exceeded performance on the MC test. Task type explained 10 percent of the variation in performance ($R^2 = .105$).

Table 6. Results of Independent Samples t test and Mann-Whitney U test to compare means from different test formats

Test	Version	Groups	t	df	Sig. (2-tailed)	Z	Asymp. Sig. (2-tailed)
1	English SAQ + MC	A + B	3.800	113	.000*		
	Hungarian SAQ + MC	C + D	1.288	121	.200		
2	Hungarian SAQ + MC	A + B	1.201	113	.232		
	English SAQ + MC	C + D				-.682	.496

* *p*<.001

Source: own creation

When comparing the difficulty level of the tests, it was found that in the case of the SAQ tests the difference between test means was significant at the $p < .001$ level, with *Test 2* being more difficult for the participants than *Test 1*. However, in the case of the MC tests: in one group it was *Test 1*, whereas in the other group *Test 2* that proved to be significantly more difficult for the students. Questionnaire data and group interviews seemed to suggest that the test takers' insufficient language knowledge did not allow them to choose the correct answer from the only slightly different options provided in the MC items, which resulted in inconsistent test-taking behaviour. In spite of these results, positive correlation was found between the students' scores on the two tests.

Next, the English and Hungarian versions of the same tests were compared to investigate native language use effect on test performance. Both the means and the facility values showed that the Hungarian versions were easier and elicited higher performance (Table 4) although in the case of the SAQ tests the difference did not reach statistical significance at the .05 level.

Table 4. Comparative data about the SAQ and the MC tests

	Test 1 SAQ		Test 2 SAQ	
	in English	in Hungarian	in English	in Hungarian
Group	Group A	Group C	Group C	Group A
N	64	60	62	60
M	13.47	13.67	9.87	10.70
Range	22	18	20	20
SD	4.8500	4.7929	4.8332	4.8198
Variability	23.523	22.972	23.360	23.231
Facility value (p)	.5625	.5688	.4113	.4458
	Test 1 MC		Test 2 MC	
	in English	in Hungarian	in English	in Hungarian
Group	Group B	Group D	Group D	Group B
N	56	66	68	58
M	6.89	8.36	6.57	7.31
Range	12	10	10	9
SD	2.4913	2.4970	2.5934	2.5902
Variance	6.206	6.235	6.726	6.709
Facility value (*p*)	.4308	.5625	.4373	.4908

Source: own creation

In the case of the MC tests, however, the test takers performed significantly better on the Hungarian version. The mean difference between the English and the Hungarian versions of *Test 1* was significant *(t_{120}=-3,245, p = .002)*, and there was a medium effect size (d = .59) (Dancey-Reidy 2004). The adjusted R squared (R^2=

computed in a Rasch analysis, and were also compared in an analysis of variance ($F_{3,234}$ = .422, p = .737). The result showed that the groups were highly comparable (Loch 2009b).

4.3 Analysing data from test results

Following the traditional line of Classical Test Theory (CTT), the scores were regarded as interval data and were processed accordingly. For the statistical analyses the Statistical Package for Social Sciences software[3] was used. As the procedures of Item Response Theory (IRT) are recommended for much larger sample sizes, their application was limited and complementary in the present study (Bachman 2004, Baker 1997, Horváth 1997).

The test takers' performances on the eight test versions were compared by using both parametric and non-parametric statistical computations because distribution on one of the eight tests was slightly skewed. The procedures applied are shown in Table 3.

Table 3. Statistical procedures employed in the data analysis

Type of analysis	Non-parametric tests	Parametric tests
Checking for distribution	Chi-square	Chi-square
Comparing means (two data sets)	Wilcoxon test Mann-Whitney U test	Paired-samples *t* test Independent *t* test
Comparing means (more than two data sets)	Kruskall-Wallis test	ANOVA
Relationship between variables	Spearman rank order correlation	Pearson product moment correlation
Relationship among variables		Regression analysis Univariate analysis of variance

Source: own creation

Inferential statistics were run at three levels. First, *Test 1* and *Test 2* versions were compared to see if they were the same difficulty level. Secondly, the English (L2) and the Hungarian (L1) versions of the same tests were compared to check them for language effect. Finally, the short answer question version and the multiple choice version of the same tests were analysed to investigate task type effect. As the four groups were highly comparable concerning language ability, group differences were excluded from the possible reasons for potential differences.

[3] SPSS Inc. (1989-2003). *Statistical Package for Social Sciences* (Versions 11.0, 12.0)

ecy Formula states the relationship between reliability and test length mathematically based on the assumption that the added items are of similar quality to other items in the test. Using the Spearman-Brown Prophecy Formula it was possible to calculate what the reliability of the MC tests would become if they contained the same number of items as the respective SAQ tests. The formula says

$$r_{ttn} = \frac{nr_{tt}}{1+(n-1)r_{tt}}$$

where, r_{ttn} = the reliability of the test when adjusted to n times its original length
r_{tt} = the observed reliability of the test at its present length
n = the number of times the length of the test is to be augmented.

By using the Prophecy Formula, in the case of *Test 1,* the estimated reliability of the MC test version was α = .8148, which corresponded to the respective SAQ test reliability (α = .8149). In the case of *Test 2* the calculated reliability for the MC test was α = .8082, which is also above the .8000 level. Although it was not possible to lengthen the MC tests to that extent, using the Spearman-Brown Formula was still relevant, and its results were reassuring. In an indirect way these results provided information about the items and confirmed their appropriateness for testing purposes.

Besides considering the reliability of the tests, the means and the facility values (calculated from the means) were also considered (Table 2). The analysis of the statistics helped to identify items which were particularly difficult or easy for the pilot population. By deleting problematic items it was possible to set the difficulty (facility value) level of the tests. After deletions, the item number of the SAQ tests was set at 24.

4.2 Forming comparable groups of participants

In order to compare performances on different test versions and draw conclusions on method effects, it was of crucial importance to set up four groups of participants, and to ensure that the groups were equivalent regarding their language proficiency.

Two-hundred and sixty-seven first-year students participated in the Main Study from Budapest Business School. On the basis of their TOEFL tests results (Phillips 1990), the participants were arranged into four groups of comparable language proficiency. As raw scores might not be regarded as interval data, the scores were converted by using the TOEFL Conversion Table. Then, the means and the standard deviations of the four groups were computed (M_A= .439.8, SD =63.6; M_B = 440.7, SD = 61.4; M_C =440.2, SD = 68.9, M_D = 441.8; SD = 63.9), and the means were compared using analysis of variance (ANOVA), which confirmed that there was no significant difference between the group means ($F_{3,234}$ = .168, p = .918), and thus, the groups were comparable. In addition, the participants' ability logits were

4. Statistical procedures

4.1. Validation of the main instrument

The validity of the tests was ensured in several ways. Besides qualitative methods, the statistical analysis of test results in two pilot studies (involving 185 and 202 students altogether) and correlating test scores with TOEFL[2] scores as a validated third measure (concurrent validity) helped ensure the validity and the reliability of the research instruments.

The data from the tests were processed using SPSS software (Version 11.0). Classical item analysis was carried out to calculate means, standard deviation, item test correlations, and reliability coefficients. The purpose of the analysis was to gain information about the tests as a whole, and to identify items for deletion or modification. Statistical results were expected to help validate the tests and decide which texts and items could be included in the final test booklets for the Main Study.

Poorly performing items were identified and modified after Pilot Study I. Besides modifying the wording of the questions, new items and new distracters were devised when necessary. After administering the tests in the second pilot stage, descriptive statistics and reliability analyses were carried out (Table 2). The results showed that the reliability of the tests increased considerably. Reliability for the SAQ test increased from $\alpha = .7399$ to $\alpha = .8398$, and for the MC test from $\alpha = .4327$ to $\alpha = .6631$ in the case of *Test 2*. In the case of *Test 1*, for the SAQ test it was $\alpha = .8149$, and for the MC test $\alpha = .7012$. The lower reliability coefficients of the MC tests were assumed to be related to the fewer number of items: the first version of both SAQ tests contained 30 items, whereas the MC tests contained 16 items only.

Table 2. The statistical analysis of the tests in Pilot Study II

Test	*M*	Facility value (%)	*SD*	Reliability (Alpha)	Adjusted reliability
Test 1 SAQ -E	21.98	73.2	4.8739	.8149	-
Test 1 MC -E	8.25	51.5	3.1057	.7012	.8148
Test 2 SAQ -E	18.87	62.8	5.6149	.8398	-
Test 2 MC -E	8.45	60.3	2.8559	.6631	.8082

Source: own creation

As reliability increases as items are added (Henning 1987, Csapó 1993), it was assumed that the reliability of the MC tests would increase if the number of items in the tests were increased to a specified length. The Spearman-Brown Proph-

[2] Testing English as a Foreign Language - the most widely accepted English language test developed by ETS (Educational Testing Service) US.

- How does the use of the native language in task rubrics, test items, and task completion influence reading comprehension performance in a short-answer questions test and in a multiple choice test?
- How do short-answer questions and multiple choice items as task types influence reading comprehension test scores when the task and the expected response are in English as the target language and in Hungarian as the native language?

3. Research method

The study compared the participants' performance on two reading comprehension tests including either short-answer questions (SAQ) or multiple choice items (MC): one in English as the target language, and one in Hungarian as the native language. Two sets of texts were selected, and four different tests were developed for each text: a short-answer questions test with rubrics and questions in English and in Hungarian, and a multiple choice test with rubrics and four options in English and in Hungarian. Thus, there were two sets of two texts and eight reading tests altogether.

Four groups of minimum fifty students each were involved in the research. Each group completed two tests. The participants in the same group worked with the same task type in the two tests, with language as the changing variable. Table 1 shows the groups and which versions of the tests they completed.

Table 1. The research matrix

	Group A	Group B	Group C	Group D
Test 1	SAQ test in English	MC test in English	SAQ test in Hungarian	MC test in Hungarian
Test 2	SAQ test in Hungarian	MC test in Hungarian	SAQ test in English	MC test in English

Source: own creation

Statistical analyses were employed at three different stages of the research:

1. in the validation procedure of the main instruments,
2. in forming comparable groups of participants,
3. in analysing data from test results and from questionnaires.

Test method facets are a set of factors that specify the actual method of the assessment procedure. They cover the following categories: testing environment, test rubric, input, expected response, and the relationship between input and response (Bachman 1991). Test method facets can be carefully designed and controlled in order to minimize their distorting effects on an individual's test results. As the present study aims to investigate the effects of test method facets on reading comprehension performance, the crucial facets are the facets of the reading text and the input. The input includes the task and the use of L1 versus L2 in the input (and the expected response).

Several comparative studies have been conducted to investigate the effects of task types and native language use on reading comprehension. In a recent study Liu (2009) compares three task types and claims that gap-filling tasks have a significantly negative effect on test-takers' reading comprehension performance. Other researchers are more cautious in their conclusions. Shohamy's (1984), Wolf's (1993), and Gordon and Hanauer's (1995) studies are particularly remarkable because besides comparing short-answer questions and multiple choice tests, they also examined the effects of native language use. Based on their findings it is clear that the items or questions in the task provide additional information for the reader that may help comprehension. The amount and the quality of this information may substantially differ in the case of different task types. Native language use in the items and the expected response resulted in higher performance in each of these studies. However, due to weaknesses in research methodology the authors could not draw any general conclusions. It is still not explained whether improvement in performance in the studies was due to better understanding the questions, or to using the additional information in the questions to better understand the text itself.

It is worth noting that, although many theoretical works discuss the criteria for validation and reliability in detail (e.g. Bachman 1991, Bárdos 2002, McNamara 2000, Popham 1990), researchers rarely provide any information about the validation methods of the research instruments in their empirical studies.

2. Research questions

The broad research area of the present study is the investigation of how two testing variables – task type and the language of task and task completion – affect reading comprehension performance. The formulation of the exact research questions was based on the taxonomies in the literature (Alderson 2000, Urquhart-Weir 1998) as well as the findings of a teacher's questionnaire (Loch 2007, 2009a) and the statistical results of a Pilot Study including 185 participants. Two task types, short-answer questions (SAQ) and multiple choice items (MC), were selected for comparison in the Main Study. Thus, the main research questions focussing on the mutual relationship of task type and language use, were formulated as follows:

How do test methods affect reading comprehension test performance?

Agnes Loch[1]

The paper describes how statistics were employed in language testing research to explore the effect of two test method variables of foreign language reading comprehension tests on test takers' reading comprehension performance. Statistical procedures were applied at three stages of the research: in the validation of the main research instruments, i.e. the reading tests (1), in grouping the participants into four comparable groups (2), and in analysing the participants' test performances on two reading comprehension tests (3).

Statistics and qualitative data analyses show that task type and native language use as test method variables, rarely have a statistically significant affect separately, but may rather exert a joint effect on performance.

Keywords: language testing, test method variables, test method effect

1. Introduction

The research explores the effect of two test method variables of foreign language reading comprehension tests - task type and native language use - on test takers' reading comprehension performance. The aim of the investigation is either to support or to reject the hypothesis that neither task type nor native (L1) or target language (L2) use influences reading comprehension performance significantly if the tasks target the same construct.

As communicative competences cannot be measured in any other way than by observing an individual's language performance, which is supposed to reflect the underlying competences, it is essential to consider all the possible factors that, besides actual reading comprehension ability, may influence performance and test results. Bachman (1991) sets up three categories to classify these contaminating factors: personal attributes (e.g. age, gender, occupation), test method facets (e.g. task type, dictionary use), and random factors (e.g. weather conditions, the test taker's physical or emotional state). Personal attributes and random factors are beyond the control of the examiner whereas method facets can be manipulated to make the assessment procedure and the results valid and reliable.

[1] Loch, Agnes, PhD, senior lecturer, Budapest Business School, Budapest, Hungary

syllabus accordingly. We should consider what and how we teach in order to make it easier for the students to acquire the necessary knowledge. We should transfer 'usable' knowledge: one that makes it easier to learn the material of the special courses of their own profession.

References

DF1, Valószínűségszámítás Tantárgyi program (MA-004) 2006: Dunaújvárosi Főiskola Matematika Intézet

DF2, Matematikai statisztika Tantárgyi program (TE-005) 2006: Dunaújvárosi Főiskola Matematika Intézet

DF3, Matematikai módszerek Tantárgyi program (TE-007) 2006: Dunaújvárosi Főiskola Matematika Intézet;

DF4, Műszaki matematika Tantárgyi program (MA-009) 2008: Dunaújvárosi Főiskola Matematika Intézet.

DF5, Műszaki matematika Tantárgyi program (MA-009) 2009: Dunaújvárosi Főiskola Matematika Intézet.

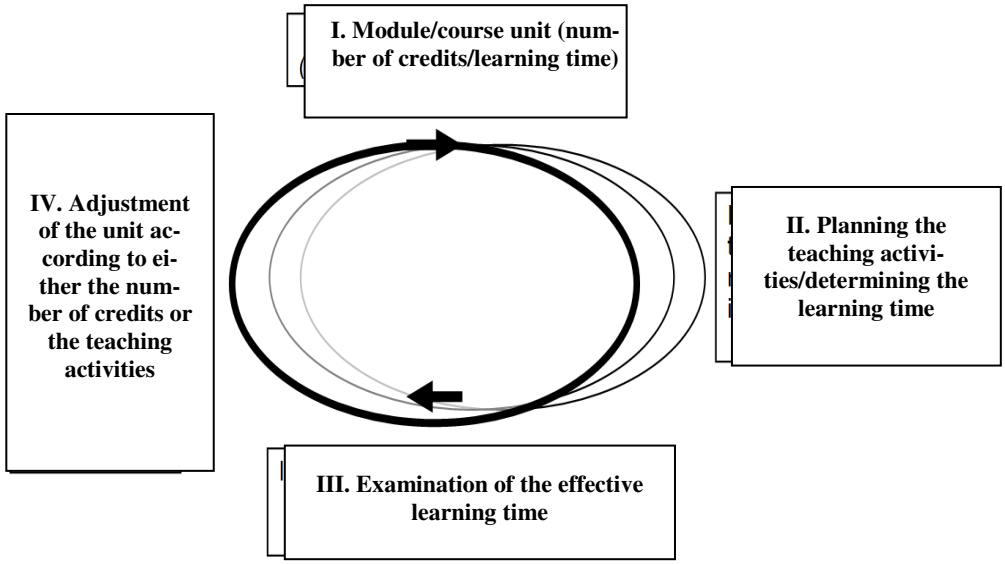
Kadocsa, L. 2004: A kreditalapú moduláris rendszere, mint a curriculumfejlesztés új minősége III. Felsőoktatási Kreditfórum, Dunaújváros 2004 jún. 3-4.

Radnóti, K. 2010: magánközlemények

must be acquired till the end of the semester. With the questionnaire, the students could put it into words more accurately with which part of the course material they have more problems, and to what would we need to devote more time.

We cannot neglect to determine the students' workload. In the future, this should be taken into account and should be reviewed constantly. A circular investigation should be started, which I represent in the following figure.

Figure 5. Determining the load and task of the students



Source: Kadocsa (2004)

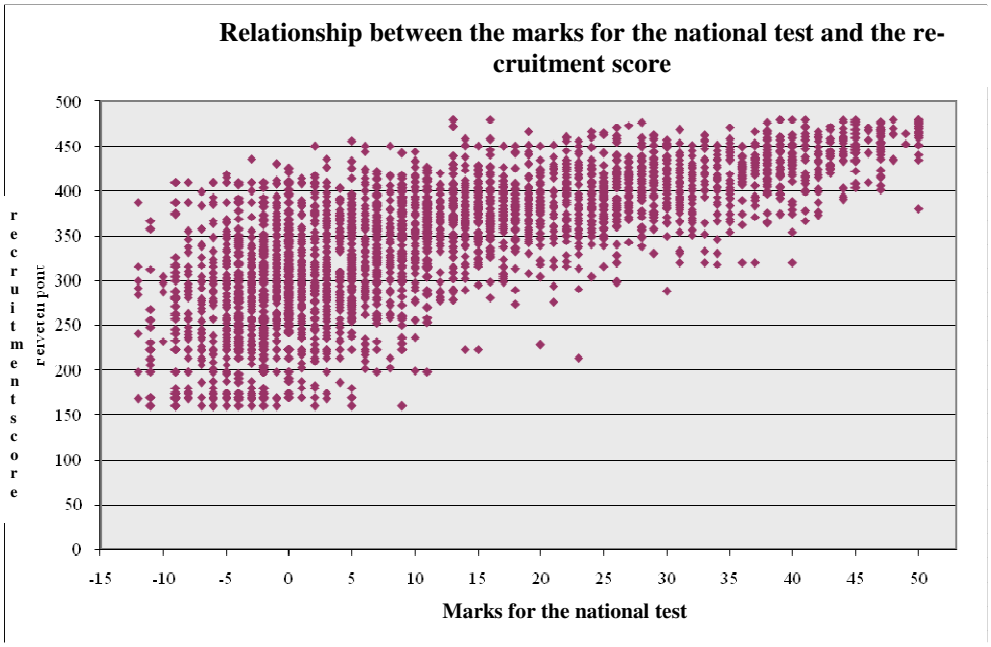
With this method, we receive a constant feedback from the students, by means of which we can redistribute the course material so as to make it more achievable for the students.

The results of the national test surveying the students should be assessed not only by us, instructors. The experts and politicians, who develop the recruitment system and determine the recruitment scores, should also draw the conclusions. Unfortunately, in many cases we should draw the conclusion that the scores achieved by the students do not reflect their real knowledge.

5. Summary

We should continue the trend that started in the country and at our College. We should take into account the students' knowledge and learning skills, and form the

Figure 4. The summarized data of the national test



Source: Radnóti (2010)

Unfortunately, now we know what we have only suspected: that the knowledge of students coming from high school is less and less. My hope is that, due to this test, we become more aware of the knowledge of the students, and by using our experience we can make the syllabus and our teaching methods better.

In the Engineering Mathematics course, the practice lessons and labs allow us to avoid the frontal – instructor centered – teaching methods, and to bring forward practice oriented – student centered, knowledge-based – methods supporting individual work.

In the future, we intend to strengthen the changes that we have already started, with the involvement of students. Within the framework of quality assurance, we already surveyed the students at the end of each semester. We asked for their opinion of the course and the instructors. We would like to make a questionnaire in which we can measure also the success of the course. Among the types of questions, there would be ones aimed at the required goals of the training, at gaining skills, and at the time necessary for acquiring these skills.

With this questionnaire, we would get better feedback on how effectively we teach. We could make it more effective with reorganizing the quizzes. Quizzes and tests provide information on the knowledge of the students, but there are skills that

fellow instructors recorded the results in a centrally prepared Excel sheet, and sent it back for process. Peer comments on the results of a centrally prepared and recorded in an Excel table, they were sent back for processing.

The test consisted of 12 multiple-choice questions, two exercises about formulas and three computational problems.

A very high percentage of students dealt only with the multiple-choice questions. They did not touch the two exercises requiring the knowledge of formulas, or used only basic formulas and did not care that they used the formulas with data different from those given in the exercises.

The third part contained the computational problems. Only ten percent of the students dealt with the last part. Even among those who made an attempt, only a few got maximum marks for them, many listed only the data gathered from the text, and made no real work.

In general, we could find numerous types of errors in the solutions.

In Figure 4, a diagram containing more than 3000 data is shown. Note how much the marks of this test and the students' recruitment scores are unrelated. I made a correlation calculation from the data where it was 0,723. One can see from the correlation coefficient that there is a correlation between the results of the qualifying examination and test but it isn't a strong connection.From the fact that someone had a high recruitment score does not follow that they got good marks on this test. Unfortunately, it seems that, among those with recruitment score between 300 and 400, we can find all the kinds of marks from the lowest to the highest. A question is raising that the process of the marticulation or perhaps the measuring isn't good. The Cronbach alpha of the measuring was calculated by the authors (Katalin Radnóti) which was 0,88 and this is a real good value. This confirms that hypotesis that the process of matriculation have to revise.

practice lessons and the laboratory exercises. We started the practice lessons with probability theory as before, but from the 9th week we continued with linear algebra, and mathematical statistics was transferred to the labs. Due to the change, we did not need computers on the practice lessons.

We decreased also the number of quizzes. There were three quizzes in practice lessons (on weeks 4, 9, and 14), and two quizzes in labs (on weeks 7 and 14); each for 20 marks. Hopefully students can achieve better results after this step (DF5 2009).

3.1. The future

Unfortunately, the lack of knowledge brought from high schools should be taken into consideration. It is therefore necessary to assess the knowledge of first-year students. This is necessary not only because of the course "Technical Mathematics", but also for the prerequisite courses Calculus I and Calculus II.

3.1.1. National survey

Due to the initiative of the Budapest University of Technology, this year was the first one in which a standardized assessment test has been written on a national level.

The objectives of the 2009 test:

- To examine the knowledge of the students entering higher education, and determine if it satisfies the requirements of their chosen degree programs.
- To investigate whether the recruitment score provides sufficient information on the student's knowledge?
- To verify our former conclusions, and support them with further empirical data.[2]

Execution of the examination:

The students wrote a sixty minute test during registration week; that is, at a time when the higher education institute had not yet "interfered" with their training. We at the College of Dunaújváros could have the test been written on September, 3rd with 500 first year students. The thing that the people preparing the test had in mind during their work was to examine the knowledge of the students important for higher education and necessary for the students' successful advancement. The test mapped especially the knowledge learnt in and brought from high school.

The questions were chosen centrally, and a solution with detailed instructions for markers was also enclosed to make marking as uniform as possible. Each institution itself organized the test and its marking, following the central instructions. Our

[2] http://members.iif.hu/rad8012/index_elemei/kriterium.htm

2007/2008 not only B.Sc. students could take it, but all others who were not able to fulfill the course requirements for at least one of the predecessor courses (Probability Theory, Mathematical Methods and Mathematical Statistics). The following rule was introduced: Those who could not pass in the predecessor courses may pass by passing in the new course. In this way, the number of students taking the course increased dramatically.

Despite the large increase in the number of students, the improvement in the quality of the grades is clearly visible. The number of students with grade 'failed' decreased, the number of the other grades increased. This increased number thanked to two things – on the one part the semester repeaters are increase the number – the another part is that those students who are in the conventional training should take up this subject instead of the older subjects which they don't completed.

Taking into account the fluctuation of the number of students, we can observe a slight decline in the number of students who passed, but after reaching a balance state, we can immediately see an improvement. According to the structure of the subject they could easier completed the curricular. In the last two semesters again there has been a decline. From what I can see the cause of it is that the students are "accommodate" to the new requirements and as they see to complete the subject easier than expend less time to learn it. Analyzing the dispersion of the marks we can see that the highest was 0,26 in the first semester, the lowest dispersion we got from the II. Semester of 2007/2008 which was 0,15. In the last semester which I investigated (I. semester in 2008/2009) I counted 0,22. This value gave exactly the average of dispersions. The data for the last semester are highlighted in the graph. We can notice that it is close to the mean values of the number of grades in the previous semesters. Fortunately, even though minimally, but the ratio of students getting 'failed' has decreased. The number of students getting 'satisfactory' has more than doubled, but, unfortunately, there has been a significant decrease in the number of better grades.

Among the causes of this decline, a major problem is that the students bring less and less knowledge from high schools. There are more and more students in higher education, and thus their level of knowledge varies on a large scale. We can do nothing else but get adapted to this new situation and shape and re-form the course syllabus so as to be able to transfer the required knowledge to the students. We have to take into consideration that some parts are omissible from the curricular or not and/or maybe we teach it according to the base of the actual theme. The students know less but this knowledge is certain.

4. The attempts made to improve performance

In the first semester of the academic year 2009/2010 we made a try with a new syllabus. We did not reduce the course material, but re-distributed the topics among the

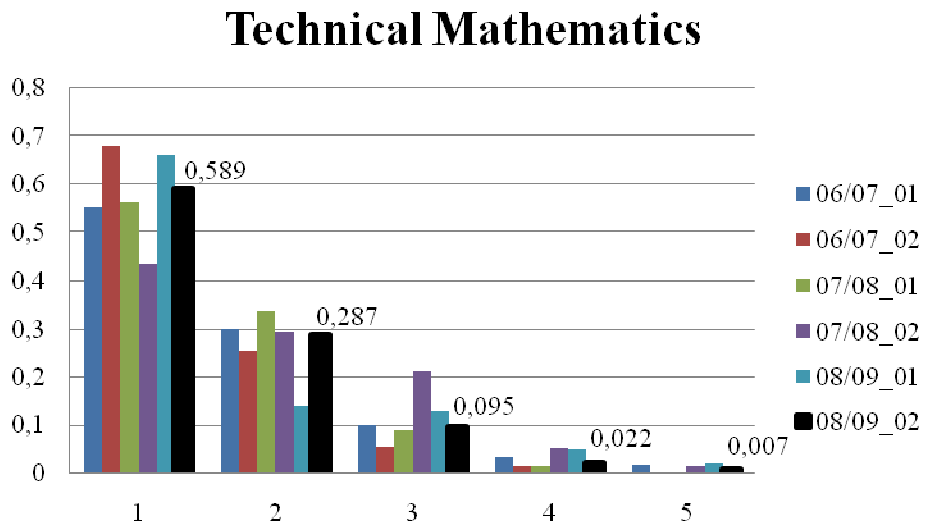
Table 5. Summary of the five completed semesters

	failed (1)	satisfac-tory (2)	mean (3)	good (4)	excellent (5)	total number of students
Semester 1 2006/07	0,55	0,3	0,1	0,03333	0,016667	60
Semester 2 2006/07	0,68	0,2533	0,0533	0,0133	0	75
Semester 1 2007/08	0,561	0,338	0,0878	0,0135	0	148
Semester 2 2007/08	0,433	0,293	0,21	0,051	0,013	157
Semester 1 2008/09	0,66	0,139	0,129	0,047	0,019	108
Semester 2 2008/09	0,587	0,287	0,095	0,022	0,007	139

Source: own creation

I graphed the data in the following figure to make it easier to compare them to one another.

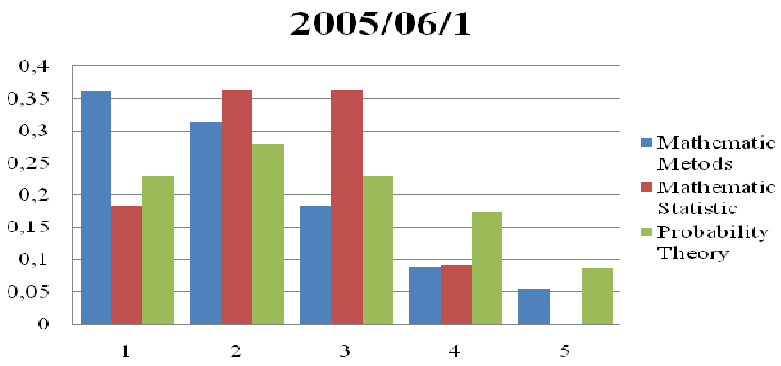
Figure 3. The graphs for the last five semesters



Source: own creation

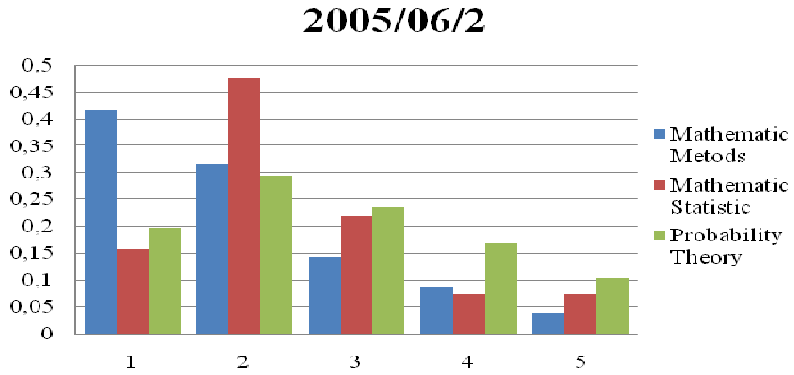
The 'total number of students' column shows that in the first academic year the course was taken only by a small number of students. The reason for this is that in that academic year only B.Sc. students were allowed to take the course. In year

Figure 1. The graphs for Semester 1 of year 2005/2006



Source: own creation

Figure 2. The graphs for Semester 2 of 2005/2006



Source: own creation

3.1 The results of Technical Mathematics

The course requirement is fairly complex, and hence the fulfillment of the course depends on various factors. In the next section I compare the grades of the students obtained since the introduction of the course.

For the course "Technical Mathematics" the five completed semesters give the base for comparison with the predecessor courses:

Table 1. Grade credits of students in Semester 1 of 2005/2006

Semester 1 of year 2005/2006	failed (1)	satisfactory (2)	mean (3)	good (4)	excellent (5)	total number of students
Mathematical Methods	159	138	80	39	24	440
Mathematical Statistics	6	12	12	3	0	33
Probability Theory	126	153	126	95	48	548

Source: own creation

Table 2. Grade credits of students Semester 2 of 2005/2006

Semester 2 of year 2005/2006	failed (1)	satisfactory (2)	mean (3)	good (4)	excellent (5)	total number of students
Mathematical Methods	217	165	74	45	20	521
Mathematical Statistics	13	39	18	6	6	82
Probability Theory	115	172	138	98	60	583

Source: own creation

One can clearly see that the courses were attended by different numbers of students. Thus the courses could be compared only with some difficulties. Therefore, we show the next graphs with the same data, but now in terms of percentage (with respect to the total number of students) for each course.

2. The introduction of the course Technical Mathematics

The course Technical Mathematics was introduced in the academic year 2006/2007. This course is one semester long and it contains those subjects which are thought separately: Probability Theory, Mathematical Methods, Mathematical Statistics. The course is composed of lectures of two hours or two-times-two hours per week (15x2 hour lectures on probability theory and statistics, 4x2 hour lectures on linear algebra), furthermore two-hour practice sessions per week (8x2 hour practice sessions on probability theory, 7x2 hour practice sessions on mathematical statistics, and 12x2 hour computer lab sessions (DF1 2006, DF2 2006, DF3 2006,).

2.1 The aim of the module "Technical Mathematics"

The main aim of the module is that the students are cognizing and acquiring the technical and practical bases of the applied mathematics subjects (linear algebra, probability theory, mathematical statistics) on such level which is indispensable for the investigation of the special subjects.

Through the using of MATLAB conceived as an aim to cognize an up-to-date mathematical software package. By the help and use of it many - related to the fields of linear algebra, analysis, and numerical methods- mathematical exercises should solve.

I won't detail the weekly exchanging and scheduling of the subject cause in 2008 I wrote it down in my gossip which issued in the publication of the Weeks of Science Conference at the College of Dunaújváros (DF4 2008, DF5 2009).

3. Experience

The fulfillment of the course depends on several factors, the course requirement is rather complex. In the next part, I introduce the average grades in the semesters before and after the introduction of the new course structure. I consider the academic year right before the introduction of the new course as the base for the comparisons for the old courses. In the following tables I summarize the distribution of the achieved grades for both semesters from that year.

The preliminaries and introduction of the course "Technical Mathematics", and the examination of the students' workload

Krisztina Lászlóné Kenyeres[1]

The course Technical Mathematics was introduced during the academic year of 2006/2007 at the College of Dunaújváros. I compare the course with those courses taught at a time period preceding the introduction of the new course. I examine what kind of changes happened in the average grades by contracting several courses and forcing the students to accomplish a similar course curriculum within a shorter time period.

With statistics, I illustrate the success of the course. I examine the syllabus of the course in terms of its achievability, and investigate and analyze the students' workload. With the help of the analysis of the results of the students – the 'learning outcomes' –, I search for solutions and compensations to provide better achievability of the course. At the time of this lecture I can present only a summary of the results of the national survey carried out during the Fall term, I supplement it later with the results themselves, which are connected to the topic of my future Ph.D. thesis.

Using the viewpoint of ECTS, I examine the structure of the syllabus, and propose innovations in methodology.

Keywords: Technical Mathematics, Learning Statistics, examination of students' workload, national survey

1. Introduction

I examine the kind of changes occurred due to the merge of several courses into one, and due to the fact that the students had to complete this course during a shorter time period.

I examine the syllabus of the course in terms of its achievability, and investigate and analyze the students' workload.

[1] College of Dunaújváros.

Nagy, F. 2009: *Büntető Jog Különös Rész*, HVG – Orac, Budapest.

Steiger, J. 2008: A jogászok, a közgazdászok és a számok; ELTE Állam- és Jogtudományi Kar – Központi Statisztikai Hivatal: A statisztika és a közigazgatás elkötelezettje, ünnepi kötet a 60 éves Katona Tamás tiszteletére, Budapest.

Belovics, E. – Molnár, G. – Sinku, P. 2007: *Büntetőjog Különös Rész*, HVG-Orac, Budapest.
Berki, E. 2008: *Munkaügyi kapcsolatok közszolgálatban*, Pólay Elemér Alapítvány, Szeged.
Blaskó, B. – Miklós, I. – Schubauer, L. 2006: *Büntetőjog Különös Rész II.,* Rejtjel Kiadó, Budapest.
Blaskó, B. – Miklós, I. – Pallagi, A. – Schubauer, L. – Zentai, Á. 2008: *Büntetőjog Különös Rész I.*, Rejtjel Kiadó, Budapest.
Complex DVD Jogtár Plusz.
Erdősy, E. – Földvári, J. – Tóth, M. 2007: *Magyar Büntetőjog Különös Rész,* Osiris kiadó, Budapest.
Farkas, Cs. – Papp, T. – Strihó, K.: *Társasági jog III./1*, A gazdasági társaságok közös szabályai, SZTE Állam- és Jogtudományi Kar, 2007. pp. 7-46.
Fehér, L. – Horváth, T. – Lévay, M. 2001: *Magyar Büntetőjog Különös Rész,* Bíbor Kiadó, Miskolc.
Földvári, J. 2006: Magyar Büntetőjog Általános Rész, Osiris kiadó, Budapest.
Görgényi, I. – Gula, J. – Horváth, T. – Jacsó, J. – Lévay, M. – Sántha, F. – Váradi, E. 2007: *Magyar Büntetőjog Általános Rész,* CompLex kiadó, Budapest.
Gyémánt, R. – Katona, T. – Szondi, I. 2005: *Demográfia*, Pólay Elemér Alapítvány, Szeged.
Hegedűs, A. 2006: *Családi jog*, Kis-ker Bt., Szeged.
http://portal.ksh.hu/pls/ksh/docs/hun/xstadat/xstadat_eves/tabl2_08_02i.html
http://www.birosag.hu
http://www.birosag.hu/engine.aspx?page=Birosag_Statisztikak
http://www.hvg.hu
http://www.ksh.hu
http://www.magyarorszag.hu
http://www.otthonbiztonsag.hu/ob/index.php/mbiztonsagrol/52-vagyelbuncselek
http://www.police.hu/elemzesek/bunuldozes/bunterkep.html?pagenum=3
http://www.police.hu/statisztika
http://www.securitron.hu/bunozes_2004_ig.htm
http://www.szmm.gov.hu/
http://www.ugyeszseg.hu
Hungarian Central Statistical Office: *Hungarian Statistical Yearbook 2007*, and CD Appendix.
Juhászné Zvolenszki, A. 2009: *Polgári eljárásjogi ismeretek*, Pólay Elemér Alapítvány, Szeged.
Kőrös, A. 2007: *A családjog kézikönyve*, HVG-Orac, Budapest.
KSH, Szociális és Munkaügyi Minisztérium: *Nők és Férfiak Magyarországon 2007*.
Nagy, F. 2005: *Az európai uniós országok fogva tartotti rátájának alakulásáról*, Börtönügyi szemle, 2005/3.
Nagy, F. 2008: *Büntető Jog Általános Rész*, HVG – Orac, Budapest.

pertinent to misdemeanor statistics if we talk about criminal statistics, but we will not touch upon further questions of this problem in this current article.)

Summing up, it shall be pointed out, if we properly apply the methods of criminal statistics and then our efforts amount to a much clearer picture on the status of crime; also -- more importantly -- we can enable students to integrate their knowledge by having an objective system at their disposal to refer to. This system should be based on accurate measurements, compiled in a way meeting strict requirements, in order to make possible the understanding of criminal legal matters e.g. why certain crimes will be classified as life threatening assault and battery and why not as attempted murder.

Criminal law without criminal statistics is blind and criminal statistics without criminal law is pointless. - said Prof. Hans-Heinrich Jeschek. We could agree with his assertion on criminology, applying it to criminal statistics. The two fields cannot exist without each other, and their close relation is important for the different methods of criminal statistics help the development of the complex legal thinking needed for those interested in criminal law to become the best professionals possible. Criminal statistics has several methods that cannot be fully applied without sufficient knowledge on the basics of criminal law, on the other hand, however, we have to consider those areas as well that are perfectly compatible with the dogmatic system of domestic criminal law.

These areas and methods shall be inserted in the current materials of legal education with intent of an extended application, and they shall also be applied in the course of seminars and lectures as well. The major fora shall also be better promoted where those interested can find statistical data and broaden their knowledge on certain aspects of e.g. criminal law. As a last point of interest, it is to be emphasized as a pivotal advantage that this branch of statistics could put those interested in possession of not only domestic but also European and international information. In light of the all-reaching integration these days, this could only serve to the benefit of the students and might lead to several different advantages for them in the course of their professional life.

References

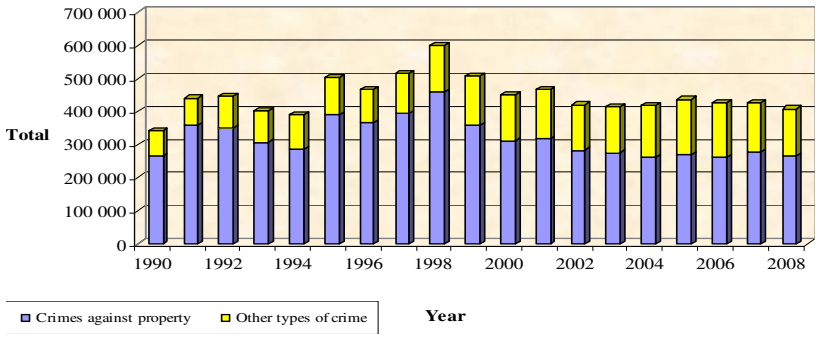
Balogh, Á. – Kőhalmi, L. 2007: *Büntetőjog I. Általános rész,* Dialóg Campus Kiadó, Budapest.

Balogh, Á. 2008: *Büntetőjog II. Különös Rész*, Dialóg Campus Kiadó, Budapest.

Bárd, K. – Gellér, B. – Ligeti, K. – Margitán, É. – Wiener, A. I. 2002: *Büntetőjog Általános Rész*, KJK Kerszöv, Budapest.

Belovics, E. – Békés, I. – Busch, B. – Molnár, G. – Sinku, P. – Tóth, M. 2006: *Büntetőjog Általános Rész*, HVG-Orac, Budapest.

Figure 15. Crimes against property within publicly prosecutes patent crimes compared to the total of other crimes 1990-2008



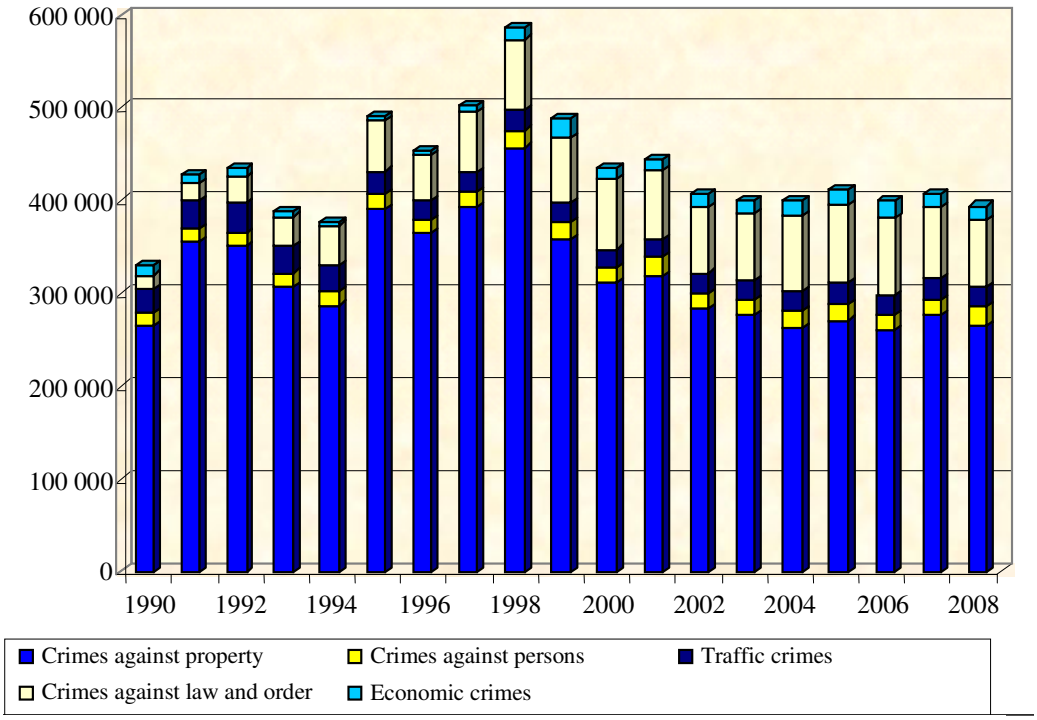
Source: http://portal.ksh.hu/pls/ksh/docs/hun/xstadat/xstadat_eves/tabl2_08_02i.html (in Hungarian)

Compared to the number of publicly prosecuted patent crimes, the rate of crimes committed against property was the highest (78.39%) in 1992. However, regarding the rates, with respect to the total of crimes committed (600.621) and also within this, the crimes against property (457.188) those were the highest in 1998. After 1998, a significant decline can be experienced. Analysis of these numbers might raise several questions in the students, regarding the emphases of the curriculum, and the importance of certain crimes. If these data are examined and used correctly by an expert in criminal statistics, it leads to objective insight, into the proportions of patent crimes.

Further analyzing the diagram above, we can only indirectly conclude, that in terms of crimes against property, the significant change in the statutory limits of value (increasing from 5.000 HUF first to 10.000 HUF, then to 20.000 HUF) had some effect on the numbers. (NB the peak points overlap with the dates of the changes in the statutory limits of value.)

This means, that crimes against property, not exceeding these limits of value, will not be governed by criminal law anymore, but by the law of misdemeanors. This of course leads to their incorporation in misdemeanor statistics, instead of criminal statistics. Pursuant to this, there appear to be fewer crimes included in criminal statistics, which the government comments, as a success in crime prevention. However, following from my arguments, this was clearly not the case, in light of the right interpretation of the numbers. It would be worth examining, how the analysis of criminal statistics disregarding misdemeanors, appears in criminal politics. [E.g. if we consider, that criminal statistics does not confirm the emphasis put on homicide (including murder and manslaughter) in the course of criminal law education.] (NB Following from the above, it is quintessential to deal with data

Figure 14. Patent (discovered) crimes (publicly prosecuted) 1990-2008



Source: http://portal.ksh.hu/pls/ksh/docs/hun/xstadat/xstadat_eves/tabl2_08_02i.html (in Hungarian)

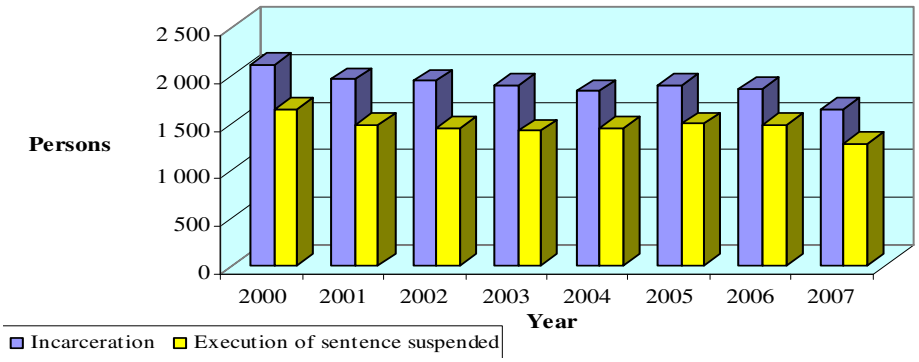
If we analyze the previous table, from a different aspect, we find that the number of crimes against property significantly exceeds the number of all other types of crime:

In terms of the „specific provisions" of the Criminal Code (part of the fifth semester curriculum) there are many more opportunities to apply criminal statistics. Law Enforcement agencies, courts and executive organs might conduct parallel research and surveys on the cases in their respective jurisdictions. Since we have no knowledge on the exact extent of crime, these analyses are of great value serving as primary guidelines in getting to know domestic law enforcement efforts. Why could this be important? It is, mostly, because the number of law books on „specific provisions" applying statistics in explanatory materials is even smaller than that of „general provisions" law books in the same respect (Nagy 2009, Balogh 2008, Belovics et al. 2007, Blaskó et al. 2008, Blaskó et al. 2006, Erdősy et al. 2007, Fehér et al. 2001).

While studying special provisions, students get to know different facta of crime and become able to analyze these in detail to the fullest extent. However, this means nothing more to them than a mere aggregate of statutory articles, paragraphs and subsections and only a few of them will be able to see the 'big picture': the system of these provisions and the different elements of those, for the first time. Statistics can help them and the others to tie the dots and provide practical aids for their understanding of the dry statutory language with data based on real life examples. How exactly would this be possible?

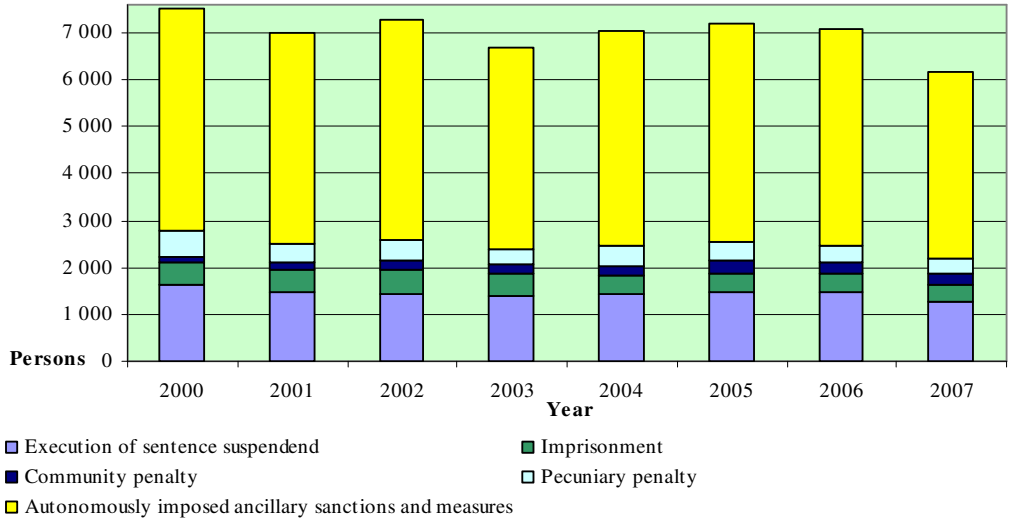
In Hungary, the material covering "special provisions" starts with the introduction to crimes against persons, which is the part of the material that shall be the most emphatic in the course of criminal law studies. However, if we step out of criminal law onto the field of criminal statistics and compare the number of crimes against persons to all crimes committed; we find that their share, compared to the whole, is virtually insignificant: it is closer in its rates to the rate of traffic crimes, which is a smaller and less emphatic part of criminal legal education. I would like to point out the detrimental effects of mass-media in this respect that leads to exaggerating occurrences of murder, bodily injury, kidnapping.

Figure 12. The ratio of suspended sentences of youth offenders to the aggregate number of imprisoners



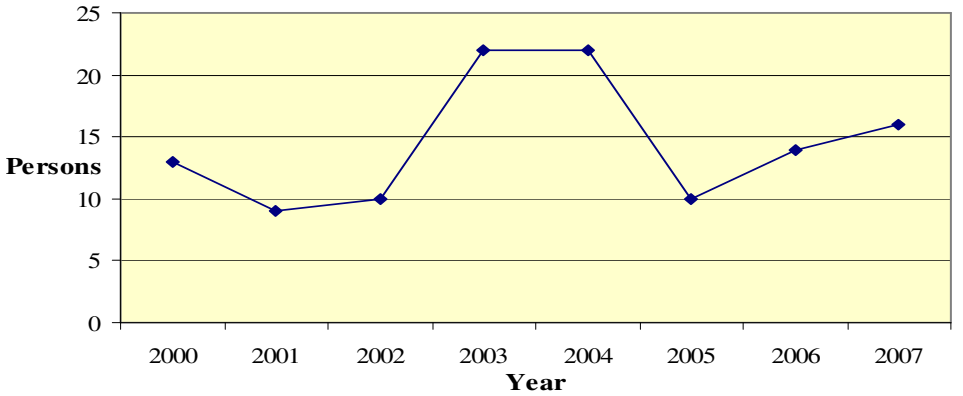
Source: Hungarian Statistical Yearbook, 2007; CD Appendix

Figure 13. Numbers of sanctions against youth offenders



Source: Hungarian Statistical Yearbook, 2007; CD Appendix

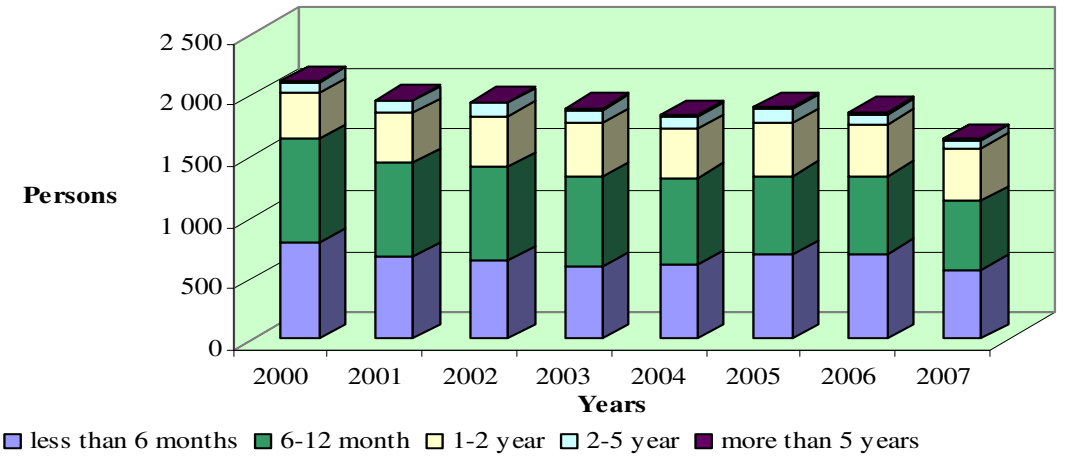
Figure 10. Change in the number of life of crimes punished by life imprisonment



Source: Hungarian Statistical Yearbook, 2007

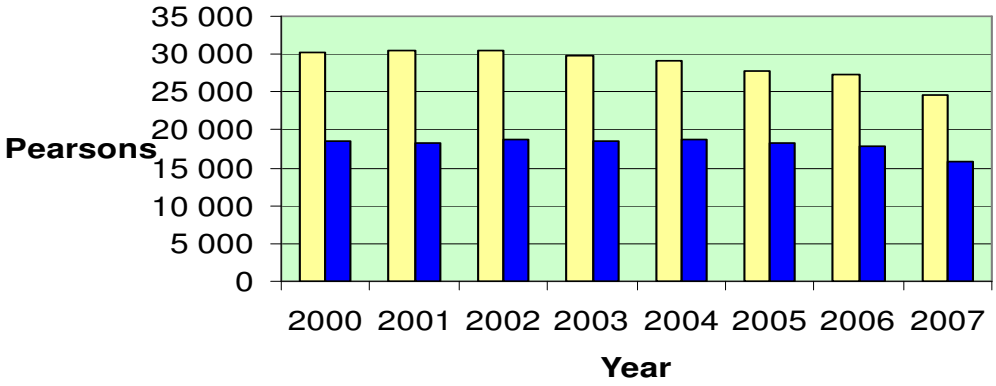
Statistical methods presented hereinbefore seem useful in representing the actual status of youth offenders as well.

Figure 11. Number of imprisonments against youth offenders



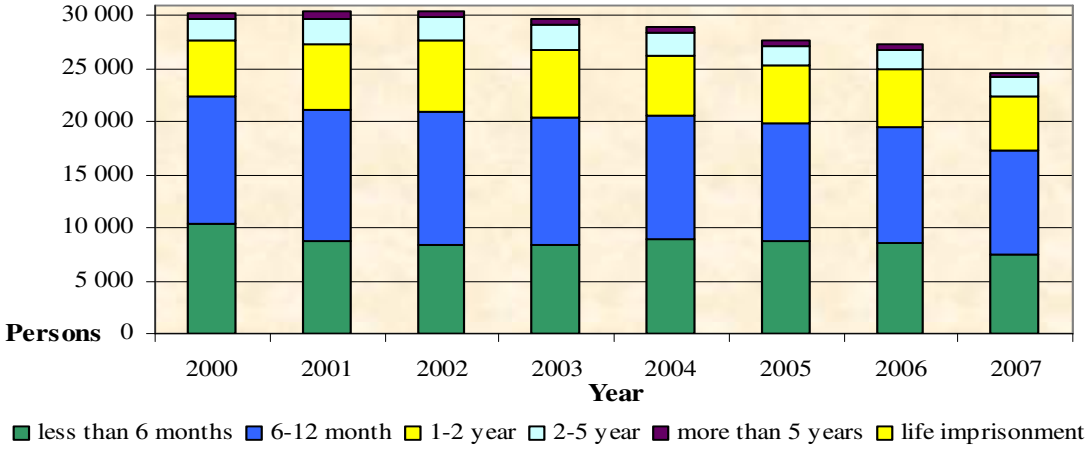
Source: Hungarian Statistical Yearbook, 2007; CD Appendix in carceration

Figure 8. Crimes punished by incarceration and the number of sentences suspended



Source: Hungarian Statistical Yearbook, 2007

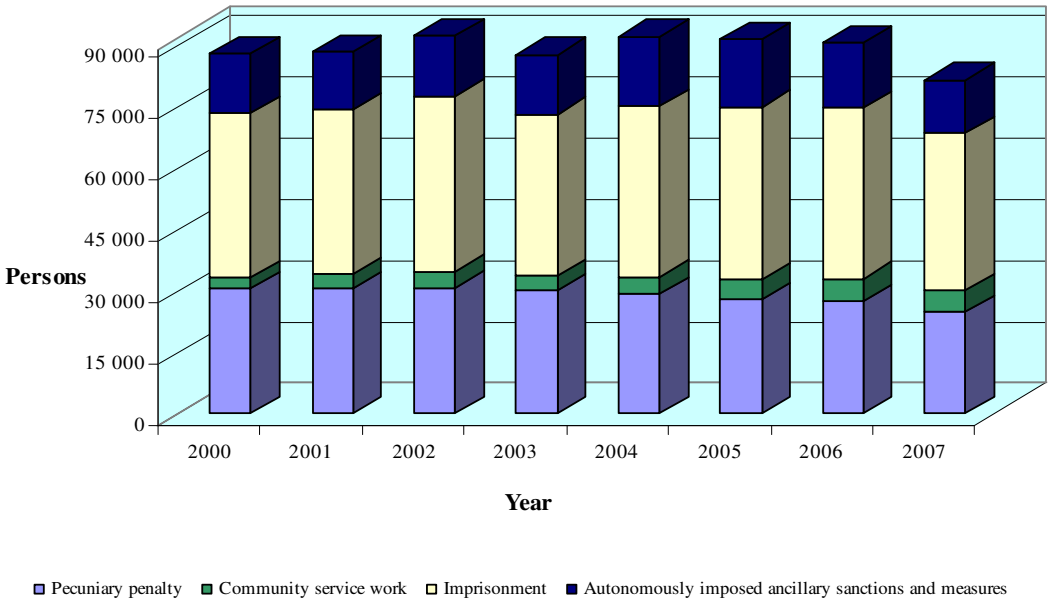
Figure 9. The number of persons sentence to incarceration



Source: Hungarian Statistical Yearbook, 2007

The number of perpetrators, who have been sentenced to life imprisonment, is significantly low. The following diagram shows the changes in the numbers from the year 2000.

Figure 7. Principal penalties 2000-2007



Source: Hungarian Statistical Yearbook, 2007

The rate of principal penalties and the proportion of the types of penalties have to be taken into account from several points of view. We can find herein useful information on how and why they change in Hungary, and this is also helpful in international comparison. However, based on international treaties, it is an obviously declared goal that the use of imprisonment shall only be exceptional; while examining the rates of incarceration the much-anticipated decline cannot be experienced (Nagy 2005). In light of this liberalization: (i) the number of those, sentenced to imprisonment (the execution of which was suspended), and (ii) the time of punishment, in case of punishments involving incarceration show the following tendencies:

(On an adjacent note, we also have to mention those institutes that conduct surveys of public opinion that provide feedback to law enforcement agencies as to their activities by asking collecting relevant data from the public.)

8. The Appearance of Criminal Statistics in Substantive Criminal Law

As for the importance of criminal statistics in legal education, in the following I would like to point out the role of statistics in legal education and its current status at my University. I will use comparison of teaching materials at different domestic universities as an explanatory tool for my reasoning, and will also refer to domestic criminal law when pointing out areas that lack statistical basis in education.

Students at the University of Szeged, Faculty of Law familiarize themselves with the basics of statistics right in the first semester and they can broaden this knowledge by applying the basics in optional courses of the curriculum throughout the remaining four years of their studies: they can choose to learn -- among others -- justice and criminal statistics as well. On the other hand, introduction to substantive criminal law is part of the third semester's curriculum, where students first learn about the „general provisions" of the domestic Criminal Code (criminal law). It is noteworthy at this point that a significantly low percentage of the currently used law books on „general provisions" apply statistical data as means to support statutory texts in spite of the fact that there are certain chapters that allow; moreover, call for the application of criminal statistics (Földvári 2006, Bárd et al. 2002, Balogh-Kőhalmi 2007, Belovics et al. 2006, Nagy 2008, Görgényi et al. 2007). One of such domains is e.g. „the Doctrine of Criminal Law Consequences", which -- among others -- gives a distinct overlook on the trends in the numbers of principal penalties domestically. (See: Diagram 1)

legislators, courts and professors when preparing laws, verdicts and or teaching strategies, or in every other aspect of day-to-day life.

The integration efforts of Hungary towards the EU were manifold both before and after 2004. Legal integration is one of many and has several branches, one of which is criminal law. Seeing through the integration in criminal law could be to a great extent assisted by examining and evaluating criminal statistics, but this field of statistics is not only important in this respect. The role of education is also to be stressed here: statistics (i) help students gain a more comprehensive outlook on the different fields, areas of law, (ii) help them achieve an integral practical knowledge-base as part of the University studies. That is why statistics must be taught in connection with main subjects in the curriculum, along with several other branches of social sciences, this way helping the students develop a complex legal way of thinking.

In the following, I wish to comprehensively present the place of criminal statistics in legal education, primarily with respect to substantive law. I will discuss the possibilities of its practical application -- while presenting the current status of criminal law in Hungary --, with particular focus on stressing the areas where statistics should be applied with increased willingness and frequency.

7. Short Outlook on the Current Situation

The Statistics Act, No. XLIV. of 1993 is safeguarding the principles referred to in the introduction and guaranteeing the accuracy and professionalism of the data collected. It contains provisions on the methods and goals of statistical data-collecting and on the organizations that are authorized to collect data for statistical purposes. As for criminal statistics, the 59/2007 (XII.23.) Order of Ministry of Justice and Law-Enforcement serves as a unified code for both law enforcement and prosecutorial statistics. Law enforcement agencies and the Prosecutor's Office are not just entitled but also legally obliged to continuously collect such data and compile criminal statistics periodically.

The appearance of the Internet and the introduction of different information systems and databases had a beneficial effect on all different areas of criminal statistics. Collecting and systematizing data became significantly quicker, and the results of the statistical analysis are only 'one click away' for those interested. The bodies mentioned above conducting statistical surveys shall forward their findings to the Statistical and Analytical Division of the Ministry of Justice and Law Enforcement's Criminal Policy Secretariat (Igazságügyi és Rendészeti Minisztérium, Büntetőpolitikai Szakállamtitkárság, Statisztikai és Elemző Osztály) who is responsible for publication of these data at crimestat.b-m.hu, also known as the Hungarian Criminal Information System, a website, which contains up-to-date domestic law enforcement information and is accessible to anyone.

because of divorces than new marriages are made. In 2008 73.800 people widowed and 40.100 got divorced. Since this process is stable, the consistency population based on family status changes considerably.

Since 1999 the number of single people has increased from 20.3 % to 31.7 % and the number of divorced people from 7,4% to 10.5%. Parallel with this the number of marriages has decreased from 61.2% to 46.5%. During the past few years it can be witnessed that among the population of the age of 15 and above married people are in minority. Since 2004 less than half of the adult population is married (KSH 2007).

Such an introduction was indispensable in order to be able to comprehend the importance of statistics, and the help it can provide in practice. Now, we would like to point out some implications relevant to the use of statistics in legal education, as a basis for practical applications, through the lens of criminal statistics.

6. Criminal Statistics in Legal Education. The Example of Criminal Law.

"The only statistics you can trust are those you falsified yourself"
(Sir Winston S. Churchill)

We have to admit: Churchill's slightly cynical motto is of perpetual truth. There might be some who nod understandingly hearing it; others might raise their voices anxiously finding its content offensive. However, it is an undisputable fact that the result of statistical analysis can always be approached from two angles: Truth of Falsehood.

It is of utmost importance that every research domain be examined, analyzed objectively, with the most accurate state-of-the-art methods possible and that the adequate conclusions be drawn from the research. This is true especially these days when -- in an extent never seen before -- chaos and pessimism characterize public opinion in economic and social terms.

Why exactly is it important to abide by this supreme moral-ethical law? If one does not abide by this principle then Churchill's words might come true as abusing data leads to chaos and pessimism detrimental not only to the individual but to the collective as well. On the other hand, however, this moral compass should be used as a mirror in order to provide answers for certain questions.

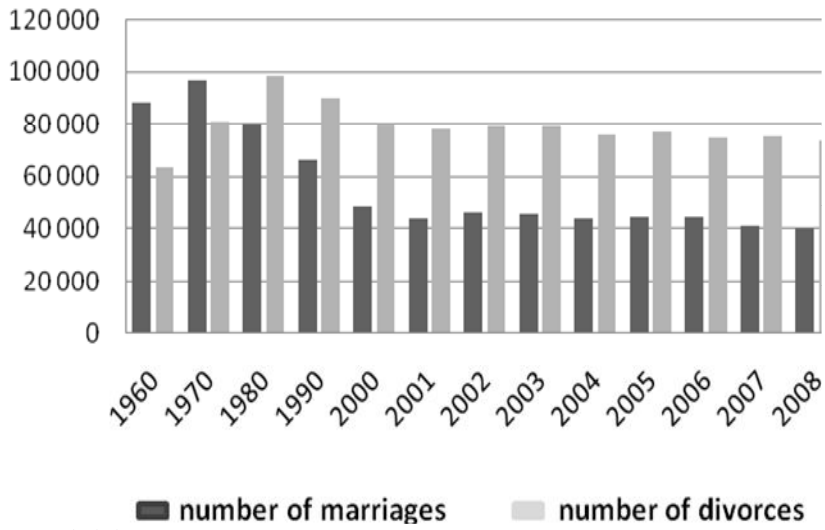
Modern day statistics are much appreciated and used in every aspect of research. Its methods shall not only provide accurate information for domestic users, but also for international users and statistical data has to comply with requirements set forth by the Eurostat and other international organizations. Given these requirements, data is collected, treated and processed from the public and private sector professionally and adjacent analyses, evaluations provide useful help to

Out of 100 labor contracts 92 indeterminate contracts were made in 2008 in Hungary. Out of the low number 8 determinate contracts 5 were made for men and 3 for women.

5. Family Law

The responsibility of the family is to secure the continuity of human race, giving birth and raising children as long as they take their places in the society. The most important legal foundation and basis of the family is the marriage regulated by the family law, which is a voluntary alliance of a man and a woman (Hegedűs 2006). Its importance can be expressed with one number. In 2008 with 40.000 more people got married than the previous year. Unfortunately we can say the same about the marriage property law. In 2008 25.000 marriages ended (Gyémánt et al. 2005).

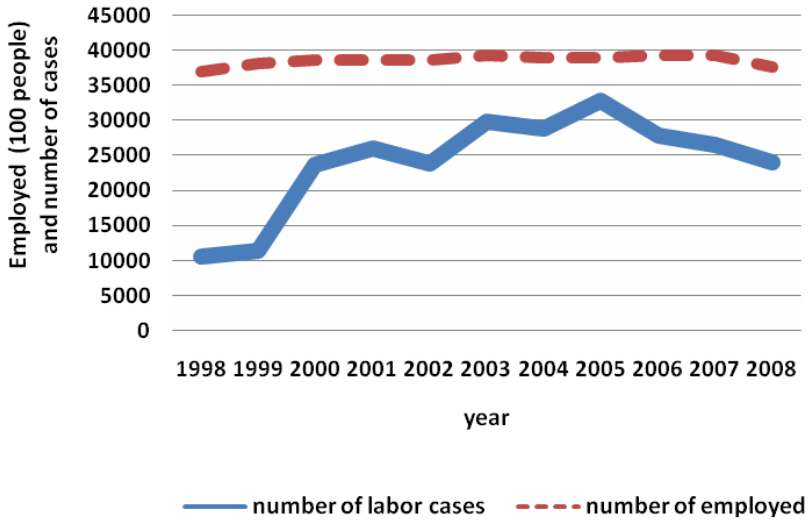
Figure 6. Marriages and divorces from 1960 through 2008



Source: www.ksh.hu

The diagram shows well the changes of the image of marriage, value judgment, and lifestyle of the society. The 44-46 thousand marriages made at the turn of the century do not reach the number of marriages made in the 70's that was 90-100 thousand. The number of marriages made in 2008 is below the previous years with 1.8 percent –apart from the I. World War- it was the lowest in the 130-year-old history of peoples' moments. The rare marriages and frequent divorces hard hit the scale of marriages (Kőrös 2007). Invariably more marriages end with becoming widower or

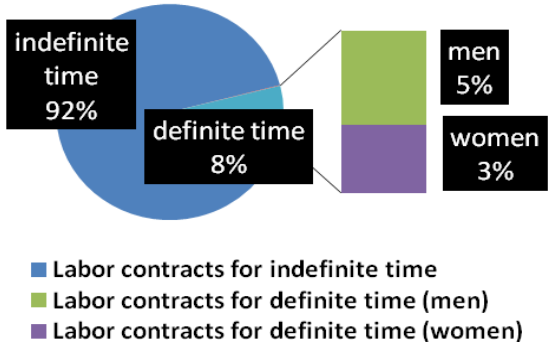
Figure 4. Labor cases and the number of employment (from 1998 through 2008)



Source: www.ksh.hu

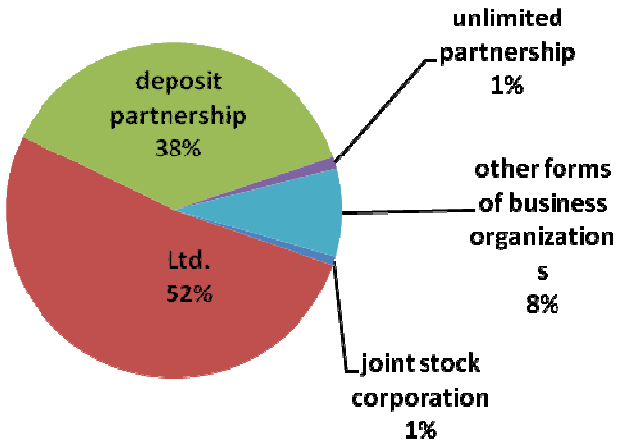
One may ask why did the number of labor cases increase between 1999 and 2005 while the number of employment did not change significantly. Law students have no time to be absorbed in all those fields of life which regulations and norms they study. In this case the greatest help is statistics with its numbers and diagrams that shows the most important characteristics of the social relations regulated by a given branch of law (in case of labor law for example the average wage, the number of unemployed people's cases).

Figure 5. Labor contracts based on duration in 2008



Source: Ministry of Labor

Figure 3. Number of business organizations in 2008



Source: www.ksh.hu

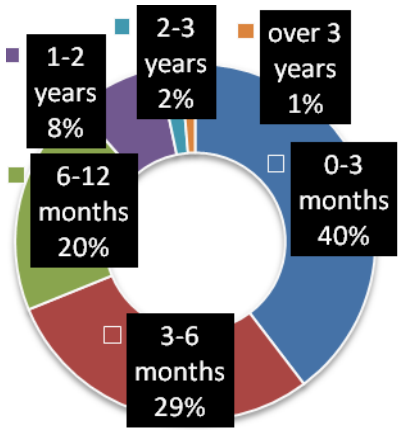
The number of joint stock corporations is relatively low beside the two ruling company forms, but it doesn't reduce the importance of it.

4. Labor Law

Law students will feel the importance of studying labor law on themselves when they have labor dispute with their employer. Students having their parents support do not really feel the usefulness of labor law. In case of facing the number of labor law cases they will study Labor Code with bigger interest (Berki 2008).

Based on the first quarter of 2009 labor force survey between the age of 15-74 4.167 million people appeared on the labor market. Among these 3.764 people were employed and 403.000 were unemployed. On national level the average net income was 121.400 HUF for blue-collar workers the average wage was 89.400 HUF and for intellectuals' the average wage was 154.200 HUF.

Figure 2. Duration of prosecution at local courts
(finished in 2008 civil cases)



Source: www.birosag.hu

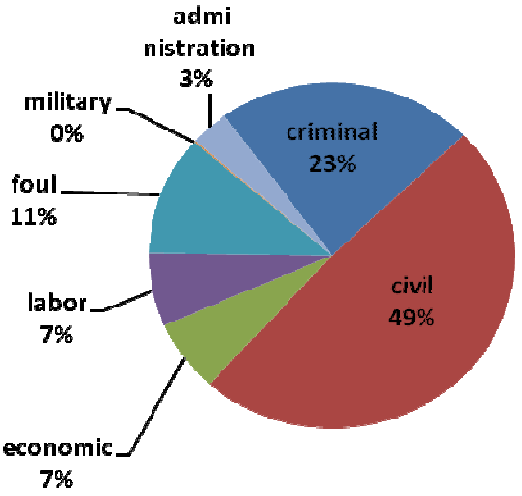
3. Corporate Law

The importance and unavoidability of norm of corporate law is clearly shown by the number of business organizations that were over 500.000 in 2008. The primary demographic data source of enterprises is based on the register of business organizations published and operated by the Central Statistics Institute. In 2008 the most popular form of enterprise was the limited liability company. Their number had exceeded 292.000. It is one and a half times larger than in 2000 (Farkas et al.). The low initial capital explains the popularity of limited liability company form. Among business organizations the second most popular form was the deposit partnership. In 2008 there were 211.000 deposit partnerships in Hungary exceeding year 2000 with 60 thousand.

the justice system via the toolbox of statistics is necessary before explaining how statistical data can improve the level of legal education.

To understand the system and function of Hungarian jurisdiction, statistics is inevitably important. People feel that like other European courts Hungarian courts are also overloaded and procedures are slow. Law students should look behind these statements to what extent the Hungarian courts are busy with cases and how long the court reaches the final judgment (Juhászné 2009). In 2008, local, county and regional courts have received 398.430 cases with the following ratio.

Figure 1. Cases arriving in 2008 to local, county and regional courts divided by cases



source: www.birosagok.hu

As it turns out, 89% of civil cases arriving to the local courts are finished in a year. The need that the cases be finished in a reasonable time is one of the basic human rights (right to a hearing within a reasonable time). To establish effective, fast, simple and cheap ways of deciding court cases has long been one of the most frequent aims of civil law codification.

providing easily and readily accessible communication channels, for professionals and lay people alike. Our lecture will attempt to yield an insight of the recent state and importance of criminal-statistics, taken into account the special circumstances of Hungary. We also would like to emphasize once more, how important is the statistical knowledge for the legal profession and legal student nowadays.

Keywords: Private Law, Criminal Statistics, Legal Education

1. Introduction

Our every-day world is characterized by economic and social difficulties. Hence, it is especially important that social sciences be able to provide a precise analysis of our circumstances. Statistics might prove as an efficient tool in this struggle for professors and students in every respect of legal education. Considering the methodology, analysis and the means system of statistics it has special relationship with social sciences, however, it often ignores the interest of university students who study law and are more prone to humanities rather than mathematics (Steiger 2008).

For decades, proficient knowledge of the fields of literature and history were the most important among admission criteria for law students. Those who could prove sufficient knowledge in these areas became eligible for admission as law students; those, who proved their proficiency in mathematics and history, chose economy as a career path. This policy resulted in loss of contact between economy and law in terms of educational methodology, and as a result of this trend law students had no useful mathematical knowledge. The system of structured studies has changed since 1989. Present law students are more aware that statistical science is helpful for their better understanding the practice, the implementation of law and the development of law.

Our Department of Statistics and Demography at the Faculty of Law of the University of Szeged considers the dissemination of knowledge in the fields of general and applied statistical methods one of its most important goals, in order to help students to acquire complex and insightful knowledge in these fields. This policy resulted in the increasing number of students, who select law and statistics for their fifth year dissertation or thesis. An average of hundred students undertakes such an enterprise each year, twenty to forty of them dissert about the connection and interrelation between law and statistics.

2. Statistics in Law. Application of Statistics in Private Law

Below, we have to provide some insight into the workings of the Hungarian justice system. The largest practical subsystem of the Hungarian legal system, besides administration, that should use statistics more often is the judicial system. Insight in

Application of Statistical Data and Methodology in Practice and Legal Education

Ildikó Szondi[1] *– Adrienn Princz*[2] *– János Mészáros*[3]

The science of statistics in regards of its methods and tools has a special relationship to the social sciences, thus sometimes escapes attention of the legal students and the members of the legal profession. In order to understand the structure of the Hungarian legal system, the knowledge of the statistical sciences is essential. For students learning private law the knowledge of statistics is helpful, showing the importance of certain legal statues or the necessity of legal rules in governing social networks. In order to show the mechanism of application of law in everyday life we also need statistical tools. Our lecture will show the statistical distribution of cases filed to the Hungarian Court System. Out of the total number of filed court cases every second one will be completed within a year. In he second part of our lecture, we will discuss some aspects of corporate law, showing the incidence of establishing business partnerships and corporations. In Hungary in 2008 the established business partnerships, corporations mainly consisted of limited responsibility and financial asset based partnerships. The science of statistics will help to better understand social life-networks by showing average wages, employment, or workplace contracts, their incidence and frequency. For our students it would be difficult to bet, to find out the real ratio and nature of work-contracts, in reality only 8% of work-contracts will be established with a pre-defined time-limit. It is also interesting to know, that out of 8 employees only 3 are women, the rest of it (5) are men. In the field of family law, in respect of the marriage data it is surprising, that in Hungary since 2004, the ratio of people older that 15 years, living in marriage, decreased, and now they are a minority. In 2008 almost 74 000 marriage ended up, by death of the partner, or divorce, and only 40 000 new marriages were registered. The followings also will show the Hungarian social and legal situation: the changing number of criminal cases, the special structure aspect of felonies, and felons, and the abovementioned and also the evaluation of the law-enforcement, court system, the work of judges, In the modern age of globalization, and connected world, it is of utmost importance that the statistical system, the criminal statistics should work with great precision and with the highest professionalism. As the Internet took over every aspect of the modern societies, it is

[1] Ildikó Szondi, Graduate of Law, Associate Professor, Department of Statistics and Demography, Faculty of Law, University of Szeged
[2] Adrienn Princz, Graduate of Law, PhD Student, Department of Statistics and Demography, Faculty of Law, University of Szeged
[3] János Mészáros, Fifth-year law student, Teachers Assistant, Department of Statistics and Demography, Faculty of Law, University of Szeged

Čobanović, K.-Sokolovska, V.-Nićin, S. 2008: Statistical Analysis of Marital Status of the Population of Vojvodina, Applied Statistics 2008, Program and Abstracts, Ribno (Bled), Slovenia, p. 48

Vojvodina in 2006. In sociology, like in many other disciplines, the use of statistical methods in practise was justified, because of its great importance in explaining the character of data, in explaining the behaviour and relations of investigated phenomena (social phenomena, demographic phenomena and so on) and in making conclusions. Both parametric and nonparametric statistical methods are suitable for use in sociology because of the nature of data. In sociology are common both quantitative and qualitative data.

References

Bethea, R. M.-Duran, B. S.-Boullion, T. L. 1985: *Statistical Methods for Engineers and Scientists*. Marcel Dekker, Inc., New York.

Blalock, H. M. 1960: *Social Statistics.* McGrow Hill Book Company, New York.

Ferguson, G. A.1966: Statistical Analysis in Psychology and Education, 2nd Edition, McGraw-Hill, London.

Greenacre, M. 2007: Correspondence Analysis in Practice, Second edition, Chapman & Hall, Boca **Raton**, FL, USA.

Hadživuković, S. 1991: Statistički metodi s primenom u poljoprivrednim i biološkim istraživanjima, Drugo prošireno izdanje, Poljoprivredni fakultet, Novi Sad.

Hinton, R. R. 2004: Statistics Explained, 2nd Edition, Routledge, New York.

Krneta, M. 1987: Statistika za sociologe, Ekonomika, Ekonomski institut

Raftery, A. E. 2000: Statistics in Sociology, 1950-2000. *Journal of the American Statistical Association*. 95, 450; ABI/INFORM Glober, pp 654-661

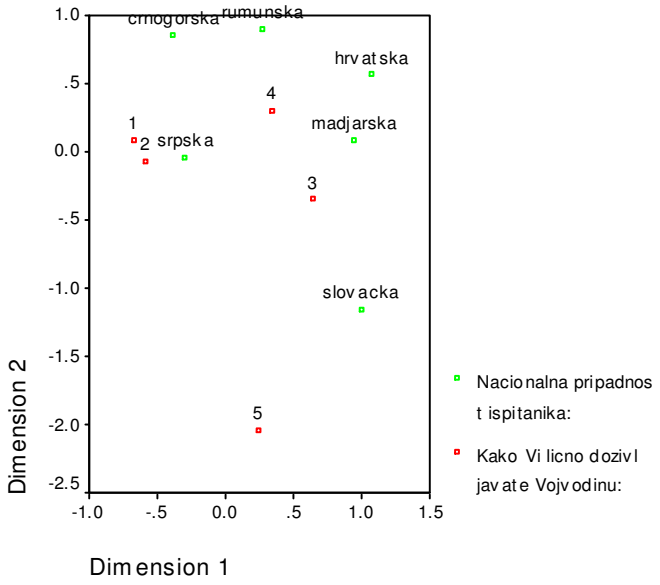
Sokolovska, V. 2009: *Acultural Processes of Ethnic groups in Vojvodina*., Doctorial Thesis, Faculty of Phylosophy , University of Novi Sad: (unpublishied material).

STATISTICA 7.0 2007: StatSoft, University Licence, University of Novi Sad

Čobanović K, -Nikolić-Đorić, E.-Mutavdžić, B. 1997: The Role and the Importance of the Application of Statistical Methods in Agricultural Investigations. *Agrieconomica*. Vol. 26. Pp. 457-470.

Čobanović, K.-Sokolovska, V.-Nikolić-Đorić, E. 2007: Statistical Analysis of Multiculturalism Research in Vojvodina, Applied Statistics 2007, Program and Abstracts, Ribno (Bled), Slovenia, p.40

Graph 3. How do the respondents perceive Vojvodina regarding the national (ethnic) affiliation of the respondents (survey of 2006)



Source: own creation

The variable "How do you personally perceive Vojvodina?" has the following modalities: 1) As a province of Serbian people. 2) As a province of Serbian people and national minorities, 3) As a province of different peoples and ethnic groups, 4) As a province of its citizens living on its territory, and 5) I do not know.

4. Conclusion

The paper states the importance of statistical methods in sociology. Due to the significance of statistical data analysis it is possible to apply adequate statistical methods, on which the value of acquired results and conclusions is based. The paper presents the experience of applying statistical method in teaching sociology students. There are also some author's results, related to applying statistical methods in the field of sociological and demographic research. Thus, the marital status of Vojvodinian population was examined on the basis of the data from the population, households and dwellings Census 2002 in Serbia. The paper also presents the results of correspondence analysis, referring to the relations of the most significant socio-economic, cultural, ethnic and other aspects of living on the territory of Vojvodina. These results are based on the sample of around 1200 respondents on the territory of

Table 1. The table of correspondence analysis for How do the respondents perceive Vojvodina regarding the national (ethnic) affiliation of the respondents (survey of 2006)

How do you personally perceive Vojvodina	national (ethnic) affiliation of the respondents						
	Serbian	Montenigrin	Croat	Hungarian	Slovak	Romanian	Active Margin
1	46	2	1	2	1	1	53
2	349	10	4	30	6	5	404
3	135	2	11	56	15	3	222
4	248	10	19	82	13	9	381
5	5	0	0	1	1	0	7
Active Margin	783	24	35	171	36	18	1067

Source: own creation

Table 2. The overview of correspondence analysis for How do the respondents perceive Vojvodina regarding the national (ethnic) affiliation of the respondents (survey of 2006)

Dimen-sion	Singular Value	Inertia	Chi Square	Sig. (20 degrees of freedom)	Proportion of Intertia		Confidence Singular Value	
					Accoun-ted for:	Cumu-lative	Standard Deviation	Corre-lation 2
1	0.280	0.078			0.898	0.898	0.027	0.030
2	0.088	0.008			0.089	0.987	0.036	
3	0.032	0.001			0.012	0.999		
4	0.010	0.000			0.001	1.000		
Total		0.087	93.016	0.000	1.000	1.000		

Source: own creation

– However, the relation between the attitude in favour of joining EU and NATO and type of settlement was not statistically significant;

– The relation between the attitude in favour of joining EU and NATO and employment status was statistically significant $\left(\chi^2 = 26.0610^{*}; df = 15; p = 0.03744\right)$;

– The relation between the category of keeping culture of the nation and the age category was statistically significant, while the relation between the category of keeping culture of the nation and national affiliation was not significant;

– The relation between the category of keeping culture of the nation and the category of confessional affiliation was statistically significant $\left(\chi^2 = 231.80^{**}; df = 70; p = 0.00000\right)$;

– The relation between the category of keeping culture of the nation and the category education level was statistically significant $\left(\chi^2 = 62.001^{**}; df = 35; p = 0.0033\right)$;

– The relations between the categories of perceiving current lifestyle and educational attainment, age groups, employment status and nationality were not statistically significant.

In the doctoral thesis "Acultural Processes of Ethnic groups in Vojvodina" – defended in May 2009 at the Department of Sociology of the Faculty of Philosophy in Novi Sad – correspondence analysis, among others, was used for analyzing the above mentioned survey in 2006. One of the analyzed attitudes is the way the respondents perceive Vojvodina. For the question "How do you personally perceive Vojvodina?" the respondents of Serbian nationality opt mostly for the reply that it is a province of Serbian people and national minorities; the Montenegrins perceive it equally as a province of Serbian people and national minorities, and as a province of its citizens living on its territory; the Croats, Hungarians and Romanians perceive it most as a province of its citizens living on its territory, while the Slovaks see it as a province of different nations and ethnic groups (Graph 3). According to the Tables 1 and 2, which present the results of the correspondence analysis, it can be observed that almost 90% of the responses to this question are assigned to the modality 1, which replaces the reply "as a province of Serbian people".

modalities); Sojourn (2 modalities); Sojourn Character (2 modalities); County (7 modalities); Previous sojourn vacancy (2 modalities); Household monthly income (20 modalities); Household size (10 modalities); Number of household members without income (10 modalities).

In this research the majority of the analyzed variables are qualitative, given descriptively, with a larger number of modalities. In order to perceive the relations of certain combinations of variables the method of correspondence analysis was applied. Correspondence analysis describes the character of the relations between the occurrences and the structure of attributive (categorical) variables (STATISTICA 7.0, Introductory Overview). The correspondence analysis in this research was based on two-dimensional contingency tables and on determining the overall chi-square value. On the basis of chi-square value the statistical significance of the examined relations, i.e. the combinations of variables, was determined. The results of the correspondence analysis were generated by the statistical programme STATISTICA 7.0. This paper states some interesting results:

- The relation between the level of education and national identification was statistically significant with the chi-square value 132.64^{**}, $(df = 50; p = 0.0000)$;
- The relationship between the national status and employment status was statistically significant $(\chi^2 = 60.473^{**}; df = 30; p = 0.00081)$. It is worth outlining that the categories of employment (employed, unemployed, economically dependants and retired) were very heterogeneous;
- The relationship between the national status and age category and the relationship between the national category and education category were not statistically significant;
- The relation between the aspect of Serbian nationality according to national minorities and employment status was statistically significant $(\chi^2 = 29.645^{**}; df = 8; p = 0.04112)$;
- The contact of national minorities to majority (Serbian) population and the category of sex was not statistically significant $(\chi^2 = 8.0816; df = 6; p = 0.23220)$;
- The relation between the attitude in favour of joining EU and NATO and the age category was not statistically significant $(\chi^2 = 31.20379; df = 25; p = 0.18240)$;
- The relation between the attitude in favour of joining EU and NATO and other groups and the category of education level was statistically significant $(\chi^2 = 50.427^{**}; df = 25; p = 0.00019)$;
- The relation between the decision to join EU and NATO and nationality was statistically significant $(\chi^2 = 117.30^{**}; df = 50; p = 0.00000)$;

$$\hat{Y}_i = 52.79073 + 0.30519X_1 + 0.05356X_2 + 0.64293X_3 - 0.51155X_4$$
$$\left(R_A^2 = 0.984\right) \quad (t = 35.61) \quad (t = 0.26) \quad (t = -4.45) \quad (t = 3.74)$$

In this case, the following explanations can be provided: by the increase of the number of urban inhabitants (X_4) the number of families without children in rural settlements significantly decreases, which can be the result of movement of young population from rural to urban settlements, while in rural settlements there are only old people left (families without children). At the same time, the increase of the number of families without children in urban settlements (X_3) has positive effect on the increase of the families without children in other settlements. This relation can be explained by the assumption that the families without children, in both urban and rural settlements, are the result of the birth and death rate decrease among the population, as well as the process of senilization of Vojvodinian population, which according to a lot of indicators, belongs to the type of old population.

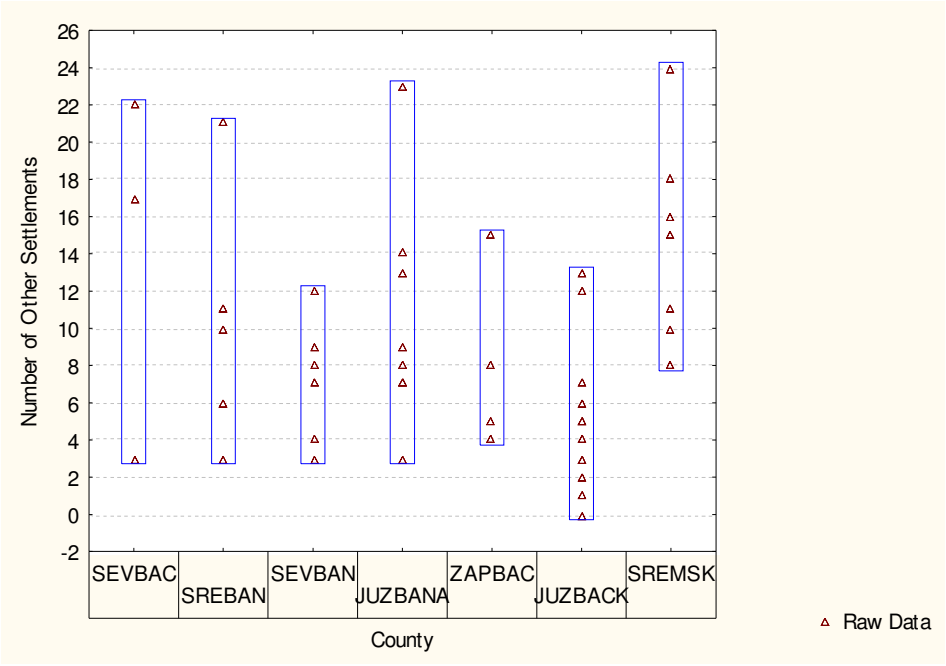
The results of the research, conducted in 2006 by the Department of Sociology of the Faculty of Philosophy in Novi Sad, on the sample of around 1,200 units of observation (respondents), and conducted on the territory of the Autonomous Province of Vojvodina, provided the material for the analysis of socio-economical and cultural aspects of multiculturality (Čobanović et al. 2007).

This research includes 21 selected variables. Some of them are defined as variables expressing grades of socio-economical situation, cultural views, ethnic aspects and the attitudes and opinions about many other features of multiculturalism in Vojvodina. Some variables refer to age, sex, national and confessional affiliation, education, occupation, activity sector, migration characteristics, county, household income, household size and the number of households with members without income. The variables are divided into 2 groups set by the author, keeping in view the assumptions on possible relations that can appear between the variables. In this way, a group of potential "independent" variables and the group of "dependent" variables were formed.

The first group of 6 dependant variables refers to the following phenomena: Evaluation of the family life quality (5 modalities); Evaluation of the most important family moments (7 modalities); Attitude towards joining EU and NATO (6 modalities), Evaluation of the culture preservation method (8 modalities); Attitude of the Serbian nationality members towards members of national minorities (7 modalities); Attitude of the national minority members towards national majority membership (7 modalities).

The second group consisting of 15 independent variables refers to the following phenomena: Sex (2 modalities); Age (6 modalities); National affiliation (11 modalities); Confessional affiliation (11 modalities); Education (6 modalities); Employment status (4 modalities); Occupation (28 modalities); Agency sector (8

Graph 2. Distribution of Other Settlements by Counties



Source: own creation

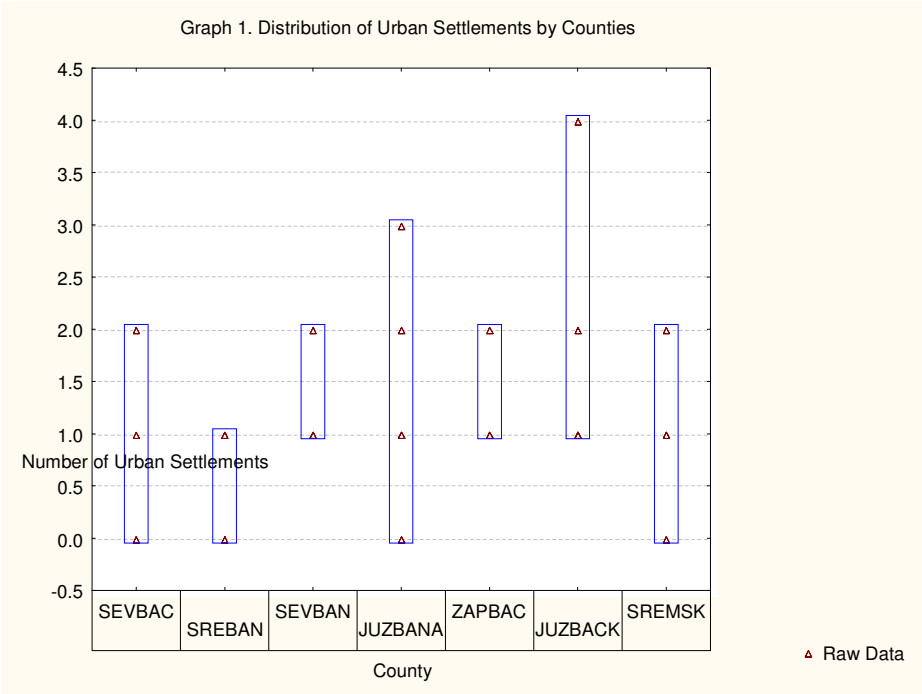
Regression analysis (for 45 municipalities) in this research, based on the model of multiple linear regression, was used for examining the influence of certain variables on the number of families with children and the number of families without children at the level of urban and rural settlements. The results of regression analysis showed that the number of families with children from urban settlements is significantly determined by the overall number of families (with the positive correlation), the number of inhabitants of rural settlements (with the negative correlation) and the number of families without children of rural settlements (with the negative correlation). The regression model is as follows:

$$\hat{Y}_i = 45.31810 + 0.54817X_1 - 0.44021X_2 - 0.31989X_3$$

$$\left(R_A^2 = 0.999\right) \quad (t = 181.14) \quad (t = 12.011) \quad (t = -2.915)$$

The regression results of the number of families without children in rural settlements on the number of families with children in urban settlements (with the positive correlation), and the overall number of families (with the positive correlation), on the number of families without children in urban settlements (with the positive correlation) and the overall number of inhabitants in urban settlements (with the negative correlation). The regression model is as follows:

Graph 1. Distribution of Urban Settlements by Counties



Source: own creation

3. Statistical methods in research

The research on marital status of Vojvodinian population according to the census results from 2002 indicates that the changes of marital status were influenced by socio-political factors, economic factors and a number of cultural factors (Čobanović et al. 2008). Two types of families were analyzed: families with children and families without children at the level of the Autonomous Province, as well as at the level of Vojvodinian counties. Furthermore, the two examined types of families were analyzed also in relation to the types of their settlements: urban settlements and other (rural) settlements. In order to determine the existence of statistically significant differences between the two types of families and types of settlements, t-test was used for 2 independent samples. The number of families according to the type (with and without children) was expressed per capita. The graphical illustration was made using a "box-plot" diagram. In this example, a "box-plot" diagram, as a means of presenting data analysis, confirmed the results of the t-test, i.e. the existence of statistically significant differences between the two types of families and the types of settlements.

This research, based on the results of the population, households and dwellings census 2002, also included the analysis of the number of municipalities according to the type of settlements in the counties of Vojvodina. The research encompassed the analysis of municipality distribution in relation to the overall number of settlements, to the number of urban settlements and to the number of other settlements in 7 counties in Vojvodina. The municipality distribution (45 municipalities) in counties is graphically presented using a "variability plot" diagram, which in this case proved to be a very suitable way of presenting and analyzing of discontinued variables (the number of settlements) (Graph 1, Graph 2).

the curriculum for the undergraduate academic studies of sociology, there are two compulsory one-semester courses in statistics on the first year of the studies.

The first course is entitled Descriptive Statistics, and its content is intended to introduce students to the problems of collecting and presenting quantitative data on the sociological issues; descriptive statistical measures; measures of central tendencies, and measures of variability; probability theory and sample method, as well as hypothesis testing. The objective of the course is mastering the basic statistical methodology used for planning quantitative research, for systematizing research results and reaching valid conclusions and decisions.

Within the second compulsory course entitled Statistical Methods, students are introduced to the basic methods of statistical analysis. Special emphasis is placed on the methods of nonparametric statistics: chi-square test, contingency coefficient, Wilcoxon rank test, Kruskal-Wallis, Friedman test, Spearman and Kendall rank correlation coefficient. Furthermore, the course encompasses simple linear regression and correlation, multiple regression and correlation analysis, time series analysis and variance analysis. The objective of the course is mastering statistical methods, including their adequate application and interpretation of obtained results.

During the academic undergraduate studies, students also participate in scientific-research projects of the Department of Sociology. Within the surveys conducted on the territory of the Autonomous Province of Vojvodina, students are engaged primarily to conduct surveys, where they gain fieldwork experience. Afterwards, they store the gathered data in software for statistical analysis and are present when preliminary research results are announced. Thereby, they can have access to a number of phases of research procedure, as well as gain insight into its complexity.

At graduate and doctoral academic studies in sociology students can opt for two optional courses in statistics. The course at graduate studies is entitled Applied Regression Analysis in Sociology, and it is intended to introduce students to the methods of applied regression analysis, as well as to point to the possibilities of its application in sociological analyses. The students also acquire knowledge on determining quantitative dependence of sociological categories, examining the validity of sociological theories hypothesis, as well as anticipating future movements of social phenomena on the basis of the assessed quantitative dependence.

At doctoral studies of sociology, students can study multivariate techniques, also as optional subjects. The course encompasses canonical correlation analysis, multivariate analysis of variance, discrimination analysis, principal components analysis, factor analysis, and grouping and correspondence analysis. Studying and research work includes gathering of adequate empirical material for thesis and its statistical analysis; examining the literature and problem studies from multivariate techniques, using statistical software for data processing, discussions on applied techniques, their scientific scope and limitation in sociological problems analysis.

are interrelated. Chi-square test is quite a common test based on determining the sum of the quotient of the square of the difference between the observed and expected frequencies and expected frequencies. Chi-square value has multiple applications, for instance, testing equality of distributions, independence test, contingency coefficient C, and so on (Hadživuković 1991, Hinton 2004); nonparametric tests for independent and dependent samples have also important application in research in sociology. These tests primarily refer to attributive variables (Hinton 2004). What should also be mentioned is the importance of nonparametric tests in application of the variance analysis method with a single factor, on the basis of ranked data. The most significant tests are Kruskal-Wallis test i Friedman test (Hadživuković 1991, Hinton 2004). Regarding nonparametric statistics, when examining the correlation and association between variables, it is important to mention Spearman rank correlation coefficient, contingency coefficient C, coefficient fi, which represents an indicator of association and is applied on discrete and dichotomous variables (Ferguson 1966, Hadživukovič 1991).

Application of quantitative methods in sociology by the World War II referred to descriptive statistics and simple methods (Raftery 2000). After the World War II, with the increase of the scope of data, more complex statistical methods started being used in sociology (Raftery 2000). For a great number of more recent statistical methods developed in 20th century it was sociology where they first found their application. **Thus, for example, correspondence analysis, a statistical technique useful in the analysis of categorical data, especially in sociological research, was emphasized in Greenacre's book (Greenacre 2007).**

The aim of this paper is to point to the importance of applying statistics in sociological research and statistics teaching. The paper presents the experience acquired while teaching statistics to sociology students at the University of Novi Sad. Moreover, the paper also presents some experience of applying statistical methods in sociological and demographic research**. Used statistical methods in the paper are results of the previous authors experiences. The authors wanted to describe the use of appropriate statistical methods in different research problems. There were used classical statistical tests in the explanation of Census data results. In the explanation of relationship between different variables, on the sample survey results base, there were used correspondence analysis and chi-square test. The authors wanted to outline the importance of use of statistical methods in social and similar sciences.**

2. Statistical methods in teaching

At the Department of Sociology of the Faculty of Philosophy in Novi Sad, the course in statistics is studied at undergraduate, graduate and doctoral studies. Within

1. Introduction

"Statistics is a branch of scientific methodology. It deals with the collection, classification, description and interpretation of data obtained by conducting surveys and experiments. Its essential purpose is to describe and draw inferences about numerical properties of populations" (Ferguson 1966).

The experience in statistics application indicates that a single statistical method can be used in different research areas for dealing with different types of problems. In this respect, we can talk about the universality of a statistical method (Bethea et al.1985; Čobanović et al. 1997). For instance, an experiment conducted in agriculture or in laboratory conditions is an equivalent to a sociological survey (Blalock 1960). Various possibilities which application of a statistical method provides in certain research areas should be considered as an adequate way of dealing with a problem of the research. However, it should be noted that statistics is not a method by which it is possible to solve all the problems in a research. Likewise, statistics is not a mere gathering of data and their presenting (Blalock 1960). Statistical texts cannot be treated as literary texts, since statistical instruments are presented in a very condensed form (Blalock 1960).

When applying a statistical method, it is common to differentiate between quantitative and qualitative features and variables. Nominal and ordinal variables and data are usually considered as qualitative (attributive), while interval variables and ratio variables are considered as quantitative (Ferguson 1966, Krneta 1987). Also, it is common to apply nonparametric statistical methods on nominal and ordinal data, while parametric methods are used for the interval and ratio data (Ferguson 1966).

Statistics and statistical methods have highly significant application in sociology. Functions of statistics are numerous: the methods of descriptive statistics have an important application for describing natural phenomena; inferential statistics is used for inductive reasoning about unknown properties of a larger group using the known indicators of the causes; hypothesis testing most frequently refers to the results of one, two or more causes, on the basis of which it is possible to draw conclusions on the problem of the research, by accepting or refuting an initial hypothesis; regression and correlation analysis, in the most simple case, examines the influence and dependence between two or more variables. If the relationship of a greater number of variables is examined, it is multiple regression and correlation. For a detailed analysis of a multiple regression model it is necessary to introduce partial correlations and explain the relations between variables using partial correlation coefficient; application of chi-square distribution and chi-square test is important in cases with qualitative variables for which it is known or assumed that

Use of statistical methods in sociology

Katarina Čobanović[1] - Valentina Sokolovska[2]

Data analysis in human sciences requires the use of statistical methods. Statistical techniques could provide explanations of how and why some statistical methods are used. The aim of the paper is oriented to the possibilities of using statistical methods both in education and in research in the field of sociology. Sociology and other human sciences, such as psychology, demography, economics and others, do require the use of statistical techniques.

This paper is focused on the analysis of data and the use of adequate statistical methods. Some statistical methods are more suitable for quantitative, some others for qualitative data. The paper is oriented to data analysis and the choice of an appropriate method. In sociology, like in other fields of science, it is necessary to deal with some simple and some more complex statistical procedures such as: descriptive statistics, inferential statistics, based on the sample survey, dealing with hypothesis testing, significance tests and analysis of variance, linear and multiple correlation and regression techniques, analysis of variance, linear and multiple correlation and regression techniques, analysis of frequency data using chi-square, nonparametric statistical procedures, multivariate methods etc.

The emphasis in the paper is placed on experience in using statistical methods in education and teaching of statistics and research work in sociology at the University of Novi Sad. The presented examples serve as an illustration of wide application of statistical methods.

Keywords: statistical methods in sociology, education, research work, types of data, analysis of data

[1] Faculty of Agriculture, Department of Agricultural Economics and Rural Sociology, University of Novi Sad, the Republic of Serbia, katcob@polj.ns.ac.yu
[2] Faculty of Philosophy, University of Novi Sad, the Republic of Serbia, valentinas@neobee.net

ACKNOWLEDGEMENT: The paper has been supported by the Ministry of Science of Republic of Serbia (Projects No. 149013D, No. 149007).

immigrated workers with Italian citizenship and other immigrants, especially for the interesting value of the study of their social integration and adaptation, after some years, to the occupational behaviour of autochthon workers.

References

Acocella - Sonnino 2003: Movimenti di persone e movimenti di capitali in Europa, il Mulino, Bologna.

Arango, J. – Hugo, G. – Kouaoci, A. – Pellegrino, A. – Taylor, J.E. 1993: Theories of international migration: a review and appraisal, Population and Development Review, Vol. 19 Issue 3.

Borjas, J. G. 1999: The Economic Analysis of Immigration. In Anshenfelter, O. - Card, D. (a cura di): Handbook of Labour Economics, vol. 3 A, North –Holland, pp. 1697-1760.

Kuijsten, A. C. 1995 : International Migrations: harmonization of data collection. In Collects et comparabilitè des donneés démografique et sociales en Europe, Chaire Quetelet 1991 : a cura di Institut de demografique dell'Universite Catholque de Luvain, Academia e L'Harmattan, Louvain-la-Neuve, pp. 593-624.

Lee 1965: Theory of migrations, University of Pennsylvania, USA.

Lutz, H. 2008: Gender in the migratory process, Conference on Theories of Migrations and Social Change, July.

Massey, D. 1990: The social and economic origins of immigration, Annals of the American Academy of Political and Social Science, Vol. 510, No. 1, 60-72 p.

Natale, M. – Strozza, S. 1997: Gli immigrati stranieri in Italia. Quanti sono, chi sono, come vivono?Cacucci Editore, Bari.

Piore, M.J. 1979: Birds of passage: migrant labour in industrial societies, Cambridge University Press.

Tronti, L. 1995: The effects of Immigrations on the Economic System. A Survey, mimeo.

Venturini, A., 1996: Extent of competition between and complementary among national and Third-World migrant workers in the labour market: an exploration of the Italian case, in AA. VV. , The jobs and the effects of migrant workers in Italy. Three essays, International Migration Papers, 11, Geneva, ILO, pp. 23-42.

Willekens, F. 1984: Comparability of Migration. Utopia or reality? Working paper, Nidi, 47, Voorburg.

predicted values of the random slopes: the negative effect of the education levels on the dependent variable is lowest for African and Asian workers (the slope passes from – 0.43 to – 0.22 for Africans and to -0.24 for Asians); it is still lower for North and Central and South Americans (the slope passes from – 0.43 to -0.34) and finally it is lower for European (mainly coming from the East, (the slope passes from – 0.43 to -0.36).These results seem to suggest a sort of hierarchic trend of the relations between instruction level and typology of profession passing from richer to poorer Countries. Moreover these numbers report important information on the different occupational decisions of immigrants offering useful instruments for the formulation of public policies and business strategies. To conclude, the hypothesis of the dual labour market theory was confirmed by the results of this study, suggesting that the division of the labour market into autochthon and immigrants opens an interesting and complex object of analysis.

6. Concluding remarks

Migration phenomenon has many common points of study with economic subjects like the socio-demographic composition of the Labour Force. This contribution tried to investigate one of those aspects consisting of the differences between natives and immigrant occupational behaviour. The methodological base from which this study moved is the "Dual labour Market Theory" which states that people from poorer countries are pulled to cover the bottom positions of the job hierarchy of richer countries. Moreover this theory motivates this assumption with some demographic trends characterizing richer countries in the last decades. In this paper some demographic variables are introduced in order to formulate the dual theory and to test it. The models have been estimated in order to test the relations between some demographic variables and the probability for a worker to be involved in a capital intensive or in a labour intensive profession and if such effects change between native and immigrated employees.

The results confirmed the initial research hypothesis consisting in the formulation of the assumption of the "Dual Labour theory" .Although, other innovative and interesting points of this contribution have arisen the first is the useful informative value of the mixed models in the study of the relations between migrations and labour market. In addition, the use of an economic database for a socio-demographic problem like the relations between migration and employment made it possible to deal with the migration phenomenon from a different perspective avoiding information limitations and ensuring the contents of the data. Another interesting result is the possibility to disaggregate labour forces in autochthons and immigrants: it showed the meaningful additive function of this demographic perspective especially when the origin of immigrants is also considered. To conclude, another interesting point to further explore is the difference in behaviour between

labour intensive segment of the labour market, is expressed by the increase or decrease in the slope of the function associated to each of the eight clusters. The predicted values obtained in the estimated models are those reported in the table 7.

Table 7. Predicted intercept and slope values for each group

CLUSTERS	PREDICTED VALUE		
	Intercept (R.I.M.)	**Intercept (R.S.M.)**	**Slope Ed.L. (R.S.M.)**
1 Italians	-1,83	0,70	-0,49
2 Foreign born and Italian citizenship	-1,37	-0,29	-0,15
3 Born in Italy and Foreign citizenship	-0,42	0,02	-0,03
4 Born abroad and European Citizenship (except Italy)	0,85	0,14	0,07
5 Born abroad and Asian citizenship	0,98	-0,19	0,19
6 Born abroad and African citizenship	0,73	-0,23	0,21
7 Born abroad and North and Central American citizenship	0,35	-0,20	0,09
8 Born abroad and South American citizenship	0,52	0,03	0,09

Source: own creation

The estimations of the predicted values gave a useful comprehension key to the analysis, adding information in terms of differences of occupational choices among workers coming from different Countries of emigration. The first reflection comes from the evident difference between the first three groups and the others. The intercept values as well as the slopes decrease if compared to the average value (shown in table 5 and in table 6) : it means that, in general, the probability of being involved in labour intensive professions is highly effected by education levels in the clusters of Italians, immigrants with Italian citizenship and Italian natives with foreign citizenship compared to the effect on the groups of European, Asian, African and American immigrants. The sense of this result is that those workers, belonging to higher levels of education, are not attracted by less satisfying jobs like labour intensive ones. On the contrary, for all the groups of immigrants without citizenship or not Italian natives, the effect of the instruction level changes: even if, as expressed by the first model, it remains negatively correlated to the labour intensive professions, in those populations it seems to be less strong. Moreover, the introduction of the random slope on the education level, diminishes the variability caused by the random intercept. The last notation refers to the single values of the

Table 5. Logistic model with random intercept

Dependent variable: Labur Intensive Profession (y=1)

Intercept and Covariates:	Beta	S. E.	T-Value	Alpha	Pr >|t|
Intercept	6,80	0.41	16.29	0.05	<.0001
Instruction level	-0.86	0.00	-95.67	0.05	<.0001
Gender (male)	0.49	0.02	23.92	0.05	<.0001
Age	-0.33	0.00	-36.13	0.05	<.0001
Variance	1,17	0.61	0.09	0.05	0.09

Source: own creation

Table 6. Logistic model with random slope (level of education)

Dependent variable: Labur Intensive Profession (y=1)

Intercept and Covariates:	Beta	S. E.	T-Value	Alpha	Pr >|t|
Intercept	4.59	0.20	22.49	0.05	<.0001
Instruction level	-0.43	0.09	-4.76	0.05	0.0031
Gender (male)	0.49	0.02	23.68	0.05	<.0001
Age	-0.35	0.00	-37.44	0.05	<.0001
Error term on intercept random effect	0.15	0.10	1.45	0.05	0.19
Covariance	-0.06	0.04	-1.36	0.05	0.22
Between clusters variance	0.05	0.02	1.87	0.05	0.11

Source: own creation

As shown in the tables, the three equations show the same effects of the covariates on the dependent variable. Although, the second model adds some more information to the study: it includes a first element of distinction among the eight groups through which the units of the sample have been classified per country of origin and country of citizenship. With the introduction of the intercept random effect, the different positions of the intercepts of each of the eight clusters can be observed and compared with the average of the overall intercept value. Finally, with the introduction of the random slope effect, for each group, the intra-clusters difference in the effect of the variable "education level" (that is the so called "between effect") on the probability of belonging to the capital intensive or to the

- β is the $1 \times m$ vector of parameters to be estimated using the data (y_{ij}, x_{ij}) and the first component of this vector is the average population intercept;
- x_{ij} is the $1 \times m$ vector of the explanatory variables and the first component is 1 due to the presence of the random intercept.

This model also implies that observations y_{ij} and y_{ik} are independent within a cluster.

The random effect represent the influence of subject (o cluster) i on its repeated observations that is no captured by the observed covariates.

Including the random effects, the expected value of the response variable y, which is related to the linear predictor via the link function, is given by: $\mu_{ij} = E\left(Y_{ij} \middle| b_i, x_{ij}\right)$. This is the expectation of the conditional distribution of the study variable given the random effects, that here equals $\Pr\left(y_{ij} = 1 \middle| b_i, x_{ij}\right)$ the conditional probability of a response given the random effects and covariates.

5.3. Results: Italian labour market. Comparisons between the behaviour of autochthons and foreigners

The following tables show the results of the statistical part of the analysis, where the research hypothesis has been tested. The logistic equation obtained in all the three models confirmed the expected relations of the effect of education, gender and age on the probability of employees to be involved in a labour intensive job. More precisely, higher instruction and age levels are negatively correlated to labour intensive jobs; on the contrary, being a male worker, increases the probability expressed by the dependent variable.

Table 4. Logistic model without random effects

Dependent variable: Labur Intensive Profession (y=1)					
Intercept and Covariates:	**Beta**	**S. E.**	**T-Value**	**Alpha**	**Pr >|t|**
Intercept	5.02	0.06	80.06	0.05	<.0001
Instruction level	-0.84	0.00	-95.20	0.05	<.0001
Gender (male)	0.47	0.02	23.26	0.05	<.0001
Age	-0.34	0.00	-37.46	0.05	<.0001

Source: own creation

that allows the predictors to affect the mean but assumes that the variance is constant will not be adequate for the analysis of binary data.

Suppose further that the *logit* of the underlying probability π_i is a linear function of the predictors $x_i'\beta$ as follows:

$$logit(\pi_i) = log\left[\frac{\mu_i}{1-\mu_i}\right] = x_i'\beta \qquad (1.2)$$

where x_i is a vector of covariates and β is a vector of regression coefficients, where β_j represents the change in the *logit* of the probability associated with a unit change in the $j-th$ predictor holding all other predictors constant.

The model defined in equations (1.1) and (1.2) is a generalized linear model with binomial response and link function *logit* with only a fixed effect.

The model used in this work is a mixed model for categorical response data including a usual fixed effects for the covariates plus the random effects on the intercept.

It is assumed that the sample get from N different clusters. Let $i = 1, \dots, N$ and $j = 1, \dots, n_i$ are respectively the indicators of the $i-th$ cluster and the $j-th$ individual observation. Let n_{ij} denote the total number of observations within the $i-th$ group. The response variable y_{ij} is binary, it can take only two values one or zero according to the presence or absence of the phenomenon of study and x_{ij} are the covariate (explanatory) variables.

Assuming that the incidence of a case of study differs from groups, an appropriate model that takes account of a random effect on intercept is a model with varied intercept from clusters, as follows:

$$\Pr(y_{ij} = 1|b_i) = \mu(b_i + \beta' x_{ij}). \qquad (1.3)$$

This is the *logit model*, a conditional probability model where:

- μ represents the probability function of y, and it is $\mu = \mu(s) = e^s/(1+e^s)$, function defined for all $s \in (-\infty, \infty)$;
- $b_i = (a_i - \alpha) \sim N(0, \sigma^2)$, where a_i is the iid random variable of the intercept, is normally distributed $a_i \sim N(\alpha, \sigma^2)$ with both parameters of population average intercept α and intercept variance σ^2 unknown;

5.1. Research hypothesis

The logistic model tests the hypothesis that age and education level have a negative effect and male gender a positive effect on the probability to perform a labour intensive job

The research question of the multilevel models relies on the hypothesis formulated on the basis of the "dual labour market theory". This principle asserts that the probability of being involved in labour intensive jobs is bigger for foreign workers than it is for autochthon workers. In addition, the study has the goal to investigate to what degree, having Italian birth and citizenship, or being immigrated (and from where), changes the effect of socio-demographic characteristics on the probability to perform a capital intensive or a labour intensive job.

5.2. Methodological description of the models

The methodological part of this research consisted on the estimation of three models:

- a simple logistic model;
- a logistic model with a random intercept effect;
- a logistic model with a random intercept and a random slope effect.

It is supposed that there are k independent observations $(y_1; \dots; y_k)$, and that the $i-th$ observation can be treated as a realization of a random variable Y_i. It is assumed that Y_i has a binomial distribution

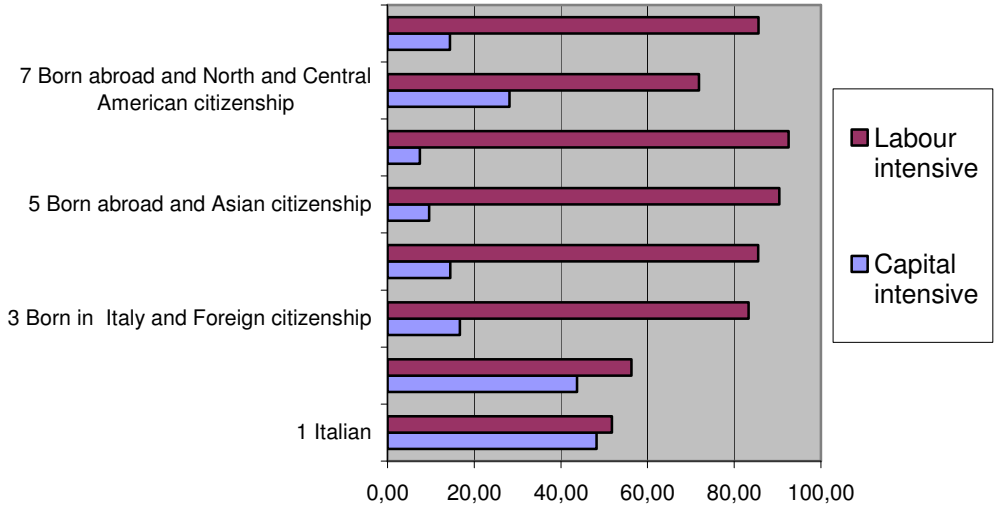
$$Y_i \sim B(n_i; \pi_i) \qquad (1.1)$$

with binomial denominator n_i and probability π_i. This defines the stochastic structure of the model. The distribution of Y_i is a *Bernoulli* distribution with parameter π_i, and can be written in a compact form as $\Pr\{Y_i = y_i\} = \pi^{y_i}(1-\pi)^{y_i}$.

The response variable y_i is binary, assuming only two values that for convenience we code as one or zero. Note that if $y_i = 1$ the probability is π_i, otherwise if $y_i = 0$ the probability is $1 - \pi_i$.

The expected value and variance of Y_i are respectively: $E(Y_i) = \mu_i = \pi_i$ and $var(Y_i) = \sigma_i^2 = \pi_i(1-\pi_i)$. Note that the mean and variance depend on the underlying probability π_i. Any factor that affects the probability will alter not just the mean but also the variance of the observations. This suggests that a linear model

Figure 3.Percentages of workers per Country of origin and typology of profession



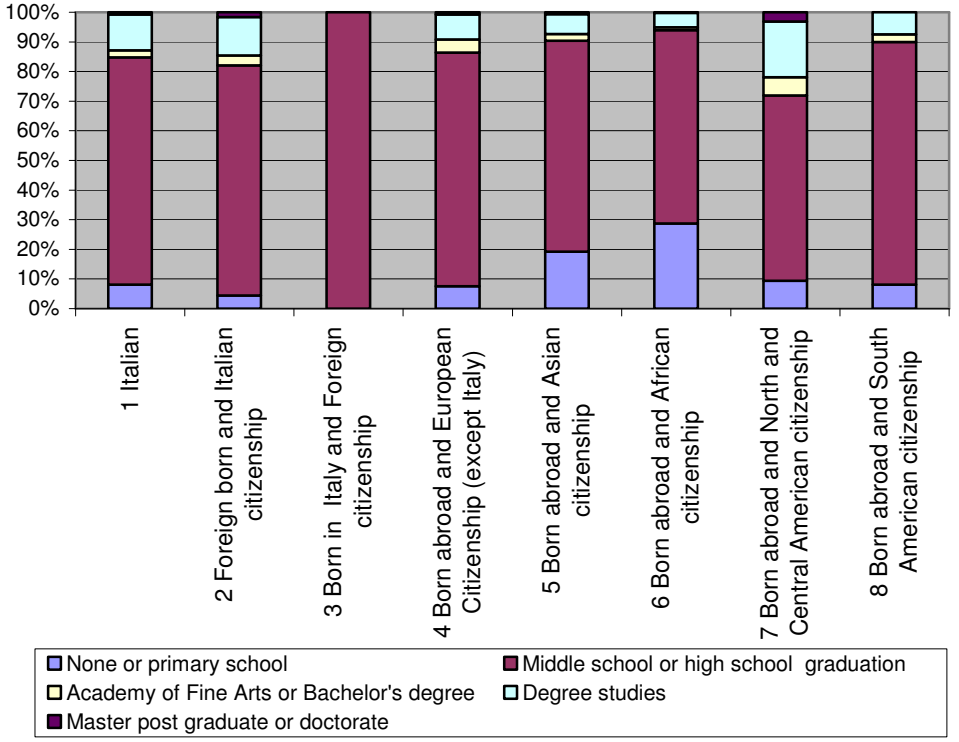
Source: own creation

The descriptive analysis pointed out that more than one third of the immigrated workers has Italian citizenship. Moreover, the composition of the sample shows that immigrants without citizenship behave in a different way with respect to the country of origin; on the contrary immigrants with Italian citizenship have more similar characteristics to Italian workers. Such differentiations are investigated in the second part of this paper with particular attention to the role of the level of instruction.

5. The professional choices of Italian Labour Forces

The challenge of determining the relations between migration and labour market is ambitious and hard for several reasons, among which, limitations due to informative and methodological nature prevail. In this contribution the effort to avoid the problem of the absence of data was reached by using stock official data; the methodological efforts to use a model able to give a robust and wide interpretation of the social problems deriving from the occupational structure of the Italian labour market, was attempted by the use of multilevel analysis. It is a specific methodology for the analysis of data with complex patterns of variability, with a focus on nested sources of variability.

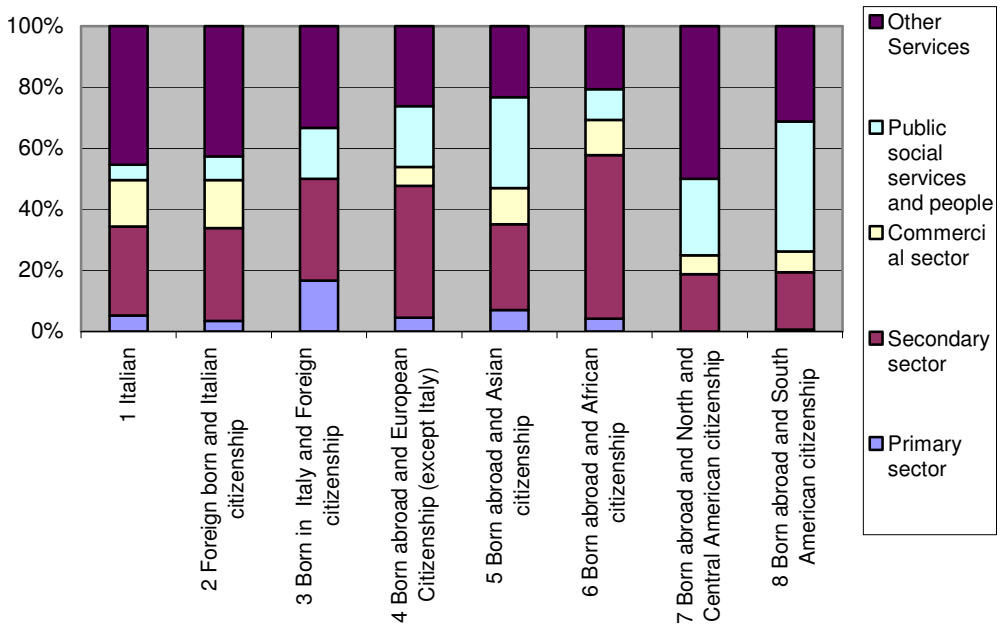
Figure 2. Percentages of workers per Country of origin and citizenship and education level



Source: own creation

The last figure shows that Italian workers strongly differ from immigrant workers considering the intensity of labour or capital: Italian native and citizens mainly work in capital intensive jobs.

Figure 1.Percentages of workers per Country of origin and citizenship and economic sector of activity



Source: own creation

The following graph shows that the proportion of workers with lower education level is bigger in the clusters of immigrants, among which especially African and Asians show the most considerable percentages associated to the modality "none or primary school" instruction level. On the contrary, immigrants particularly from North and Central America are characterized by the biggest percentage of graduated and post-graduated workers.

Table 3 shows the male presence ratios that refers to the percentage of males out of the females. It is very interesting to notice that immigrants from South and North America are mostly female, on the contrary, Italians, Europeans, Africans and Asians immigrated workers are more often males .

Table 3. Percentages of workers per Country of origin and citizenship

Origin	**Male gender ratio**
1 Italian	150,4
2 Foreign born and Italian citizenship	113,3
3 Born in Italy and Foreign citizenship	500,0
4 Born abroad and European Citizenship (except Italy)	113,4
5 Born abroad and Asian citizenship	192,5
6 Born abroad and African citizenship	358,3
7 Born abroad and North and Central American citizenship	52,4
8 Born abroad and South American citizenship	61,6
Total	149,1

Source: own creation

Figure 1 represents the distribution of the Italian labour forces per sector of activity; it shows that the larger proportion of immigrants are in the sectors of services to the persons and in the secondary sector. Both the sectors regard intensive labour jobs and the higher frequency of the variable has probably to be read considering the professions of cleaning ladies, old and baby sitters, and, in the secondary sector, mainly those in the construction services.

Table 1. Number and percentages of workers per country of provenance and citizenship

Origin/citizenship	**N. of presence**	**% of presence**
1 Italian	54195	93,7
2 Foreign born and Italian citizenship	1427	2,5
3 Born in Italy and Foreign citizenship	6	0,0
4 Born abroad and European Citizenship (except Italy)	1238	2,1
5 Born abroad and Asian citizenship	313	0,5
6 Born abroad and African citizenship	495	0,9
7 Born abroad and North and Central American citizenship	32	0,1
8 Born abroad and South American citizenship	160	0,3
Total	**57866**	**100,0**

Source: own creation

Table 2 illustrates a cross classification of the units on the basis of the following two variables: age class and the Country of origin and of citizenship. It shows that the largest proportion of workers concentrates in the central classes. The survey collects data on the national labour forces: by consequence, most of the units refer to workers not younger than 15 and not older than 64 , which is the conventional occupational age interval.

Table 2. Percentages of workers by classification of origin and age class

Origin	**15 - 24**	**25 - 34**	**35- 44**	**45- 54**	**55- 64**	**65- 74**	**75 and over**	**Total**
1 Italian	6	21	31	28	12	2	0	100
2 Foreign born and Italian citizenship	4	21	44	23	7	1	0	100
3 Born in Italy and Foreign citizenship	33	33	17	17	0	0	0	100
4 Born abroad and European Citizenship (except Italy)	10	33	32	21	4	1	0	100
5 Born abroad and Asian citizenship	7	24	40	23	5	0	0	100
6 Born abroad and African citizenship	7	28	46	18	2	0	0	100
7 Born abroad and North and Central American citizenship	9	28	28	28	6	0	0	100
8 Born abroad and South American citizenship	4	31	33	22	10	1	0	100
Total	6	21	31	28	12	2	0	100

Source: own creation

- master post graduate or doctorate.

Finally the last variable is the classification variable used for the aggregation of the observations in clusters. It aggregates the observations on the basis of Country of birth and Country of citizenship and counts the following eight modalities:

4. origin;
    - Italians;
    - born abroad and Italian citizenship;
    - Italian birth and foreign citizenship
    - born abroad and European citizenship (except Italy);
    - born abroad and Asian citizenship;
    - born abroad and African citizenship;
    - born abroad and North and Central American citizenship;
    - born abroad and South American citizenship.

4. Description of the sample

In this section, a brief description of the analysed sample on the basis of the distribution of the socio-demographic variables is proposed, in order to trace a first representation of the relations tested in the models. The demographic composition of the analysed sample provenance and citizenship is illustrated in the following table 1. It shows that the presence of immigrants in the sample is around 6.3%, 2.5 of which could be defined as “long term immigrants” being immigrants with Italian citizenship. The remaining percentage summing up to 3.9 %, is represented by “ more recently immigrated people”; this definition derives from the fact that they are those immigrated that still did not acquire the Italian citizenship. This last typology of immigrants was disaggregated by country of birth in order to observe if their behaviour differs throughout the clusters.

moves from the consideration that this specific database gives us *a-posteriori* information on employees' choices. In this sense our final considerations on the one hand can be or cannot be a confirmation of the hypothesis of the cited theory, on the other hand could add some other interesting points of research.

3. Informative instruments of the analysis

The study was conducted on the data of the Istat database of Italian Labour Force survey of the third trimester of 2006. The statistic unit of this survey is the family of the interviewed worker, for this reason the database also contains data of the other members of the worker's family and the observations refer to each family member in professional conditions and currently working. The final database[6] contains 57866 cases. The variables and the relative modalities included in the model are reported here.

The dependent variable:

1. profession (dichotomised) having the following modalities:
   - labour intensive;
   - capital intensive.

The following variables are included in the model as covariates and are reported here with the relative modalities:

2. age classes:
   - 15 – 24 years old;
   - 25 -34 years old;
   - 35 - 44 years old;
   - 45 – 54 years old;
   - 55 – 64 years old;
   - 65 – 74 years old;
   - 75 and over;
3. gender:
   - male;
   - female;
   - level of education:
   - none or primary school;
   - middle school or High school graduation;
   - academy for Fine Arts and Bachelor's degree;
   - degree studies;

[6] The database obtained after the elimination of all the units with at least one missing value in the analysed variables.

the model as covariates, three models were estimated: the first is a classic logistic model aimed at investigating which socio-demographic variable, among those considered, has an affect on the probability to be involved in a labour intensive profession (which is complementary to the probability of a capital intensive job); the second and the third models add to the previous information the explicative function of the random intercept and the random slope respectively. Those models are aimed at observing the intra-clusters variability. The goal of the generalized mixed model estimation is to observe if the effect of the covariates on the dependent variable changes when the origin of workers is used as variables of aggregation.

2. "Dual labour" migration theory

The "Dual labour" theory proposes a different approach for the explanation of international migration in alternative to neo-classical theories. Its basic assumption is that demographic and social changes currently affecting modern societies of developed countries - such as the decline in t birth rates, the increase in divorce and legal separation, the ageing phenomenon, educational expansion, the economic level of development and the well-being of people - generate job vacancies in the lower positions of the job hierarchy (Piore 1979).

The dual labour theory describes the occurring of the migration phenomenon as the effect of such demographic and social dynamics. More precisely the decrease in fertility levels in industrialized countries might have led to smaller inflows of teenagers who are more willing to accept less satisfying job positions; the increase in the emancipation of women and the rise in divorce rates could have led to the increase in female employment rates causing the lack of workers able to substitute the female role in the household services. Finally, a higher instruction level created an increase in the supply of more qualified personnel profiles and, as a consequence of a higher perception of the life social status, the decrease in the availability to accept economically and socially less satisfying jobs.

This theory divides the labour market into a primary segment, characterized by a capital-intensive method of productions and a secondary segment characterized by labour intensive production process. By consequence, it states that skilled workers, holding social status, higher income and employment conditions in higher consideration, are in the first segment; on the contrary, unskilled ones are in the secondary segment. The "dual labour" theory hypothesizes that foreign workers migrate for the dual labour market structure of the receiving country where they are pulled to cover the lower positions of the job hierarchy. By using the Labour Force Istat Dataset this paper aims to test this basic assumption of the "Dual Market Theory" and, at the same time, to clarify the effects of some demographic characteristics on the final job decision of workers. This methodological approach

Migration is one of the most urgent phenomenons as demonstrated by the most recent data issued by Istat on the international presence in the Italian labour population, showing once again in 2008 and in the first three months of 2009, the decrease in autochthon workers and the increase in immigrants. Although, the study of this social phenomenon recalls definitional and motivational questions. The first ones regard the difficulties in the individuation of the definition of "migrants" (Willekens 1984) mainly due to the following factors: quality of data; system of data collection; differences among countries; missing and unknown information (Kuijsten 1995). Moreover, different kinds of problems occur when dynamic (Natale - Strozza 1997) or stock data are analysed; in fact both approaches face meaningful limits whose nature changes on the basis of the specific object of analysis.

Many researchers have been dealing with the topic of international migration and their implications on local occupational levels. Those studies are mainly directed towards the analysis of income with the goal of analysing the effect of immigrants on local level of employment and the question point is if their presence can be considered as complementary or competitive. Even the estimation of the salary differentials and the determinants of the relative causes has been treated by several international (Borjas 1999) and Italian authors (Acocella - Sonnino 2003). Other studies concern the impact of the presence of international workers on the occupational opportunities of local population in terms of employment rates and living conditions and also the role of illegal immigration (Tronti 1995, Venturini 1996). The object of this contribution differs from those here reported; it can be contextualized in the demographic and economic research area of "migration theories" (Piore 1979, Massey 1990, Arango et al 1993). Those theories were formulated with the aim of specifying a socio-economic motivation with a scientific basis, towards the occurring of migration flows.

In this paper the definition of migrant is based on the concept of permanent or semi-permanent change of residence[5] (Lee 1965). The focus of the study relies on the analysis of the composition of the Italian Labour Force moving from the following assumption of the "Dual Labour Market Theory": foreign workers cover the shortages of the labour market of the receiving country in the lowest positions of the job hierarchy. For this reason, the main variable object of this analysis is the typology of profession and it was aggregated in two modalities: capital intensive and labour intensive professions. In order to explicit the socio-demographic factors influencing the decision to accept a capital intensive or a labour intensive profession, a specific preliminary descriptive analysis has been undertaken on the variables already tested in previous researches (Acocella - Sonnino 2003, Lutz 2008). The following variables were introduced in the analysis: gender, education level and age class. After the descriptive analysis of the socio-demographic variables included in

[5] Coherently with the Lee's definition, in this study, all foreigners (for birth or for citizenship) were considered immigrants for professional reasons.

Comparing native versus immigrants' occupational choices of the Italian labour force: a generalized linear mixed model approach

Luca Bartoli[1] -MariaChiara Pagliarella[2] -Laura Palombo[3] -Renato Salvatore[4]

In this paper a short-term analysis of the occupational behaviour of the Italian labour force is proposed by analysing official Istat micro-data of 2006 through a logit model with subject-specific intercept. The aim is to compare the labour choices of autochthon workers to those of immigrants also considering the native country. Many national and international authors have been dealing with this interesting topic mainly in an economic key, treating the problem of defining the phenomenon of labour migration as a cost or as a benefit for the destination Country. Moving from the assumption that the complexity of the phenomenon and the lack of qualitatively acceptable data create such big limitations and delays in the formulation and the elaboration process of meaningful solutions, our proposal is an "a posteriori" analysis of labour choices taking into consideration the intrinsic demographic characteristics of native and non native individuals. The idea is to clarify if native and non-native workers share the same labour choices and expectations in terms of qualification grade as a proxy of their satisfaction. To this extent, in this paper firstly the hypothesis that native and non native workers show the same expectation will be tested, secondly the occupational behaviour of immigrants and autochthons will be compared on the basis of derivation country.

Keywords: Immigrants, labour, mixed model

1. Introduction

This contribution analyses the sample of the Italian Labour Force with specific emphasis on the distribution of workers per country of origin. The goal is to highlight the presence of systematic relations between the socio-demographic characteristics of workers and their occupational choices; each research question of the study is investigated, underlying the differences in the behaviours of immigrant and autochthons.

[1] Luca Bartoli, University of Cassino
[2] Maria Chiara Pagliarella, University of Cassino
[3] Laura Palombo, University of Cassino
[4] Renato Salvatore, University of Cassino

Weller, S. C. - Romney, A. K. 1988: *Systematic Data Collection.* Qualitative Research Methods, Volume 10, Sage Publications.

5. Conclusions

We have introduced the theoretical background of the systematic data collection in our paper, we approached the issue only from a methodological perspective in our study. This method, developed in the US, is used to define the concepts of social sciences, including economic sciences, that are difficult to be defined.

The advantage of this method is that it decreases the necessary sample and, at the same time, the level of reliability of the data is the same as that of those research techniques using large samples.

We have developed the research plan for employment policy and unemployment, and we intend to use the systematic method of data collection in our future research.

We regard the method as acceptable, as Eszter Siposné Nándori has tested systematic data collection for poverty surveys and she received similar findings as using the traditional data collection. We suggest that the method of systematic data collection may make researchers' work easier, although the focus of the research is of key importance.

References

Bánfalvy, Cs. 1997: *A munkanélküliségről.* Magvető könyvkiadó, Budapest.

Cazes, S. – Nesporova, A. 2003: *Labour markets in transition: Balancing flexibility and security in Central and Eastern Europe.* International Labour Office, Geneva.

Galasi, P. 2003: *Munkanélküliségi indikátorok és az állásnélküliek munkaerő-piaci kötődés.* Budapesti Munkagazdaságtani Füzetek.

Global 2002: *Global Competitiveness Report*, The World Bank, Washington DC.

Itelszon, L.B. 1967: *Matematikai és kibernetikai módszerek a pedagógiában*. Tankönyvkiadó Vállalat, Budapest.

Maxfield, M. – Babbie, E. 2009: *Basics of research methods for criminal justice and criminology* (2nd ed.). Belmont, CA: Wadsworth.

Nunally, J. C. 1978: *Psychometric Theory*. McGraw-Hill, New York.

Roberts, J. M. 1964: The Self-Management of Cultures In Goodenough, W.H. (ed.): *Explorations in Cultural Anthropology.* McGraw-Hill, New York. 433-454. p.

Romney, A. K. -Weller, S. C. – Batchelder, W. H. 1986: *Culture as Consensus: A Theory of Culture and Informant Accuracy*. American Anthropologist, New Series, 88/2. 313-338. p.

Siposné Nándori, E. 2008: Judgement of Poverty in Borsod-Abaúj-Zemplén County According to the Method of Systematic Data Collection. *MicroCad 2008*, 187-193. p.

Vajda, Zs. 1999: Somlai Péter: Szocializáció. *Szociológiai Szemle*, Vol 2., 166. p.

competence of the governmental, local-governmental and business sphere, in which the government relies on the contribution of the actors of the economy in order to accomplish the employment goal. The tasks of the employment policy can be illustrated by the enumeration below. Employment policy is an external regulatory system of the labour market aiming to influence the labour market in a way that:

- its operation becomes more flexible and harmonic,
- it fosters the establishment of labour market equilibrium,
- the implementation of social functions also receives attention,
- the size and professional/educational structure of the labour supply entering the labour market meets the requirements of the demand,
- the labour supply decreases, e.g. the increase of the number of participants in education or the length of education,
- it continuously adjusts labour supply to demand as perfectly as possible,
- it helps develop the size and structure of labour market demand that is capable of employ workforce appearing as labour supply to the greatest extent possible,
- it supports the living, on the basis of the principle of social care and solidarity, of those who exited the labour market through no fault of their own.

4.4. Questions asked in the interview

The method of systematic data collection can start with the questions below. Since the objective of our study is to introduce the methodology, we have not conducted the actual survey yet; we are going to elaborate on the findings of our future research in a subsequent paper.

- Who do you think is unemployed?
- What is the minimum period of time spent unemployed by those regarded as permanently unemployed?
- Up to what age can someone regarded as of working age?
- What does “black work” mean for you?
- Up to what age do you consider someone to be an entrant in the labour market?
- Who do you think are bereft of hope in terms of employment?
- What do you mean by the application of active employment policy tools?
- What are the characteristics of the traditional forms of employment?
- What is the minimum working time of full-time employees in terms of hours?
- Below how many hour of working time per week can be somebody regarded as part-time employee?

leads to the insufficient efficiency of the economic policy measures because the population cannot make use of these measures due to misunderstanding. The axiom taken from business management suggests that every strategy is worth the extent to which it is implemented.

The indirect result we wish to reach is to explore information that can help economic policy actors communicate their measures more precisely to the population. The direct result will be the population's interpretation of the labour market definitions.

4.2. Definitions relating unemployment

The literature knows and uses many and basically similar definitions for both employment and unemployment. Below give an incomplete overview of the most well known ones.

"Unemployment, as an overall social phenomenon has existed in the world since capitalism became a dominant form of production" – suggests Bánfalvy (1997) in the introduction of his book.

Unemployment has gradually reached its modern meaning which demonstrates the evolution of the concept: "That person can be regarded as unemployed who would occupy a job, but currently (s)he has not got a paid job". (Galasi 2003)

Those bereft of hope are also worth mentioning who, although unemployed, do not seek job because of the lack of employment opportunities. Hidden unemployed often get out of the traditional unemployed category and are regarded as inactive then.

Theoretically an individual is unemployed if (s)he, among given labour-market conditions, is willing but unable to undertake a paid job. The situation is somewhat more complicated from an empirical perspective. A person is unemployed, according to ILO, if the following conditions apply:

- (s)he did not have a job during the week before the survey,
- although (s)he actively sought job and
- would have been available for work. If (s)he spent at least 1 hour work for wage or salary the week before the survey then (s)he is regarded as employed. If (s)he had a paid job for shorter than 1 hour or did not seek job, then (s)he is inactive.
- A person is registered unemployed if (s)he has been registered with a job-agency. The one who is eligible for unemployment allowance is called insured unemployed. Their proportion might show a distorted picture because this category may include those who do not want to undertake job or may exclude those who seek job but do not register.

4.3. Range of tasks of employment policy

Employment is, on the one hand, shaped by the market and, on the other hand by the governmental policy influencing it. Making employment policy work is the joint

- Effect of employment policy tools on national and regional development

While conceptualization means refining and specifying abstract concepts, operationalizing means developing those concrete research procedures that take us to real empirical observations using these concepts, that is, we define concrete empirical procedures leading to measuring the variables.

Having set the main objectives, we have to identify the most appropriate method to reach them. In this case we chose the previously introduced method of systematic data collection. The research is new also in that this method has mainly been used for sociological researches, and not for labour market and economic analyses.

We used Itelszon's hypothetic-deductive method in our research, this method includes six phases. (Itelszon 1967)

1. We assume that the economic performance of the national economy relates to the efficiency of regional employment policies. The central tools do not get to the heavily problematic segments in the case of specific countries.
2. We are going to examine the unsuccessfulness of employment policy.
3. According to our hypotheses the weight of regional employment policies has to be significant and the have to respond to individual problems. Communication faults play role in the efficiency of the employment policy tools. The labour market definitions do not cover the population's concepts. The concept of employment, activity and unemployment are undefined in the society.
4. We can critically argue that regional competitiveness is improved by strategies focusing on individual problems. Its success requires strategy-awareness.
5. In a crucial experiment we are going to make a comprehensive comparison using regional employment policy tools.
6. As a theorem, we wish to assert that the various regions in Europe are in different labour-market situations; therefore, not only EU and national employment policy are necessary but also more explicitly defined regional employment policy.

In ordinary people's thinking employment means that one has a job or not. On the contrary, unemployment means the lack of job – in over-simplified terms. That is why we hold it important to ask, by way of the method introduced above, the man in the street about the most important questions and definitions relating employment and unemployment.

The purpose of our research is to match the meaning of the concepts emerged in employment policy with the colloquial language of the man in the street. Our reason for establishing this purpose was that, according to our hypothesis, the professional definition and the population's definition concepts mismatch. It also

and proportion of poverty calculated on the basis of the various theoretical concepts are different. The essence of the systematic data collection methods is that every interview respondent receives the same questions. The objective of her research with these types of questions was to identify those factors the respondents would consider to have the closest (also less close or no) relationship with poverty (Siposné 2008). She conducted a traditional data collection as well in order to test the reliability of the estimation and hypothesis testing, in the course of which she used a sample size generally accepted in social sciences. She examined, by way of contrasting the findings of the two data collections, whether the systematic data collection using small sample can reach the same level of reliability as the traditional methods.

Research findings: the sample she used was extremely small; she tested the reliability of the method by a similar research conducted on a larger sample. Having contrasted the two research findings it has become clear that they have lead to the same result in the case of all variables examined. Therefore, the assumption behind the systematic data collection, that is one can draw reliable conclusions from smaller samples than usual if the average competences of the respondents are taken into consideration, is reasonable. Average competences of respondents mean here the extent to which they think in a similar way of the concepts used in the survey.

A great advantage of systematic data collection regarding the clarification of employment-related issues, is that it decreases the necessary sample size, and thus, it reduces the time and cost to be spent on the research. Whether or not the method can really be applied in surveying employment and unemployment could be ascertained once we have the result of the, so called, free-listed enumeration. The average competence of data providers, that has to reach a critical value so that later phases of the method are applicable, can be established only on the basis of this.

4. Research plan aiming at the clarification of the concepts of employment policy

4.1. Concptualization, operationalization and Itelszon's hypothetic-deductive method

By our research plan we use the conceptualization and operationalization. Conceptualization is a process by way of which we define what mean by particular expressions. In the course of conceptualization we specify what mental pictures are related to our concepts and we choose the observations and measurements are appropriate for our research purposes.

The following concepts need precise definition in our research:

- Unemployment
- Employment
- Economic meaning of employment policy tools

After establishing the list of items, triad data may be collected by first enumerating all sets of size three for the items. The next step is to randomize the triads, both by position within the triad and by triad order. With a triad form ready, you can collect data either orally or with a questionnaire. Informants are asked to order the items within triads from "most" to "least" on some attribute. As for unemployment, informants would be asked to choose the item most related to unemployment. Then, from the remaining two items they are asked to choose again the one that is most related to unemployment. The total number of triads, however, goes up very rapidly with an increase in the number of items, so this data collection format is only practical with 10 or fewer items. The total number of triads can be reduces by using a balanced incomplete block design (Weller - Romney 1988).

Balanced-incomplete block (BIB) designs systematically compare subsets of items. The designs work by controlling the number of times that each pair is compared. By reducing the number of times each pair is compared to other items, the total number of subsets is reduced, while still maintaining comparisons among all items. BIB designs are identified with three parameters: n, the number of items; lambda, the number of times each pair occurs; k, the number of items in each set or block; and b, the number of sets or blocks. Even using this data collection format, the method becomes impractical with more than 25 items (Weller - Romney 1988).

Comparison of a large number of items can be done with quicksort. It minimizes the number of paired comparisons by assuming transitivity. It means that if an informant has judged A to be greater than B and B greater than C, then you assume that A is greater than C and you do not ask any questions about that pair. Names of items are written on cards which are first randomized and a card is selected as a 'standard'. All cards are compared to the standard and are divided into two piles: the cards 'greater than' and those 'less than' the standard. This process is repeated for each pile, until all items are ordered (Weller - Romney 1988). In case of unemployment related items, informants would be asked to divide them into two piles: items that are related more and related less to unemployment.

To answer the second question, namely the exact meaning of short and long term unemployment, rating scales can be used. They are the most widely used methods to collect data in written format as they work best with literate informants. Scales are usually expressed as four- to eleven-point scales. The more points a scale has, the more reliable it is said to be (Weller - Romney 1988). To find out the general belief about long term unemployment, informants may be asked the following question "What is the minimum time period for which somebody has to be unemployed in order to be considered long term unemployed?". A similar question can be asked for short term unemployment ("For how long can you be unemployed at longest to be considered short term unemployed?") to control for the potential inconsistency in the answers.

Eszter Siposné Nándori tested the model referring to poverty. In fact, there is no single exclusive interpretation of poverty and that is the reason why both the size

specific question (Romney et al. 1986). It is assumed that the correspondence between the answers of any two informants is a function of the extent to which each is correlated with the truth, for example with the culturally correct answers (Nunally 1978). Cultural competence is the result of the socialization process, which refers to the incorporation of social effect and beliefs to the individual's behaviour and personality (Vajda 1999). Roberts (1964) states that in any culture information is stored in the minds of its members and in artefacts. This study focuses on the part of the culture that is stored in the minds of its members.

According to consensus theory, the number of informants needed depends on the average level of competence, the confidence level and a minimum rate of questions we would like to classify correctly (Weller - Romney 1988). The average level of competence can be defined on the basis of free listing and then consensus theory can help to define the sample size.

If at least 90% of the questions should be classified correctly at the 0.95 confidence level, the reference to Table 1 shows the sample size.

Table 1. Minimal Number of Informants Needed to Classify a Desired Proportion with a Specified Confidence Level for for Different Levels of Competence

Proportion of Questions	*Average Level of Competence*				
	0.5	0.6	0.7	0.8	0.9
0.95 Confidence Level					
0.80	9	7	4	4	4
0.85	1	7	4	4	4
0.90	13	9	6	4	4
0.95	17	11	6	6	4
0.99	29	19	10	8	4
0.99 Confidence Level					
0.80	15	10	5	4	4
0.85	15	10	7	5	4
0.90	1	2	7	5	4
0.95	3	14	9	7	4
0.99	*	0	3	8	6

* Well over 30 informants needed

Source: Weller, S. C - Romney, A. K.: Systematic Data Collection. p 77.

Informants then can be selected using the method of multistage cluster sampling with stratification (Maxfield - Babbie 2009).

The general belief about unemployment and social aid can be investigated with the methods of triadic comparisons, quicksort or balanced-incomplete block designs. The choice of the method to be used is determined by the number of items elicited in free listing.

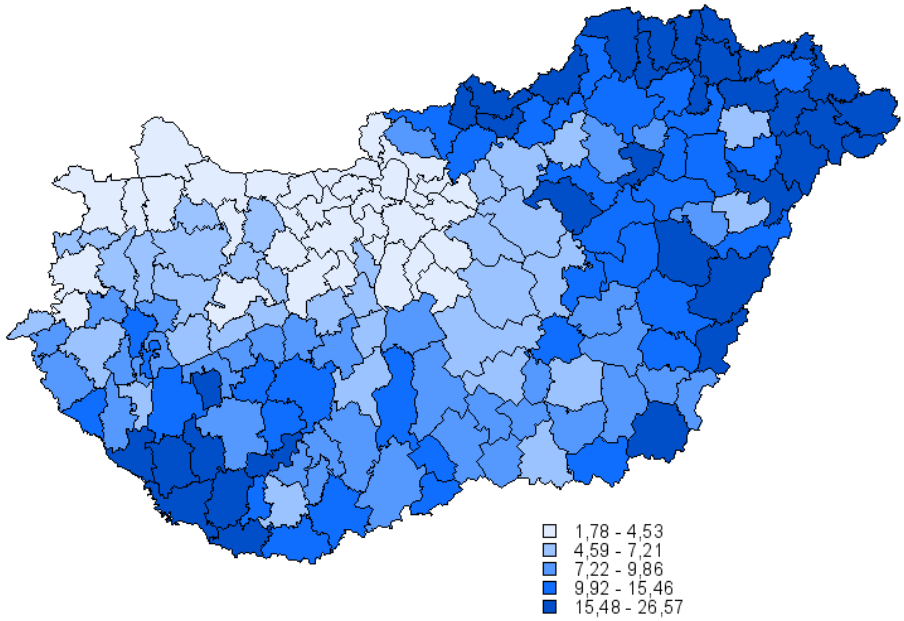
longing to the domain of interest. Free listing is a good way to ensure that the domain and the items are culturally relevant. The most important use of free listing is to ensure that the researcher is dealing with relevant items and to find the boundaries of the domain (Weller - Romney 1988). In this study, free listing is going to be used to get a list of items for further research, and to obtain understanding on what people think of the terms of employment. The sample size necessary for free listing was determined by taking into account the amount of agreement in the responses of the informants. Stability in order usually can be found at around 20 or 30 informants, so a sample size of 20 or 30 is in general sufficient. Informants are going to be chosen using multistage cluster sampling. The list of unemployment-related items is going to be elicited by asking the following questions: 'Who do you consider to be unemployed in general?', 'Do you know unemployed people?', 'Why do you think they are poor?'. Similarly, items related to social aids can be collected by asking: 'What kind of social aids do you know?', 'Do you know of anybody who is entitled to social aids', 'What kind of aid does he/she get?', 'Why do you think he/she is entitled to it?', 'Do you think that social aid can help him/her in solve his/her problem?'. Based on their answers, further answers can be asked if necessary.

This study has two aims: finding out (1) what the factors most related to the term unemployment and social aid are; and (2) what the exact meaning of long and short term unemployment are. The first question can be answered by ranking the items from free listing. As ranking a large number of items could be difficult and could take a lot of time, special methods are used to simplify the task. The method of the analysis is thus systematic data collection, developed in the 1980s by the anthropologist, Susan C. Weller, and the mathematician, A. Kimball Romney in the United States. This method helps researchers in the social sciences collect better interview or questionnaire data. The goal of social sciences, that is the better understanding of experimental and observational data, requires careful analysis of data. Increased understanding requires systematic observation, classification, analysis and evaluation. Structured interviewing formats help to reach this goal (Weller - Romney 1988).

The qualitative research method of systematic interviewing – where each informant is asked the same set of questions – diminishes the sample size required in social science research in a revolutionary way while the reliability of the results is still as high as in case of traditional techniques. This is possible by taking into consideration cultural competence of the population in defining the sample size.

Consensus theory is used when the researcher does not know what the answers are or what they should be and, instead, tries to discover the "culturally correct" responses to the questions. This theory helps to determine the number of informants necessary to get reliable answers in cases when the answers to the questions are not known ahead of time. It gives the possibility to measure the cultural competence of informants (the probability that the informant knows the answer to a given question) and it allows to reconstruct the "culturally relevant" answers to a

Figure 2. Rate of unemployment in Hungarian subregions (2007)



Source: Resource map, Hungarian Academy of Science

First of all, significant changes are necessary in the institutional system of labour-market policies, in the inter-institutional co-ordination, and in the integrated operation of the organisation that puts the employment policy into practice. It is common that the institutional system of the labour policies is, similarly to other countries, scattered in Hungary and there is much to improve in the area of inter-institutional co-ordination as well. The various organisations use different concepts, and the impacts of the various measures, which are well known and properly interpreted by experts, fail because of the social incomprehension and misunderstandings. The various concepts are not only mixed up in the surveys conducted in the specific member states but also in the international comparative researches.

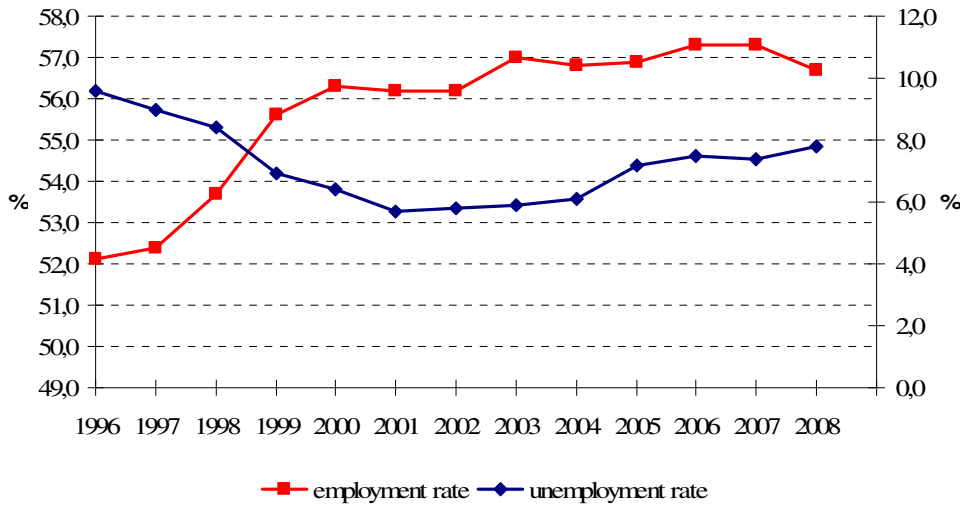
3. Methodology

Items expressing the general belief about the terms of employment can be found using qualitative research methods. First, a list of items related to employment is elicited from informants, and then they are asked to rank these items to find the ones most related to the field of interest.

The first step in any study is to get a clear understanding and boundaries of what is being studied. To do this, informants have to be asked to list the items be-

even greater than that of the external demand, entailed considerable real sacrifices. The rate of unemployment was also at its lowest point, the number of long-term unemployed significantly grew.

Figure 1. Labour market situation in Hungary



Source: own compilation on the basis of Eurostat data

The value of the two main labour market indicator, in the chart displayed in Figure 1, has showed a similar trend to the other ex-socialist countries. Putting the rates of unemployment and employment in the same system of co-ordinates demonstrates that the trend changed to opposite between 1998 and 1999.

Hungary is different from other countries of the region in two respects: the job-seeking intensity of unemployed men (with special regard to those with low level of education) is extraordinarily low; some sources mention that the costs of launching new enterprises are high. Apart from the general European objectives (supporting the disabled, fostering life-long learning and fighting tax-avoidance), employment policy has to concentrate on the above areas in Hungary including making efforts to find out, using more profound data-collection and analysis, that whether the picture, developing from the elementary statistics and the available research findings, is precise enough.

The rate of unemployment well demonstrates the regional differences in Hungary (Figure 2).

thus his/her sense of uncertainty decreases; (s)he helps economy get over the crisis by way of optimistic anticipations. We argue that the Hungarian employment policy, its institutional system, the official databases and the researchers use numerous definitions, while the society interprets the particular phases differently. It is the task of the legal system and the profession to use the concepts developed in a given era.

In this study we introduce a method for labour market research that enables faster reaction at lower investment in terms of time and money. It means a fast and flexible feedback between the hypothesis and research findings and formulating a more comprehensible economic policy.

It leads to stories, we can generate processes, which are effective and can correct the hypotheses, extremely far from reality, of theory and politics; or at least warn those people who try to seek connections between the professional theorems to the concrete processes that they are wasting their time.

The reason for choosing the method we use – briefly introduced in chapter 3 – is to attempt to introduce a fast, flexible and cost-effective process for labour market analyses. The method basically is a sampling survey, its information are not deeper, however it provides opportunity for quick report-like surveys.

2. Employment situation

One of the mid-term priorities of the Hungarian employment policy, aiming at enhancing the long-term stagnating employment, is that the expansion of employment should be facilitated by way of letting inactive population return to the labour market and preventing unemployment to become permanent.

The token of success in the case of the European and thus the Hungarian employment policy is that the actors of the economy can understand and operate it. The labour market problem of the core member states of the European Union and those problems experienced in new joiners are similar. European labour markets – as opposed to the American one – are usually regarded as inflexible and sclerotic because the employment is lower, the regional and industrial crises are more permanent, it takes longer for the unemployed to find job, and the wages and salaries adapt to changes in the demand over a longer period of time. Hungary occupies an average position in the region according to the labour market indicator used in the comprehensive World Bank study (2002) and the Cazes-Nesporova (2003) book. Similarly to other countries in the region, the employment of those completed primary school is extraordinarily low, the regional differences are high and permanent, and the taxes on labour are relatively high.

The Hungarian economy got into a slump in the fourth quarter of 2008, as a result of the recession of the world economy that began to improve only in the second half of 2009. The nadir in 2009 brought about a decrease in living standard, it also increased income discrepancies. The setback of the internal demand, which was

Judgement of employment by the methods of systematic data collection

Dr. Zsuzsanna Dabasi Halász[1] *– Katalin Lipták*[2] *– Eszter Siposné Nándori*[3]

Our research aims at examining the opinion of the Hungarian population about employment. We would like to define what factors are mainly associated with employment and what the exact meaning of these factors is. To carry out our research we have used the techniques of systematic data collection.

This method, developed in the 1980s by anthropologist Susan C. Weller, and mathematician A. Kimball Romney in the United States, help researchers in the social sciences collect better interview or questionnaire data. This slowly spreading qualitative research method of systematic interviewing – where each informant is asked the same set of questions – diminished the sample size required in social science research in a revolutionary way.

In our paper we present a research overview about how to examine employment with the method of systematic data collection.

Key words: systematic data collection, interview techniques, employment, sampling

1. Introduction

In spite of many favourable macro-economic processes the man in the street feels that the nadir of global recession is still ahead. The Eurobarometer survey conducted in the EU 27 member states indicates that citizens regard unemployment as the main menace. At the same time, their anxieties concerning the economic situation have become somewhat less.

That is why communicating the objectives and tools of employment policy in the most comprehensible way possible receive heightened significance. It is important that the man in the street understands and makes use of the opportunities and

[1] Dr. Zsuzsanna Dabasi Halász Ph.D., assistant professor, University of Miskolc, Faculty of Economic, World- and Regional Economic Institute
[2] Katalin Lipták, Ph.D. student, University of Miskolc, Faculty of Economic, World- and Regional Economic Institute
[3] Eszter Siposné Nándori, junior assistant professor, University of Miskolc, Faculty of Economic, World- and Regional Economic Institute

The research of Katalin Lipták has been supported by the "Közösen a Jövő Munkahelyeiért" Foundation.