Statistika

Vol. 48 (2) 2011

Economy and Statistics Journal

ČSÚ published by the Czech Statistical Office

ISSN 1804-8765 (Online)
ISSN 0322-788X (Print)

EDITOR-IN-CHIEF

EDITORIAL BOARD

ASSOCIATE EDITORS

MANAGING EDITOR

CONTENTS

ANALYSES

METHODOLOGIES

About Statistika

The journal of Statistika has been published by the Czech Statistical Office since 1964. Its aim is to create a platform enabling national statistical and research institutions to present the progress and results of complex analyses in the economic, environmental, and social spheres. Its mission is to promote the official statistics as a tool supporting the decision making at the level of international organizations, central and local authorities, as well as businesses. We contribute to the world debate and efforts in strengthening the bridge between theory and practice of the official statistics. Statistika is published quarterly in English only.

Publisher

The Czech Statistical Office is an official national statistical institution of the Czech Republic. The Office main goal, as the coordinator of the State Statistical Service, consists in the acquisition of data and the subsequent production of statistical information on social, economic, demographic, and environmental development of the state. Based on the data acquired, the Czech Statistical Office produces a reliable and consistent image of the current society and its developments satisfying various needs of potential users.

Contact us

Journal of Statistika | Czech Statistical Office | Na padesátém 81 | 100 82 Prague 10 | Czech Republic
e-mail: statistika.journal@czso.cz | web: www.czso.cz/statistika_journal

Factors Influencing the Rating of Regional Economic Performance

or Reasons why Prague has Become the 6th Best Economically Performing Region of the EU

Miloslav Chlad[1] | *Czech Statistical Office, Prague, Czech Republic*
Jaroslav Kahoun[2] | *Czech Statistical Office, Prague, Czech Republic*

Abstract

Regional Gross Domestic Product per capita is a key indicator for the distribution of financial resources within the Structural Funds in the EU. With regard to this fact, an increasing attention is also paid to this indicator in the Czech Republic (especially after its accession to the EU). Regional GDP per capita is often incorrectly presented as the indicator of economic well-being of the population residing in the region. Some factors with significant impact on the value of regional GDP per capita and/or the ranking of regions in this international comparison are neglected. These factors include, for example, commuting to work, NUTS classification and/or units of measure (PPS versus euro) etc. The analysis shows the influence of these factors.

Keywords

Gross domestic product (GDP), nomenclature of territorial units for statistics (NUTS), purchasing power standards (PPS), purchasing power consumption standards (PPCS), net disposable income of households, exchange rate deviation Index (ERDI)

JEL code

E01, R11, R13, R19

INTRODUCTION

On 24 February 2011, the Statistical Office of the European Commission (Eurostat) issued a press release on the generated gross domestic product (GDP) per capita for the NUTS 2[3] regions of the European Union member countries in 2008.

Although the limits of the indicator in terms of its content and interpretation have been described rather unambiguously in the conclusions of discussions held particularly on international conferences[4],

[1] Quarterly and Regional Accounts Section, e-mail: mchlad@czso.cz.

[2] Quarterly and Regional Accounts Section, e-mail: jaroslav.kahoun@czso.cz.

[3] The NUTS 2 regional level is considered and recommended as the basis for statistical comparisons, which are used when re-distributing of resources within Structural funds.

[4] E.g. OECD, Forum on the Role of Statistics in Democracies, Palermo, 10.–13.11.2004 a 2nd World Forum — Measuring and Fostering the Progress of Societies, Istanbul, 27.–30.6.2007; Beyond GDP: Measuring progress, true wealth and the well-being of nations, Brussels, 19.–20.11.2007.

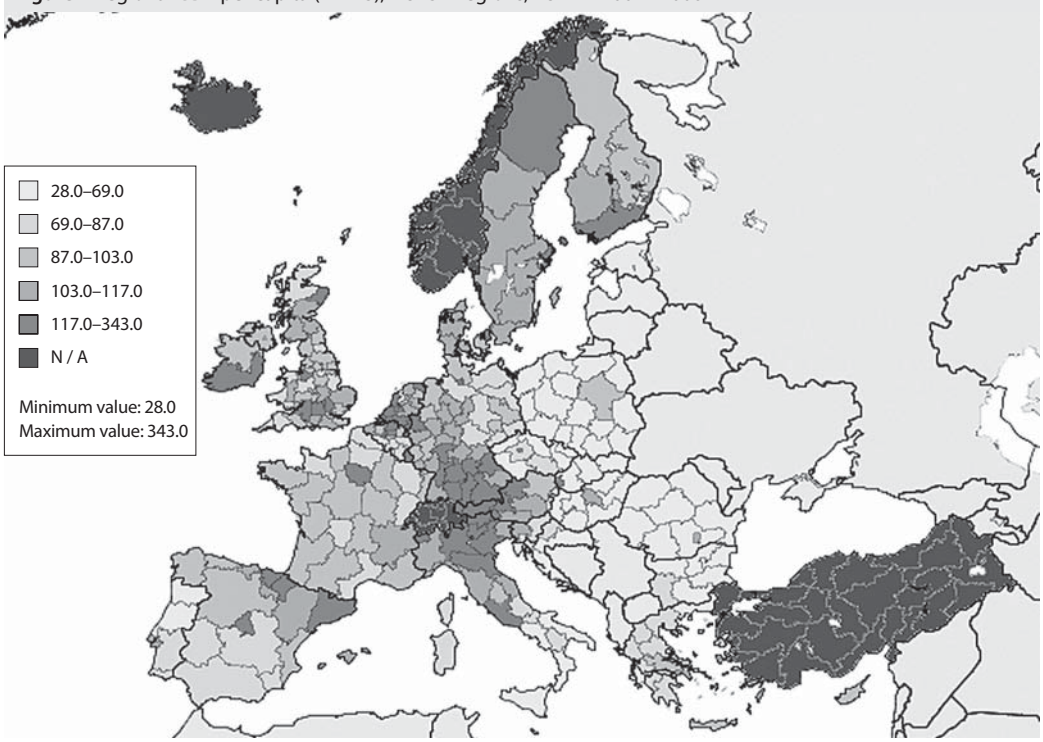
GDP remains to be the most commonly used indicator to measures the economic performance of an economy. It has become and in the upcoming years will certainly continue to be the basis for political and economic decisions.

When being interpreted in the media, statistical data are very often presented incorrectly (most commonly, generation of GDP is mistaken for the well-being of the region), which may consciously or unconsciously shape public opinion as well as the decisions of competent authorities concerning further development of a region, etc. Given these facts, the data should be accompanied by methodical explanations, commentaries and other information.

1 EU REGIONAL POLITICS

Regional politics is a very broad term, in a broader context it can be characterized as politics concentrating on the development of regions, i.e. the growth of socioeconomic and environmental potential and competitiveness of regions leading to the increase of well-being and quality of life of its inhabitants (MMR, 2006). EU regional politics is aimed at balancing, i.e. it is based mainly on the solidarity principle — supporting regions which have not benefited from the integration process so far and also regions that have been negatively affected by this process. Pursuing this approach the EU provides support to underdeveloped regions, supports the conversion of troublesome industrial areas, implementation of research and scientific projects, modification of agriculture and redevelopment of towns and municipalities. In this way, EU allows all regions to contribute to the increase of the European Union's competitiveness. The aim of EU regional politics is to strengthen economic and social solidarity between the Communities, mainly through measures in the area of regional, structural, social and agricultural

Figure 1 Regional GDP per capita (in PPS), NUTS 2 regions, EU27 = 100 in 2008

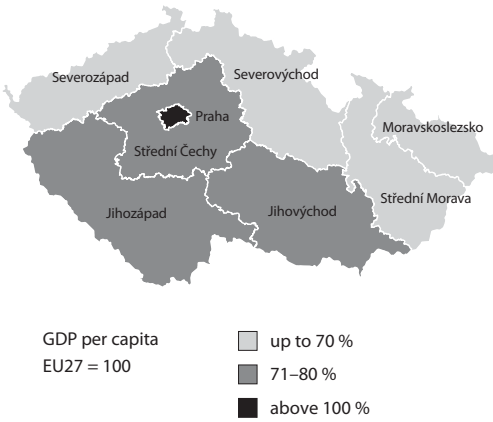


Source: Eurostat

politics and measures supporting employment. When providing aid to underdeveloped regions, it is necessary to define its objective and also select the eligible regions. These "problematic" regions are determined on several levels, however, it can be said that EU regional politics focuses mainly on NUTS 2 administrative regions.

Figure 1 shows the disparities of the proportion of GDP per capita in PPS for NUTS 2 regions in EU member and candidate countries in 2008. Figure 2 a 3 present the results for 2009 for Czech NUTS 2 and NUTS 3 regions compared to the EU27 average.

Figure 2 GDP per capita according to NUTS 2 regions, EU27 = 100, year 2009



Source: Czech Statistical Office, own figure

Figure 3 GDP per capita according to NUTS 3 regions, EU27 = 100, year 2009

Source: Czech Statistical Office, own figure

2 EVALUATION OF NUTS 2 CAPITAL REGIONS

Eurostat's periodical press releases on regional GDP only compare the position of regions pertaining to the NUTS 2 classification level, which is crucial for the allocation of resources from EU structural funds. In most EU member countries, regional self-administration units correspond to this level.[5]

Different results would be obtained in a comparison of NUTS 3 regions, which correspond to the regional self-governing units of the Czech Republic (kraje — regions). In this group, Prague ranked the 47th in 2008, compared to the 6th position among the NUTS 2 regions (Bratislava region — 56th position compared to the 9th position at the NUTS 2 level). The reason for this significant difference is mainly the fact that "purely urban regions" of the EU countries are normally classified at the *NUTS 3 regional level*. Similarly as Prague, Brussels, Vienna, and Berlin are classified at both NUTS 2 and NUTS 3 regional level. These cities have the nature of "purely" urban regions. On the contrary, the capitals of Bratislava, Stockholm, Madrid, Athens and Valletta are also classified as NUTS 2 and NUTS 3 regions, however, unlike Prague; these regions also include their surroundings. A third specific group is represented by Nicosia and Luxembourg, which are not classified according to the NUTS classification due to their size and only the respective national data are applied to them.

[5] The quoted Eurostat press release on regional GDP in EU27 regions can be viewed at the Eurostat Statistical Office web site: <http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/1-24022011-AP/EN/1-24022011-AP-EN.PDF>, or at the Czech Statistical Office web site: <http://apl.czso.cz/pll/rocenka/rocenka.indexnu_reg>.

Table 1 Regional comparison of the area and number of inhabitants

Territory	NUTS2			NUTS3		
	$\bar{x}$	x_{min}	x_{max}	$\bar{x}$	x_{min}	x_{max}
EU27						
Area in km²	15 869	13	153 439	3 300	13	98 249
Number of inhabitants in thousands	1 819	27	11 360	378	10	5 218
Czech Republic						
Area in km²	9 658	485	17 068	5 519	485	10 808
Number of inhabitants in thousands	1 281	1 127	1 641	732	304	1 252
Capital city of Prague						
Area in km²	485	485	485	485	485	485
Number of inhabitants in thousands	1 176	1 176	1 176	1 176	1 176	1 176

Source: Eurostat

In general, we can say that the territorial units classified at NUTS 2 level are smaller than the EU27 average in their size and number of inhabitants, which is one of the causes of the extraordinary position of Prague among NUTS 2 regions in the EU, see Table 1.

Table 2 Comparison of EU27, Czech Republic and the Capital of Prague

Territory	in the years						
	1995	2000	2004	2005	2006	2007	2008
EU27							
GDP per capita in PPS	14 700	19 100	21 700	22 500	23 700	25 000	25 100
Czech Republic							
GDP per capita in PPS	10 700	13 000	16 300	17 100	18 200	19 900	20 200
Capital city of Prague							
GDP per capita in PPS	18 300	26 000	33 400	35 600	38 300	42 800	43 200

Source: Eurostat

Table 2 compares the average GDP per capita in PPS for EU27, the Czech Republic and the Capital city of Prague, showing the long-term development from 1995 to 2008. The data are taken from the EUROSTAT's database. If we wish to make a more homogeneous comparison of NUTS 2 regions, i.e. comparing for example only regions which include the capital of the given country, the ranking of regions and member countries would be as shown in Table 3.

Table 3 shows the ranking of regions and member countries according to the economic performance of these NUTS 2 regions. In the case of Estonia, Cyprus, Lithuania, Latvia, Luxembourg and Malta, the data are equal to the data obtained at the national level; in these cases, the NUTS 1 and NUTS 2 regions are identical territorial units. In the group of regions defined in this way, Prague is the fourth most successful region. In this case, however, it must be taken into account that individual regions according to the NUTS classification are very different in terms of their size and number of inhabitants.

Table 4 compares urban regions corresponding to the NUTS 3 classification level. Possible differences were described above. The comparison of economic performance of regions at the NUTS 2 and NUTS 3 level shows that Prague has been overtaken by the capitals of France, the Netherlands, Ireland and Denmark. Despite this fact, the 9th position among the 27 capitals of the EU demonstrates the economic potential of our capital. It can be said that this statistics proves Prague's extraordinary position among EU regions.

Table 3 Gross domestic product per capita in NUTS 2 regions comprising the capital city (in PPS)

NUTS 2 region / period	1995	2000	2004	2005	2006	2007	2008
Inner London	41 000	59 400	72 900	75 900	80 300	83 200	85 800
Luxembourg	32 700	46 700	54 800	57 300	64 000	68 600	70 000
Région de Bruxelles-Capitale / Brussels Hoofdstedelijk Gewest	38 400	48 800	52 100	53 300	53 900	55 200	54 100
Praha	**18 300**	**26 000**	**33 400**	**35 600**	**38 300**	**42 800**	**43 200**
Île de France	25 800	34 400	36 700	38 600	39 600	42 400	42 000
Stockholm	24 100	33 800	37 700	38 100	39 600	42 500	41 900
Bratislavský kraj	14 900	20 700	27 900	33 000	35 000	40 200	41 800
Wien	28 500	35 300	37 500	37 800	39 600	40 700	40 900
Noord-Holland	20 800	29 500	33 100	34 600	35 900	37 900	38 200
Hovedstaden	:	:	:	35 300	36 300	37 800	38 000
Southern and Eastern	16 500	27 600	34 000	35 800	38 200	40 900	37 000
Etelä-Suomi	18 100	26 200	29 000	29 500	31 100	33 600	33 600
Comunidad de Madrid	17 600	25 200	28 700	29 900	32 300	34 000	33 500
Lazio	20 800	25 800	28 000	28 500	29 200	30 600	30 800
Bucuresti-Ilfov	6 900	10 700	14 800	17 300	19 800	23 000	28 300
Attiki	13 300	18 200	24 000	24 200	26 200	27 300	28 200
Zahodna Slovenija	12 900	18 000	22 500	23 600	24 900	26 600	27 300
Lisboa	15 700	21 700	23 500	25 200	26 100	27 400	27 200
Közép-Magyarország	10 700	16 200	22 000	23 200	24 900	25 900	26 800
EU27	**14 700**	**19 100**	**21 700**	**22 500**	**23 700**	**25 000**	**25 100**
Berlin	19 200	20 800	21 500	22 600	23 600	24 700	24 700
Kypros/Kibris	13 000	16 900	19 600	20 400	21 400	23 100	24 400
Mazowieckie	8 000	13 900	16 700	18 300	19 600	21 800	22 200
Malta	12 700	15 900	16 700	17 600	18 600	19 300	19 500
Yugozapaden	5 700	7 100	11 100	12 200	14 200	16 600	18 200
Eesti	5 500	8 600	12 400	13 800	15 600	17 300	17 000
Lietuva	5 200	7 500	10 900	11 900	13 100	14 700	15 300
Latvija	4 600	7 000	9 900	10 900	12 200	13 900	14 100

Source: Eurostat

The Czech capital has a specific position within the country and there is no other similar centre in the Czech Republic. Many countries have several natural centers of this kind (such as Hamburg, Munich, Frankfurt upon Mohan in Germany, Amsterdam, Groningen and Utrecht in the Netherlands or London, Aberdeen and Oxford in the United Kingdom). Some of these centers show better economic performance than the capitals of the respective countries, such as Germany, Italy, and in some cases also Spain. The territory of the Prague region is determined by the city boundaries and does include the near surroundings that represent the immediate catchment area. The capital of Prague is the seat of most of state administration authorities and national and international companies' headquarters. This leads to an extraordinary concentration of gross added value created by the general government sector and a major part of the services sector, including the fields of financial, insurance and telecommunications services.

This fact is influenced by many other factors (i.e. a different structure of sectors, commuter inflows and outflows from and to other regions of the Czech Republic, demographic influences, professional and

Table 4 GDP per capita in NUTS 3 regions comprising the capital (in PPS, EU27 = 100)

NUTS1, NUTS2, NUTS3	1995	2000	2004	2005	2006	2007	2008
European Union (27 countries)	**100**	**100**	**100**	**100**	**100**	**100**	**100**
Inner London	279	311	336	337	339	333	342
Paris	315	329	299	304	291	305	300
Luxembourg	222	245	253	255	270	274	279
Arr. de Bruxelles-Capitale / Arr. van Brussel-Hoofdstad	261	255	240	237	227	221	216
Groot-Amsterdam	182	202	206	208	204	203	203
Dublin	137	172	197	204	207	210	193
Byen København	:	:	:	:	177	178	179
Hlavní město Praha	**124**	**136**	**154**	**158**	**162**	**171**	**172**
Stockholms län	164	177	174	169	167	170	167
Bratislavský kraj	101	108	129	147	148	161	167
Miasto Warszawa	126	137	143	153	157	166	166
Wien	194	185	173	168	167	163	163
Uusimaa	141	165	157	155	155	160	159
Budapest	90	109	131	135	140	138	143
Madrid	120	132	132	133	136	136	133
Roma	155	150	144	140	135	132	133
Osrednjeslovenska	102	110	124	125	126	127	129
Grande Lisboa	122	131	129	133	130	130	128
Bucuresti	48	59	70	80	85	95	117
Attiki	90	95	111	108	111	109	112
Põhja-Eesti	:	66	88	93	103	106	103
Sofia (stolitsa)	48	47	66	72	81	92	100
Berlin	131	109	99	100	100	99	98
Kypros/Kibris	88	88	90	91	90	92	97
Riga	:	63	83	88	89	96	96
Vilniaus apskritis	43	54	74	79	85	92	94
Malta	:	85	78	80	80	79	79

Source: Eurostat

educational structure of the population and workforce, the fact that many crucial companies have their headquarters there, etc.). Some of these influences will be analyzed further in this article.

2.1 EVALUATION OF NUTS 2 REGIONS OF EU MEMBER COUNTRIES

The available data relating to capital regions of EU member countries clearly show that in 2008 most of these regions (19 out of 27 EU capital cities) achieved a higher GDP per capita in PPS than was the EU27 average. Table 3 of the Annex shows the existing disparities between these regions.

If we extend the analysis of NUTS 2 regions to the entire group of 271 EU regions, we obtain the following characteristics. In 2008, GDP per capita in PPS in NUTS 2 regions of EU27 ranged from 28 % of the EU27 average registered in the Bulgarian region Severozapaden to 343 % registered in Inner London in Great Britain. The factor between these extreme values is 12.1:1. In the preceding year, this ratio for the same regions was 13.1:1. Despite this strong disproportion, we can conclude that the convergence objectives are being fulfilled. The dispersion (Eurostat, 2010a) of GDP per capita in PPS for NUTS 2 regions

decreased from 32.7 % to 28.3 % for the EU27 average between years 2000 and 2007. There was practically no change in 2008. If we focused on the dispersion for each MS of the EU we cannot disregard that in most of the "old" EU15 member states the decreasing dispersion confirms the increasing convergence of economic development of NUTS 2 regions. On the contrary, the trend in the "new" member countries is the opposite. For example, in the Czech Republic, the dispersion between years 2000 and 2007 in NUTS 2 regions increased from 22.7 % onto 26.5 % (Eurostat, 2010b), which result is strongly influenced by a faster growth of GDP in Prague. The strongest increase of dispersion was recorded in Bulgaria (from 17.6 % to 35.4 %). In terms of convergence, the decrease of share of inhabitants by 4.1 percent points in the group of regions with more than 125 % of the average GDP per capita on the EU27 level is also positive. The share of inhabitants decreased also in underdeveloped regions (from 27.2 % to 24.5 %), whose GDP per capita does not exceed 75 % of the EU average (Eurostat, 2010b).

3 NON-EXISTENCE OF REGIONALLY DISTINGUISHED PURCHASING POWER PARITY

The comparison between regions is limited by the fact that although the PPS excludes the influence of the differences between price levels in individual countries, it does not, *however, take into consideration the difference between price levels in the regions* within individual countries. These differences between price levels in capitals and other regions are significant, mainly due to the price of rentals and some other kinds of services. This leads to a situation where capitals usually produce a nominal gross added value which is higher than the value that may actually be used at the given place (there is a transfer of income from the region of production to the place of actual use), and therefore, when compared with other regions, the real income of the inhabitants is lower than it may seem. Higher prices of selected items are compensated by higher income.

The Czech Republic is a typical example of this situation, as its capital (due to many specific particularities compared to other regions of the Czech Republic) is a separate NUTS 2 region delimited by the city boundaries. If the city's wider surroundings were included, the inter-regional price difference would not be so significant (due to a dilution). At the same time, Prague is also defined as a NUTS 3 and LAU 2 (Local Administrative Unit) region at the level of municipalities, which is in accordance with Regulation EC 1059/2003.

An important factor that can also significantly affect the ranking of regions in terms of their economic performance would be the use of another unit, i.e. expressing the GDP per capita in Euros instead of PPS. The Eurostat report presents this indicator not only in PPS, but also in Euros. Possible changes can be determined by comparing the development of the exchange rate of the Czech Crown to Euro or as the respective coefficient to PPS. By comparing the exchange rates and the PPS coefficient we obtain another indicator (ERDI — Exchange Rate Deviation Index), which reflects the relationship between the purchasing power and the exchange rate, i.e. the level of underestimation or overestimation of the national currency exchange rate.

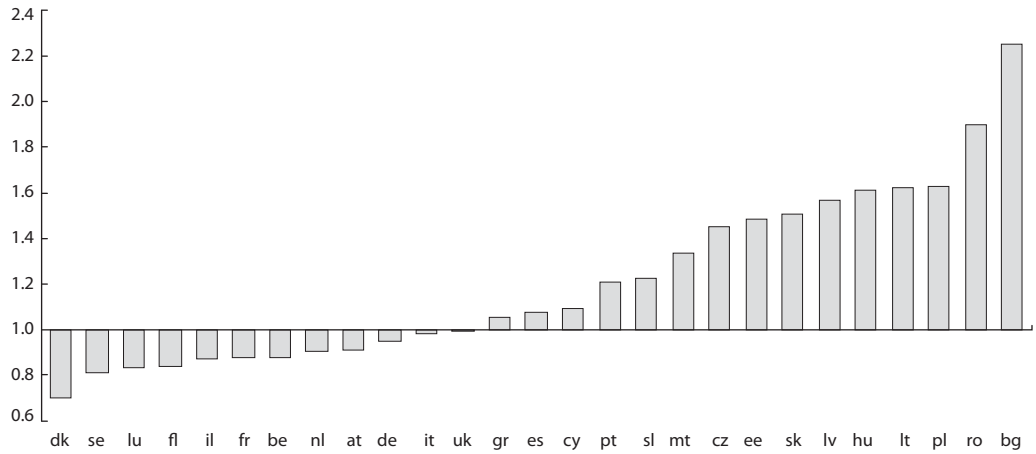
Table 5 Development of the exchange rate and purchasing power standard in relation to CZK

	1995	2000	2004	2005	2006	2007	2008	2009	2010
Exchange rate CZK / EUR	34.6960	35.5990	31.8910	29.7820	28.3420	27.7660	24.9460	26.4350	25.2840
Coefficient CZK / PPS	13.2005	16.3432	16.9600	17.0961	17.2313	17.1703	17.5384	17.9382	17.4258
ERDI	2.62839	2.17822	1.88037	1.74204	1.64480	1.61710	1.42237	1.47367	1.45095

Source: Eurostat (exchange rate and PPS), ERDI – own calculation

Figure 4 shows the comparison of ERDI of EU27 member states for 2010. Although the value of this indicator for the Czech Republic has decreased, it is still obvious that the exchange rate of the Czech

Figure 4 ERDI in 2010 for EU27 countries



Note: Share of the exchange rate and purchasing power standard.
Source: Own calculation

Crown remains to be significantly overestimated in comparison to its purchasing power as a consequence of the measures adopted by the federal government consisting in three devaluations in the 1990s. The development of the Czech Crown's exchange rate demonstrates that there remains a significant space for consolidation and approximation to its actual purchasing power.

4 INTER-REGIONAL COMMUTER INFLOWS AND OUTFLOWS

The fact that regional disparities may be significantly affected by commuter inflows and outflows is mentioned by Eurostat (2011a) in its press release reporting the data on regional GDP for NUTS 2 regions: *"It should be noted, however, that in some regions the GDP per capita figures can be significantly influenced by commuter flows. Net commuter inflows in these regions push up production to a level that could not be achieved by the resident active population on its own. The result is that GDP per capita appears to be overestimated in these regions and underestimated in regions with commuter outflows."* The data obtained in a selective workforce survey showed that in Prague in 2008, inflowing commuters accounted for 18.5 % of all persons employed in the capital.

Exact data on commuter inflows and outflows of economically active inhabitants to work and pupils and students to schools are only gathered within the census of population, houses and flats. Their values are influenced by social, age and professional structure of the population, the rate of economic activity, type and distribution of job opportunities and the nature of the residential structure and infrastructure in the region. In the period between censuses, these data are not updated at the nationwide level. The informative value of these data decreases with the time lapsed from the end of the census and with the increasing dynamics of the development of regional economies. Given these facts, we use the output of the selective survey of workforce, which gathers data on the residents' place of work. Based on these data it was ascertained that there is a significant flow of workforce between the Středočeský kraj (region of Central Bohemia) and the capital of Prague (the balance between commuter inflows and outflows in the Central Bohemia is around 12 % in favour of outflows, while in Prague, on the contrary, this balance ranged from 15 to 19 % in favour of inflows). Certain decrease in commuter outflow is indicated by the data collected in regions with lagging economies, such as the Karlovarský, Ústecký, Olomoucký, Zlínský and Moravskoslezský regions.

5 EMPLOYMENT RATE OF FOREIGNERS

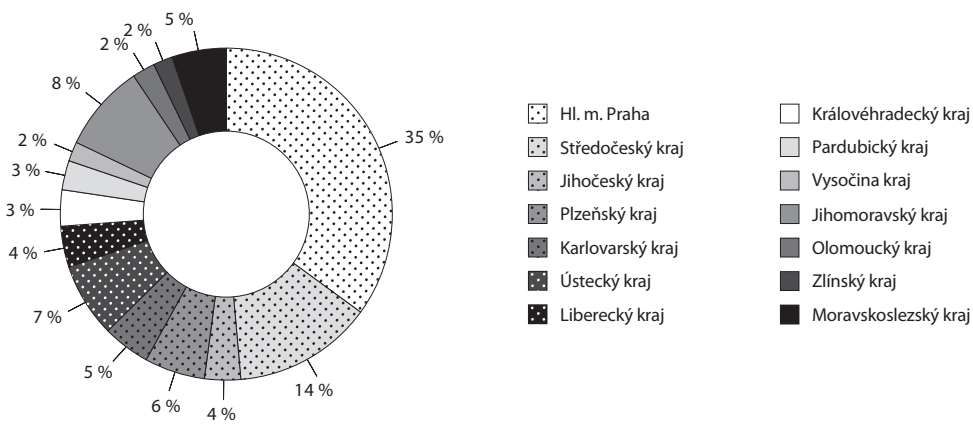
This factor is related to the previous point and its significance has been increasing mainly in last years, when the Czech Republic has seen a growth of the proportion of foreigners in its population. The increasing unbalance between regions in terms of non-resident foreigners must be also taken into account. Since 2000, the share of foreigners in the population of the Czech Republic has more than doubled, while this increase was significantly higher in Prague than in the rest of the country. The influence of foreigners and their economic activity may be one of the decisive factors for faster development of the GDP in the capital compared to the disposable household income, in which the share of Prague on the overall result of the Czech Republic has been rather stagnating. This influence was stronger

Table 6 Number of foreigners in NUTS 3 regions

	1996	2000	2005	2007	2008	2009	2010
Czech Republic	**199 151**	**200 951**	**278 312**	**392 315**	**437 565**	**433 305**	**425 301**
Hl. m. Praha	61 203	57 583	89 997	129 002	141 841	148 398	148 815
Středočeský kraj	22 413	26 993	35 304	50 273	60 123	58 544	57 815
Jihočeský kraj	9 175	8 088	10 595	15 171	16 560	27 580	15 051
Plzeňský kraj	8 670	8 913	13 206	20 986	27 636	15 429	25 198
Karlovarský kraj	7 670	10 439	14 437	19 419	20 321	19 647	19 621
Ústecký kraj	13 985	14 427	22 130	33 053	35 451	32 086	30 705
Liberecký kraj	8 446	8 794	11 675	15 288	17 320	17 359	16 831
Královéhradecký kraj	8 418	8 117	11 294	15 512	16 517	15 326	14 914
Pardubický kraj	5 335	5 648	6 418	10 562	12 588	11 981	12 077
Vysočina kraj	3 739	4 198	6 160	8 729	9 771	8 589	8 029
Jihomoravský kraj	14 723	16 813	24 234	32 606	35 619	37 050	36 107
Olomoucký kraj	6 037	6 197	7 497	10 322	9 909	9 455	9 481
Zlínský kraj	6 558	7 057	5 926	7 639	8 413	8 147	8 048
Moravskoslezský kraj	22 779	17 684	19 337	22 962	25 496	23 714	22 609

Source: Czech Statistical Office

Figure 5 Share of foreigners in regions in 2010 (in %)



Source: Own construction

at the time of economic recession in 2009, when the proportion of foreigners in Prague's population further grew (from 11.6 % to 11.9 %).

The numbers of foreigners living in individual regions of the Czech Republic are shown in Table 6 and Figure 5, which illustratet the regional structure of foreigners according to regions. The data again confirm the extraordinary position of the capital of Prague, with 35 % of the total number of 425 thousand foreigners, whose share in the total number of employed persons cannot be neglected. The 14 % of the total number of foreigners living in the region of Central Bohemia region is also worth mentioning. In the remaining regions, the share of foreigners ranged from 2 to 7 %.

Table 7 characterizes the regional differences in the share of foreigners in the overall employment rates. The extraordinary position of the capital of Prague is demonstrated again, with a 9.2 % share of foreigners in the region's overall employment rate in 2009. Other regions with higher foreigner employment rates are Středočeský, Plzeňský, Jihomoravský and Liberecký regions. On the contrary, the lowest long-term employment rates of foreigners have been reported by the Olomoucký and Ústecký regions.

Table 7 Share of foreigners in the overall employment rates (in %)

	1995	2000	2004	2005	2006	2007	2008	2009
Czech Republic	**3.6**	**3.9**	**4.5**	**4.7**	**5.4**	**4.6**	**5.4**	**4.4**
Hl. m. Praha	7.4	7.5	9.5	9.1	10.0	8.7	10.4	9.2
Středočeský kraj	5.0	5.3	5.9	6.4	7.2	6.6	8.0	6.8
Jihočeský kraj	2.2	2.3	3.1	3.2	3.6	3.0	3.2	2.2
Plzeňský kraj	2.4	2.8	4.2	5.5	7.0	7.0	7.5	4.7
Karlovarský kraj	4.1	4.8	5.8	6.3	6.7	2.7	3.2	2.6
Ústecký kraj	2.1	2.8	2.7	3.0	3.5	1.8	2.1	1.7
Liberecký kraj	3.6	3.9	3.7	4.7	5.7	4.5	4.7	4.0
Královéhradecký kraj	2.6	2.6	2.9	3.5	4.1	4.0	4.5	3.4
Pardubický kraj	2.1	2.7	2.2	3.0	4.1	5.9	6.6	3.8
Vysočina kraj	2.8	2.7	2.5	2.6	3.2	2.9	3.0	1.6
Jihomoravský kraj	3.0	3.3	3.4	3.6	4.3	3.9	5.2	4.7
Olomoucký kraj	1.6	2.0	1.7	1.9	2.0	1.5	1.6	1.2
Zlínský kraj	3.1	3.0	2.4	2.5	2.5	1.9	2.2	1.9
Moravskoslezský kraj	3.0	3.2	3.4	3.0	3.3	2.2	2.5	2.2

Source: Own calculation

6 PUBLIC RESOURCES REDISTRIBUTION RATE AND NET DISPOSABLE HOUSEHOLD INCOME

Unlike the production of gross domestic product and some factors influencing the amount of GDP, this chapter focuses more on the use of GDP, which is partially reflected in the net disposable household income indicator (NDHI) (Ježdík and Chlad, 2009). The limitations of GDP and derived indicators have been often discussed recently; special attention being paid to GDP per capita, whose informative value is limited by the above-mentioned factors, such as commuter inflow and outflow, absence of regional consumer price indexes, inter-regional transfers, etc. To evaluate the welfare of the inhabitants of regions, we selected the above-described indicator NDHI per capita with permanent residence in the given region. It is basically the balance between income and expenditure recorded on the account of secondary income distribution. With a certain degree of simplification, it can be said that this indicator characterizes the level of the wealth of households (expressed by value of their income, which does not indicate the value of property) residing in the given region. For the purpose of possible in-

ternational comparison, the net disposable household income is expressed in PPCS units (Purchasing Power Consumption Standards) based on the purchasing power derived from the final consumption of households. As some member countries do not provide the values of this indicator, the average value for the European Union is not estimated either and therefore only limited comparisons can be made for NUTS 2 regions of the countries that do record this indicator.

Table 8 Net disposable household income per capita (in PPCS)

NUTS 2 regions comprising a capital	1995	2000	2004	2005	2006	2007
London	13 875	21 114	22 602	23 834	25 332	24 733
Paris	12 914	16 558	19 020	19 545	19 704	21 072
Vienna	14 298	17 056	17 700	18 280	18 977	19 485
Athens	9 536	11 820	15 277	15 673	16 502	18 251
Madrid	9 902	14 255	15 742	16 542	17 315	17 638
Rome	12 568	14 586	16 161	16 515	16 515	17 115
Stockholm	10 657	13 334	15 145	15 517	16 182	17 071
Amsterdam	10 122	12 986	14 361	14 832	15 237	16 488
Dublin	8 329	11 711	14 308	15 146	15 600	16 293
Brussels	11 991	14 263	14 258	14 378	14 983	15 438
Berlin	12 540	13 421	14 032	14 447	14 830	15 049
Lisbon	8 641	11 674	12 852	13 733	14 111	14 374
Helsinki	7 931	10 463	12 593	12 802	13 238	14 202
Bratislava	5 050	8 005	9 983	11 867	12 195	13 749
Copenhagen	:	11 445	:	12 549	13 265	13 518
Prague	6 738	8 827	10 578	11 225	12 247	13 181
Ljubljana	:	9 712	11 415	12 016	12 498	13 065
Budapest	:	7 650	10 328	11 079	10 868	10 506
Warsaw	5 020	7 450	8 547	8 722	9 166	10 248
Bucharest	:	4 120	5 699	6 891	7 708	9 296
Vilnius	2 693	4 350	6 122	6 839	7 549	8 092
Tallinn	2 666	4 053	5 447	6 101	6 938	7 857
Riga	2 195	3 658	5 265	5 801	6 776	7 736
Sofia	:	2 695	4 229	4 250	4 495	5 541
Nicosia	:	:	:	:	:	:
Luxembourg	:	:	:	:	:	:
Valletta	:	:	:	:	:	:

Source: Eurostat

Another limitation of this indicator for international comparison consists in the fact that there are significant differences between member countries in the redistribution of resources between sectors. The regional disparities measured in this way are strongly influenced by social and other transfers carried out by the sector of governmental institutions. The effort to eliminate the differences between social levels influences also the elimination of regional disparities in the given country and vice versa (this is illustrated by the relatively neutral position of Stockholm and Vienna compared to the national aver-

age, and, on the contrary, by the strengthening position of Prague or the Bratislava region compared to the nationwide average values of GDP per capita). Unfortunately, as this indicator is not generally used as a criterion for the redistribution of resources from the European funds, member countries are not forced to remove these disparities. Given these limitations for international comparisons, the possibilities of applying this indicator are reduced. For the sake of maximum homogeneity of data, this indicator should be mainly used to compare the regions within the given country and asses their development in time.

Table 8 reveals certain disparity between the ranking of NUTS 2 regions comprising the capital of the given country depending on whether we evaluate the production of regional GDP or the net disposable household income. As it is explained above, this ranking is strongly influenced by the redistribution processes and the differences between these processes in individual EU27 countries. While Prague ranked the 4th in terms of regional GDP, in terms of net disposable household income it ranked the 16th or 17th, as the data for Luxembourg are not available.

The discrepancy between the production and use of resources is obvious also when assessing the regional values for the regions of the Czech Republic in relation to the national average. This relationship is indicated in Figure 6. The significant decrease of net disposable household income compared to the GDP in Prague is influenced mainly by commuter inflows and outflows (Kahoun, 2010).

Table 9 Comparison of regional GDP and NDHI per capita with the national average

Territory	Regional GDP per capita, CR = 100				NDHI per capita, CR = 100			
	1995	2000	2005	2009	1995	2000	2005	2009
Czech Republic	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Hl. m. Praha	170.6	199.6	208.9	220.4	123.6	133.2	135.1	129.1
Středočeský kraj	86.3	94.1	92.1	91.8	102.7	104.6	106.3	104.2
Jihočeský kraj	93.7	91.9	90.1	86.2	98.4	97.4	96.7	99.0
Plzeňský kraj	96.6	93.8	94.4	86.8	101.0	100.0	100.3	99.9
Karlovarský kraj	93.5	83.7	75.3	67.6	96.2	97.6	89.7	90.2
Ústecký kraj	94.8	81.5	81.2	79.8	95.8	91.4	88.1	90.8
Liberecký kraj	90.7	89.5	83.8	69.5	94.9	95.9	93.9	93.1
Královéhradecký kraj	93.3	94.6	87.6	84.3	100.5	100.4	97.4	96.2
Pardubický kraj	89.8	85.4	82.4	82.9	93.5	92.0	95.0	95.4
Vysočina kraj	85.5	83.8	84.8	78.3	91.4	91.1	93.8	94.4
Jihomoravský kraj	95.8	92.4	91.0	95.5	97.5	96.6	97.6	98.0
Olomoucký kraj	83.3	79.6	75.8	75.4	91.9	91.4	90.6	92.4
Zlínský kraj	89.6	83.6	80.7	83.0	94.5	93.8	93.4	93.6
Moravskoslezský kraj	87.6	78.3	84.7	81.5	95.8	90.9	91.0	93.2

Source: Czech Statistical Office

The standard deviation for the production of regional GDP per capita in comparison with the national average in 2009 was four times higher than for NDHI per capita for NUTS 3 regions in comparison with the national average. When excluding Prague, the variability rate shall decrease to one half. These characteristics again confirm the exclusive position of the capital of Prague in comparison with other regions of the Czech Republic.

The indicated trend — exceeding of the average national NDHI per capita by NUTS 2 regions comprising a capital city is obvious in the majority of member countries, as illustrated in the Table 10. The only exceptions in 2007 were Brussels and Berlin.

Table 10 Net disposable household income per capita (in PPCS)

Territory / period	1995	2000	2004	2005	2006	2007
Belgium	11 890	14 300	14 687	15 072	15 670	16 181
Brussels	11 991	14 263	14 258	14 378	14 983	15 438
Bulgaria	:	2 296	3 500	3 499	3 740	4 228
Yugozapaden	:	2 695	4 229	4 250	4 495	5 541
Czech Republic	5 450	6 630	7 744	8 309	8 933	9 765
Prague	6 738	8 827	10 578	11 225	12 247	13 181
Denmark	9 407	10 789	11 746	12 037	12 757	13 096
Copenhagen	:	11 445	:	12 549	13 265	13 518
Germany	12 800	15 102	16 615	17 192	17 646	18 060
Berlin	12 540	13 421	14 032	14 447	14 830	15 049
Estonia	2 666	4 053	5 447	6 101	6 938	7 857
Eesti	2 666	4 053	5 447	6 101	6 938	7 857
Ireland	8 016	11 340	13 767	14 410	14 959	15 708
Southern and Eastern	8 329	11 711	14 308	15 146	15 600	16 293
Greece	8 371	10 184	12 128	12 767	13 401	14 817
Attiki	9 536	11 820	15 277	15 673	16 502	18 251
Spain	8 318	11 467	12 981	13 718	14 349	14 793
Comunidad de Madrid	9 902	14 255	15 742	16 542	17 315	17 638
France	10 464	13 397	15 278	15 924	16 496	17 326
Île de France	12 914	16 558	19 020	19 545	19 704	21 072
Italy	11 725	13 835	14 676	15 001	15 494	16 055
Lazio	12 568	14 586	16 161	16 515	16 515	17 115
Cyprus	:	:	:	:	:	:
Kypros/Kibris	:	:	:	:	:	:
Latvia	2 195	3 658	5 265	5 801	6 776	7 736
Latvija	2 195	3 658	5 265	5 801	6 776	7 736
Lithuania	2 693	4 350	6 122	6 839	7 549	8 092
Lietuva	2 693	4 350	6 122	6 839	7 549	8 092
Luxembourg	:	:	:	:	:	:
Luxembourg	:	:	:	:	:	:
Hungary	:	5 721	7 337	7 739	8 081	8 052
Közép-Magyarország	:	7 650	10 328	11 079	10 868	10 506
Malta	:	:	:	:	:	:
Malta	:	:	:	:	:	:
Netherlands	9 525	12 281	13 582	13 988	14 423	15 569
Noord-Holland	10 122	12 986	14 361	14 832	15 237	16 488
Austria	12 489	15 291	16 565	17 442	18 345	19 022
Wien	14 298	17 056	17 700	18 280	18 977	19 485
Poland	4 071	5 791	6 703	6 844	7 247	8 095
Mazowieckie	5 020	7 450	8 547	8 722	9 166	10 248
Portugal	7 105	9 188	10 059	10 655	11 060	11 215
Lisboa	8 641	11 674	12 852	13 733	14 111	14 374
Romania	:	2 859	4 023	4 067	4 491	5 200
Bucuresti-Ilfov	:	4 120	5 699	6 891	7 708	9 296

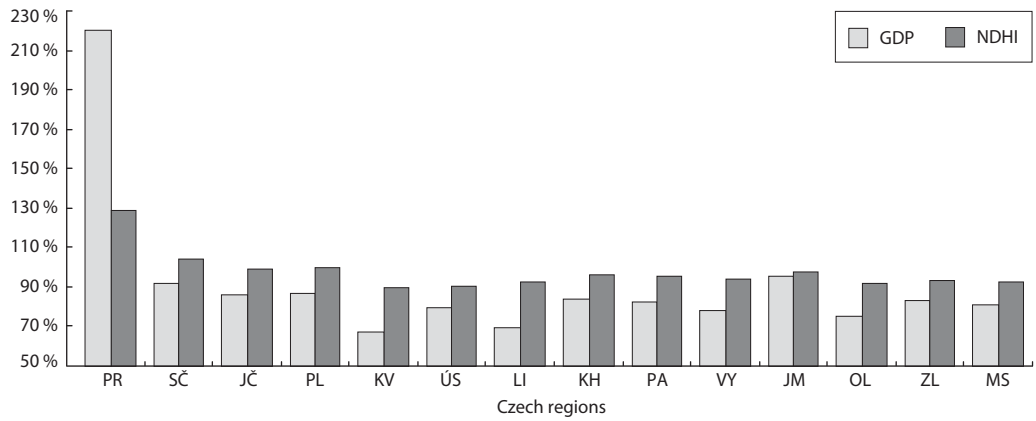
Source: Eurostat

Table 10 Net disposable household income per capita (in PPCS) — continuation

Territory / period	1995	2000	2004	2005	2006	2007
Slovenia	:	8 952	10 602	11 215	11 701	12 289
Zahodna Slovenija	:	9 712	11 415	12 016	12 498	13 065
Slovakia	3 786	5 449	6 462	7 259	7 832	8 905
Bratislavský kraj	5 050	8 005	9 983	11 867	12 195	13 749
Finland	7 620	9 803	11 781	11 964	12 480	13 454
Etelä-Suomi	7 931	10 463	12 593	12 802	13 238	14 202
Sweden	9 226	11 389	13 150	13 450	14 060	14 983
Stockholm	10 657	13 334	15 145	15 517	16 182	17 071
United Kingdom	10 889	14 561	16 731	17 219	17 726	17 440
Inner London	13 875	21 114	22 602	23 834	25 332	24 733

Source: Eurostat

Figure 6 Comparison of the GDP and NDHI per capita in 2009, CR = 100



Source: Own construction

CONCLUSION

The aim of this article was to respond to the misleading interpretation of statistical data in certain media and to demonstrate the complexity of their objective evaluation and drawing of unambiguous conclusions. Certain rules and principles must be observed to secure correct comprehension of data presented. One of these principles is to study the methodology of used indicators and the classification defining the corresponding structures before the interpretation.

Another issue covered in this article is the difference between the "generated" and "used" GDP, i.e. between the production of GDP and the net disposable household income, which is similar to the analysis of regional differences according to "wealth" in terms of the amount of income of the household sector. It is obvious from the analysis that GDP or GDP per capita reflects the rate of the overall economic activity in the given region. GDP is usually used to compare the rate of economic development of regions and not the amount of income or wealth of the residents in the region (Eurostat, 2011b).

Last but not least, the article demonstrates the extraordinary position of Prague within the Czech Republic as well as among the EU regions. On the other hand, to achieve an objective evaluation, it is necessary, taking into account the nature of the given territory, to create homogeneous groups and fil-

ter some internal discrepancies, particularly administrative, although they comply with the applicable supranational legislation.

Compared to the NUTS 2 regions, the capital city of Prague has the character of a purely urban region, without the adjacent surroundings. Similar is the case of Brussels, Vienna and Berlin, while the other NUTS 2 regions comprising a capital city, such as Paris, Madrid and Budapest, include the city surroundings. This fact leads to lower commuter inflows in the capital city, and therefore the GDP per capita in such regions is lower than that of Prague. In addition, the inclusion of the wider surroundings in the capital region may also affect the structure of the produced gross added value. A significant change is observed if we compare NUTS 3 regions, i.e. territorial units with higher homogeneity — regions having urban character. In these comparisons, Prague falls from the 6th position among the NUTS 2 regions to the 47th position, or from the 4th to the 9th position if we only compare regions comprising the capital of the given member country. Last but not least, significant changes may take place if we change the measuring units, i.e. if we express the GDP per capita in Euros instead of purchasing power standard units. In this case, the ranking will be influenced by the level of underestimation or overestimation of the national currencies exchange rate to Euro compared to the standard of the national currency.

ANNEX | Methodical Definition of Selected Terms and Indicators

Classification of territorial units for statistical purposes — NUTS

Geographic division of a territory is the basis for all regional analyses of phenomena and processes. It is used to secure a single unified structure of territorial units in the EU countries. The "Classification of Territorial Units" has been proposed and gradually elaborated for statistical purposes. Individual levels of this territorial classification represent certain size groups of territorial units. Their size is usually given by the number of inhabitants and square area. Until 2003, the classification of territorial units within the NUTS was a result of a bilateral "gentlemen's" agreement between member or candidate countries and EUROSTAT. The regulation distinguishes between two basic types of territorial units — *administrative*[6] and other, *non-administrative units*. When defining individual levels, the administrative division of a state is preferred, applying the complementarities principle, meaning that higher levels consist of a certain number of lower-level units and a group of all regions of one level covers the entire territory of the state. The basic structure of the NUTS should be the state's territory divided without remains. Individual regional levels correspond to the respective degrees of territorial administrative division for regional levels NUTS 1 to NUTS 3:

- *NUTS 1* — territorial units corresponding to major regions, smaller states, macro regions, federal states of countries like Germany or autonomous regions;
- *NUTS 2* — generally the medium administration level, which also serves as the basis for application of EU regional policies;
- *NUTS 3* — usually the lowest level of territorial administration, usually corresponding to the division of the given state, e.g. groups of districts or regions of a state (Chlad, 2007).

In 2003, the principles of a unified methodology for establishing the European NUTS classification were issued in the form of *Regulation (EC) No 1059/2003 of the European Parliament and of the Council.* The European NUTS classification is binding upon all EU member countries. Based on the said regulation 1059/2003, EUROSTAT generated the European NUTS classification. Its relatively fixed criteria should not exclude the existence of other classifications[7] and possibly of another more detailed division of territorial units.

The NUTS classification was created mainly to satisfy the needs of statistical data users, particularly in order to harmonize the data available in the European Union and ensure their comparability. This norm affects the collection, processing, transmission, publishing and presentation of statistical data of individual member countries and Communities. Therefore it became the prerequisite for spatial comparison of regional statistics, which are the basic feature of the European statistical system and serve for a wide range of purposes. One territorial unit can be classified at several NUTS levels. However, two different territorial units cannot be classified under the same name at one NUTS level. Member countries may have more NUTS levels at their own discretion.

---

[6] European Commission — Methods and Nomenclatures — Regions Nomenclature of territorial units for statistics NUTS — 2003.

[7] For example LAU — Local Administrative Unit, which uses two levels, of which LAU 2 is particularly important as it represents the level of municipalities.

For the sake of functionality of this classification, it is necessary to define regions based on relatively objective criteria, in order to ensure impartiality when processing and applying regional statistics. The current administrative units of member states are the first criterion for the definition of territorial units. In this context, "administrative units" are geographical areas with an administrative body which has the competence to adopt administrative or political decisions for this area within the legal and institutional framework of the given member country.

The NUTS classification is limited to the economic territories[8] of member states, which also includes the so-called extra-regio territories consisting of aerial space, territorial waters and continental shelves, territorial enclaves, particularly embassies, consulates and military bases, and reservoirs of oil and natural gas in international waters used by resident units.

Changes to the NUTS classification are not made more often than every three years. When changing the NUTS classification, the given member state presents to the Commission the time series of the new regional division replacing the previously submitted data. The list of time series and their duration is determined using a regulatory procedure, considering whether they may be submitted at all. These time series must be submitted within two years following the changes made to the NUTS classification.

Gross domestic product (GDP) is the monetary representation of the total value of goods and services newly created in the given period in a certain territory; it is used to assess the performance of an economy. It can be calculated using the following three methods:

- production method,
- expenditure method,
- and income method.

Gross domestic product per capita[9]

Economic performance of a territorial unit is usually characterized by the formation of gross domestic product per capita. This indicator represents the relationship between the total volume of formed GDP and the number of all inhabitants residing in the given territory (regardless of their age). The advantage of this indicator is the fact that it uses the number of inhabitants which is relatively easy to detect also in international context. Its disadvantage subsists in the fact that it also includes the performance of citizens commuting to work in the given region and contributing to the production of GDP and does not take into account the number of commuters outside the region.

The GDP per capita shows the productivity or economic performance of the given region. This indicator was adopted as the basis for international comparisons and also as a criterion for the redistribution of resources from EU structural funds, favouring underdeveloped regions (Ježdík and Chlad, 2009).

Dispersion of regional gross domestic product (GDP) per capita[10]

The dispersion of regional GDP (at NUTS level 2 and NUTS 3) is measured by the sum of the absolute differences between regional and national GDP per capita, weighted with regional share of population and expressed in percent of the national GDP per capita.

The indicator is calculated from regional GDP figures based on the regional accounts of the European System of Accounts (ESA95).

---

[8] See Commission Decision 91/450/EEC.

[9] <http://www.czso.cz/csu/redakce.nsf/i/hruby_domaci_produkt_(hdp)>.

[10] Eurostat (2011b).

Gross disposable household income per capita

Gross disposable household income represents the part of formed gross domestic product that households have available for their use (final spending and savings) and characterizes an important aspect of the population's quality of life. Sometimes this indicator related to inhabitants is used to express the level of the inhabitants' economic well-being. The amount of net disposable household income depends on a number of factors, particularly on the volume of formed GDP and the method of its distribution in the form of primary and secondary income.

Gross disposable household income is the result of the balance of income and expenditure recorded on the secondary income distribution account. It reflects how the net primary income (compensation to employees, mixed income, operating surplus and income from ownership) is allocated through redistribution in the form of taxes, social benefits and contributions and other usual transfers (Kahoun, 2010).

Although gross disposable household income per capita is a more accurate indicator of the economic well-being of the capitals of individual regions, it is used much less often than GDP per capita. This limited use is mainly due to the different rates of redistribution in individual EU member countries, which illustrates the fact that the area of taxes and social transfers is not harmonized and individual countries are not willing to unify these areas.

Purchasing power parity (PPP) — Purchasing power standard (PPS)[11]

The differences between GDP values in individual countries, also after converting them using the respective exchange rates to euro, cannot be attributed only to different volumes of goods and services. The "level of price" is another major factor. Exchange rates are influenced by many factors relating to the demand and offer on currency markets, foreign trade, inflations prognoses and interest differential. Conversion using exchange rates has only a limited significance for international comparison. To be able to obtain a more accurate comparison by individual countries, it is necessary to use special conversion coefficients to remove the impact of price levels. Purchasing power parity is a conversion coefficient converting GDP as an economic indicator expressed in national currencies to an artificial common currency, the so-called purchasing power standard (PPS). It is used to convert GDP and other economic aggregates (i.e. spending on the consumption of certain groups of products) from different countries to comparable volumes of spending expressed in the purchasing power standard (PPS).

After the introduction of euro, we can compare the prices in individual countries of the Eurozone. However, the purchasing power of euro in individual countries inside the Eurozone is different, which is related to the level of prices in individual countries. Therefore, PPP must be used continually in the Eurozone to calculate the value of macroeconomic aggregates for member states in PPS.

To put it simple, PPP represents the price of one type of goods or services in the national currencies for individual countries (e.g. a loaf of bread costs EUR 2.30 in France, EUR 1.90 in Germany, GBP 2.40 in Great Britain, etc.). Consumer baskets with comparable goods and services are used in the price statistics of consumer prices. They are selected so as to represent an entire range of goods and services, taking into account the structure of consumption in different countries. These simple rates of prices of goods and services are then aggregated in PPP according to groups of products, then for the total consumption and finally also for GDP. In order to obtain comparable values for the calculation of purchasing power parity, one country is usually selected and used as a reference country for which the value is equal to 1. This selection of a single country is not an optimal method for the European Union. Therefore PPS is an artificial common currency unit of comparison used in the

[11] Eurostat (2010b).

European Union to express the volume of economic aggregate indicators for the purpose of spatial comparisons in real values.

Unfortunately, due to the high costs, it will not be possible to consider and use regional conversion coefficients in the near future. If such regional parities of purchasing power were available, then the GDP in PPS for many periphery or rural areas of the EU would be probably higher than the GDP calculated using purchasing power parity at the national level.

Regions may be evaluated differently when expressing the value of indicators in PPS instead of euro. For example, in 2007, the Swedish region Östra Mellansverige achieved a GDP per capita of 31 300 EUR, thus outperforming Madrid, whose GDP per capita was 30 600 EUR. However, when expressed in PPS, the GDP per capita in Madrid was 34 100, which is more than the 26 500 recorded in Östra Mellansverige.

The macroeconomic aggregate of GDP per capita expressed in PPS is a key indicator to determine whether the given region will receive aid within the EU structural policies at the NUTS 2 regional level.

Purchasing power standard defined according to the final household consumption

Purchasing power standard defined according to the final household consumption, abbreviated as PPCS, is analogical to the purchasing power standard (PPS), which is used to compare the regional values of the macroeconomic aggregate of gross domestic product (GDP). It is also an artificial conversion unit for the calculation of final consumption expenditure of households.

Purchasing power parities and international volume comparisons

The differences in GDP values between countries, even after conversion by means of exchange rates to a common currency, cannot be attributed solely to differing volumes of goods and services.

The 'level of prices' component is also a major contributory factor. Exchange rates are determined by many factors related to demand and supply in the currency markets, such as international trade, inflation forecasts and interest rate differentials. Conversions using exchange rates are therefore of only limited relevance for international comparisons. To obtain a more precise comparison, it is essential to use special conversion rates which eliminate the effect of price-level differences between countries. Purchasing power parities (PPPs) are conversion factors of this kind which convert economic indicators from national currencies into an artificial common currency, called the purchasing power standard (PPS). PPPs are therefore used to convert GDP and other economic aggregates (e.g. consumption expenditure on certain product groups) of various countries into comparable volumes of expenditure, expressed in purchasing power standards.

With the introduction of the euro, prices can now, for the first time, be compared directly between countries in the euro area. However, the euro has different purchasing power in the different countries of the euro area, depending on the national price level. PPPs must therefore also continue to be used to calculate pure volume aggregates in PPS for the Member States within the euro area.

In their simplest form, PPPs are a set of price ratios between the prices in national currency of the same good or service in different countries (e.g. a loaf of bread costs EUR 2.30 in France, EUR 1.90 in Germany, GBP 2.40 in the UK, etc.). A basket of comparable goods and services is used for price surveys. These are selected so as to represent the whole range of goods and services, taking account of the consumption structures in the various countries. The simple price ratios at product level are then aggregated to PPPs for product groups, then for overall consumption and finally for GDP. In order to have a reference value for the calculation of PPPs, one country is usually chosen and used as the reference country, and set to 1. For the European Union, the selection of a single country as a base is inappropriate. Therefore, PPS is the artificial common reference currency unit used in the European Union to express the volume of economic aggregates for the purpose of spatial comparisons in real terms.

Unfortunately, for reasons of cost, it will not be possible in the foreseeable future to calculate regional conversion factors. If such regional PPPs were available, the GDP in PPS for numerous peripheral or rural regions of the EU would probably be higher than that calculated using national PPPs.

The regions may be ranked differently when calculating in PPS instead of euros. For example, in 2007 the Swedish region of Östra Mellansverige had per capita GDP of EUR 31 300, putting it well ahead of Madrid at EUR 30 600. However, in PPS, Madrid at 34 100 PPS per capita is ahead of Östra Mellansverige at 26 500 PPS per capita.

In terms of distribution, the use of PPS rather than the euro has a levelling effect, as countries with a very high GDP per capita also generally have relatively high price levels. The range of GDP per capita GDP in NUTS level 2 regions in the EU27 thus falls from 93 400 in euros to 76 900 in PPS.

Per capita GDP in PPS is the key variable for determining the eligibility of NUTS level 2 regions under the European Union's structural policy.

References

CZECH STATISTICAL OFFICE. *Co ovlivňuje pořadí ekonomické výkonnosti regionů aneb příčiny postavení Prahy jako 6. ekonomicky nejvýkonnějšího regionu EU* [online]. Prague: Czech Statistical Office, 2011. <http://apl.czso.cz/nufile/Postaveni_Prahy_mezi_regiony_Evropy.pdf>.

EUROSTAT. *Eurostat Quality Profile, Dispersion of regional gross domestic product per inhabitant (Indicator definition)* [online]. Luxembourg: Eurostat, 2010a. <http://epp.eurostat.ec.europa.eu/portal/page/portal/sdi/files/Dispersion%20of%20regional%20GDP.pdf>.

EUROSTAT. *Methodologies and working papers. Regions in the European Union / Nomenclature of territorial units for statistics NUTS 2006/EU27* [online]. Luxembourg: Eurostat, 2007. <http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-020/EN/KS-RA-07-020-EN.PDF>.

EUROSTAT. *News release, GDP per inhabitant ranged from 28 % of the EU27 average in Severozapaden in Bulgaria to 343 % in Inner London* [online]. Luxembourg: Eurostat, 2011a. <http://epp.eurostat.ec.europa.eu/cache/ITY_PUBLIC/1-24022011-AP/EN/1-24022011-AP-EN.PDF>.

EUROSTAT. *Eurostat regional yearbook 2010* [online]. Luxembourg: Eurostat, 2010b. <http://epp.eurostat.ec.europa.eu/portal/page/portal/publications/regional_yearbook>.

EUROSTAT. *Sustainable development indicators, Socio-economic development — Eurostat Quality Profile* [online]. Luxembourg: Eurostat, 2011b. <http://epp.eurostat.ec.europa.eu/portal/page/portal/sdi/indicators/theme1>.

CHLAD M. *Regionální aspekty makroekonomických ukazatelů.* Prague: University of Economics, 2007.

JEŽDÍK V. and CHLAD M. *Rozdíly mezi kraji ČR — vývoj a příčiny* [online]. Prague: Czech Statistical Office, 2009. <http://czso.cz/csu/2009edicniplan.nsf/p/1382-09>.

KAHOUN, J. *Regionální disparity v ČR — HDP versus disponibilní důchod.* Prague: CES, 2010.

MINISTRY OF REGIONAL DEVELOPMENT. *Strategie regionálního rozvoje — Zásady regionální politiky ČR.* Prague: Ministry of Regional Development, 2006.

Price Setting Behaviour in the Czech Republic, Micro Data Evidence

Robert Murárik[1] | *Czech National Bank (CNB), Prague, Czech Republic*

Abstract

The aim of this analysis was to obtain information on the strategies of retailers of consumer goods and services in terms of changes to final prices. From detailed data on many price changes in all the monitored stores we have evaluated, for example, how often the prices of specific items change or rise or fall, and by how much on average, how these indicators change during the year, whether downwards price rigidity exists and so on.

The average price change frequency for all the selected items came to 0.26, which means that approximately one in every four prices was changed compared to the month before. A typical characteristic of the prices of regulated items was that these prices mainly rose and this usually by steps of higher percentages and mostly at the start of the year. Food prices had a higher frequency of price changes, mainly in the case of unprocessed foods, which is a consequence of the volatile development of the prices of agricultural commodities. The prices of tradables excluding food and fuels continuously fell for the whole of the monitored period and only had a low price change frequency. The prices of non-regulated non-tradables continuously and smoothly rose and, with the exception of hypothetical rent and package holidays, this subgroup demonstrated the lowest frequency of price changes. The prices of fuels changed most frequently while on average these price changes were the lowest in size, as they react relatively quickly to changes in the prices of raw materials and the koruna exchange rate.

Keywords

Price setting, micro data, price frequency, rigidity, price changes, individual consumer price

JEL code

D40, M31

INTRODUCTION

The aim of this analysis was to obtain information on the strategies of retailers of consumer goods and services in terms of changes to final prices. Our approach was based on similar analysis, calculated on the data of Spain (Álvarez, 2004) and Portugal (Dhyne, 2005). For this purpose we performed a detailed analysis of an extensive database describing price movements for individual items in individual stores. For this purpose, the Czech Statistical Office provided us with the necessary data on all the prices found for the selected items in all the stores where it had obtained such data.

[1] Na Příkopě 28, Prague 1, Czech Republic, e-mail: robert.murarik@cnb.cz.

In view of the significant quantities of data and calculations we selected only part of all the items in the consumer basket. However, we did this in such a way as to ensure that this selection best represented the whole consumer basket and its subgroups. We selected such a quantity of items in the consumer basket that represented 60 % of the constant weight of the whole consumer basket as well as all evaluated consumer basket subgroups.

We first quantified and evaluated the necessary statistics for all the selected items on an individual basis and then we merged these results using their weight in the consumer basket into various groups and evaluated the results for those groups. However, to keep this text to a manageable size, we will mainly present only the results for the whole consumer basket and its main subgroups (as understood by the Czech National Bank).

We used data from the period from January 2001 to December 2005 in the form of prices in koruna for the relevant weight or volume unit. During this time period there were no changes to the consumer basket. There were usually around 190 time series for each item. To put it more simply, each of these time series represented the development of the price of a specific item in a specific store. It was first necessary to numerically clean the data of various "defects" (for example to ensure that a change in price is not considered to be a change if there was also a change to the store at the same time).

The detailed data on price investigations obtained from the Czech Statistical Office enabled us to calculate and analyse the following factors for individual items in the consumer basket:

- *The frequency of price changes,*[2] the frequency of price reductions, the frequency of price rises,
- A comparison between the frequency of price rises and reductions (the existence of downward price rigidity),
- The seasonality of price changes (the frequency of price changes at retailers in a specific month),
- The numbers of variously sized price steps,
- *The average sizes of price changes (as %),* the average price fall, the average price rise,
- A comparison of the sums of rises and falls (the existence of downwards price rigidity),
- The seasonality of the sizes of price changes,
- *The average length of periods without price changes* (directly established[3] and implicitly derived from frequencies[4]),
- The average length of periods without changes after a price fall,
- The average length of periods without changes after a price rise.

1 SELECTIONS OF CONSUMER BASKET ITEMS AND THEIR SPLIT INTO SUBGROUPS

In view of the difficulty and scope of the calculations, we decided to restrict the quantity of consumer basket items used. There are around 800 items in the consumer basket and generating all the data for all

[2] The frequency of price changes is defined as the number of prices that were changed compared to the total number of prices discovered. For example, for Edam block cheese there were 9 331 valid data items and price changes were recorded in 3 003 cases, giving a frequency of 3003 / 9331 = 0.32. This means that frequency can be represented by a value between 0 and 1.

[3] By direct calculation we mean, for the item in question, finding all periods in all stores in which no change in price occurred, and then calculating their average lengths in months (performed using our own programme).

[4] An implicit calculation is performed using the formula (1) *Average length of period without a change in price* = 1 / *frequency of changes* or the formula (2) *Average length of period without a change in price* = –1 / ln (1 – *frequency of changes*). The difference in the results from these two formulas is small. We will present the results using formula (2), the results of which are closer to the directly calculated data. These formulas are based on the assumption that the more frequent the price changes are, the shorter the average length of the period without price changes.

the consumer basket items would be extremely demanding in terms of both capacity and time. It was therefore necessary to select only part of the items from the consumer basket but do this in such a way as to ensure that this selection best represented the whole consumer basket and its subgroups.

We decided on the following procedure for the selection of the items. We started with a basic 12 groups in the consumer basket (the main division according to the Classification of Individual Consumption by Purpose — COICOP):

- Food and non-alcoholic beverages,
- Alcoholic beverages and tobacco,
- Clothing and footwear,
- Housing, water, electricity, gas and other fuels,
- Furnishings, household equipment and routine maintenance of the house,
- Health,
- Transport,
- Post and telecommunications,
- Recreation and culture,
- Education,
- Restaurants and accommodation,
- Other goods and services.

We combined this division with the basic division of the consumer basket used by the Czech National Bank (CNB calculates inflation development of other subgroups of consumer prices for analytical reasons — due to various factors which influenced them. See footnotes 7 and 9):

- Regulated prices (administered prices),
- Food prices,
- Other non-tradables (i.e. unregulated),
- Other tradables without fuels (i.e. without food prices),
- Fuel prices.

We further divided each of the 12 basic groups (when appropriate) according to the division used by the CNB. In this way we divided the consumer basket into around 30 subgroups.

From each of these subgroups we selected enough items so that their constant weights made up 60 % of the constant weight of the group in question (we considered 60 % to be sufficient in order to obtain information about the development of the whole subgroup). In the great majority of cases we selected from the subgroups items with the largest constant weights. Exceptions to this were, for example, the prices of food, where we also used some items with lower weights for the reason that they were typical representatives of some type of food. We made other exceptions in cases where the selected items with the highest weights were, in terms of their types, too uniform. For example, in the regulated housing prices group, the selection would include practically only several electricity and gas price items, and so we selected less of them, meaning that items of a different type also made it into the selection. In addition, we can assume that all the electricity charges and also all the gas charges would change together at the same time, and so restricting the selection in this way is also pragmatic in view of the aim of our analysis.

This approach led to the selection of 257 items, for which we performed all the calculations. We thus covered 60 % of the constant weight of the whole consumer basket as well as a sufficient quantity of subgroups necessary to calculate the results for all the regularly used items in the consumer basket.

To keep the text a manageable size, in later chapters we will comment in greater detail only on the aggregated results for several subgroups of the consumer basket.[5] These subgroups were determined in view of the subgroups used during the analysis of the structure of inflation by the CNB. In addition, the prices of foods were split into processed and unprocessed and a subgroup was created — other non-tradables without holidays and hypothetical rent price items (the given excluded items had significantly different results than the whole group of non-tradables).

Specifically, in this text we will evaluate the results in the following subgroups of the consumer basket:

- Processed foods[6] (including alcoholic beverages and cigarettes),
- Unprocessed foods,
- Regulated prices (according to the CNB definition),
- Fuels,
- Other non-tradables,[7]
- Other non-tradables without the holidays and hypothetical rent price items,[8]
- Other tradables without fuels,[9]

and for the whole consumer basket (all the selected items together).

We aggregated the results of the individual items into subgroups using a weighted arithmetic average (using the weights in the consumer basket) as well as using a simple arithmetic average. By comparing the results of the calculated weighted and simple arithmetic averages we obtained another evaluation — whether the pricing strategy for items with higher and lower weights significantly differed. The items with the highest weight are actually the most frequently purchased items (for example for foods these are beer, cigarettes and bread). *In this text we will usually comment on the results obtained by using the weighted arithmetic average.* We will mention the results obtained by using a simple arithmetic average only if they are of interest.

2 RESULTS FOR THE WHOLE CONSUMER BASKET (ALL THE SELECTED ITEMS)

The frequency of price changes for the whole consumer basket calculated using a weighted arithmetic average from the frequency of all the selected items came to 0.26. This practically means that one in every approximately four prices was changed compared to the previous month. This gives a *frequency of price increases as* 0.16 and a *frequency of price reductions as* 0.10. The higher frequency of price increases cor-

---

[5] In the event of interest in detailed results calculated for the individual items of the consumer basket or any of its subgroups not listed here, it is possible to contact the author and request the results. The unabridged original text (in Czech language) is also available and gives a more detailed description of the procedures, methodology and many more results. However, results cannot be provided in cases when this would mean a breach of the fundamentals of the Czech Statistical Office relating to not providing individual company data (i.e. data and results through which something could be deduced about a specific company).

[6] The division into processed and unprocessed foods arises from the methodology used by the Czech Statistical Office (respectively Eurostat) for calculating the HICP (the Harmonised Index of Consumer Prices).

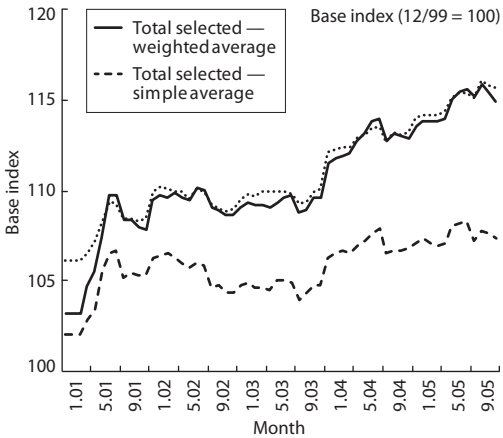
[7] Non-tradables are items for which there is mainly no possible foreign competition (these are mainly service items). Other non-tradables = non-tradables without regulated prices.

[8] Holidays and hypothetical rent prices showed significantly different results from the majority of other non-tradables and in view of their high weight distorted the results of the whole subgroup, and therefore we excluded them here.

[9] Tradables include items for which there can be foreign price competition (these are mainly consumer goods items). Other tradables = tradables without food price items.

responds to the fact that consumer prices were exhibiting a rising trend in the monitored period (see Figure 1).

Figure 1 The price development of the selected items

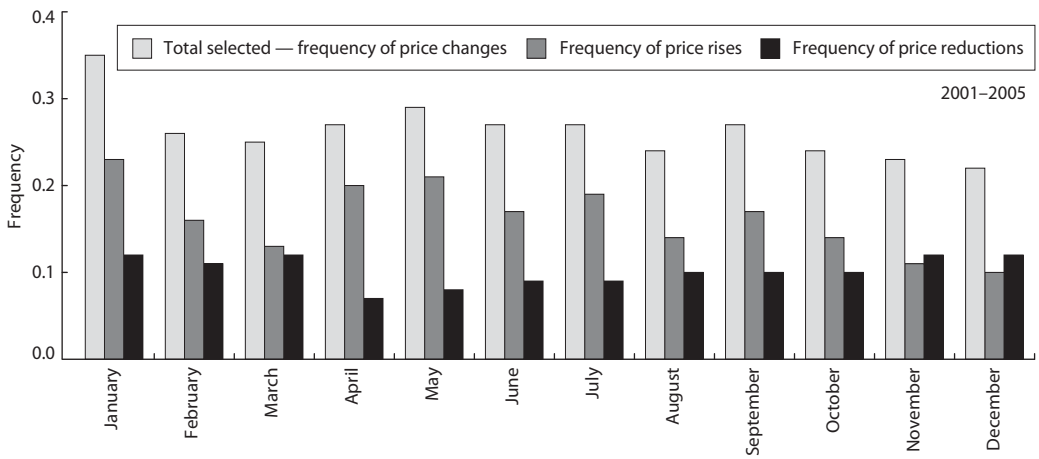


Source: Own construction, Czech Statistical Office

NOTE: *In this analysis the results obtained are in direct connection with inflation for the item in question. From the Figure 1 above we see that in the monitored period total inflation exhibits a rising trend. The significant accord between official inflation and inflation for the selected items (weighted arithmetic average) shows that the selected items are a good approximation for the whole consumer basket. The fact that the simple arithmetic average for inflation for the selected items is positioned lower shows that the prices of the items with a higher weight in the consumer basket rose in sum over the whole period more than the prices of items with a lower weight (mainly through the influence of regulated prices, the prices which usually rise at an above-average rate and have a high weight in the consumer basket).*

We also calculated the frequency for individual months and this gave us an idea of the *seasonality of the frequency of price changes.* Most price rises were in January and in May (see Figure 2), and the least in November and December. The largest number of price reductions was in January, March, November and December. In general, the seasonality of the frequency of reductions in price is less significant than the seasonality of the frequency of price rises. The highest number of price rises in January corresponds to the seasonality of consumer prices, which exhibit the highest seasonal rises precisely in January. In general, however, the seasonality of the frequency and inflation cannot be compared, as they also depend on the size of the price changes.

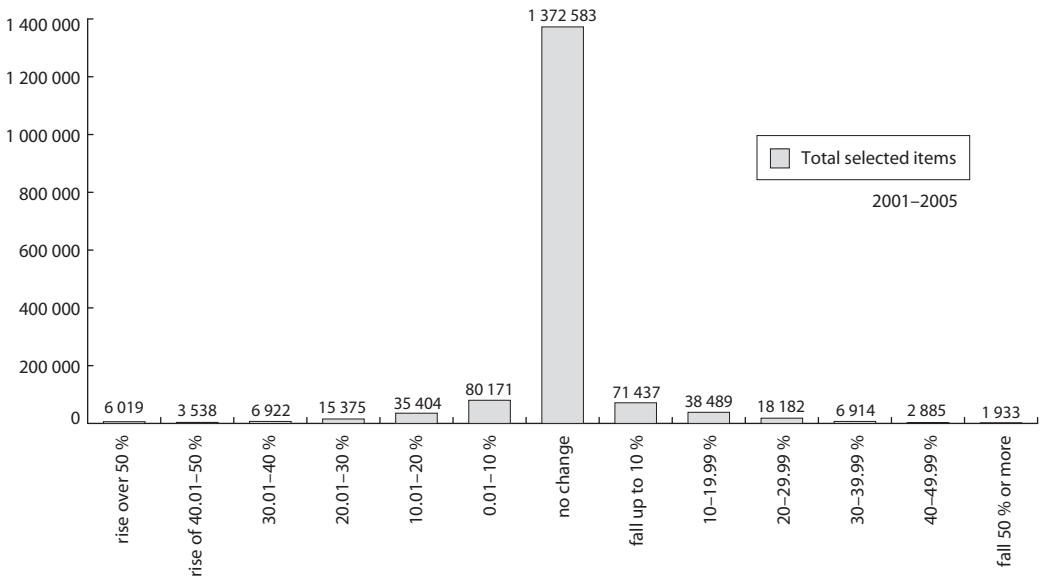
Figure 2 Frequency of price changes in individual months (weighted arithmetic average)



Source: Own construction

We also analysed the *numbers of price changes differentiated according to their sizes.* The Figure 3 shows the differentiation of price changes according to the size of the change, initially with a wider range including the number of zero changes (see Figure 3). We san see that the majority of changes are up to 20 % in size, however the numbers of larger price changes are not insignificant either. A similar differentiation of numbers of changes larger than 20 %, i.e. a similar tempo of fall in the number in terms of larger changes, was displayed by all the described subgroups of the consumer basket except for fuel price items, where there was only one price change of over 20 %.

Figure 3 Numbers of changes in price according to size



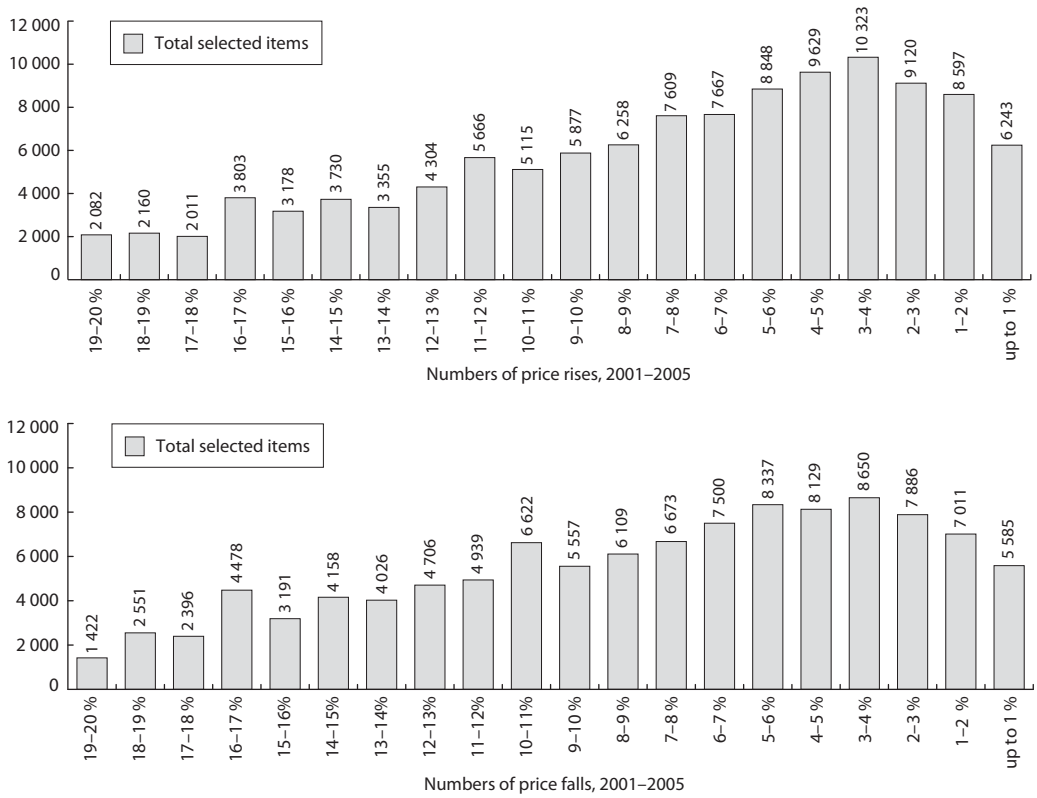
Source: Own construction

We also split price changes of up to 20 % more finely. The first graph below (see Figure 4) gives the numbers of price rises and the second graph the numbers of price falls. As you can see, the division of the numbers of price changes according to size is similar for price rises and price falls. The majority of the changes are between 3 % and 4 %. Thereafter, the numbers of changes fall smoothly on both sides with the exception of slight deviations around 11 % and 16 %.

We intuitively expected higher numbers of changes with "round" percentage sizes (5 %, 10 % and 15 %) and for many items this was confirmed. It was also frequently the case that the highest number of changes was not by these round percentages, but by one per cent more, and so there were higher numbers of changes for example by 11 % and 16 %, and this was the case for all the selected items in total. The reason for this phenomenon could be the rounding of prices. In this case it would mean that retailers more commonly round in the direction of the price change — meaning that they round up when increasing a price and round down when decreasing a price.

NOTE: *The results on the numbers of variously sized changes for the whole basket were less showing than similar results for individual items from which it was possible to discover different pricing strategies. For example, raising prices in small steps, dropping prices in large steps of round percentages and so on. For the sake of clarity we will present one example — the division of price changes according to size for Edam cheese*

Figure 4 Numbers of price rises and falls between 0 % and 20 %



Source: Own construction

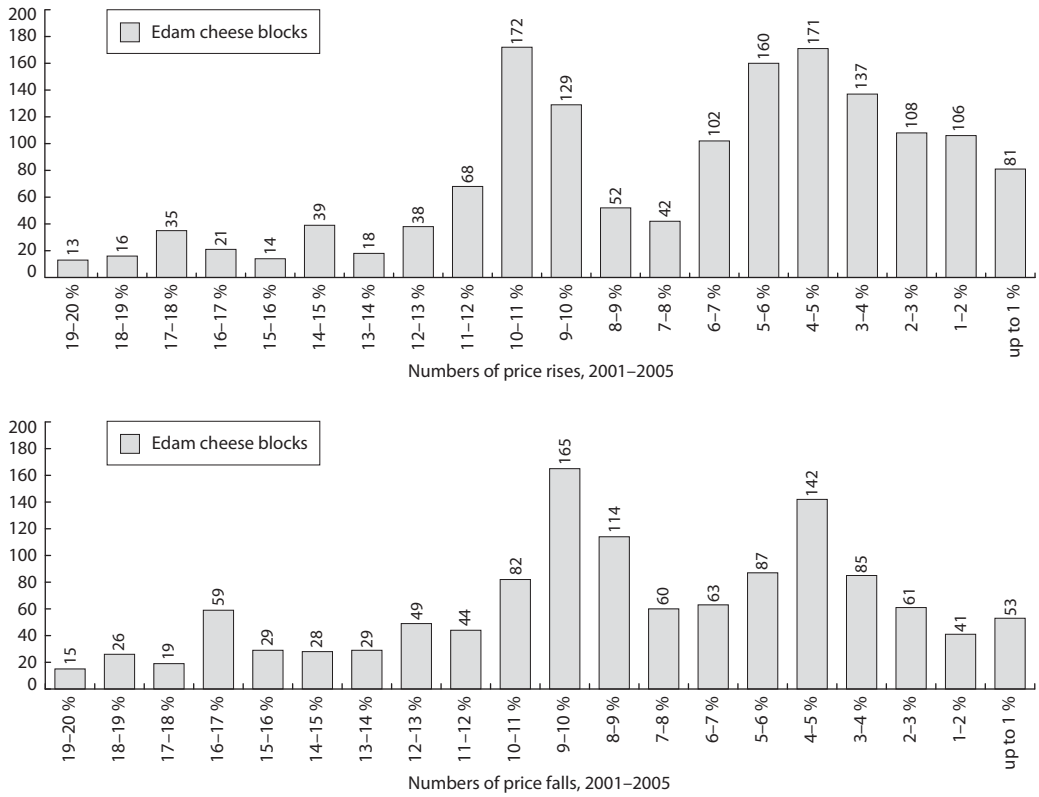
blocks (total number of observations 9 331, number of price changes 3 003, frequency 0.32, frequency of price increases 0.18, frequency of price reductions 0.14), see Figure 5.

A more detailed division shows that for Edam cheese blocks the most frequent price rises and falls are by around 5 % and 10 %. In terms of larger changes, those around 15 % are the most popular. This unequivocally shows a certain rationality in the behaviour of the retailers. On the one hand this will is connected with minimising the costs for price changes and most likely also with the expectations of the retailers regarding the psychological effect on customers (customers do not like price rises but the size of the change is not so important – this means that it is better to change the prices less often but by larger amounts). These graphs (Figure 5) also enable us to understand the differences between strategies when increasing and decreasing prices (specifically here there is no significant difference).

The overall *average price rise* (for the whole consumer basket) was *10.7 %*, and fall *–9.1 %*. The rise in consumer prices in the monitored period was thus implemented through a higher frequency of price rises as well as a higher average size of price rise than price fall. A comparison with the results obtained using the simple arithmetic average shows that on average *the prices of items with lower weight change less frequently, but in larger steps.*

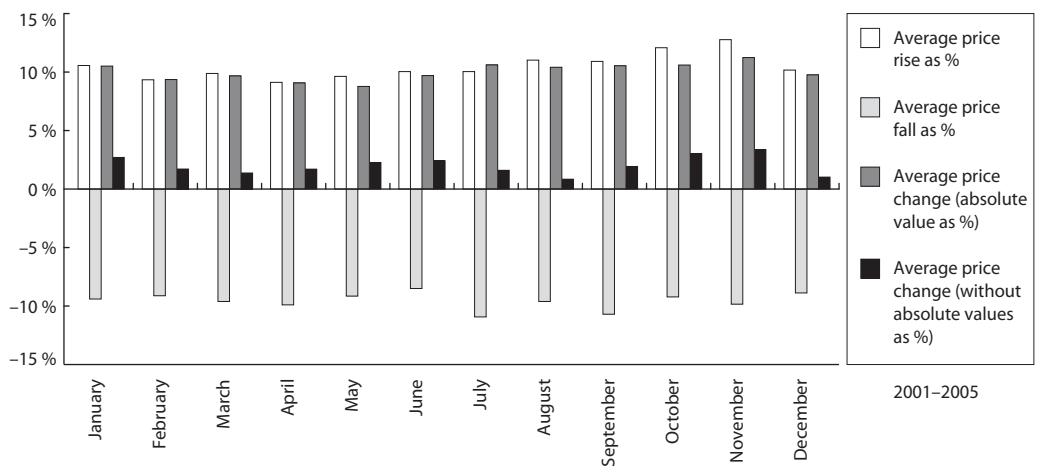
We also calculated the average sizes of price changes for individual months and thus obtained the *seasonality of average sizes of price changes* (see Figure 6). The average size of price rises smoothly increased

Figure 5 Numbers of price rises and falls between 0 % and 20 % for Edam cheese blocks



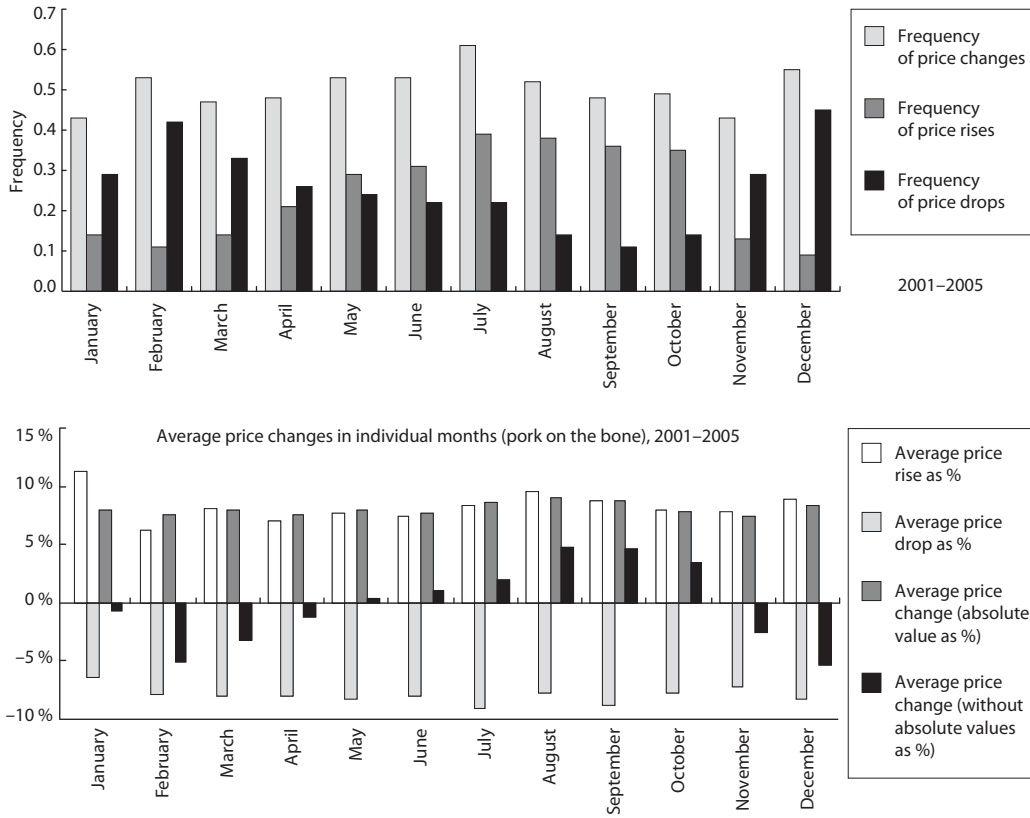
Source: Own construction

Figure 6 Average price changes in individual months (weighted arithmetic average)



Source: Own construction

Figure 7 Frequency of price changes and average price changes for item: pork on the bone



Source: Own construction

between June and November. Higher values are also achieved in January. The average size of price falls is more stable during the year — there are higher values in July and September (in July there is the perceptible influence of the higher seasonal drop in the prices of food and in September the influence of seasonal drops in the price of holidays and related items; these seasonal influences also significantly influence the seasonality of overall inflation).

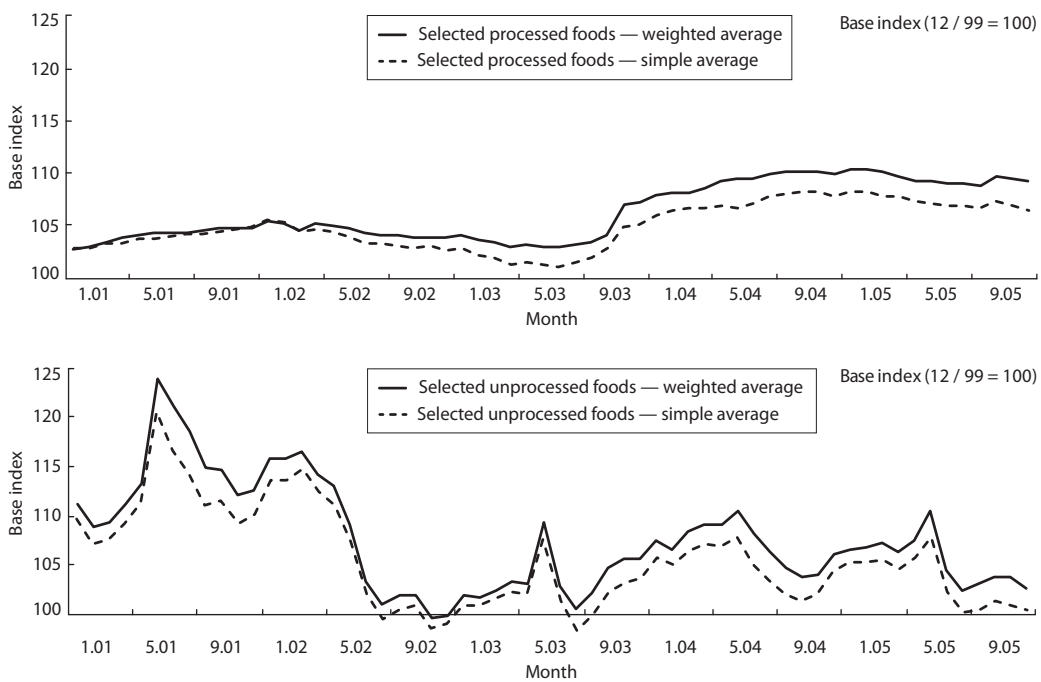
NOTE: *The results of the average sizes of price changes illustrated different pricing strategies for various items: for example either the average size of price changes was stable during the year (the change in consumer price was achieved through a higher or lower frequency of price changes) or on the contrary the average size of price changes during the year significantly changed and the frequency of price changes was stable (changes in consumer prices were achieved mainly through variously sized price changes with a stable number of price changes). As an example we can present the results for the item: pork on the bone (total number of observations 9 239, number of price changes 4 667, frequency 0.51, frequency of price increases 0.25 and frequency of price reductions 0.26). The Figure 7 shows the seasonality of frequency and the seasonality of average sizes of price changes.*

These graphs show the relatively significant seasonality of the frequency of price changes together with a relatively stable average size of price changes during the year (with the exception of January and February).

Seasonal price drops are most in evidence from November to April, while price rises preponderate from May to October. There are no strong reasons for this seasonality at a pork meat plant and neither was the cause of this seasonality any fundamental phenomenon (developments in global prices etc.) This means that the causes will be pricing strategies, also very probably taking into account the influence of the spread of demand during the year. Further, it is possible to read from the results that with the exception of January and February retailers on average change prices by relatively stable large steps and change the frequency of these steps according to need. In January and February they have a different strategy. In January compared to other months the average size of price increases is significantly higher, yet their frequency is lower. In February the frequency of price increases rises but the price increases themselves are below average in size. The anomaly in January could be connected to repricing at the start of the year — setting prices at new levels (not complying with the usual size of steps).

The last group of parameters that we calculated were *average lengths of periods without price changes* (a change after one month = length 1). For *the whole consumer basket* (all the selected items) the directly calculated *average length of periods without price changes came to 5.7 months, 5.96 months after a price rise and 5.12 months after a price fall.* However, only 39.1 % of all the input data could be used for the calculation, and therefore these results are significantly distorted downwards (we provide a more detailed explanation in chapter 3.3). The actual situation will be closer to the implicitly calculated data. *The implicitly calculated length of periods without price changes for all the selected items came to 10.72 months.* We calculated the lengths of periods without price changes for all the selected items (as well as for other groups of items) as the weighted arithmetic average of the lengths of periods without price changes for individual items, and therefore these results differ from the result that we would obtain through an implicit calculation from the overall frequency.

Figure 8 The development of the prices of processed and unprocessed food differs significantly



Source: Own construction

3 RESULTS FOR SELECTED SUBGROUPS OF THE CONSUMER BASKET

3.1 Price development for selected subgroups

As we indicated in the introduction, in order to keep the size of this text manageable we will only describe in detail the results for the subgroups of the consumer basket used during the analysis of the structure of inflation by the CNB. In addition, food prices have been divided into processed and unprocessed and a subgroup — other non-tradables without the price of holidays and hypothetical rent items — was created (the list of subgroups, more information on their selection and their definitions are given in chapter 1).

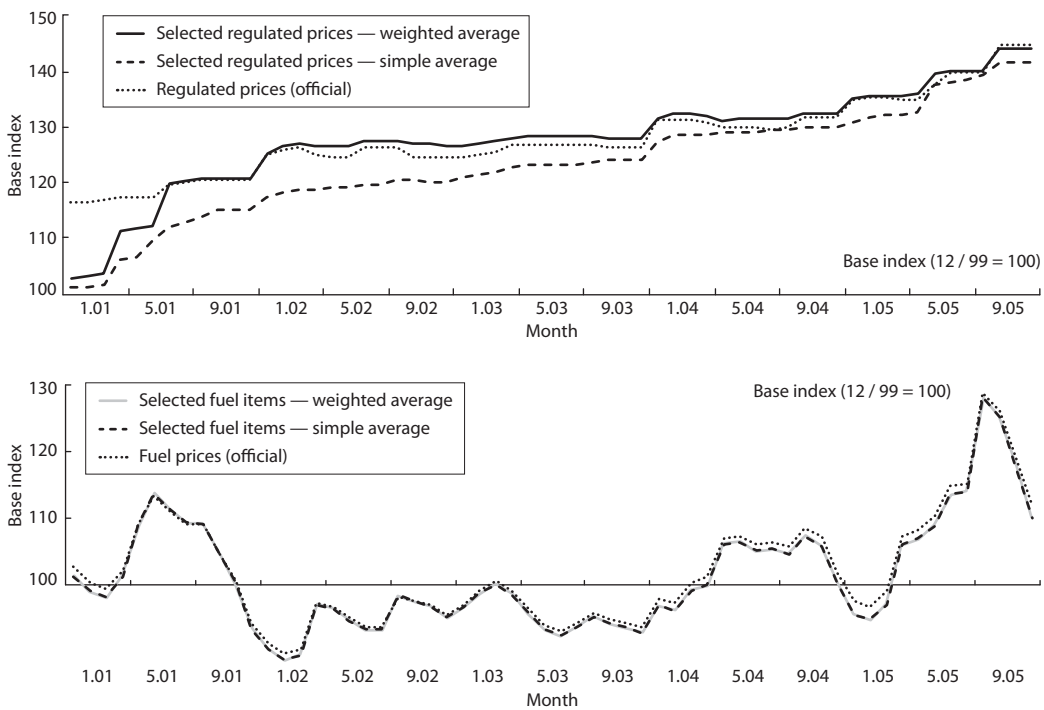
First we will acquaint ourselves with the development of the consumer prices of selected subgroups of items (see Figure 8–11). If this is normally published data, the graphs also give official figures. The closeness between the series of these official figures and the weighted arithmetic average for the selected items shows that the selected items represent the whole group sufficiently well.

For regulated prices the weighted arithmetic average of the base indexes of the selected items differs from the official figures only in 2001, when a change in the basket meant that detailed data for electricity and water prices, which were subject to high rises in that period, were missing from the database of selected items. For fuel prices the figures differ only minimally (see Figure 9).

After excluding holidays and hypothetical rent prices from the development of the prices of other non-tradables (i.e. unregulated), both the seasonality also the difference between the weighted and simple arithmetic average disappear (see Figure 10).

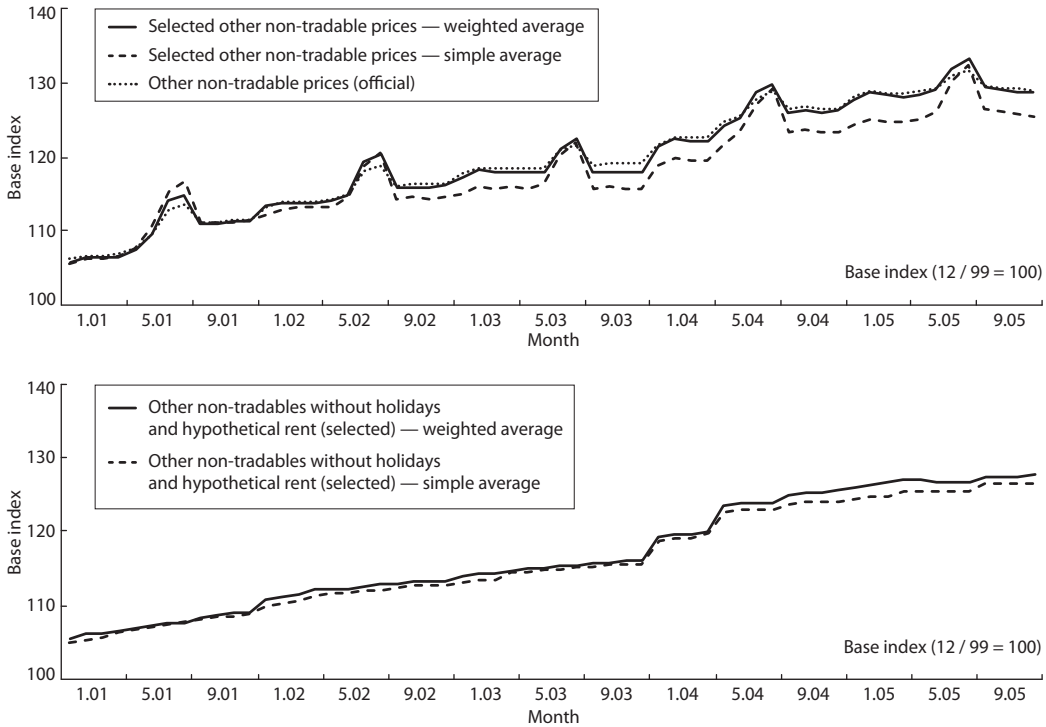
For the prices of other tradables without fuels (i.e. without food prices), on average the prices of items with higher weight fell faster (mobile telephones, personal computers, video recorders, women's footwear and so on) and therefore the prices of the selected items are lower than the prices for the whole group (see Figure 11).

Figure 9 The development of regulated and fuel prices



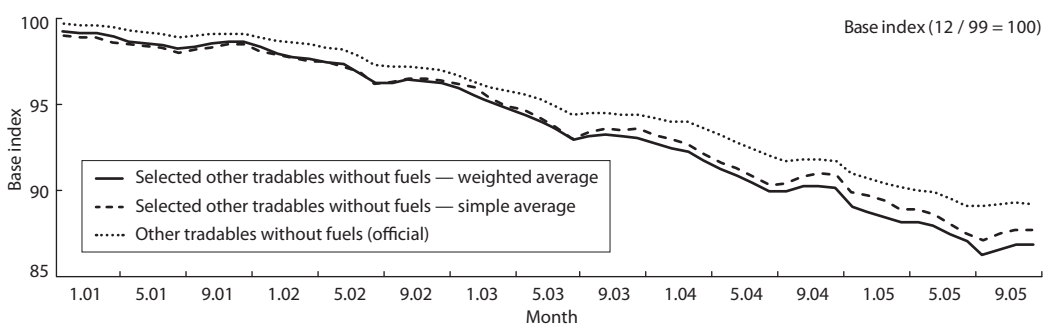
Source: Own construction, Czech National Bank

Figure 10 The development of other non-tradables prices



Source: Own construction, Czech National Bank

Figure 11 The development of other tradables prices without fuels



Source: Own construction, Czech National Bank

From these graphs of the price development, it is also very clear, how the development of prices among individual subgroups differs as it is one of the reasons why the consumer basket has been divided precisely in this way for the purposes of the analysis. Regulated prices and the prices of other non-tradables continuously rose in the monitored period. Non-tradables (mainly services) are generally dependent in particular on the development of domestic cost pressures and the level of domestic demand. The basis for their continuous rise is mainly the gradual rise in domestic salaries. On the

Table 1 Frequency of price changes and average price changes

January 2001 to December 2005	Unprocessed food	Processed food	Regulated prices	Fuels	Other non-tradables	Other non-tradables without holidays and hypothetical rent	Other tradables without fuels	Whole consumer basket
Frequency of price changes	0.42	0.19	0.10	0.85	0.39	0.06	0.12	0.26
Frequency of price rises	0.22	0.11	0.07	0.32	0.31	0.05	0.05	0.16
Frequency of price falls	0.21	0.09	0.03	0.53	0.09	0.02	0.07	0.10
Average price change as %	1.60	2.28	5.97	–0.21	5.02	8.25	–2.04	2.57
Average price change in absolute value as %	14.23	9.95	7.83	4.34	9.55	13.21	12.69	10.18
Average price rise as %	15.44	10.65	9.00	5.50	9.90	13.89	12.55	10.74
Average price fall as %	–13.03	–8.77	–3.99	–3.64	–9.08	–12.27	–12.73	–9.11

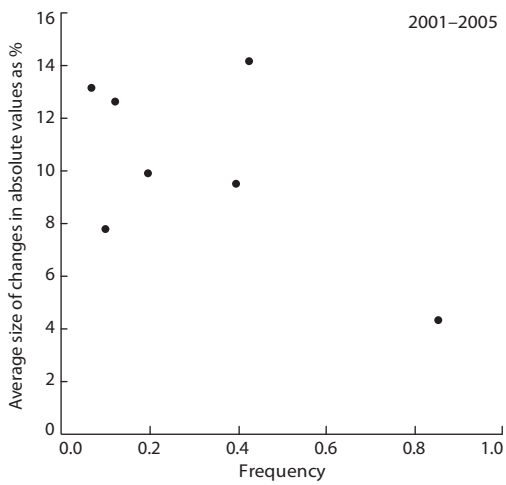
Source: Own costruction

other hand, for the development of the prices of other tradables, the main roles are played by the high level of competition, the wide range of imported consumer goods and the long-term strengthening of the koruna exchange rate. Thus in the monitored period these prices continuously fall even under the conditions of rising domestic costs.

3.2 Frequency and average size of price changes

The comparison table (see Table 1) shows that of the monitored subgroups of the consumer basket, the prices of fuels clearly change most frequently, which was the expected result (the prices of fuels react relatively quickly to changes in the global prices of oil and developments in the koruna exchange rate). Unprocessed food is some way back in second place, where frequent price changes are caused in particular by the volatile prices of fruit and vegetables. Then follow the prices of other non-tradables, for which, however, the higher frequency of price changes was caused mainly by hypothetical rent and holidays prices. Both have a high weight in the consumer basket and their prices change very frequently. The frequency of price changes of other non-tradables without hypothetical rent and holidays is, on the contrary, the lowest of all the monitored subgroups in the consumer basket, even lower than for regulated prices. Other tradables without fuels also have a low frequency of price changes.

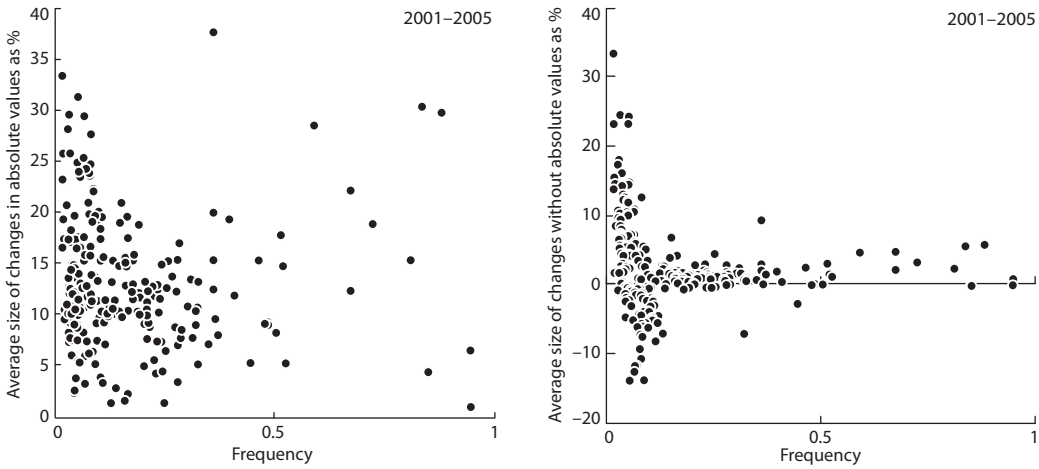
Figure 12 Relation between frequency and average size of changes



Source: Own costruction

Unprocessed food has the highest average sizes of price changes and this is once again caused by the prices of fruit and vegetables, for which high rises and falls in price of tens of percentage points are normal. If we disregard non-market regulated prices and the very volatile prices of unprocessed

Figure 13 Relation between frequency and average size of changes



Source: Own costruction

food, it is possible to see a certain regularity in the case of the average sizes of changes for the evaluated subgroups of the consumer basket: *the lower the frequency for a subgroup of price changes the larger the average size of the price change.* This relationship is shown in the Figure 12.

With the exception of the unprocessed food point (at the top in the middle) and the regulated prices point (bottom left) the points almost lie on a straight line.

Fuel prices have the highest frequency and the lowest average size of price changes. These are followed by other non-tradables with the second-highest frequency and the second-smallest average length of price changes. Then we have processed food, followed by other tradables without fuels. Other non-tradables without holidays and hypothetical rent have the lowest frequency and highest average size of price changes.

In the case of individual items, however, the situation is different. We cannot expect any uniform pricing strategies across all the items. The Figure 13 (on the left) shows the relationship between frequency and the average size of price changes in absolute values. The one on the right gives the average size of price changes without absolute values.

The image in the graph on the right shows the picture where the points form an image of a capital "T" lying on its side. This means that for many items, irrespective of the frequency of the changes, the average size of price changes (without absolute values) hovers around zero, respectively slightly above zero. Also here, however, we have a group of items with a very low frequency of price changes, but with a significantly positive or negative average change in price (their prices change infrequently, but by large steps mainly in a single direction). The points that deviate most significantly in the upwards direction are mainly regulated price items: television fees, nursery school fees, higher education fees, drawing up wills, but here there are also some unregulated prices in the area of services. The points that deviate most significantly in the downward direction are for clothing and footwear items followed by electronics.

The average sizes of price changes calculated for the whole consumer basket using the simple arithmetic average were 13.4 %, respectively –11.8 %, which together with the information that the simple arithmetic averages of the frequency of price rises and falls were 0.09, respectively 0.08, implicates that on average the prices of items with lower weight change less frequently, but by larger steps (in comparison with the results obtained using a weighted arithmetic average).

3.3 Lengths of periods without price changes

The calculation of the average length of periods without price changes has shown itself to be the most problematic. Due to many defects in the data that hinder the coherency of the investigation of the prices of a specific item at a specific place (for example a change of store, change of the specific type of item) a large percentage of input data was excluded from the direct calculation (see footnote 3). Deducting the length of a period without a change was logically correct only in the case when no such defect arose between two price movements. This interruption in time series was relatively frequent and thus a lot of input data had to be excluded from the calculations (see the indicated percentage of data used in Table 2). If this was not done, only a "false" price change would arise in the time series (for example when making the deduction in a different store, the price there could be different to that in the original store, while in both stores no price change might actually have taken place).

This therefore meant that only data sections when there was no defect between two price changes could be used (so data from the start and the end of the time series was not used). The lower the frequency of changes an item had, the larger the quantity of unusable data — sections without price changes were long and thus it was more probable that some form of defect would occur in the data during this longer period, and this was frequently confirmed in practice.

This means that the lengths of periods without price changes directly calculated from the data are very probably shorter than in reality, as the longer sections were more likely to be excluded (because of defects in data). On the other hand, the implicitly calculated lengths of periods without changes (see footnote 4) could be somewhat longer than in reality. For example, in a situation in which two sections in the data without price changes are separated by a section for which there is no available data, the calculation will

Table 2 Lengths of periods without price changes

January 2001 to December 2005, in months	Unprocessed food	Processed food	Regulated prices	Fuels	Other non-tradables	Other non-tradables without holidays and hypothetical rent	Other tradables without fuels	Whole consumer basket
Implicit average length of period without a price change	2.51	7.03	16.65	0.52	12.19	21.50	12.21	10.72
Average length of a period without a price change discovered from data								
– total*)	2.36	5.71	9.12	1.18	6.47	10.76	4.19	5.70
– after a rise	2.55	6.03	9.61	1.09	6.51	11.05	4.56	5.96
– after a fall	2.16	5.78	7.07	1.22	6.12	9.98	3.80	5.12
*) *useable quantity of data for its calculation, as %*	*70.70*	*57.70*	*40.10*	*91.40*	*39.40*	*30.50*	*23.00*	*39.10*

Source: Own construction

actually see this as a single longer section without a price change (as no price change occurred during it, and so the calculated frequency does not increase and therefore the implicitly calculated length of the period without price changes is not reduced). While a price change could have taken place in the period for which data is missing.

The results for the lengths of periods without price changes (Table 2) show that for many items the individual prices change relatively infrequently. With the exception of fuel prices and unprocessed food prices, individual prices change on average only after around five months and, for example, for the prices

of other non-tradables without holidays and hypothetical rent prices, the estimated change is less than once a year on average.

From the table it can be seen that the larger the percentage of data used, the more the directly calculated length of the period obtained from that data without price changes corresponded to the implicitly calculated length of the period without price changes. The implicitly calculated length of a period without price changes is therefore very probably closer to the real situation.

NOTE: *Please also note the result that the length of a period without price changes after a price increase directly calculated from the data is longer than that after a price decrease. This applied for all the described subgroups of the consumer basket with the exception of fuel prices. This could be because some price reductions are temporary special offers for a limited period, i.e. that the price will return to its original level soon after the special offer. However, it could also indicate some type of downwards price rigidity — reducing prices is less pleasant for the retailer, and therefore there is another price correction sooner than after a price increase.*

3.4 Some other results for subgroups or individual items

From the *processed food* subgroup (including alcoholic beverages and cigarettes) the following products had the highest frequency of changes: eggs at 0.41, butter at 0.37 and granulated sugar at 0.33. The products with the lowest frequency were cigarettes at 0.05 to 0.06, white wheat bread products at 0.08 and regular caraway-seed bread at 0.13. The seasonality of frequencies was not marked and the majority of the price changes were 4 % to 5 % in size. In August and October processed food displayed an above-average level of price rises.

The prices of cigarettes themselves had a very low frequency of changes (0.06), with a frequency of price increases 0.05 and price decreases 0.01. The prices of cigarettes were most often increased by 10 %, followed by an increase of 3 %, while they were most often decreased by 3 %. The low frequency of changes in cigarette prices is also influenced by the fact that the price change is bound by the amendment of price stamps which are affixed to cigarette packages.

The prices of *unprocessed food* are much more closely connected with the fluctuating developments of the prices of raw materials. Unprocessed food include (according to the methodology used for the HICP): all raw meat, smoked meats, tinned meat, fruit, potatoes, vegetables and pulses. In our selection of items, fruit and vegetables had the highest frequency of price changes (0.6 to 0.9) and Hungarian salami and pork livers the lowest at 0.15 and 0.16. With a breakdown of price changes according to the size of the price changes a different strategy was visible for increasing prices and decreasing prices. The prices were increased most frequently by 3 % to 4 % and decreased most frequently by 5 % to 6 %. The average size of price changes was stable for most of the year.

Regulated prices in the monitored period rose practically continuously and often in steps rather than smoothly, which is typical for them. Among the regulated price items, regulated rent and the waste removal had the lowest frequency of changes (0.04). Items with higher frequencies included in particular medical items. The seasonality of frequency is given for regulated prices by the fact that their prices change most from the start of the year and then from the start of the quarters. The only exception to this rule was May, and this because from May 2004 there was a reduction in the basic VAT rate from 22 % to 19 %, which affected many regulated price items.

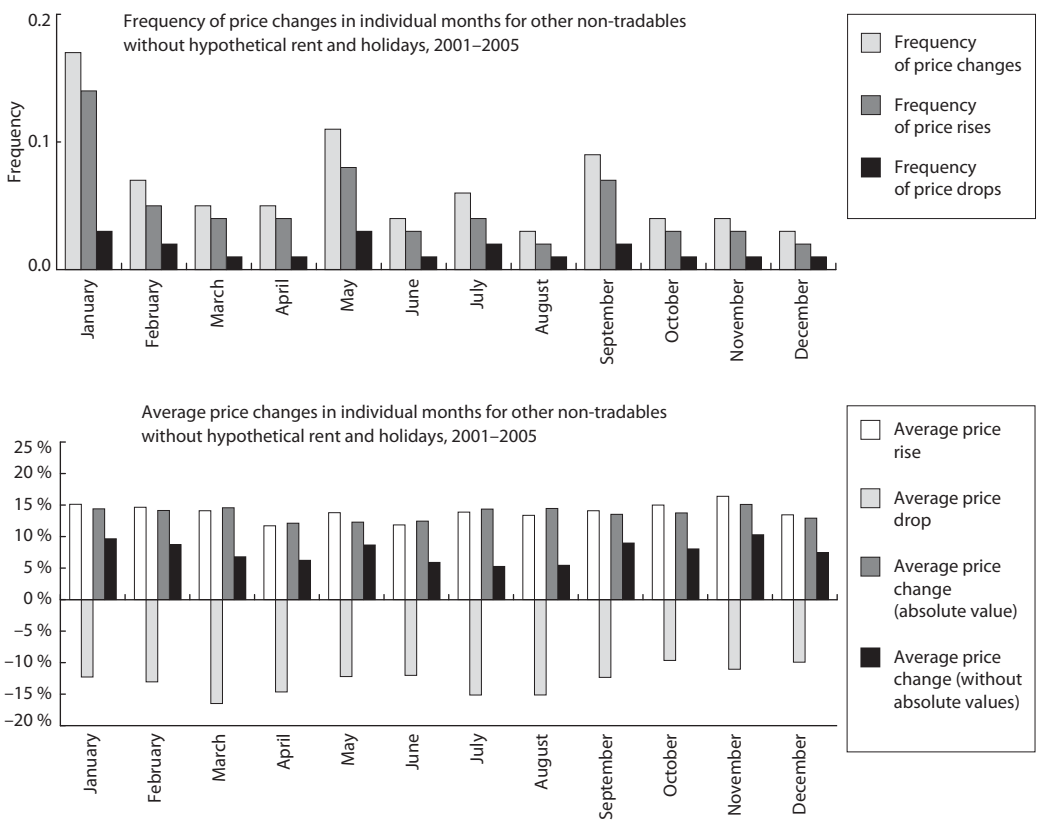
With regulated prices the division of price changes according to size for price increases and price decreases differed greatly. Regulated prices were reduced significantly less often than they were increased and significantly the most reductions in their prices were of sizes up to 1 %, which confirms that the "willingness" to reduce their prices was low, and when they were actually reduced, then this was only by a small amount. For regulated prices there was thus a clear downward price rigidity.

From *other non-tradables* the following items had the lowest frequency at 0.02: rent for garages, the payment of services connected with the rental of an apartment and a contractually rented 2-room apart-

ment. The hypothetical rent (0.95) and holidays price (up to 0.95) items had the highest frequency. Other non-tradables had in sum over all the months a significant preponderance of price rises compared to price reductions. The most price rises occurred in January and in July. Price lists are changed in January and holidays prices are increased in July. When breaking down the price changes according to size, the largest number of price changes was price increases and decreases between 1 % and 2 %.

In the results for this group, hypothetical rent and holidays prices had a fundamental influence because of their high weight and different results. *Hypothetical rent* is the theoretical level of rent that would be paid by people living in their own houses. It is thus a type of summary price calculated from several items and so it is logical that the level of hypothetical rent is always changing and thus has a high frequency of price changes. Hypothetical rent had a frequency of changes of 0.95. Of this, the frequency of price increases was 0.78 and the frequency of price reductions 0.17. *Holidays prices* also displayed a higher than average frequency. The reason for this is the variability of their prices during the year — during the tourist season prices initially increase and then decrease once again. Winter holidays items are represented in the consumer basket far less than summer holidays items and thus the results valid for the whole holidays group are given in particular by the development of summer holidays prices. *Other non-tradables without the holidays and hypothetical rent items* had a relatively stable size of average changes in individual months while on the other hand their frequency of price changes displayed a significant seasonality (see Figure 14).

Figure 14 Frequency of price changes and average price changes for other non-tradables without hypothetical rent and holidays



Source: Own construction

The great majority of selected items in the *other tradables without fuels* group have a frequency of up to 0.2. Cut flowers had the highest frequency of price changes with a frequency of around 0.5. Fireworks and wedding rings (both 0.02) had the lowest frequency of price changes from this group, followed by gold chains and woven carpets (both 0.03). From September to November other tradables without fuels had a higher frequency of price increases than price decreases, otherwise the frequency of price reductions predominated. The number of price rises was in total lower and they also had lower percentages, while there were many more price drops and these were also by large steps. The situation corresponds to the fact that the prices of other tradables without fuels continuously fell in the monitored period (in particular as a consequence of the long-term strengthening of the koruna exchange rate).

CONCLUSION

The aim of this analysis was to obtain knowledge about the strategies of consumer goods and services retailers in terms of changes to final prices. We evaluated various indicators, the frequency of price changes, the average size of price changes and the average lengths of periods without price changes. We calculated these indicators for selected individual consumer basket items, for different subgroups of the consumer basket as well as for the whole consumer basket.

The knowledge gained has shown that there are various strategies for price changes not only for different types of items but also for items of the same type. In spite of this it has been possible to indicate prevailing characteristics of pricing strategies in selected subgroups of the consumer basket.

The typical characteristics of the prices of regulated items were that these prices mainly only rise and also do this in larger-sized steps and mostly at the start of the year. Food prices display a higher frequency of price changes mainly in the case of unprocessed food, which is a consequence of the volatile development of the prices of agricultural commodities. Tradables without the prices of food and fuels continuously fell for the whole of the monitored period and displayed only a low frequency of price changes. The prices of non-regulated non-tradables continuously rose smoothly and, with the exception of hypothetical rent and holidays items, the lowest frequency of price changes was recorded in this subgroup. On the contrary, hypothetical rent and holidays price had a high frequency of price changes. The definition of hypothetical rent means that it changes practically every month and holidays prices often change due to the fluctuations of their prices during the year. Fuel prices changed most frequently and on average displayed the lowest size of price changes, as they react relatively quickly to rapid changes in the prices of raw materials and the koruna exchange rate.

The prices of the whole consumer basket mainly rose in the monitored period. The frequency of price changes for all the selected items in sum was 0.26, which means that approximately one in every four prices was changed compared to the month before. The average size of the price changes was 10.18 %. For the presented subgroup the rule applied that the lower the frequency of price changes, the higher the average size of those price changes was. The average length of periods without price changes was calculated using various methods with different results. We estimate that the actual average length of periods without price changes was approaching the value of 10.7 months from below.

References

ÁLVAREZ L. J. and HERNANDO, I. *Price setting behaviour in Spain: stylized facts using consumer price micro data.* Banco de España, 2004.

DHYNE E., ÁLVAREZ, L. J. et al. *Price setting in the Euro Area: some stylized facts from individual consumer price data.* Banco de Portugal, 2005.

Gini Index in Czech Republic in 1995–2010

Luboš Marek[1] | *University of Economics, Prague, Czech Republic*

Abstract

The author calculates the Gini index in Czech Republic over the period of 1995–2010 in this paper. This index is further calculated in accordance to sex and age (3 age groups: under 30, 30–50, over 50). The author compares all groups and shows that there is a significant difference between them. The Gini index is also calculated for Czech regions over the years 2000–2010 and this calculation is followed by the comparison of these regions. The values of the Czech Gini index are published in this form for the first time. At the end there is a comparison of the Czech Gini index with the values of this index in some selected countries.

Keywords

Gini index, sex, age, region

JEL code

C10, E64

INTRODUCTION

The author computes the values of the Gini index in Czech Republic over the period of 1995–2010. The Gini index is also computed over the same period separately for men and women. Author compares the value of index for both sexes and shows that there is a significant difference between values for men and women. The values for men are greater and the scissors between men and women are more and more opened over time. The Gini index is further computed for three age groups (till 30, 30–50, over 50). Very similar values are reached for two "older" groups, for the youngest group the values are dramatically smaller.

The author computes the values of the Gini index for each region in Czech Republic in the next part of the article. The time period for analysis is shorter because the new defininition of regions in Czech Republic started in year 2000. The values are very similar for all of the regions except the capitol. Values for Prague are considerable and they are comparable with some developed European countries.

Data was gathered by Trexima. This company observes wages of employees (more precisely wages derived from the average hourly earnings) for the second quarter of the relevant year. The second quarter was chosen because of the stablest working time. The size of the analysed sample was constantly growing from more than 300 000 observations in 1995 to more than 2 millions in 2010. Because of the revision and update of the 2007 and 2008 data, the author worked with the new updated data. This fact is important to bear in mind when comparing with some older results. A table of Gini index values for some chosen countries is published in the last part of the article. The author compares the values of Czech Republic with values of these countries.

[1] Department of Statistics and Probability, Faculty of Informatics and Statistics, e-mail: marek@vse.cz.

The whole analysis is realized in MS Excel 2003. Earlier results from previous years can be found in the articles of authors Marek and Vrabec (see references).

1 GINI INDEX

The construction of the Gini index is based on a database around 3 million records with wage values in Czech Republic. For the analysis is used the base definition of the Lorenz curve and definition of the Gini index. There is a very detailed interval distribution of wages for the purposes of analysis. The length of intervals is 500 Czech crowns (approximatelly 20 Euro) and the data are in the form:

Wage interval	Frequency
…	…
9 500–10 000	38 869
10 000–10 500	40 711
10 500–11 000	42 825
11 000–11 500	43 666
11 500–12 000	45 960
…	…

Source: Own construction

This table is a small illustration of the real dataset. The wage intervals are very small and the reached results are very accurate.

2.1 GINI INDEX IN CZECH REPUBLIC

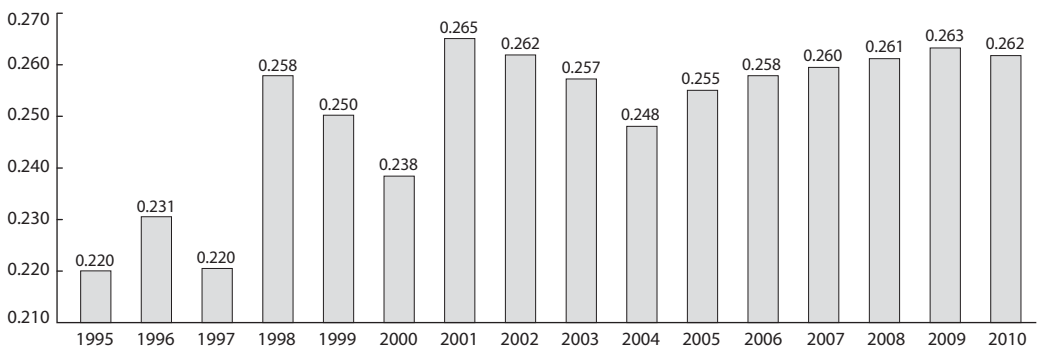
The values of the Gini index in Czech Republic are in the Table 1 and in the Figure 1.

Table 1 Gini Index — Czech Republic

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Gini	0.220	0.231	0.220	0.258	0.250	0.238	0.265	0.262	0.257	0.248	0.255	0.258	0.260	0.261	0.263	0.262

Source: Own construction

Figure 1 Gini index — Czech Republic



Source: Own construction

From the Table 1 and from the Figure 1 we can see that maximum of the Gini index is in year 2001 (0.265). Then the values are decreasing to value 0.248 in year 2004, and in the next years the values are increasing to value 0.263 in year 2009.

2.2 Gini index in accordance to sex

The values of the Gini index in accordance to sex are in the Table 2.

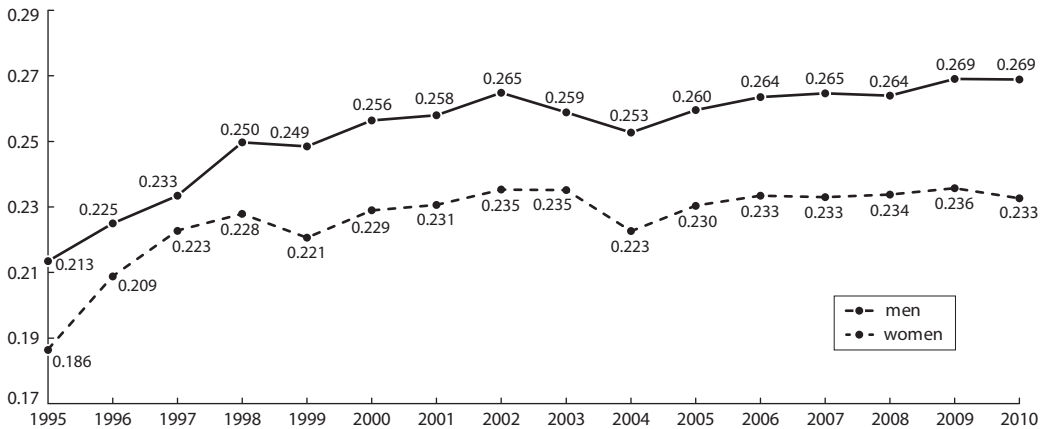
Table 2 Gini index in accordance to sex

Sex \ Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Men	0.213	0.225	0.233	0.250	0.249	0.256	0.258	0.265	0.259	0.253	0.260	0.264	0.265	0.264	0.269	0.269
Women	0.186	0.209	0.223	0.228	0.221	0.229	0.231	0.235	0.235	0.223	0.230	0.233	0.233	0.234	0.236	0.233
Diff.	0.027	0.017	0.011	0.022	0.028	0.028	0.027	0.030	0.024	0.030	0.029	0.030	0.032	0.030	0.034	0.036

Source: Own construction

We can see at the first look that the values of the Gini index are greater for men than for women. It signs that the measure of inequality is smaller for women than for men.

Figure 2 Gini index in accordance to sex



Source: Own construction

The values of the Gini index for women copy the trend for men. The difference is in the level. The average of the Gini index (we used geometric mean) is equal:

- men: 0.253,
- women: 0.226.

2.3 Gini index in accordance to age

The values of the Gini index for three age groups are in the Table 3 and in the Figure 3.

Table 3 Gini index in accordance to age

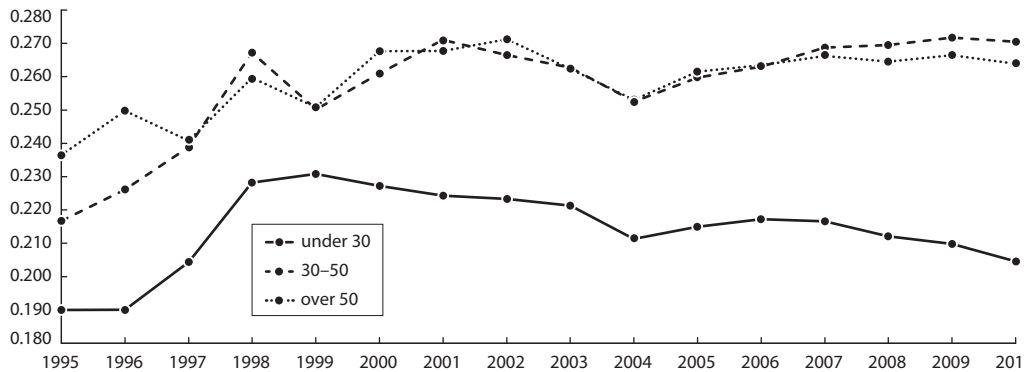
Age \ Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
under 30	0.190	0.190	0.204	0.228	0.231	0.227	0.224	0.223	0.222	0.211	0.215	0.217	0.217	0.212	0.210	0.204
30–50	0.217	0.226	0.239	0.267	0.250	0.261	0.271	0.267	0.263	0.253	0.260	0.263	0.269	0.269	0.272	0.270
over 50	0.236	0.250	0.241	0.260	0.251	0.268	0.268	0.271	0.262	0.253	0.262	0.263	0.266	0.265	0.266	0.264

Source: Own construction

From the table and the graph we do some conclusions for three age groups:

- The age groups 30–50 and over 50 are very similar.
- The age group under 30 is quite different — not only in the level but in the trend, too.
- The average value of the Gini index (geometric mean) is:
  - under 30: 0.214,
  - 30–50: 0.257,
  - over 50: 0.259.
- For the better quality of results it will be more suitable to use the detailed age classification, but we have no available data.

Figure 3 Gini index for age groups



Source: Own construction

2.4 Gini index for regions in CR

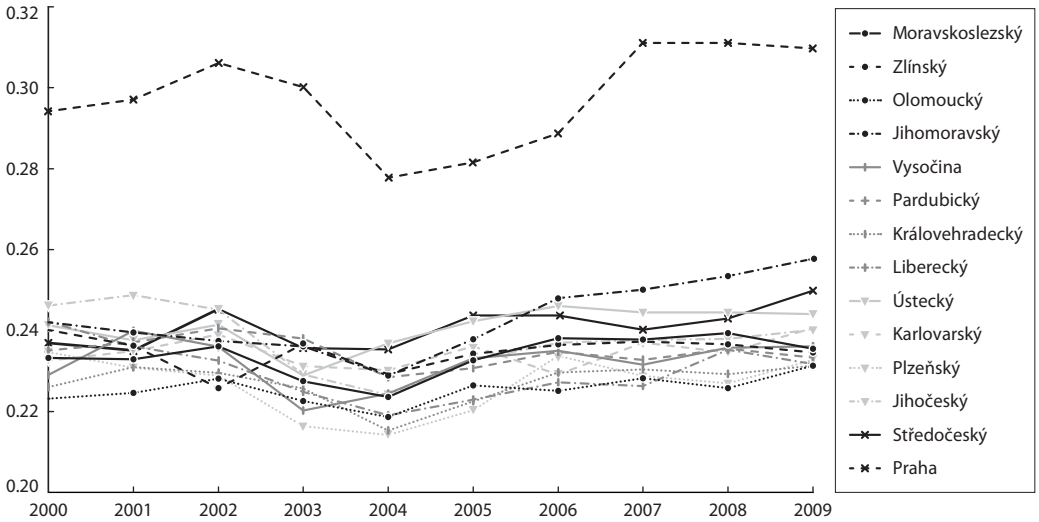
The values of the Gini index for 14 regions of CR are in the Table 4 and in the Figure 4.

Table 4 Gini index for Czech Republic regions

Region \ Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Praha	0.294	0.297	0.306	0.300	0.278	0.281	0.289	0.311	0.311	0.310	0.304
Středočeský	0.237	0.235	0.245	0.236	0.235	0.244	0.244	0.240	0.243	0.250	0.247
Jihočeský	0.246	0.249	0.245	0.229	0.224	0.233	0.237	0.238	0.238	0.240	0.242
Plzeňský	0.235	0.231	0.229	0.216	0.214	0.220	0.234	0.229	0.227	0.233	0.232
Karlovarský	0.233	0.235	0.239	0.231	0.230	0.236	0.229	0.237	0.235	0.241	0.238
Ústecký	0.241	0.238	0.241	0.229	0.237	0.242	0.246	0.245	0.245	0.244	0.246
Liberecký	0.242	0.236	0.233	0.225	0.219	0.223	0.227	0.226	0.235	0.232	0.233
Královehradecký	0.226	0.231	0.230	0.226	0.215	0.222	0.230	0.230	0.229	0.231	0.229
Pardubický	0.235	0.237	0.240	0.238	0.228	0.231	0.235	0.233	0.236	0.233	0.235
Vysočina	0.229	0.240	0.236	0.220	0.224	0.233	0.235	0.232	0.236	0.236	0.236
Jihomoravský	0.242	0.240	0.237	0.236	0.229	0.238	0.248	0.250	0.253	0.258	0.256
Olomoucký	0.223	0.225	0.228	0.223	0.219	0.226	0.225	0.228	0.226	0.231	0.230
Zlínský	0.240	0.236	0.226	0.237	0.229	0.234	0.236	0.237	0.237	0.235	0.236
Moravskoslezský	0.233	0.233	0.236	0.227	0.224	0.233	0.238	0.238	0.239	0.235	0.238

Source: Own construction

Figure 4 Gini index for Czech Republic regions



Source: Own construction

From the computations, table and graph we can make some conclusions:

- Quite different from the other regions is Prague. The Gini index in year 2010 in Prague reaches the value 0.304. None from the other regions is near this value.
- The Gini index is decreasing in most of the regions in year 2010.
- The highest growth is in the region Jihomoravský and Středočeský in the last years (excepting 2010).

Table 5 Average values of the Gini index for Czech Republic regions

Region \ Year	Average
Praha	0.298
Jihomoravský	0.244
Středočeský	0.241
Ústecký	0.241
Jihočeský	0.238
Karlovarský	0.235
Zlínský	0.235
Pardubický	0.235
Moravskoslezský	0.234
Vysočina	0.232
Liberecký	0.230
Královehradecký	0.227
Plzeňský	0.227
Olomoucký	0.226

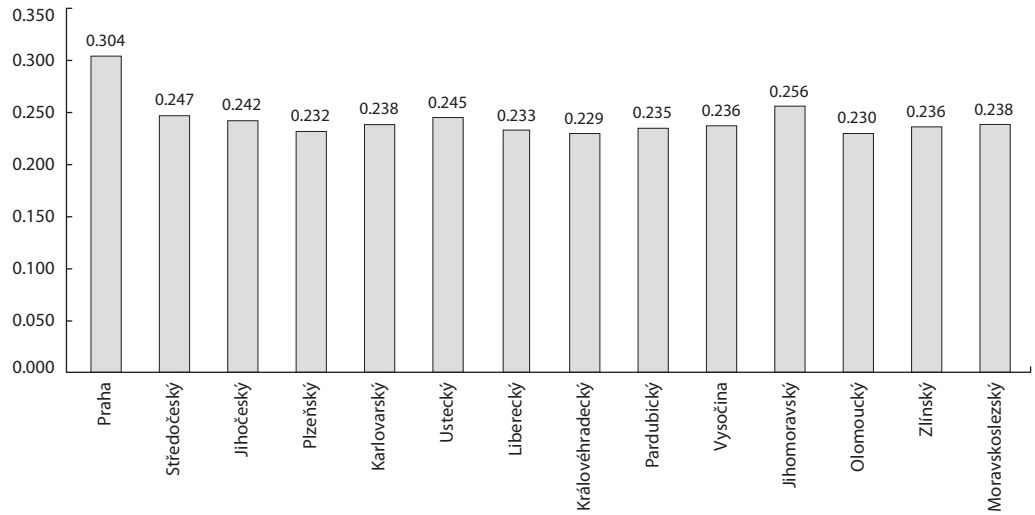
Source: Own construction

The average values of the Gini index over all observed periods are in the Table 5 (using geometric mean). The values are sorted in a descending order:

The values of the Gini index in 2010 are in the Figure 5.

It is clear, with the exception of Prague, that there are very small differencies between the regions. Between the second value 0.256 (region Jihomoravský) and the last value 0.229 (region Královehradecký) is the difference 0.027 only. The difference between the first value (Prague) and the second value (region Jihomoravský) is 0.052, almost double. The difference between the first (Prague) and the last region (Královehradecký) is 0.071! This confirms the fact, that Prague is an extraordinary region in CR. This is valid for not only the Gini index but for all the economic data.

Figure 5 Gini index — regions CR in 2010



Source: Own construction

Table 6 Gini index — selected European countries and USA in 2007

Denmark	0.232	Rumania	0.315
Sweden	0.250	Bulgaria	0.316
Norway	0.258	Ireland	0.320
Czech Rep.	**0.260**	Spain	0.320
Slovakia	0.262	Italy	0.330
Luxemburg	0.268	Greece	0.330
Austria	0.268	Estonia	0.340
Finland	0.269	Great Britain	0.340
Belgium	0.280	Lithuania	0.360
Hungary	0.280	Poland	0.360
Germany	0.280	Latvia	0.377
France	0.287	Portugal	0.385
Cyprus	0.290	USA	0.450
Netherlands	0.309		

Source: Own construction

2.5 International comparison for year 2007

For the international comparison we used the data from the Table 6.

The data in the Table 6 are from 2007 because we cannot find newer data for all the selected countries. Because of that we compare data for year 2007, eventhough we have newer data for Czech Republic. The source of data is OECD (2009). The data for the same year are slightly different when we use an other source such as CIA (2008). Because of that, the comparison is only approximate.

We can see, that Czech Republic belongs to the group of countries with a lower value of the Gini index. It means, that the amount of redistributing is small in CR, with comparison with the rest of the European countries. Smaller values of the Gini index are only in three north countries. The other values in table are higher.

CONCLUSION

The final conclusion is a summary of partial conclusions only. We can generally confirm, that the values of the Gini index in CR for the last 10 years are approximately constant. The differencies in single years are small and are not significantly different from the average value. A similar situation is for the Gini index in accordance to sex and in accordance to age. But there are significant differencies between single groups. When we compare the values of the Gini index in regions, Prague is quite

different from the other regions. The values of the Gini index of Prague are comparable with values of other European countries. In the international comparison the value of the Czech Gini index does not reach the OECD average.

References

CIA. *Central Intelligence Agency — The World Factbook* 2007. CIA, 2008.

MAREK L. Analýza vývoje mezd v ČR v letech 1995–2008. *Politická ekonomie,* 2010, no. 58/2, pp. 186–206. ISSN 0032–3233.

MAREK L. and VRABEC M. Gini Index in Czech Republic. Portorož 5.6.2010–9.7.2010. In *ISBIS — 2010 International Symposium on Business and Industrial Statistics.* Prague: UK MFF, 2010, pp. 60.

MAREK L. and VRABEC M. Porovnání příjmů v krajích ČR. Znojmo 20.11.2008–21.11.2008. In *Nové trendy — Nové nápady 2008* [CD-ROM]. Znojmo: SVŠE, 2008, pp. 363–371. ISBN 978-80-903914-8-2.

MAREK L. The Trend of Incomes in Czech Republic. Porto 24.8.2008–29.8.2008. In *Compstat'2008.* Porto: Copyright, 2008, pp. 269. ISBN 978-972-95396-3-3.

OECD. *Factbook 2006: Economic, Environmental and Social Statistics.* Paris: OECD, 2006.

OECD. *Factbook 2008: Economic, Environmental and Social Statistics.* Paris: OECD, 2008.

VRABEC M. and MAREK L. Forecast of the Income Distribution in the Czech Republic in 2011. Ras Al Khaimah 29.11.2010–3.12.2010. In *ICABR 2010 — VI. International Conference on Applied Business Research* [CD-ROM]. Brno: Mendel University in Brno, 2010, pp. 142. ISBN 978-80-7375-462-4.

Relationship between the Land Rent and Agricultural Land Prices in the Czech Republic

František Střeleček[1] | *University of South Bohemia in České Budějovice, České Budějovice, Czech Republic*
Ladislav Jelínek[2] | *Institute of Agricultural Economics and Information, Prague, Czech Republic*
Jana Lososová[3] | *University of South Bohemia in České Budějovice, České Budějovice, Czech Republic*
Radek Zdeněk[4] | *University of South Bohemia in České Budějovice, České Budějovice, Czech Republic*

Abstract

The aim of this article is the evaluation of factors and the importance of their impact on the land rent and land price. A great share of rented land in the Czech Republic requires an assessment of relationship between the land rent and the price of land. Price of land is primarily influenced by location, size and purpose of use of the purchased land. An average market price of agricultural land regardless its location, size and purpose of the land use showed significant differences year by year. The ratio between the land rent and market price of land referred to as the capitalization rate has been increasing continuously since 2003. The payback period that is reciprocal to the capitalization rate corresponds to the standards of European developed states. An average growth rate of the land rent should not exceed 6 % neither should it lead to a mismatch between the development of the land rent and land prices or to a disproportionate growth of land price.

Keywords

Price of land, land rent, capitalization rate, payback period

JEL code

Q15

INTRODUCTION

Total agricultural area covers 4 244 thousand hectares in the Czech Republic. Agricultural land forms 54 % of the total land area as 38 % of the total is represented by arable land. The share of arable land has been slowly decreasing from approximately 75 % in 1991 to about 71 % in 2008. Property rights related to agricultural area are consolidated with the exception of the state owned land. Total area is divided into 17.5 million parcels of land with an average area of 0.52 ha (Ministry of Agriculture, 2009). Major part of agricultural area is owned by individuals, less is owned by state or private companies and associations of different type.

[1] Faculty of Economics, Studentská 13, 37 005 České Budějovice.
[2] Mánesova 1453/75, 120 56 Prague 2.
[3] Faculty of Economics, Studentská 13, 37 005 České Budějovice, corresponding author: e-mail: lososova@ef.jcu.cz, phone: +420 387 772 457.
[4] Faculty of Economics, Studentská 13, 37 005 České Budějovice.

There is a great number of small owners in the Czech Republic while majority of them do not manage the land they own. Compared to the EU in the Czech Republic, the share of rented land is more than twice as much. The land market has been revived by sales of agricultural land owned by state.

There are two types of land price in the Czech land market. Administrative price is set according to land valuation published in price regulations of the Ministry of Finance of the Czech Republic. Market price is regulated by supply and demand. An average rent of farm land is low in the Czech Republic, compared to the EU; however the growth rate is increasing.

1 LITERATURE SURVEY AND METHODS

The market land price setting is based on three basic approaches specifying current method of valuation. The cost approach is based on the premise that the informed buyer would not pay more than for a property with comparable features. The comparison approach is based on comparing current market prices of land with comparable characteristics. The income approach is based on capitalization of the land income.

Gwartney (2004) adjusted the methodology of the land price valuation by other specific methods: Sales Comparison is based on analysis of vacant parcels and their comparison in order to provide the price of the assessed land. The Proportional Relationship is based on comparison of the size of a parcel with standard size. The difference is expressed as a ratio adjusting the price of a standard parcel. The Land Residual Technique assumes that the land is improved to its best use. All operating expenses and the return attributable to other agents of production are deducted, and the net income is capitalized. The allocation divides the price into two parts expressing the land value and its improvements. The extraction estimates the land value by subtracting estimated value of depreciated improvements from the known sales price of the property. The Ground Rent Capitalization is used when the land rent and market price data are available. The Subdivision Development is based on the assumption that uncultivated land is of the same value as the cultivated and sold land. Cultivating costs and other charges are subtracted from the sales price, and the net income projection is discounted over the estimated period required for market absorption of the cultivated sites.

Huang et al. (2006) discussed the impact of factors that are not directly related to the production. Explanatory variables included land productivity, parcel size, and distances to large cities, an urban–rural index, farm density measures, income, and inflation. They proved that farmland price increases with soil productivity and population density and declines with parcel size, country character of the district, and distance from large cities. A hedonic price model of forest land prices in Northern Minnesota is presented in Snyder et al. (2007). They included economic and social factors as well as recreational features and some commodity variables as explanatory. Access to roads and density, proximity to population centres, presence and proximity to a water body, and the use of contract financing had showed the most positive influence. Chavas and Shumway (1981) model land price as function of economic rent. The economic rent includes the land rent as well as the maximum profit. The land price is expressed as annual discounted flow. To this end, a single Gordon's model with a constant growth rate is specified. Gwartney (2004) compares the land rent and the market price of land. The above mentioned method is based on the following relations:

$$\textit{Capitalization rate} = \textit{Land rent} / \textit{Market price of land}. \quad (1)$$

The following results imply:

$$\textit{Market price of land} = (\textit{Land rent} - \textit{Land tax}) / \textit{Capitalization rate}, \quad (2)$$

$$\textit{Land rent} = \textit{Market price of land} \times \textit{Capitalization rate} + \textit{Land tax}. \quad (3)$$

The assessment in the analysis has been based on the above mentioned relations. The capitalization rate is a very sensitive index requiring special abilities to assess it. To this end, payback period in years is used as well since it is more instructive indicator.

$$\textit{Payback period} = 1 / \textit{Capitalization rate}. \quad (4)$$

Both static and dynamic approach can be employed in the calculation. The statistic approach calculates with the number of years to pay the land rent as a reciprocal value of the capitalization rate. The dynamic approach calculates the time value of money that allows calculating the number of years within required interest rate.

Changes at the land market, land price and land rent after the EU enlargement are discussed in Buday (2007), Němec and Kučera (2007), Hamza and Miskó (2007). The impact of Single Area Payment scheme on the land market and the land rent is analysed by Boinon et al. (2007), Patton et al. (2008) showing that the distributional impact of different types of payment provides a space for further research. Their study revealed that direct decoupled payments are directly connected with land and they directly influence the land rent. The results of simulations prepared up to 2030 with dynamic model of partial balance revealed that the GDP growth resulted to a stronger effect on changes in the land use than the CAP (Ciaian, 2007).

Assessing farmland price in the Czech Republic employs different types of analysis. The research of the Czech Statistical Office covered the whole area. The data are sourced from the land price specified in commercial (purchase) agreements recorded by the Ministry of Finance for purposes of the real estate transfer tax assessment. The research of prices of purchased state land is also monitored within the whole area according to Act 95/1999 Sb. This research is based on the records of the Land Fund of the Czech Republic on purchased farmland classified by the purchase type:

a) according to Section 7 specification — transfers to self-employed farmers and farmland owners, business companies partners, members of cooperatives with the price resulted from the competition; and
b) according to Section 8 — purchase to other individuals (land that was not sold under the previous paragraph).

Land transfers according to paragraphs 5 and 6 are not considered due to their low coverage. The price recorded by the SGAFF (the Support and Guarantee Agricultural and Forestry Fund) is based on commercial agreements between the seller and the buyer selling private land, where the part of interest was paid from the PGRLF to buyers. The programme was announced in 2004 and the submission of requests finished in 2010.

The land rent is based on the FADN (The Farm Accountancy Data Network) database consisting of monetary and in-kind payments per 1 ha of rented ("external") farmland. Classification into production areas is based on localization of a parcel to the land plan within the appropriate area. Currently, there are three types of agricultural area in the Czech Republic: production areas, less favoured areas and vulnerable areas. This paper deals with the land rent in the first and second above mentioned type.

2 RESULTS AND DISCUSSION

2.1 Land price and its development according to different research types

In the Czech Republic, both market price of farmland and official price are used for tax purposes and for the sale and purchase of state-owned land. An average official price amounted to 52 400 CZK/ha ranging from 7 000 CZK/ha to 148 100 CZK/ha. An average market price is similar to the official price as reported by the Czech Statistical Office. The Czech Statistical Office does not specify or analyze (classification, the influence of factors) the data. The land purchased for further non-agricultural use is not distinguished in any way (Table 1).

Table 1 Growth of farmland price in the Czech Republic (CZK / ha)

	2003	2004	2005	2006	2007	2008	I2008 / 2003	Av. growth rate
Farmland price	48 481	49 791	48 279	46 806	51 848	59 257	1.222	1.041

Source: CZSO

Sample survey of prices based on a selection of commercial agreements is presented in Table 2. The data were sourced from commercial agreements and their survey. Agreements were not filtered from the land purchased for non-agricultural purposes (in 2004, the category of land for building purposes was defined as the "virgin building land" before transferring into the category of building land) with the market price up to ten times bigger compared to the average market price of land for agricultural purposes.

Table 2 Sales price growth according to the size of a parcel (CZK / ha)

	2000	2001	2002	2003	2004	2005	2006	2007
Up to 1 ha	921 360	1 087 776	971 424	1 166 803	1 042 360	957 673	1 470 704	1 594 934
1–5 ha	135 994	199 450	129 042	132 286	136 285	107 641	204 698	215 576
Above 5 ha	41 971	55 664	34 020	35 742	37 511	37 094	36 128	35 875

Source: Institute of agricultural economics and information

According to the survey conducted by the Institute of Agricultural Economics and Information (FADN, 2010), the price is significantly influenced by location, size and purpose of purchased land. An average market size of agricultural land regardless location, size and purpose of the purchase differs significantly year by year. The price of land of less than 1 ha has increased by 73 % in 2007 compared to 2000. The price of land from 1 to 5 ha has increased by 58.5 %. Market price of farmland is significantly influenced by the way of land use. Land parcels of less than 1 ha are purchased for different purpose than farming in up to 95 % (Ministry of Agriculture, 2009). It increases their market price considerably. Land parcels of the size ranging from 1 to 5 ha are used for different than farming purposes in 40–50 %. On the contrary, land parcels of more than 5 ha are purchased mainly for agricultural production (approximately 85 %). The development of sales prices is presented in Table 3. The price is defined separately in two categories — arable land and permanent pastures.

Table 3 Sales price of state-owned farmland in the Czech Republic (CZK / ha)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Purchase according to §7	65 943	59 448	46 184	40 877	43 383	42 230	43 992	47 569	52 705	54 058
Arable land	78 093	76 544	62 908	57 763	58 302	54 564	55 923	57 875	65 556	66 877
Permanent pastures	32 447	24 817	22 040	22 302	24 389	27 316	27 580	32 969	36 838	37 878
Purchase according to §8	66 370	24 193	43 430	37 872	42 395	44 527	52 899	68 291	72 407	91 099
Arable land	68 707	28 641	50 730	45 671	50 957	50 842	56 831	78 113	83 244	105 730
Permanent pastures	19 085	14 634	24 945	25 448	28 267	31 901	42 966	51 814	55 876	70 941

Source: Land Fund of the Czech Republic

Table 3 revealed that there has been a change in prices of land purchased according to Section 7 and Section 8 since 2005. By that time, an average price did not differ between the above mentioned categories (with the exception of 2001). Since 2005, the price of land purchased to other people (according to Section 8) has been significantly higher (by 68.5 % in 2009). The greatest difference occurred in the category of permanent pastures which was sold for prices bigger by approximately 15 % in 2002–2005. After 2005, the price of permanent pastures increased by 52–87 % (within the comparison of Section 7 and 8) — see Table 3.

The survey of the SGAFF is based on purchased land not owned by state as the SGAFF extends loan for such purchases. The purchase of 47 738 ha of land was supported within the framework of the Land Purchase Programme to 2008. This type of land is largely used for farming purposes. Comparing different ways of the average land price valuation in the Czech Republic revealed the highest price of land up to 1 ha in which different use than farming is very likely.

Table 4 Price of farmland supported by the SGAFF (CZK / ha)

	2004	2005	2006	2007	2008	2009
Support to purchase of private land	38 484	43 504	42 693	52 741	46 851	69 938

Source: Support and Guarantee Agricultural and Forestry Fund (SGAFF)

Prices of land purchased with the SGAFF support in 2005–2007 with differences smaller than 10 % were the most similar to the survey of the Czech Statistical Office that consider neither the land use nor the parcel size (Table 4). The differences between the results of survey conducted by the Czech Statistical Office and state-owned land purchases ranged between 6 and 15 %.

2.2 AVERAGE LAND RENT CLASSIFIED ACCORDING TO PRODUCTION AREA

Compared to the EU, the rented land in the Czech Republic constitutes a big share in the total area. In 2008, the share amounted to 87 % of farm land and 95 % of arable land. The situation of land owners not farming their land is specific for the Czech Republic with no similar situation in any other state of the EU (Ministry of Agriculture, 2009) was reported. In the Czech Republic, the land rent is established by law and unless the owner and the renter agree differently the land rent amounts to 1 % of the official price of farm land.

Table 5 Growth of land rent according to production area (CZK / ha)

Production area	2003	2004	2005	2006	2007	2008	2009	I2009 / 2003	Av. growth rate
Maize	1 058	1 136	1 273	1 356	1 491	1 624	1 747	1.65	1.09
Beet	1 100	1 247	1 434	1 452	1 534	1 600	1 845	1.68	1.09
Potato	456	508	617	673	752	879	1 011	2.22	1.14
Potato and oats	354	419	508	538	625	765	885	2.50	1.16
Mountain	211	329	376	443	451	510	670	3.17	1.21
CZE	693	782	890	963	1 058	1 134	1 307	1.88	1.11

Source: Farm Accountancy Data Network (FADN)

An average land rent based on the FADN survey amounted to 1 307 CZK/ha in 2009. There were significant differences in the land rents in different production areas. There was only a small difference between maize and beet area while the land rent in potato area amounted to 43–58 % of land rent in maize area, land rent in potato and oat area amounted to 31–51% of land rent in maize area and land rent in mountain areas amounted to 18–38 %. The comparison of the land rent growth rates revealed that the increase was due to worse production area decreasing the difference of the land rent in various production areas (Table 5). The land rent in maize and beet area increased by 65 % and 68 % compared to 2003 while it went up more than three times in mountain areas. Faster increase of the land rent in mountain areas is influenced by compensatory payment in the LFA.

Table 6 The land rent according to LFA (CZK / ha)

LFA type	2001	2002	2003	2004	2005	2006	2007	2008	2009	I2009 / 2001	Av. growth rate
Mountain	231.7	252.0	246.6	339.5	394.6	446.0	499.9	571.3	764.8	3.30	1.16
Other than mountain	396.2	412.6	463.2	503.0	542.6	576.7	685.7	752.1	922.0	2.33	1.11
Partial	792.9	640.3	697.2	674.4	737.2	762.0	843.4	857.0	1 067.0	1.35	1.04
Outside LFA	959.8	1 019.0	1 042.0	1 196.0	1 373.0	1 417.0	1 492.0	1 608.0	1 805.0	1.88	1.08

Source: Farm Accountancy Data Network (FADN)

Discussion on the LFA land rent revealed the same trend as in classification by production area. An increase of the land rent amounted to 88 % outside the LFA in 2001–2009, while in the mountain LFA it amounted to 230 % with double average growth rate in the mountain area (Table 6).

2.3 Relationship between the land rent and market price of land

The ratio between the land rent and market price of rent is called the capitalization rate of farmland. The payback period that is reciprocal to the capitalization rate is more instructive defining the number of years necessary to pay the price of land in the land rent.

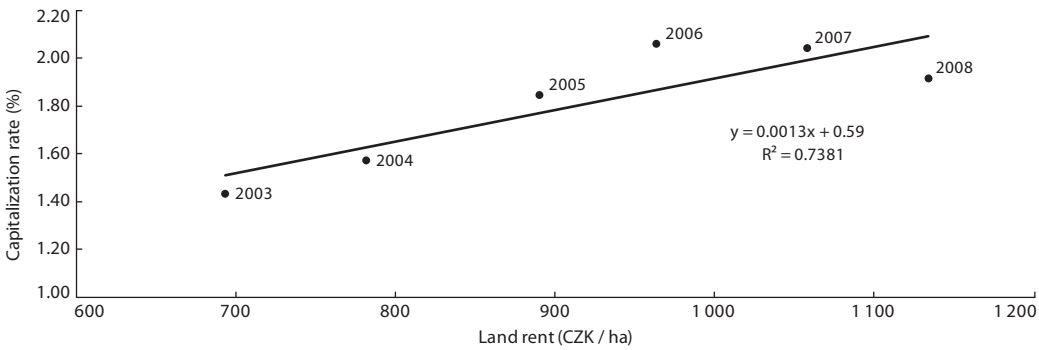
Table 7 Capitalization rate of farmland and payback period

	2003	2004	2005	2006	2007	2008
Price of FL (CZK / ha)	48 481.00	49 791.00	48 279.00	46 806.00	51 848.00	59 257.00
Land rent (CZK / ha)	693.00	782.00	890.00	963.00	1 058.00	1 134.00
Capitalization rate (%)	01.43	01.57	01.84	02.06	02.04	01.91
Payback period	69.91	63.69	54.23	48.58	49.01	52.26

Source: Price of Land — CZSO, Land Rent — FADN

In the Czech Republic, the capitalization rate ranged between 1.43 % and 2.06 % (Table 7). The impact of the land rent on average capitalization rate is presented in Figure 1. The degree of this linear relation expressed as the correlation coefficient of 0.86 has revealed significant statistical dependence.

Figure 1 Relation of capitalization rate and land rent



Source: Price of Land — CZSO, Land Rent — FADN

2.4 Time value of money and the capitalization rate

The price of a parcel should express the bearing interest of the land rent as well as the land rent in a number of years. It is important to compare interest rates of long-term loans with the capitalization rate. The same interest and capitalization rate mean efficient purchase of land. Long-term loans are quite high in different states which do not fit the creation of land price as presented below. An adequacy of bank interest rate and the capitalization rate can be assessed by the real discounted payback period. Discounted payback period:

$$n = \frac{\log \frac{P_0}{P_0 - r \,.\, CP_0}}{\log (1 + r)}, \tag{5}$$

with P_0 = land rent in the period 0, r = interest rate, CP_0 = land price in the period 0 (per 1 ha). The above mentioned equation can be solved only if the capitalization rate will be greater than the interest rate. This condition is connected with many restrictions. If it is required when calculating the average payback period to use the average interest rate for each year, the interest rate in every year must be less than the rate of capitalization in that particular year (Table 8).

Table 8 Number of years necessary to pay the price of farmland by the land rent

	2003	2004	2005	2006	2007	2008
Capitalization rate	1.43	1.57	1.84	2.06	2.04	1.91
Payback period for 0.5 % interest rate	86.00	77.00	63.00	56.00	56.00	61.00
Payback period for 1 % interest rate	121.00	102.00	79.00	67.00	68.00	74.00
Real interest rate	4.12	4.82	3.54	3.80	4.30	4.63
Maximum interest rate that is possible to calculate	1.43	1.57	1.84	2.06	2.04	1.91
Payback period for maximum interest rate	574.00	585.00	336.00	271.00	419.00	333.00

Source: CZSO

Table 8 shows that in any year is not possible to use long-term interest rate as the capitalization rate is significantly lower. In addition, the discounted payback period is unreal even for the applicable maximum return on the appropriate level of capitalization rate. Payback period, the discounted payback period of purchased agricultural land, respectively, ranged from 49 to 70 years in the monitored years. Payback period discounted by one percent ranged from 67 to 121 years, i.e. it is almost twice as long.

The domain of the discounted payback period can be defined by these limits. The lower limit of discounted payback period is determined by the number of years, corresponding to a zero interest rate and thus the inversed capitalization rate. Upper limit of the discounted payback period for the capitalization rate is the interest rate approaching the capitalization rate from the left. Determination of discounted payback period is a question of subjective decision-making. In this respect, you can program the price of land so that for any given level of capitalization rate the expected return can be achieved (Table 9).

Table 9 Discounted payback period in relation to the capitalization rate and interest rate

Capitalization rate (Land rent / market price) in %	Payback period for zero interest rate (years)	Selected interest rate in % (CR — 0,1)	Payback period for selected interest rate (years)
1	100.00	0.9	257.0
2	50.00	1.9	159.2
3	33.33	2.9	119.0
4	25.00	3.9	96.4
5	20.00	4.9	81.8
6	16.67	5.9	71.4
7	14.29	6.9	63.7
8	12.50	7.9	57.6
9	11.11	8.9	52.8
10	10.00	9.9	48.8
1.4304 (2003)	69.90	1.329	201.4
2.0584 (2006)	48.60	1.957	156.0

Source: Own calculation

2.5 Use of the Gordon's model to assess the discounted payback period

Other models based on capitalization rate, which can be used are the Gordon models. Their use is proposed for example by Chavas and Shumway (1981). In terms of international comparisons, the single Gordon model with one continuous growth rate of rents is suitable.

$$CP_n = \frac{P_0 (1 + g)^n}{r - g}. \tag{6}$$

By this model a discounted payback period as well as the land price is possible to asses,

$$n = \frac{\log\left(CP_n (r - g) / P_0\right)}{\log (1 + g)}, \tag{7}$$

or acceptable interest rate:

$$r = \frac{P_0 (1 + g)^n}{CP_n} + g, \tag{8}$$

where CP_n = land price after n years; P_0 = land rent in the period 0; g = growth rate of land rent; r = interest rate; n = discounted payback period in years. The above mentioned analysis revealed that it is impossible to reach appropriate interest rate within mentioned prices and land rent. The calculation of the real interest rate in 2004–2008 is presented in Table 10.

Table 10 Real interest rate based on the land price and land rent

Land rent	2003	693.5
Growth rate of land rent (*g*)	2003–2008	0.1033
Land price	2004	49 791
	2005	48 279
	2006	46 806
	2007	51 848
	2008	59 257
Maximum interest rate *r* (%)	2004	11.87
	2005	12.08
	2006	12.32
	2007	12.32
	2008	12.25

Source: Own calculation

Real interest rate created by comparing the land rent growth rate and land price showed inappropriately high interest rates. It is very likely that the growth rate of price will always be less flexible compared to the growth rate of price which will always be connected with inappropriate growth of interest rate.

To assess the relation between the land rent dynamics and the land price, the capitalization ratio can be used. Its value and dynamics in relation to selected interest rate show if it is profitable to sell the land. Regarding the analysis, it is useful to assess the appropriateness of the land rent growth rate from the following views:

1. To what extent does the real adjustment of land price affect the capitalization rate of given growth rate within certain interest rate?
2. How does the inflation in each year influence this relation?

2.6 The influence of the land rent growth rate on adjusted land price

As a criterion for the assessment of the first task, adjusted real land prices based on the Gordon model is possible to use. This criterion assumes that the growth rate affects the return on rents, but rents will offset dynamics of the dynamics of land prices. Verification of the reality of this condition was comparison of the actual price with the modified price. The average growth rate was calculated for the period of 2003–2008.

$$UCP = \sum_{n=1}^{N} \frac{P_{08}\,(1+g)^n}{(1+r)^n}\,. \tag{9}$$

Table 11 Influence of the land rent increase on the adjusted land price

Land rent 2008	P_{2008}	1 134
Growth rate	g	0.103
Interest rate of long-term loans 2008	r	0.046
Land price 2008	CP_{2008}	59 257
Adjusted land price	UCP	324 270
Difference	$UCP–CP$	265 013
Ratio	UCP / CP	5.47

Source: Own calculation

In terms of adjusted discounted land price (UCP), compliance with the higher growth rates is unrealistic, since the adjusted discounted price of land is 5.5 times higher than real price in 2008 (Table 11). It is becoming clear that the price of land is not able to follow the dynamics of growth rate of rents. The high growth rate of rent has been caused by excessively low rents in 2003. As a result of subsidies and price changes, the profit has improved influencing the growth rate of the land rent. Unbalanced dynamics of land prices leads to the fact that cases of high growth rate of rent is the advantage of the buyer, whereas low rents is an advantage to the seller. This discrepancy may also cause a reluctance to sell land.

2.7 The influence of inflation on adjusted land price

The minimum requirement for the discounted price of land is that the inflation rate should not affect the pricing adversely. It turns out that when the average inflation rate (i) is less than 5 %, the static payback period may not increase the price of land. Inflation rate is reliably covered with the level of capitalization. The impact of inflation is assessed by comparing the land price in 2008 in relation to the modified price discounted by fixed-rate loans. Discounted adjusted price of an average inflation rate is equal to:

$$UCP = \sum_{n=1}^{N} \frac{P_{08}\,(1+i)^n}{(1+r)^n}\,. \tag{10}$$

Table 12 Influence of inflation on adjusted land price

Land rent 2008	P_{2008}	1 134
Average inflation	i	0.0257
Interest rate of long-term loans 2008	r	0.046
Land price 2008	CP_{2008}	59 257
Adjusted land price	UCP	36 322
Difference	$UCP–CP$	–22 935
Rate	UCP / CP	0.61

Source: Own calculation

Application of the average rate of inflation for the years 2003–2008 for a static payback period does not interfere with calculation of land prices, as the average rate of inflation is less than the level of capitalization. Applying the growth rate of inflation for the other constant conditions causes a decrease of land prices. Compared to the price of land in 2008, adjusted discounted price of land is by 39 % lower than the price in 2008 (Table 12). This reduction is significant. Low average inflation, along with lower price also means lower rates of capitalization and consequently longer discounted payback period.

2.8 Acceptable growth rate of land rent including the inflation rate for different price models

The impact of the two previous criteria raised the following question: what is an acceptable average growth rate of land rents for the capitalization rate at which the price of land would not change or increase in prices would be acceptable. Table 13 shows the acceptable growth rate of rent for the price of land in 2008 and 1.1 multiple up to 1.5 multiple of that price.

Table 13 Average growth rate of land rent in 2008 price and its multiples

	g	Capitalization rate	Payback period
For 2008 price	0.0465		
For 1.1 multiple of 2008 price	0.0502	0.0211	47
For 1.2 multiple of 2008 price	0.0535	0.0228	44
For 1.3 multiple of 2008 price	0.0565	0.0244	41
For 1.4 multiple of 2008 price	0.0592	0.0258	39
For 1.5 multiple of 2008 price	0.0617	0.0272	37

Source: Own calculation

Assuming normal development, i.e. that the price of land in 2008 could increase by 50 %, the capitalization rate from 2.1 to 2.7 % and the discounted payback period is also relatively constant. The average growth rate of rent higher than 5 %–6 % would have induced an excessive increase in prices of land, or would lead to a mismatch between the development of rents and land prices. Such a situation would be acceptable only for a transitional period and could lead to a slowdown in trade with land.

CONCLUSION

Many institutions have been dealing with establishing a market price of land in recent years. Unfortunately, their estimates of market land prices vary considerably. The estimated average market price of land in addition to a targeted survey requires a detailed classification of land prices, not only from the aspect of the quality of the land and its size; but also considering the type of its future use, see e.g. Snyder et al. (2007), Chavas and Shumway (1981) and others. A big share of rented land in the Czech Republic (87 % of agricultural land and 95 % of arable land) requires an assessment of the relation of rent and land prices. This relation in addition to other factors affects market with the rented land. The average annual growth rate of rent outside the LFA was 108 %, 116 % in mountain area and 111 % of other LFAs.

Capitalization rate of agricultural land has been increasing continuously since 2003 (1.43) to 2006 (2.06) showing a slight drop in the following years. In 2008, the capitalization rate was 1.91 %. The payback period, which is the reciprocal of the capitalization rate in the range of 52–69 years and corresponds to the standard of developed European states. The discounted payback period due to low level of capitalization is not the solution within higher interest rates, keeping unreasonable price of land with low interest. The high rate of the land rent growth in Gordon's model cannot be reconciled with the level of capitalization that adjusted land prices were real.

The above relations result in the fact that regarding the capitalization rate it is currently advantageous to buy farmland, but it is disadvantageous to sell. On the other hand, there may be other reasons to sell land not mentioned in the paper. The settlement of relations between rents and capitalization rate by increasing the land price is limited while stabilization of the growth rate of land rent is more acceptable.

ACKNOWLEDGMENTS

This paper was prepared as an output of the National Agency for Agriculture Research QH82162 project "The possibilities of farmland ownership defragmentation".

References

BOINON J. P. et al. Enforcement of the 2003 CAP reform in 5 countries of the West European Union: Consequences on land rent and land market. In *Agricultural Economics-Zemedelska ekonomika,* 2007, 53, pp. 173–183.

BUDAY S. Agricultural land market in Slovakia. In *Agricultural Economics-Zemedelska ekonomika,* 2007, 53, pp. 146–153.

CHAVAS J. P. and SHUMWAY, C. R. A Pooled Time-Series Cross-Selection Analysis of Land Prices. In *American Journal of Agricultural Economics,* 1981, 63, pp. 1049–1049.

CIAIAN P. Land use changes in the EU: Policy and macro impact analysis. In *Agricultural Economics-Zemedelska ekonomika,* 2007, 53, pp. 565–579.

GWARTNEY T. *Estimating Land Values* [online]. April 2004. <www.henrygeorge.org/ted.htm>.

HAMZA E. and MISKÓ K. Characteristics of land market in Hungary at the time of the EU accession. In *Agricultural Economics-Zemedelska ekonomika,* 2007, 53, pp. 161–168.

HUANG H. et al. Factors Influencing Illinois Farmland Values. In *American Journal of Agricultural Economics,* 2006, 88, pp. 458–470.

INSTITUTION AGRICULTURAL ECONOMIES AND INFORMATION. *Results of the survey FADN CZ* [online]. 2010. [cit. 13.5.2010]. <http://www.vsbox.cz/fadn>.

MINISTRY OF AGRICULTURE. *Situation and perspective report Soil.* Ministry of Agriculture: Prague, 2009.

NĚMEC J. and KUČERA J. Land market development after the accession to EU. In *Agricultural Economics-Zemedelska ekonomika,* 2007, 53, pp. 154–160.

PATTON M. et al. Assessing the influence of direct payments on the rental value of agricultural land. In *Food Policy,* 2008, 33, pp. 397–405.

SNYDER S. A. et al. Determinants of Forest Land Prices in Northern Minnesota: A Hedonic Pricing Approach. In *Forest Science,* 2007, 53, pp. 25–36.

The Determinants of ICT Expenditures by Households: a Micro Data Analysis[1]

Montagnier Pierre[2] | *Organisation for Economic Co-operation and Development — OECD, Paris, France*
Spiezia Vincenzo[3] | *Organisation for Economic Co-operation and Development — OECD, Paris, France*

Abstract

This paper examines the determinants of ICT expenditures in 17 OECD countries, including the Czech Republic, looking at Household Expenditures Surveys data. Previous OECD work has shown that average ICT expenditures vary with the gender and the educational attainments of the family head and with the lifecycle and the income of the household. Average data, however, hide the interactions among these factors that occur at the level of households. In order to control for these interactions, this paper analyses the determinants of ICT expenditures based on the household-level data (micro data).

Using a double-hurdle model (a probit model to estimate whether a consumer will spend on a certain good or service or not, followed by a truncated regression model that estimates how much to spend on that good or service) applied on micro data collected by households budget surveys from 16 OECD countries, the paper analyses the effects of determinants on ICT expenditures. The paper shows that the effects of various determinants between communication services and information technology goods are markedly different. Some determinants, in particular income and presence of children, have significant and relatively similar effects, at the level of more elementary expenditures components (IT goods, IT services, communication goods and communication services). It also shows that the effects of other determinants (such as education level, geographical area, age, or life cycle stage) are less similar across countries or between ICT goods and services.

Keywords

ICT expenditures by households, determinants, micro data analysis

JEL code

D12, C81, O33

INTRODUCTION

Over the last decade, there has been a clear increase in household ICT expenditure, as well as the ICT budget coefficient (OECD, 2007). It has also been observed that ICT consumption is affected by

[1] Disclaimer: the opinions expressed and arguments employed in this paper do not necessarily reflect the official views of the Organisation or of the governments of its member countries. An earlier version of this paper has been presented at the OECD's Working Party on Indicators for the Information Society at its April 2009 meeting.

[2] ICT Unit, Economic Analysis and Statistics Division, e-mail: pierre.montagnier@oecd.org.

[3] ICT Unit, Economic Analysis and Statistics Division, e-mail: vincenzo.spiezia@oecd.org.

the level of household income, with low-income households spending proportionally more on ICT than high-income households. In addition, low-income households tend to devote a higher share of their ICT expenditure to ICT services compared with high-income households. Following on from these observations, the use of micro-data would make it possible to account for all the determinants of ICT expenditure.

Yin et al. (2005), in one of the few analyses of the impact of the socio-economic factors on ICT expenditure, explored some of the determinants of household expenditure on computer hardware and software. They used the US Bureau of Labour Statistics 2000 Consumer Expenditure Survey (CES) micro-data, applied to the double-hurdle model proposed by Cragg (1971), to analyse consumers' expenditure on durable goods.

Likewise, the present paper is based on micro-data collected by the household budget surveys from selected OECD countries, and using the double-hurdle model, constitutes an initial attempt to shed light on some determinants of household expenditure on ICT goods and services.

1 RESEARCH QUESTIONS, METHODOLOGY AND DATA

1.1 Research questions

Most of the literature concerning the diffusion and use of ICT among households refers to determinants of equipment and use, and tends not to focus specifically on ICT expenditure.

As pointed out by Yin et al. (2005), since most consumers buy computers for self-use, computer ownership can be assumed to be equivalent to computer spending behaviour. For the household, it can be similarly assumed that most of them buy ICT goods and services for self-use. The main determinants on ICT expenditure, although not necessarily the same, should be closely related to those on equipment and use.

For ICT equipment and use, the literature usually focuses on income, age, education, occupation, gender (Bigot, 2006, McKeown et al., 2007, OECD, 2007), marital status and children and, less frequently, on specific variables linked to literacy (Veenhof et al., 2005), cultural capital or attitudes (CERI, 2009, Horrigan, 2007). Some of those determinants are clearly associated with the life cycle stage of the household. The age of the household's reference person, marital status, and presence of children have been frequently used to identify the life cycle stages.

Based on examination of the data on computer ownership, Yin et al. (2005) suggested four groups of households: married with children, married without children, single persons, and others (single parents and other types of households). They also suggested a negative quadratic relationship between age of the household's reference person (positive for age and negative for age-squared) and both the probability of spending and the amount spent.

Income is also one of the important determinants of the household expenditure as generally, the more income a household has, the more goods or services it can afford to purchase. The relationship between income and many types of expenditures has been found to be positive. Computer ownership and Internet access rates are positively related to income. Income is expected to influence positively expenditure on ICT goods and services.

Education may be another factor that affects ICT expenditure. Computer and Internet use are influenced by the level of education of the household's reference person (OECD, 2007). Differences in ICT use and familiarity according to education level have been also pointed out in many countries (Veenhof et al., 2005). US studies (Yin et al., 2005) have shown, for instance, that health and personal care expenditure is positively related to the level of education of the household's reference person. Reading material and occupational expenses are also positively related to the level of education of individuals. We can assume a positive relationship between education and spending on ICT goods and services.

Geographical area may be another factor that affects ICT expenditure. Several indices show that Internet access and use, and mobile access and use, are relatively heterogeneous according to where the household is living. A rural location remained one of the barriers to Internet use in Canada in 2005 (McKeown et al., 2007). In France, in 2006, computer equipment and frequency of use, mobile equipment, use of SMS, Internet access and frequency of use all differed greatly according to the population density of the place of residence (CREDOC, 2007). We can assume a positive relationship between the level of the population density and spending on ICT goods and services.

Many factors impact women's access to and use of ICT, including ICT infrastructures, social norms, time-budget allocation, education, employment, and available content and cultural constraints. Many studies have found gender differences in patterns of computer and Internet use (Veenhof et al., 2005, Montagnier and Van Welsum, 2006). Attitudes toward technology are also not the same according to gender. At the beginning of the 2000s, US female-headed households were found to be less confident about information technology than male-headed households (Yin et al., 2005). It can be expected that gender will have a similar effect on ICT purchase and the amount spent.

From the above, we can formulate the following hypotheses:

Life cycle stage | Households whose reference person is married without children are less likely to spend on ICT than households whose reference person is married with children.

Of households that spend on ICT, households whose reference person is married with children are likely to spend more on ICT than households whose reference person is married without children.

Age | A positive relationship between the household reference person's age and the likelihood of spending on ICT is observed. There is a negative relationship between age-squared and the likelihood of spending on ICT.

Of households that spend on ICT, the effect of age on amount spent is positive and negative for age-squared.

Income | There is a positive relationship between household income and the likelihood of spending on ICT. Of households that spend on ICT, the effect of income on amount spent is positive for income.

Education attainment | Households whose reference person has a low level of education are less likely to spend on ICT than households whose reference person has a high level of education.

Of households that spend on ICT, households whose reference person has a low education are likely to spend less on ICT than households whose reference person has a high level of education.

Density of population (or rural and urban) | Households whose reference person is resident in a rural — or low densely populated — region are less likely to spend on ICT than households whose reference person is resident in an urban — or highly densely populated — region.

Of households that spend on ICT, households whose reference person is resident in a rural — or lowly densely populated — region should spend less on ICT than households whose reference person is resident in an urban — or highly densely populated — region.

Gender | Households with a male reference person are more likely to spend on ICT than households with a female reference person.

Of households that spend on ICT, households with a male reference person are likely to spend more on ICT than households with a female reference person.

1.2 Methodology

In expenditure studies, and this is especially valid for durable goods, it is common for a large number of households not to have purchased anything in a particular category during the survey period. In this case, the dependent variable will be zero for a significant number of observations, and no conclusion can be drawn for the population as a whole. In econometrics, this is referred to as the limited dependent variable problem. In order to take into account this bias, Cragg (1971) proposed a double-hurdle model: it is made first of a probit model, which estimates whether a consumer will spend on a certain good or not, and second of a truncated regression model, in order to estimates how much to spend on that good. The regression model takes into account the selection bias and incorporates it into the regression, so that the results yield for the population as a whole. This correction is known as the Heckman correction, or two-stage method (Heckman, 1979). Due to the "truncated" nature of the dependent variable, the traditional estimation method of OLS (Ordinary Least Square) is not appropriate and the maximum likelihood estimation method is used instead.

Dependent variables

Probit and truncated regression models have been estimated for information and communication expenditure with various combinations of information technology and communication goods and services (see definitions below).

The dependent variable in the probit model is whether to spend on ICT goods and services (and various combinations). It is coded 1 if the household spent on ICT, and coded as 0 otherwise.

The dependent variable in the truncated regression model is the logarithm of the amount spent on ICT goods and services (and various combinations). If there is no selection effect, the simple OLS model is used instead of the Heckman correction for the regression.

Independent variables

Following the approach from Yin et al. (2005), a simple model has been developed and tested in order to see the influence of the main socio-economic factors of households on their ICT expenditure.

According to the research questions mentioned above, income, age, life-cycle stage, education level of the household's reference person, geographical location and gender of the household's reference person should have significant influence on both the probability of spending on ICT goods and services and the level of ICT expenditure.

1.3 Data sources

Data are from the household's budget surveys. For the European countries (except Czech Republic), data are from the Eurostat Database on Household Budget Survey, and refer to the latest collection round in the reference year 2005.

For European countries (except Czech Republic), the authors did not have direct access to the micro-data. Eurostat provided the OECD with an initial sample of data with the selected variables. This sample was used by the OECD to prepare a SAS programme. This programme was then implemented and applied to the micro-data by Eurostat[4] and the aggregated results provided to the OECD. It was therefore only possible to test the assumptions through a limited number of interactions.

For Canada, the data come from the Survey of Household Spending (SHS), and for Czech Republic and Switzerland, the data come from the Household Budget Survey.

---

[4] The authors thank Peter-Paul Borg and Guillaume Osier from Eurostat (F-3 Living Conditions and Social Protection statistics) for their co-operation and support in providing the results of the programme prepared by the OECD and applied to Eurostat micro-data.

1.4 Data

Definition of ICT expenditure

Households' expenditure survey generally use classifications related to consumption functions. For instance, European countries use the United Nation Classification of Individual Consumption According to Purpose (COICOP) classification.

For comparison purpose, ICT goods and services expenditures have been defined using the United Nation Classification of Individual Consumption According to Purpose (COICOP).

ICT components

ICT expenditures have been split according to information technology and communication on the one hand, and goods and services on the other, forming four groups: IT goods, IT services, communication goods and communication services. The detailed selected items are provided in the Annex.

In four countries (Austria, Greece, Hungary and Norway), the variable expenditure has been miscoded: both zero values (i.e. no expenditures) and missing values have been coded as "missing". In order to keep these countries in the analysis, we have decided to treat all missing codes as zero values (i.e. no expenditures). As a consequence of this choice, the effect of all independent variables (i.e. their coefficients) are likely to be underestimated. Therefore, we will report the regression coefficients for these four countries but we will not compare them with those of other countries.

Definition of the independent variables

A first round of preliminary tests for the European countries with the four categories adopted by Yin *et al.* (2005) to define the life cycle stage of the household — married with children, married without children, single persons, and others (single parents and other types of households) — did not lead to significant results, due to strong interactions between the age of the household's reference person and the household life cycle. It was not possible to isolate the respective effects of age and different types of households. It can be also considered that married households without children include two different types of households: young households as yet without children, and more senior households whose children have already left. Similarly, a negative quadratic relationship between age and ICT expenditure could not be clearly established.

It was therefore decided to select as independent variables related to life cycle of the household the presence of children, and if the household's reference person was living in a couple or not.

For the same reason, we specified a linear relationship between ICT expenditure and age. Based on previous research (e.g. ARCEP, 2008) we expect both the probability to spend on ICT and the level of expenditure to decrease with age.

Therefore, the hypotheses regarding life cycle stage and age have been revised as follows:

Life cycle stage — Children

Households with children are more likely to spend on ICT than households without.

Of households that spend on ICT, households with children are likely to spend more on ICT than households without children.

Life cycle stage — Couples

Couples are more likely to spend on ICT than other households.

Of households that spend on ICT, couples are likely to spend more on ICT than other households.

Age

There will be a negative relationship between the household reference person's age and the likelihood of spending on.

Of households that spend on ICT, the effect of age on amount spent will be negative.
The independent (or explanatory) variables are therefore the following:

- *Income of the household:* the logarithm of the income will be selected. For European countries, the equivalent income has been selected (see the methodology).
- *Level of education of the household's reference person:* medium and high level will be compared to low level. This variable is not available for the United Kingdom.
- *Population density of the area where the household is living:* medium and highly densely populated area will be compared to lowly densely populated area. This variable is not available for Ireland and the Netherlands. For Canada, urban will be compared to rural areas.
- *Children:* presence of children (coded 1) will be compared with absence (coded 0). This variable is not available for the Netherlands. For Sweden and the United Kingdom, data could not be exploited properly.
- *Age:* age of the household's reference person. For Canada, age was provided using 5 year bracket intervals. A proxy for age value has been calculated, using the middle of the age interval (i.e. if the age was between 25 and 29, the age value has been put to 27 — see the methodology).
- *Couples:* households living in a couple (coded 1) will be compared with other households.
- *Gender:* households whose reference person is man (coded 1) will be compared with households whose reference person is a woman (coded 0).

Definition of the dependent variables
Preliminary tests, when using total ICT expenditure as a whole as dependent variable, could not lead to any coherent conclusion with respect to the independent variables. It was decided to focus specifically on each of the components of the ICT expenditure: information technology goods, information technology services, communication goods and communication services. The independent variables have specific effects on each of those ICT components which are not observable at a more aggregated level.

The dependent variables are therefore the following:

- Whether the household spends on information technology goods (coded 1) or not (coded 0) for the selection, and the logarithm of the amount of information goods expenditures, for the regression.
- Whether the household spends on information technology services (coded 1) or not (coded 0) for the selection, and the logarithm of the amount of information services expenditures, for the regression.
- Whether the household spends on communication goods (coded 1) or not (coded 0) for the selection, and the logarithm of the amount of communication goods expenditures, for the regression.
- Whether the household spends on communication services (coded 1) or not (coded 0) for the selection, and the logarithm of the amount of communication services expenditures, for the regression.

Descriptive statistics of the dependent and independent variables by country are provided in the Table A2 (see the Annex).

The effects of the independent variables on each of the dependent variable, mirrored by the coefficients in the tables, are analysed and discussed in the next section. An empty cell indicates that the cor-

responding variable could not be included in the regression, either because of strong collinearity with other variables or due to a low quality of the data.

Differences observed between communication goods and communication services should also be interpreted bearing in mind that the supply of communication services, in many countries, may incorporate the supply of communication goods (provided as package), reflecting a blurring of the frontier between goods and services.

2. RESULTS AND DISCUSSION

2.1 IT goods and services

IT goods — selection

The probability of spending on IT goods increases with the income in all the countries. The effect of income on the probability is particularly strong in Spain and Slovak Republic, and weak in the Netherlands, and seems to be relatively similar in countries such Canada, Finland, France.

The probability of spending on IT goods also increases — generally monotonically — with the level of education of the household's reference person: the higher the level of education of the household's reference person, the higher the probability of spending on IT goods. In Finland however, a household whose reference person has a medium level of education has the highest probability of spending on IT goods, followed by households whose reference person has a high level of education.

Living in a densely populated area generally increases the probability that the household will spend on IT goods, except in France and Belgium. The effect of the population density is generally monotonic.

Households with children have a higher probability of spending on IT goods compared with households without children. This is in line with what has been generally observed concerning the adoption of ICT within households.

In all the countries, the probability of spending on IT goods decreases with the age of the household's reference person.

The effect of living in a couple is somewhat different among countries: households living in couple have a higher probability of spending on IT goods in five countries, but a lower probability in six others.

If the household's reference person is a man, it generally increases the probability of spending on IT goods, except in the Slovak Republic.

IT goods — regression

The income elasticity varies from 0.187 in Sweden to 0.83 in Switzerland. In most of the countries, the higher the level of education of the household's reference person, the more the household spends on IT goods. In Denmark, by contrast, households whose reference person has a high level of education spend less on IT goods compared to households whose reference person has a lower level of education.

Overall, the more densely the area is populated, the more the household spends on IT goods. By contrast, this relation is the reverse in France. And in Denmark and Finland, the households living in medium densely populated area spend more than those living in highly densely populated area.

Households with children spend more on IT goods, compared with households without children, in all the countries but France.

IT goods expenditures decrease with the age of the household's reference person.

As for the probability of spending on IT goods, the impact of living in a couple on the level of IT goods expenditure is somewhat different between countries: households living in a couple spend more on IT goods in Canada, Czech Republic and Sweden, but less in five other countries. When the household's reference person is a man, it increases the expenditure on IT goods.

Table 1 IT goods selestion[1]

		intercept	ln income	d_edu_high	d_edu_med	d_geo_high	d_geo_med	d_child	age	d_cple	d_male
Austria	Coef.	–0.814	0.017	0.570	0.312	0.146	0.067	0.484			
	SE	0.0028	0.0002	0.0025	0.0018	0.0017	0.0019	0.0016			
Belgium	Coef.	–2.734	0.288	0.149	0.040	0.017	–0.101	0.088	–0.012	–0.051	0.168
	SE	0.0176	0.0017	0.0026	0.0027	0.0048	0.0049	0.0026	0.0001	0.0022	0.0022
Canada[2]	Coef.	–2.583	0.455	0.620	0.276	0.017		0.266	–0.030	0.165	0.046
	SE	0.0076	0.0007	0.0015	0.0013	0.0015		0.0017	0.00003	0.0012	0.0011
Czech Rep.	Coef.	–3.946	0.327	0.2853*	0.1669*	0.261	0.1685*	0.378	–0.021	0.733	0.1730**
	SE	0.8722	0.0580	0.0953*	0.0579*	0.0626	0.0685*	0.0586	0.00202	0.1023	0.1006**
Denmark	Coef.	–2.842	0.278	0.388	0.303	0.248	0.132	0.294			
	SE	0.01092	0.00105	0.00260	0.00200	0.00217	0.00224	0.00230			
Finland	Coef.	–4.636	0.436	0.407	0.427	0.142	0.194	0.465			
	SE	0.01428	0.00142	0.00239	0.00207	0.00201	0.00247	0.00245			
France	Coef.	–3.625	0.427	0.221	0.174	–0.099	0.007	0.112	–0.020	–0.087	0.079
	SE	0.0046	0.0005	0.0008	0.0006	0.0006	0.0008	0.0007	0.0000	0.0006	0.0006
Greece	Coef.	–4.582	0.454	0.180	0.089	0.258	0.146	0.503	–0.017	–0.149	
	SE	0.0121	0.0012	0.0021	0.0017	0.0015	0.0039	0.0018	0.0001	0.0016	
Hungary	Coef.	–4.965	0.462	0.402	0.235	0.078	0.075	0.328			
	SE	0.0107	0.0012	0.0020	0.0017	0.0018	0.0018	0.0016			
Ireland	Coef.	–2.847	0.349	0.199	0.113	n.a.	n.a.	0.119	–0.015	–0.011	0.151
	SE	0.0179	0.0016	0.0030	0.0029	n.a.	n.a.	0.0027	0.0001	0.0025	0.0024
Netherlands	Coef.	–0.6336†	0.179	0.535	0.3267*	n.a.	n.a.	n.a.	–0.018	0.2565*	0.367
	SE	0.4406†	0.0416	0.1306	0.1134*	n.a.	n.a.	n.a.	0.00264	0.0930*	0.0923
Norway	Coef.	–0.4225*	0.1144	0.303	0.1607*	0.0829*	0.0859**	0.368	–0.0194		
	SE	0.1419*	0.0124	0.0536	0.0491*	0.0387*	0.0503**	0.0432	0.0013		
Slovak Rep.	Coef.	–6.838	0.570	0.564	0.468	0.205	0.090	0.176	–0.003	0.083	–0.052
	SE	0.0259	0.0026	0.0067	0.0061	0.0033	0.0030	0.0029	0.0001	0.0026	0.0028
Spain	Coef.	–5.476	0.632	0.167	0.146	0.281	0.052	0.166	–0.017	–0.218	0.137
	SE	0.0079	0.0008	0.0011	0.0011	0.0009	0.0010	0.0010	0.00003	0.0008	0.0010
Sweden	Coef.	–2.179	0.292	0.324	0.024	0.024	0.015	n.a.	–0.019	0.235	0.126
	SE	0.0098	0.0010	0.0020	0.0019	0.0016	0.0020	n.a.	0.0000	0.0016	0.0014
Switzerland	Coef.	–3.612	0.389	0.530	0.454	n.a.	n.a.	0.292	–0.012	–0.069	0.074
	SE	0.0134	0.0015	0.0029	0.0027	n.a.	n.a.	0.0017	0.0000	0.0019	0.0019
United Kingdom	Coef.	–3.607	0.373	n.a.	n.a.	–0.0360†	–0.0311†	n.a.	–0.014		
	SE	0.1356	0.0118	n.a.	n.a.	0.0328†	0.0383†	n.a.	0.000547		

[1] The coefficients have all a p value < .001, except: * p < .05; **p < 0.1; † p > =0.1.

[2] Income instead of equivalent income. Urban instead of d_geo_high and d_geo_med. See methodology.

Source: OECD, based on data from the Czech Statistical Office, Eurostat, Statistics Canada and the Swiss Federal Statistical Office

Table 2 IT goods regression[1]

		intercept	ln income	d_edu_high	d_edu_med	d_geo_high	d_geo_med	d_child	age	d_cple	d_male
Austria	Coef.	6.231	0.015	0.254	0.129	0.073	–0.0029†	0.048			
	SE	0.0152	0.0003	0.0052	0.0035	0.0024	0.0025†	0.0039			
Belgium	Coef.	3.240	0.222	0.275	0.039	0.117	–0.157	0.152	–0.009	–0.065	0.304
	SE	0.0552	0.0043	0.0046	0.0046	0.0077	0.0080	0.0040	0.0002	0.0035	0.0041
Canada[2]	Coef.	0.053	0.614	0.271	0.089	0.151		0.178	–0.017	0.025	0.177
	SE	0.0065	0.0006	0.0013	0.0013	0.0012		0.0009	0.00003	0.0009	0.0008
Czech rep.	Coef.	6.374	0.290	0.524	0.2473*	0.2133*	0.1187†	0.279	–0.023	0.617	0.4253*
(ols)	SE	1.1707	0.0775	0.1115	0.0771*	0.0827*	0.0911†	0.0628	0.0029	0.1569	0.1621*
Denmark	Coef.	0.682	0.455	–0.449	0.069	0.110	0.190	0.224			
	SE	0.0333	0.0027	0.0053	0.0045	0.0045	0.0046	0.0041			
Finland	Coef.	2.210	0.345	0.289	0.313	0.124	0.170	0.145			
	SE	0.0265	0.0021	0.0030	0.0029	0.0022	0.0026	0.0026			
France	Coef.	1.569	0.433	0.268	0.159	–0.138	–0.137	–0.250	–0.006	–0.092	0.045
	SE	0.0138	0.0013	0.0012	0.0011	0.0009	0.0012	0.0010	0.0001	0.0009	0.0009
Greece	Coef.	–0.406	0.550	0.208	0.161	0.025	–0.124	0.344	–0.017	–0.208	
	SE	0.0248	0.0022	0.0029	0.0025	0.0022	0.0056	0.0031	0.0001	0.0024	
Hungary	Coef.	1.165	0.365	0.224	0.100	0.022	–0.038	0.018			
	SE	0.0296	0.0025	0.0030	0.0025	0.0025	0.0025	0.0024			
Ireland	Coef.	3.320	0.278								
(ols)	SE	0.2909	0.02577								
Netherlands	Coef.	3.918	0.238	0.1916†	–0.0009†	n.a.	n.a.	n.a.	–0.011		
(ols)	SE	0.3946	0.0358	0.1250†	0.1183†	n.a.	n.a.	n.a.	0.0024		
Norway	Coef.	5.189	0.092	0.297	0.1344**						
(ols)	SE	0.2207	0.0163	0.0820	0.0797**						
Slovak Rep.	Coef.	–0.7335†	0.560								
(ols)	SE	1.2264†	0.1305								
Spain	Coef.	–0.345	0.576	0.484	0.365	0.151	0.054	0.072	–0.026	–0.105	0.133
	SE	0.0190	0.0018	0.0017	0.0017	0.0016	0.0018	0.0015	0.0001	0.0015	0.0017
Sweden	Coef.	4.492	0.187	0.087	0.052	0.244	0.023	n.a.	–0.008	0.043	0.112
	SE	0.0176	0.0016	0.0025	0.0022	0.0017	0.0021	n.a.	0.0001	0.0018	0.0016
Switzerland	Coef.	–4.398	0.835	0.669	0.654	n.a.	n.a.	0.200	–0.021	–0.180	0.268
	SE	0.0280	0.0030	0.0062	0.0059	n.a.	n.a.	0.0032	0.0001	0.0036	0.0034
United Kingdom	Coef.	3.523	0.317	n.a.	n.a.	–0.1484†	–0.0261†	n.a.	–0.0052*		
(ols)	SE	0.4356	0.0382	n.a.	n.a.	0.1068†	0.1247†	n.a.	0.0020*		

[1] The coefficients have all a p value < .001, except: * p < .05; ** p < 0.1; † p > =0.1.

[2] Income instead of equivalent income. Urban instead of d_geo_high and d_geo_med. See methodology.

Source: see Table 1

IT services — selection

The probability to spend on IT services increases with the income in all the countries, and the effect of income on the probability is particularly strong in France, Ireland and Slovak Republic.

The education level of the household's reference person always has a positive, generally inverse u-shaped, effect on the probability that this household will spend on IT services. In a significant number of countries, the strongest effect is provided by the medium level of education, followed by the highest level of education.

The household probability of spending on IT services generally increases with the population density of the area where the household lives. However, the relation is inverted in Sweden. And in that country and in the Slovak Republic and the United Kingdom, households living in a highly densely populated area have the lowest probability of spending on IT services compared with households living in other areas.

Table 3 IT services — selection[1]

		intercept	ln income	d_edu_high	d_edu_med	d_geo_high	d_geo_med	d_child	age	d_cple	d_male
Austria	Coef.	–1.349	0.010	0.179	0.159	0.197	0.091	0.069	0.003	0.022	0.089
	SE	0.0047	0.0003	0.0029	0.0021	0.0019	0.0022	0.0020	0.0001	0.0018	0.0018
Belgium	Coef.	–1.12434	0.208099								
	SE	0.0141	0.0014								
Canada[2]	Coef.	–3.070	0.383	0.495	0.191	0.112		0.281	–0.0003	0.508	–0.191
	SE	0.0083	0.0008	0.0017	0.0015	0.0017		0.0019	0.00004	0.0014	0.0012
Czech Rep.	Coef.	–2.7465*	0.1763*	–0.0924†	0.339	–0.1146†	0.0148†	0.0755†	0.027	0.736	
	SE	1.1917*	0.0798*	0.1305†	0.0940	0.0956†	0.1099†	0.0719†	0.00307	0.0927	
Denmark	Coef.	–1.771	0.340								
	SE	0.0125	0.0012								
Finland	Coef.	–1.010	0.226								
	SE	0.0179	0.0017								
France	Coef.	–7.076	0.734								
	SE	0.0039	0.0004								
Greece	Coef.	0.588	0.192	0.161	0.284	0.749	3.3740†	0.040	–0.001		
	SE	0.0582	0.0054	0.0128	0.0104	0.0113	9.4331†	0.0093	0.0002		
Hungary	Coef.	–3.689	0.347	0.323	0.310	1.009	0.600	–0.090			
	SE	0.0100	0.0011	0.0020	0.0017	0.0018	0.0017	0.0017			
Ireland	Coef.	–5.411	0.580	0.320	0.257	n.a.	n.a.	0.523			
	SE	0.0186	0.0018	0.0041	0.0037	n.a.	n.a.	0.0039			
Netherlands	Coef.	–1.811	0.0858**	0.1771*							
	SE	0.4824	0.0473**	0.0785*							
Norway	Coef.	–1.536	0.0274*	0.1531*	0.1732*	0.1179*	0.207				
	SE	0.1633	0.0122*	0.0606*	0.0573*	0.0436*	0.0552				
Slovak Rep.	Coef.	–4.366	0.527	0.428	0.464	–0.239	0.172	0.226	0.008	–0.112	–0.074
	SE	0.0228	0.0025	0.0048	0.0035	0.0032	0.0029	0.0032	0.0001	0.0028	0.0028
Spain	Coef.	–4.742	0.445	0.036	0.075	0.245	0.204	0.061	–0.008	–0.190	0.198
	SE	0.0077	0.0007	0.0010	0.0010	0.0008	0.0010	0.0009	0.00003	0.0008	0.0010
Sweden	Coef.	–1.685	0.151	0.031	0.314	–0.214	–0.081	n.a.	0.028	0.412	0.056
	SE	0.0102	0.0010	0.0031	0.0031	0.0022	0.0029	n.a.	0.0001	0.0023	0.0020
Switzerland	Coef.	–3.756	0.531	0.202	0.174	n.a.	n.a.	–0.028	0.011	0.375	–0.126
	SE	0.0214	0.0025	0.0043	0.0037	n.a.	n.a.	0.0034	0.00008	0.0034	0.0031
United Kingdom	Coef.	0.764	0.314	n.a.	n.a.	–0.1297*	0.0216†	n.a.	–0.046	0.132	
	SE	0.1698	0.0148	n.a.	n.a.	0.0452*	0.0533†	n.a.	0.0009	0.0285	

[1] The coefficients have all a p value < .001, except: * p <. 05; ** p < 0.1; † p > =0.1.

[2] Income instead of equivalent income. Urban instead of d_geo_high and d_geo_med. See methodology.

Source: see Table 1

Households with children have a higher probability of spending on IT services, except in Hungary and Switzerland. The positive effect of a child's presence is the strongest in Ireland.

The effect of age on the probability of households spending on IT services varies according to the country, contrasting with its systematic negative orientation with respect to IT goods.

Households living in a couple have generally a higher probability of spending on IT services, except in Slovak Republic and Spain.

The gender of the household's reference person does not have a similar effect in all the countries on the probability of the household spending on IT services. This contrasts with the positive effect on the probability of spending on IT goods when the household's reference person is a man.

IT services — regression

The income elasticity is varies from 0.07 in Finland to 0.6 in France.

Table 4 IT services — regression[1]

		intercept	ln income	d_edu_high	d_edu_med	d_geo_high	d_geo_med	d_child	age	d_cple	d_male
Austria	Coef.	6.840	0.0096*								
(ols)	SE	0.0435	0.0044*								
Belgium	Coef.	4.123	0.090								
	SE	0.0073	0.0007								
Canada[2]	Coef.	2.779	0.323	0.077	0.078	0.083		0.064	0.001	0.119	–0.024
	SE	0.0036	0.0003	0.0007	0.0007	0.0007		0.0005	0.00002	0.0005	0.0004
Czech Rep.	Coef.	8.071	0.179	0.0605†	0.0289†	0.106	0.0930*	0.097	0.0007†	0.208	
	SE	0.3752	0.0248	0.0377†	0.0242†	0.0260	0.0285*	0.0212	0.0009†	0.0264	
Denmark	Coef.	4.621	0.129								
	SE	0.0049	0.0005								
Finland	Coef.	4.664	0.070								
	SE	0.0027	0.0003								
France	Coef.	–1.356	0.603								
	SE	0.0033	0.0003								
Greece	Coef.	0.679	0.316	0.183	0.082	0.064	0.0148†	0.121			
(ols)	SE	0.1475	0.0147	0.0275	0.0216	0.0188	0.0494†	0.0206			
Hungary	Coef.	3.429	0.151	0.092	0.104	0.233	0.125	0.036			
	SE	0.0101	0.0008	0.0011	0.0010	0.0020	0.0016	0.0009			
Ireland	Coef.	3.353	0.228	–0.0038*	0.089	n.a.	n.a.	0.040			
	SE	0.0155	0.0013	0.0016*	0.0017	n.a.	n.a.	0.0015			
Norway	Coef.	5.726	0.166	–0.985	–0.739	–0.3228*					
(ols)	SE	0.3987	0.0306	0.1717	0.1621	0.1026*					
Slovak Rep.	Coef.	2.287	0.179	0.160	0.098	0.447	0.300	0.016	–0.007	0.0024*	–0.013
	SE	0.0110	0.0011	0.0021	0.0017	0.0013	0.0011	0.0012	0.0000	0.0010*	0.0011
Spain	Coef.	–0.8908†	0.470								
(ols)	SE	0.7401†	0.0713								
Sweden	Coef.	4.582	0.108	0.072	0.132	–0.042	–0.057	n.a.	–0.002	0.076	0.083
	SE	0.0043	0.0004	0.0009	0.0008	0.0007	0.0008	n.a.	0.0000	0.0007	0.0006
Switzerland	Coef.	2.511	0.125	–0.092	–0.053	n.a.	n.a.	0.022	0.003	0.045	0.034
	SE	0.0054	0.0006	0.0009	0.0008	n.a.	n.a.	0.0006	0.0000	0.0007	0.0006
United Kingdom	Coef.	3.563	0.227	n.a.	n.a.	0.0840*	0.0409†	n.a.	–0.003	0.0248†	
(ols)	SE	0.1462	0.0127	n.a.	n.a.	0.0393*	0.0459†	n.a.	0.0007	0.0225†	

[1] The coefficients have all a p value < .001, except: * p < .05; ** p < 0.1; † p > =0.1.

[2] Income instead of equivalent income. Urban instead of d_geo_high and d_geo_med. See methodology.

Source: see Table 1

In most of countries, when the household's reference person has a high level of education, the household has a higher expenditure on IT services, compared to households whose reference person has a low level of education. By contrast, for households whose reference person has a high level of education, the level of expenditure is lower in Ireland and in Switzerland.

Households living in densely populated areas generally spend more on IT services, except in Norway and Sweden. Household expenditure on IT services decreases with age in three European countries, contrasting with Canada and Switzerland where they increase.

Households with children spend more on IT services compared with households without children. Similarly, households living in a couple also spend more on IT services.

The gender of the household's reference person does not have a homogenous effect on the level of IT services expenditure of that household.

IT goods and services

Overall, the effects of the various explanatory variables seem to be more homogeneous, as far as both selection and regression are concerned, for IT goods compared with IT services. In addition, the effects between selection and regression seem to be much more similar for IT goods as for IT services.

For IT goods, income, education, child, population density in the area where the household live and household's reference person being a man are all factors which generally increase both the probability of spending and the level of expenditure of the household. And both the probability and the level decrease with age. By contrast, the fact that the household lives in a couple does not lead to similar effects.

For IT services, income, child, and the fact that households live in a couple, are factors that have a positive effect on both probability and level of expenditures. Education generally has a positive effect on the probability of spending on IT services but the effect is less homogeneous as regards the level of expenditure. And age, population density in the area where the household live and household's reference person being a man are all factors which have a rather heterogeneous effect on both the probability of spending and on the level of expenditure.

2.2 Communication goods and services

Communication goods — selection

The probability of spending on communication goods increases with the income in all the countries. The positive effect of income on the probability is relatively strong in Belgium and Spain, weaker in France and Denmark, and seems to be relatively similar among countries such as Canada and Finland, or Czech Republic and Switzerland.

The probability of spending on communication goods also increases with the level of education of the household's reference person: the higher the level of education of the household's reference person, the higher generally is the probability of spending on communication goods. In Finland, Norway, and the Slovak Republic however, a household whose reference person has a medium level of education has the highest probability of spending on ICT, followed by households whose reference person has a high level of education. In Ireland, by contrast, for households whose reference person has a high level of education, the probability of spending in communication goods is the lowest.

Living in a densely populated area tends to have a more heterogeneous effect, depending on the country, on the probability that the household will spend on communication goods.

As observed for IT goods, households with children generally have a higher probability of spending on communication goods, compared with households without children. And in all the countries, the probability of spending on communication goods decreases with the age of the household's reference person.

The impact of living in a couple is generally positive, except in Ireland and Spain.

If the household's reference person is a man, the effect varies according to the country.

Table 5 Communication goods — selection[1]

		intercept	ln income	d_edu_high	d_edu_med	d_geo_high	d_geo_med	d_child	age	d_cple	d_male
Austria	Coef.	−2.236	0.008	0.182	0.125	−0.059	0.0013†	0.289			
	SE	0.0061	0.0005	0.0055	0.0041	0.0036	0.0040†	0.0032			
Belgium	Coef.	−5.521	0.408								
	SE	0.0208	0.0019								
Canada[2]	Coef.	−2.000	0.174	0.130	0.116	−0.007		0.079	−0.011	0.061	−0.046
	SE	0.0059	0.0006	0.0012	0.0012	0.0012		0.0009	0.00003	0.0009	0.0008
Czech rep.	Coef.	−4.090	0.227	−0.0130†	0.0504†	0.1917*	0.0359†	0.274	−0.007	0.514	
	SE	0.8605	0.0567	0.0836†	0.0550†	0.0589*	0.0657†	0.0470	0.0019	0.0613	
Denmark	Coef.	−2.482	0.118	0.103	0.095	0.053	0.043	0.298			
	SE	0.0165	0.0016	0.0030	0.0024	0.0025	0.0026	0.0025			
Finland	Coef.	−2.376	0.163	0.077	0.211	0.051	0.032	0.590			
	SE	0.0141	0.0014	0.0024	0.0021	0.0020	0.0024	0.0022			
France	Coef.	−2.200	0.101	0.360	0.301	−0.094	0.052	0.280			
	SE	0.0048	0.0005	0.0008	0.0007	0.0006	0.0009	0.0007			
Greece	Coef.	−3.429	0.178	0.214	0.031	0.068	0.187	0.073	−0.007	0.040	−0.068
	SE	0.0219	0.0022	0.0036	0.0032	0.0026	0.0063	0.0032	0.0001	0.0029	0.0033
Hungary	Coef.	−3.202	0.282	0.093	0.049	−0.071	0.089	0.191	−0.013		
	SE	0.0146	0.0015	0.0024	0.0021	0.0023	0.0022	0.0020	0.0001		
Ireland	Coef.	−2.998	0.291	−0.088	0.045	n.a.	n.a.	0.421	−0.018	−0.243	0.177
	SE	0.0185	0.0017	0.0031	0.0030	n.a.	n.a.	0.0027	0.0001	0.0026	0.0026
Netherlands	Coef.	−2.150	0.163								
	SE	0.4240	0.0412								
Norway	Coef.	−0.766	0.0218*								
	SE	0.1297	0.0101*								
Slovak Rep.	Coef.	−5.202	0.307	0.0258*	0.050	0.093	−0.061	−0.074	−0.0004**	0.103	0.108
	SE	0.0551	0.0057	0.0124*	0.0106	0.0068	0.0065	0.0065	0.0002**	0.0057	0.0064
Spain	Coef.	−4.921	0.429	0.173	0.068	0.070	−0.053	0.079	−0.011	−0.098	0.073
	SE	0.0092	0.0009	0.0011	0.0012	0.0010	0.0012	0.0010	0.0000	0.0009	0.0011
Sweden	Coef.	−1.290	0.138	0.087	0.047	−0.063	0.014	n.a.	−0.017	0.093	−0.026
	SE	0.0109	0.0011	0.0022	0.0021	0.0017	0.0020	n.a.	0.0001	0.0017	0.0015
Switzerland	Coef.	−3.473	0.229	0.182	0.174	n.a.	n.a.	0.028	−0.010	0.110	−0.037
	SE	0.0237	0.0027	0.0052	0.0049	n.a.	n.a.	0.0028	0.0001	0.0033	0.0031
United Kingdom	Coef.	−3.635	0.207	n.a.	n.a.	−0.0346†	0.0615†	n.a.	−0.007		
	SE	0.2593	0.0223	n.a.	n.a.	0.0617†	0.0706†	n.a.	0.0011		

[1] The coefficients have all a p value < .001, except: * p < .05; ** p < 0.1; † p > =0.1.

[2] Income instead of equivalent income. Urban instead of d_geo_high and d_geo_med. See methodology.

Source: see Table 1

Communication goods — regression

The income elasticity varies from 0.13 in Ireland to 0.52 in Denmark.

The level of education of the household's reference person has a very heterogeneous effect on the level of expenditure that a household devotes to communication goods.

Households living in a high densely populated area spend less on communication goods in European countries but more in Canada.

Households with children spend more on communication goods, compared with households without children, in all the countries except Hungary.

As with IT goods expenditure, communication goods expenditure decreases with the age of the household's reference person.

Table 6 Communication goods — regression[1]

		intercept	ln income	d_edu_high	d_edu_med	d_geo_high	d_geo_med	d_child	age	d_cple	d_male
Belgium	Coef.	2.749	0.256								
	SE	0.1078	0.0071								
Canada[2]	Coef.	1.634	0.295	−0.030	0.035	0.133		0.089	−0.005	−0.101	0.016
	SE	0.0312	0.0018	0.0020	0.0020	0.0014		0.0013	0.0001	0.0012	0.0010
Denmark	Coef.	−1.6075†	0.518								
(ols)	SE	1.5393†	0.1418								
Finland	Coef.	2.325	0.239	−0.046	0.097	−0.021	0.075	0.160			
	SE	0.0351	0.0022	0.0030	0.0032	0.0024	0.0028	0.0054			
France	Coef.	2.199	0.210	0.085	−0.0235*	−0.126	−0.021	0.0266*			
	SE	0.1137	0.0037	0.0131	0.0111*	0.0037	0.0027	0.0101*			
Greece	Coef.	1.4471**	0.315								
(ols)	SE	0.8526**	0.0808								
Hungary	Coef.	2.350	0.209	0.104	0.069	−0.0066*	0.037	−0.097	−0.004		
	SE	0.0334	0.0027	0.0032	0.0028	0.0031*	0.0029	0.0029	0.0001		
Ireland	Coef.	3.510	0.136	−0.159	−0.139	n.a.	n.a.	0.133			
(ols)	SE	0.2259	0.0206	0.0397	0.0400	n.a.	n.a.	0.0330			
Netherlands	Coef.	0.9976*	0.236								
	SE	0.4997*	0.0460								
Spain	Coef.	1.2264†	0.2635*								
(ols)	SE	0.8771†	0.0839*								
Sweden	Coef.	4.269	0.147	−0.0973†	0.0390†	−0.0318†	−0.1959*	n.a.	−0.013		
(ols)	SE	0.3618	0.0349	0.0975†	0.0954†	0.0750†	0.0890*	n.a.	0.0024		
		<.0001	<.0001	0.319	0.6829	0.6718	0.0281	.	<.0001		
Switzerland	Coef.	−0.5026†	0.513								
(ols)	SE	1.3844†	0.1506								

[1] The coefficients have all a p value < .001, except: * p < .05; ** p < 0.1; † p > =0.1.
[2] Income instead of equivalent income. Urban instead of d_geo_high and d_geo_med. See methodology.
Source: see Table 1

Communication services — selection

The probability of spending on communication services increases with the income in all the countries. The positive effect of income on the probability is particularly strong in Spain, contrasting with its weakness in Denmark.

The education level of the household's reference person always has a positive effect on the probability that this household will spend on communication services. In a significant number of countries though, the strongest effect is provided by the medium level of education, followed by the highest level of education.

The household probability of spending on communication services generally increases with the population density in the area where the household lives. In Denmark and Canada, households living in a highly densely populated area have the lowest probability of spending on IT services compared with households living in other areas.

Households with children have a higher probability of spending on communication services, except in Denmark, the Slovak Republic and Switzerland. The positive effect of a child's presence is the strongest in Canada.

The effect of age on the probability of households spending on communication services is generally positive, except in Austria and Ireland. This overall positive effect contrasts with the generally negative effect observed as far as IT goods and communication goods are concerned.

The impact of living in a couple is positive on the probability of spending on communication services, except in Finland. The impact is the strongest in Ireland.

If the household's reference person is a man, it decreases the probability of spending on communication services in all the countries.

Communication services — regression

The income elasticity varies from 0.09 in Denmark to 0.7 in the Slovak Republic.

In most of countries, when the household's reference person has a high level of education, the household has a higher expenditure on communication services, compared with households whose reference person has a low level of education. In more than half the countries, the higher the level of education of the household's reference person, the more the household spends on communication services. By contrast, in Finland and Ireland, for households whose reference person has a high level of education, the level of expenditures is the lowest.

Households living in a densely populated area generally spend more on communication services, except in Canada and France.

Households with children spend more on communication services, compared with households without children.

Table 7 Communication services — selection[1]

		intercept	ln income	d_edu_high	d_edu_med	d_geo_high	d_geo_med	d_child	age	d_cple	d_male
Austria	Coef.	−0.316	0.015	0.047	0.144	0.153	0.042	0.113	−0.004	0.057	
	SE	0.0040	0.0002	0.0026	0.0018	0.0017	0.0019	0.0018	0.0000	0.0015	
Belgium	Coef.	−1.124	0.153	0.375	0.278	0.215	0.114	0.137			
	SE	0.0169	0.0016	0.0028	0.0029	0.0053	0.0054	0.0028			
Canada[2]	Coef.	−2.356	0.374	0.540	0.265	−0.032		0.793	0.008	0.138	−0.238
	SE	0.0136	0.0013	0.0032	0.0026	0.0032		0.0053	0.0001	0.0026	0.0022
Czech Rep.	Coef.	−2.8671†	0.3896*	0.1494†	0.646	0.1138†	0.0585†	0.2436†	−0.025	0.818	
	SE	1.9711†	0.1343*	0.2150†	0.1584	0.1380†	0.1494†	0.1712†	0.0048	0.1371	
Denmark	Coef.	1.731	0.010	0.348	0.182	−0.048	0.293	−0.032			
	SE	0.0245	0.0024	0.0059	0.0042	0.0043	0.0052	0.0047			
Finland	Coef.	−2.012	0.396	0.183	0.312	0.361	0.120	0.319	0.003	−0.029	−0.398
	SE	0.0409	0.0043	0.0066	0.0057	0.0062	0.0064	0.0087	0.0001	0.0061	0.0051
France	Coef.	−4.161	0.565	0.177	0.0024*	0.173	0.043				
	SE	0.0068	0.0007	0.0015	0.0010*	0.0010	0.0014				
Greece	Coef.	−8.141	1.054	5.3070†	0.772	0.390	−0.195	0.677			
	SE	0.0484	0.0052	0.0000†	0.0082	0.0065	0.0090	0.0086			
Hungary	Coef.	−5.537	0.790	0.555	0.650	0.158	0.031	0.055			
	SE	0.0157	0.0019	0.0052	0.0039	0.0031	0.0028	0.0034			
Ireland	Coef.	−0.117	0.252	0.292	0.207	n.a.	n.a.	0.285	−0.008	0.856	−0.356
	SE	0.0332	0.0030	0.0103	0.0086	n.a.	n.a.	0.0114	0.0002	0.0109	0.0063
Slovak Rep.	Coef.	−3.554	0.470	0.299	0.379	0.029	0.157	−0.187			
	SE	0.0181	0.0021	0.0045	0.0032	0.0030	0.0026	0.0027			
Spain	Coef.	−6.956	0.901	0.072	−0.082	0.321	0.043	0.132			
	SE	0.0152	0.0016	0.0038	0.0030	0.0020	0.0023	0.0025			
Sweden	Coef.	0.106	0.123	0.102	0.075	0.331	0.0061†	n.a.	0.011	0.521	−0.292
	SE	0.0148	0.0014	0.0046	0.0042	0.0040	0.0041†	n.a.	0.0001	0.0037	0.0032
Switzerland	Coef.	1.651	0.033	0.203	0.309	n.a.	n.a.	−0.120	0.012	0.660	−0.417
	SE	0.0557	0.0066	0.0112	0.0102	n.a.	n.a.	0.0088	0.0002	0.0091	0.0084
United Kingdom	Coef.	−2.7957	0.4065	n.a.	n.a.	−0.0543†	−0.0479†	n.a.	0.0110	0.2949	−0.1993
	SE	0.1674	0.0149	n.a.	n.a.	0.0631†	0.0735†	n.a.	0.0009	0.0399	0.0327

[1] The coefficients have all a p value < .001, except: * p < .05; ** p < 0.1; † p > =0.1.

[2] Income instead of equivalent income. Urban instead of d_geo_high and d_geo_med. See methodology.

Source: see Table 1

Table 8 Communication services — regression[1]

		intercept	ln income	d_edu_high	d_edu_med	d_geo_high	d_geo_med	d_child	age	d_cple	d_male
Austria	Coef.	6.837	0.0114*	0.223	0.178	0.099					
(ols)	SE	0.0550	0.0047*	0.0502	0.0357	0.0285					
Belgium	Coef.	3.558	0.265	0.237	0.142	0.107	0.046	0.061			
	SE	0.0127	0.0011	0.0020	0.0020	0.0035	0.0036	0.0016			
Canada[2]	Coef.	3.469	0.317	0.040	0.069	–0.034		0.087	–0.004	0.055	–0.092
	SE	0.0030	0.0003	0.0006	0.0006	0.0006		0.0005	0.00001	0.0004	0.0004
Czech Rep.	Coef.	8.894	0.222	0.1318*	0.110	0.122	0.0086†	0.183	–0.008	0.598	
	SE	0.4593	0.0304	0.0465*	0.0302	0.0325	0.0358†	0.0261	0.0011	0.0325	
Denmark	Coef.	5.199	0.089	0.033	0.134	0.135	0.071	0.313			
	SE	0.0076	0.0007	0.0017	0.0013	0.0014	0.0015	0.0014			
Finland	Coef.	1.542	0.504	–0.119	–0.0363†	0.0140†	0.0374†	0.268	–0.009	–0.0728*	
(ols)	SE	0.1922	0.0195	0.0299	0.0275†	0.0238†	0.0289†	0.0282	0.0007	0.0252*	
France	Coef.	3.163	0.397	0.096	0.058	–0.117	–0.0485*	0.071	–0.014	–0.169	
(ols)	SE	0.1240	0.0125	0.0221	0.0174	0.0157	0.0223*	0.0190	0.0005	0.0158	
Greece	Coef.	0.217	0.600	0.340	0.300	0.069	–0.030	0.176			
	SE	0.0050	0.0005	0.0009	0.0007	0.0006	0.0017	0.0007			
Hungary	Coef.	0.490	0.560	0.393	0.264	0.155	0.029	0.020			
	SE	0.0054	0.0006	0.0010	0.0009	0.0009	0.0009	0.0008			
Ireland	Coef.	2.156	0.541	–0.143	–0.0010†	n.a.	n.a.	0.152	–0.022	–0.178	–0.008
	SE	0.0106	0.0009	0.0019	0.0018†	n.a.	n.a.	0.0017	0.0001	0.0016	0.0015
Netherlands	Coef.	5.221	0.178	0.1197*	0.1040*	n.a.	n.a.	n.a.	–0.012		
(ols)	SE	0.1884	0.0175	0.0546*	0.0502*	n.a.	n.a.	n.a.	0.0010		
Norway	Coef.	6.239	0.095								
	SE	0.2092	0.0164								
Slovak Rep.	Coef.	–1.181	0.693	0.305	0.251	0.184	0.145	0.020			
	SE	0.0144	0.0014	0.0024	0.0021	0.0014	0.0013	0.0012			
Spain	Coef.	–0.729	0.665	0.148	0.176	0.102	0.082	0.019			
	SE	0.0036	0.0004	0.0005	0.0005	0.0004	0.0005	0.0004			
Sweden	Coef.	5.313	0.165	0.106	0.119	0.109	0.061	n.a.	–0.014	0.059	0.040
	SE	0.0043	0.0004	0.0010	0.0009	0.0008	0.0009	n.a.	0.0000	0.0008	0.0007
Switzerland	Coef.	2.577	0.312	0.026	0.037	n.a.	n.a.	0.226	–0.015	0.068	0.012
	SE	0.0064	0.0007	0.0013	0.0011	n.a.	n.a.	0.0008	0.0000	0.0009	0.0009
United Kingdom	Coef.	4.127	0.308	n.a.	n.a.	0.0085†	–0.0165†	n.a.	–0.016	–0.068	–0.0184**
	SE	0.0738	0.0064	n.a.	n.a.	0.0179†	0.0209†	n.a.	0.0003	0.0112	0.0109**

[1] The coefficients have all a p value < .001, except: * p < .05; ** p < 0.1; † p > =0.1.
[2] Income instead of equivalent income. Urban instead of d_geo_high and d_geo_med. See methodology.
Source: see Table 1

Households' expenditure on communication services decrease with age.

The effect of living in a couple differs among countries: a household living in a couple have a higher probability of spending on IT services in four countries, but a lower one in four others.

Similarly, if the household's reference person is a man, the effect varies according to the country.

Communication goods and services

Overall, the various explanatory variables have diverging effects according to whether we are referring to communication goods or communication services:

- Concerning the probability of spending, the population density in the area where the household is living generally has a positive effect where services are concerned, whereas the effect is much more variable across countries where goods are concerned. Similarly, if the household's reference

person is a man, this has a negative effect on services, whereas the effect is much more variable across countries where goods are concerned.
- Concerning the level of expenditure, it is generally higher for households whose reference person has a high level of education, where services are concerned, whereas this is not so frequently the case where goods are concerned.
- Age has opposite effects on goods and services in the selection (negative for goods, more positive for services). Geographical densely populated area also has opposite effects in the regression (more positive for services, more negative for goods).

In addition, where both communication goods and communication services are concerned, the effects of variables are not always the same on the probability of spending and the amount spent.

- *For communication goods,* the education level of the household's reference person generally has a positive effect on the probability of spending. But this effect on the amount spent is much more heterogeneous, depending on the country.
- *For communication services,* the probability of spending increases with age, but the amount spent decreases with age.

2.3 Main pattern of determinants

Determinants could be classified according to the homogeneity of their effect on ICT expenditures.

- First, income and child have positive effect on the probability to spend and on the level of expenditures in all types of ICT expenditures.
- Second, education level has generally positive and monotonic effect for IT goods and services and communication services. But the effect is more variable across countries for communication goods.
- Third, geographical area has positive effect — monotonic for the probability, but not for the amount spent — on IT goods and communication services. But the effect is more variable across countries for communication goods and IT services.
- Fourth, age decreases both the probability to spend and the amount spent on goods. The effect is more variable for services.
- Fifth, if the household's reference person is a man, it has an increasing effect for both the probability to spend and on the level of expenditures on IT goods. For IT services and communication goods and services, the effect is more variable.
- Sixth, to live in a couple has an increasing effect for the probability to spend and a decreasing effect on the amount spend for communication goods and services. The effect is positive for IT services but more variable for IT goods.

Two main patterns of determinants of Household's expenditures on ICT seem to emerge, both relatively similar across countries (Figure 1): one for IT goods, one for communication services. They distinguish from each other for age and gender effects. Age decreases the probability to spend on IT goods but increases the probability to spend on Communication services. And if the household's reference person is a man, it has an increasing effect for both the probability to spend and on the level of expenditures on IT goods, but a more variable effect for communication services.

Otherwise, both for IT goods and communication services, income, education level, geographical area and child have positive effects on the probability to spend and on the level of expenditures.

Communication goods could fall under the "communication" umbrella for couple effect and under the "goods" umbrella for age and gender effect. IT services is compatible both with IT goods and communication services features.

Figure 1 The effects[1] of selected determinants on households ICT expenditure

n = number of countries where results are available

	Information Technology					Communication				
Goods		probability	*n*	level	*n*		probability	*n*	level	*n*
	income	+	17	+	17	income	+	17	+	12
	education level	+	16	+	14	education level	+	14	+ / –	6
	geographical area	+	14	+	12	geographical area	+ / –	12	–	5
	child	+	14	+	11	child	+	12	+	5
	age	–	13	–	10	age	–	10	–	3
	couple	+ / –	11	+ / –	8	couple	+	8	–	1
	gender (male)	+	10	+	7	gender (male)	+ / –	7	+	1
Services		probability	n	level	n		probability	n	level	n
	income	+	17	+	16	income	+	17	+	17
	education level	+	12	+	9	education level	+	14	+	15
	geographical area	+ / –	10	+ / –	8	geographical area	+	13	+	13
	child	+	9	+	7	child	+	12	+	12
	age	+ / –	9	+ / –	6	age	+	8	–	8
	couple	+	8	+	5	couple	+	8	+ / –	8
	gender (male)	+ / –	6	+ / –	4	gender (male)	–	6	+ / –	5

[1] Main effect observed across countries. The sign "+ / –" mirrors no dominant trend across countries.
Source: see Table 1

CONCLUSION AND NEXT STEPS

A first step into the analysis of determinants of household's ICT expenditures has shown that some determinants, at the level of more elementary expenditures components (IT goods, IT services, Communication goods and communication services) have significant and relatively similar effects, as expected in the research questions. This is the case of income and presence of children, which have generally a positive impact on both the probability to spend and the level of expenditures.

The effects of other determinants are less similar across countries or between goods and services, but do not invalidate the assumptions made in the initial questions.

The marked different effects between communication services and information technology goods also mirror the fact that ICT goods and services do not necessarily follow a uniform pattern of consumption. Looking at elementary components is certainly useful for a better understanding of the mechanisms at stake.

Those findings may also call for revisiting in detail the existing ICT expenditure categories in the consumption surveys, especially for cases where the frontier between goods and services is blurring.

This version will be completed by *i)* including the United States in the analysis, *ii)* looking at the effects of the determinants on the share of ICT expenditure devoted to ICT goods, *iii)* looking at the effects of the determinants on hardware and software that can be isolated in the case of United States and Canada, and *iv)* looking at the effects of the existing computer and mobile phone equipment of the households on their ICT goods and services expenditures.

ANNEX | ICT Expenditures Definition and Country Data

1 ICT EXPENDITURES: DETAILS AND COMPONENTS

1.1 ICT expenditures: COICOP items

COICOP Categories for detailed ICT goods and services:

Communication expenditures

08.2.0 Telephone and fax equipment:

- Purchases of telephones, radio-telephones, telefax machines, telephone-answering machines and telephone loudspeakers.
- Repair of such equipment.

Excludes: telefax and telephone-answering facilities provided by personal computers (09.1.3).

08.3.0 Telephone and telefax services:

- Installation and subscription costs of personal telephone equipment.
- Telephone calls from a private line or from a public line (public telephone box, post office cabin, etc.); telephone calls from hotels, cafés, restaurants and the like.
- Telegraphy, telex and telefax services.
- Information transmission services; Internet connection services.
- Hire of telephones, telefax machines, telephone-answering machines and telephone loudspeakers.

Includes: radio-telephony, radio-telegraphy and radiotelex services. Excludes: telefax and telephone answering facilities provided by personal computers (09.1.3).

Audio-visual, photographic and information processing equipment

Definitions COICOP:

- 09.1 Audio-visual, photographic and information processing equipment,
- 09.1.1 Equipment for the reception, recording and reproduction of sound and pictures,
- 09.1.2 Photographic and cinematographic equipment and optical instruments,
- 09.1.3 Information processing equipment,
- 09.1.4 Recording media,
- 09.1.5 Repair of audio-visual, photographic and information processing equipment.

Information and Communication technologies expenditures are made from the sum of communication expenditures and audio-visual, photographic and information processing equipment expenditures.

- 09.4.2 — Cultural services (*).

Services provided by:

- cinemas, theatres, opera houses, concert halls, music halls, circuses, sound and light shows,
- museums, libraries, art galleries, exhibitions,
- historic monuments, national parks, zoological and botanical gardens, aquaria,
- hire of equipment and accessories for culture, such as television sets, video cassettes, etc.,
- television and radio broadcasting, in particular licence fees for television equipment and subscriptions to television networks,
- services of photographers such as film developing, print processing, enlarging, portrait photography, wedding photography, etc.

Includes: services of musicians, clowns, performers for private entertainments.

(*) When available, the only part which will be included in the ICT expenditures will be "hire of equipment and accessories for culture, such as television sets, video cassettes, etc." and "television and radio broadcasting, in particular licence fees for television equipment and subscriptions to television networks".

1.2 ICT expenditures items selected by countries

European countries (including Czech Republic)

For the European countries, according to the available detail level (Eurostat, 2006), the following items have been selected:

- Telephone and telefax equipment (HE082),
- Telephone and telefax services (HE083),
- Audio-visual, photographic and information processing equipment (HE091),
- Television and radio taxes and hire of equipment (HE09423).

The various ICT components are defined as follow:

- *IT goods* include Audio-visual, photographic and information processing equipment (HE091) less Repair of audio-visual, photographic and information processing equipment (HE0915),
- *IT services* include Television and radio taxes and hire of equipment (HE09423) and Repair of audio-visual, photographic and information processing equipment (HE0915),
- *Communication goods* include Telephone and telefax equipment (HE082),
- *Communication services* include Telephone and telefax services (HE083).

Canada

For Canada, according to the available detail level (Statistics Canada, 2008a), the ICT expenditures and their components are defined in the Table A1.

Table A1 ICT expenditures and their components in Canada

Communication		IT	
Goods	**Services**	**Goods**	**Services**
Purchases of communications equipment H004	Telephone services H005	Computer equipment and supplies M110	Internet sevices H009
	Cellular services H008	Photographic goods (M116-M199)	On-line services H070
		Audio equipment M150	Photographic services M199
		Pre-recorded media M151	Cablevision and satellite services M165
		Blanc media M187	
		Televisions and other video equipment M186	
		Rental of DVDs, video tapes and video games M156	
		Rental of home entertainment equipment M157	

Source: OECD, based on Statistics Canada SHS Data Dictionary 2006 Data Model Entity (PUMF)

Switzerland

For Switerland, the most detailed level (level 5) of the classification used by the Household Budget Survey has been used. Consumptions expenditures are classified according to COICOP, as established by EUROSTAT. ICT expenditures items are in line with the other European countries.

2 COUNTRY DATA SPECIFICITIES

2.1 European countries

Data source

The source of the data is the Eurostat database on Household Budget Survey. The reference year is 2005 (Eurostat, 2006).

Income and equivalent income

Income refers to the net income (total income from all sources including non-monetary components minus income taxes).

Equivalent income of the household is used instead of income of the household.

Equivalent income of the household has been calculated by dividing the income of the household by the equivalent size of the household and multiplying the result by the household size:

Equivalent income = income / household equivalent size × household size. (1)

Household size refers to the sum of household members.

Household equivalent size is established by allocating weighting coefficients to the household's members according to their demographic characteristics. Given the existence of big differences in the sizes and structures of households, comparability can be improved by using expenditure or income by adult equivalent. The OECD scale is used, which consists in allocating the following weightings to persons in the calculation of the "equivalent household's size":

- first adult in the household[1] = 1.0,
- each adult thereafter (aged over 13) = 0.7,
- each child (13 or under) = 0.5.

Calculation rule:

Household equivalent size = 0.3 + (0.7 × *A*) + (0.5 × *B*), (2)

A = Sum of household members where MB03 > 13,

B = Sum of household members where MB03 < 14.

[1] The first adult of the household counts by 1 because of the addition of the constant term 0.3, assuming that each household must have at least one adult.

Education level of the reference person

The education level of the reference person corresponds to the level of studies completed by the reference person, using the ISCED (International Standard Classification of Education) nomenclature:

d_edu_high: higher education (ISCED = 5, 6),

d_edu_med: upper secondary education and post-secondary non-tertiary education (ISCED = 3, 4),

d_edu_low: none or primary education and lower secondary education (ISCED = 1, 2).

In selection and regression, d_edu_low is used as reference.

Geographical area

Geographical area refers to the population density domain. It has been divided in 3 categories:

d_geo_high: densely populated (at least 500 inhabitants / km^2),

d_geo_med: intermediate (between 100 and 499 inhabitants / km^2),

d_geo_low: sparsely populated (less than 100 inhabitants / km^2).

In selection and regression, d_geo_low is used as reference.

Reference person

As stated in Eurostat (2006), "reference person" is a European concept, which usually differs from the national concept of "head of household". The reference person is the adult (16+) contributing most to the total income of the household. In that sense, the reference person can also be designated as "main income earner".

2.2 Czech Republic

The source of the data is the *Czech Household Budget Survey (HBS),* with 2006 as reference year. The variables are in line with the Eurostat definitions. The methodology of the Czech Household Budget Survey (HBS) is provided on the Czech Statistical Office website.[5]

2.3 Switzerland

The source of the data is the *Household Budget Survey (HBS).* The reference year is 2005. The geographical area (population density domain) is not available. Education level of the reference person has been aligned with the European countries, based on the education short classification (EWL) and using a mapping of national educational educational programmes with ISCED. Income refers to household income before taxes. It includes income for household from earnings, investment, and from transfer payments. The methodology of the Household Budget Survey (HBS) is provided on the Swiss Federal Statistical Office web site.[6]

2.4 Canada

Data source

The source of the data is the *Survey of Household Spending PUMF 2006,* from Statistics Canada (Statistics Canada, 2008b and 2008c).

Age of the reference person

Age of the reference person is provided by groups. The variable has been modified as follows:
For age group of less than 25 year, age value has been put to 22. For age group 25–29, age value has been put to 27. For age group 30–34, age value has been put to 32 (and similarly for other age groups). For age group of 85 and over, age value has been put to 92.

Income

Income refers to household income before taxes. It includes income for household from earnings, investment, government transfer payments and other sources. It excludes personal income tax refunds (Statistics Canada, 2008a).

Education level of the reference person

Statistics Canada provided a concordance table between original codes used in the Canadian micro-data file and the 3 levels of ISCED used by Eurostat for the European countries, as follow:

Figure A1

Initial code	Description	ISCED	Education level
1	No degrees, certificates or diplomas	0,1,2	Low
2	Secondary (high) school diploma or equivalent	3	Medium
3	Trade / vocational certificate	4	Medium
4	Apprenticeship certificate	4	Medium
5	Community college, CEGEP or nursing school diploma	5B	High
6	University certificate or diploma below Bachelor's	5B	High
7	Bachelor's degree (B.A., B.Sc., B.Ed.)	5A	High
8	University degree, certificate or diploma above a Bachelor's	5A/6	High

Source: Statistics Canada, Culture, Tourism and the Centre for Education Statistics Division

[5] <http://www.czso.cz/eng/redakce.nsf/i/home>.

[6] <http://www.bfs.admin.ch/bfs/portal/en/index/infothek/erhebungen__quellen/blank/blank/habe/02.html>.

In selection and regression, d_edu_low is used as reference.

Urban-rural

Instead of the 3 different levels of population density provided for the European countries, the urban area indicator is used. Urban area is defined as follow (Statistics Canada, 2008c):

For the Survey of Household Spending (based on the LFS sampling frame), urban areas include:

- All large metropolitan areas (even though they do contain some rural areas);
- Most small metropolitan areas (also called census agglomerations). In some cases, where a census agglomeration contains a large rural population, only the urban portion is considered urban;
- Urban areas based on the census definition: "Urban areas have minimum population concentrations of 1 000 and a population density of at least 400 per square kilometer based on the previous census population counts".

Rural area

All territory outside urban areas is considered rural. Taken together, urban and rural areas cover all of Canada.

In selection and regression, d_rural is used as reference.

Reference person

The household member being interviewed chooses which household member should be listed as the reference person after hearing the following definition. "The household reference person is the member of the household mainly responsible for its financial maintenance (e.g. pays the rent, mortgage, property taxes, and electricity). This person can be either male or female. When all members of the household share equally, any member may be shown as the reference person." This person must be a member of the household at the time of the interview (Statistics Canada, 2008c).

Table A2 Variables by country

AUSTRIA

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	3 830	6.584796	21.7163567	10 105 606.3	–0.787144	10.59144
ln_IT_serv	1 698	6.93203325	11.0802667	5 036 033.22	3.94019	9.47311
ln_com_goods	215	6.6212685	22.9137602	539 090.786	3.07269	8.83054
ln_com_serv	3 835	7.12990086	16.7945892	1 1075 188.3	1.20683	9.64398
ln_eq_inc	8 400	9.37619	63.47003	32 722 895	0	12.20797
d_edu_high	7 719	0.13003	6.87484	419 300	0	1
d_edu_med	7 719	0.6377	9.825	2 056 398	0	1
d_geo_high	8 400	0.39996	9.98614	1 395 870	0	1
d_geo_med	8 400	0.2361	8.65699	824 000	0	1
d_child	8 400	0.28799	9.23066	1 005 102	0	1
age	8 400	50.51617	346.34064	176 301 428	18	99
d_cple	8 400	0.41706	10.05103	1 455 554	0	1
d_male	8 400	0.64429	9.7586	2 248 572	0	1

	N Non spending (=0)	N Spending (=1)
d_IT_goods	4 570	3 830
d_IT_serv	6 702	1 698
d_com_goods	8 185	215
d_com_serv	4 565	3 835

BELGIUM

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	1 600	5.99260592	30.6841199	4 702 133.29	1.79176	10.03764
ln_IT_serv	3 093	5.18243459	13.8418797	8 058 731.89	1.94018	11.90543
ln_com_goods	406	5.73104568	15.1197954	1 159 543.97	1.09861	7.83634
ln_com_serv	2 926	6.616369	18.91293	9 492 543.43	1.79176	9.38568
ln_eq_inc	3 550	10.46827	17.26379	18 358 154	0	12.91155
d_edu_high	3 496	0.43405	11.02494	750 622	0	1
d_edu_med	3 496	0.28224	10.01189	488 089	0	1
d_geo_high	3 550	0.59753	10.90109	1 047 894	0	1
d_geo_med	3 550	0.35734	10.6526	626 667	0	1
d_child	3 550	0.29713	10.15861	521 074	0	1
age	3 550	50.0539	346.87104	87 779 294	19	86
d_cple	3 550	0.45079	11.06065	790 556	0	1
d_male	3 550	0.65296	10.58171	1 145 101	0	1

	N Non spending (=0)	N Spending (=1)
d_IT_goods	1 950	1 600
d_IT_serv	457	3 093
d_com_goods	3 144	406
d_com_serv	624	2 926

CANADA

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	12 359	6.4182683	41.7447695	69 953 797.3	0.693147	10.1973
ln_IT_serv	13 647	6.56624926	23.9093233	77 683 206.9	0.693147	9.15736
ln_com_goods	4 369	4.71529179	26.9094043	18 014 131	0.693147	8.16052
ln_com_serv						
ln_inc	14 618	10.82036	24.92811	137 807 166	5.29832	15.03929
d_edu_high	14 635	0.44691	14.67833	5 700 643	0	1
d_edu_med	14 635	0.37537	14.2958	4 787 998	0	1
d_urban	14 018	0.87686	9.88864	11 131 235	0	1
d_child	14 635	0.2901	13.39799	3 700 353	0	1
agebis	14 635	49.68506	487.5916	633 761 084	22	92
d_male	14 635	0.50217	14.76163	6 405 452	0	1
d_comp	14 635	0.75387	12.71747	9 616 002	0	1
d_mob	14 635	0.67709	13.80488	8 636 660	0	1

	N Non spending (=0)	N Spending (=1)
d_IT_goods	2 272	12 346
d_IT_serv	983	13 635
d_com_goods	10 254	4 364
d_com_serv	208	14 410

CZECH REPUBLIC

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	2 222	10.642461	1.70823586	22 470.4018	5.64706	15.00042
ln_IT_serv	2 793	10.9536781	0.59414348	30 670.1048	3.34448	13.24215
ln_com_goods	899	10.1979024	1.42524059	8 593.39484	3.34448	13.47027
ln_com_serv	2 919	12.2471653	0.81714881	35 493.5891	8.21201	14.70736
ln_eq_inc	2 967	14.4183297	0.65712332	42 753.8668	11.5511	17.6938
d_edu_high	2 967	0.11234	0.31574	333.1058	0	1
d_edu_med	2 967	0.37329	0.48362	1.11E+03	0	1
d_geo_high	2 967	0.37649	0.48444	1 116	0	1
d_geo_med	2 967	0.24209	0.42829	717.8559	0	1
d_child	2 967	0.46749	0.79371	1 386	0	4
age	2 967	52.09347	15.72286	154 470	20	90
d_cple	2 967	0.63774	0.48059	1 891	0	1
d_male	2 967	0.71147	0.45302	2 110	0	1
d_comp	2 967	0.43338	0.49548	1 285	0	1
d_mob	2 967	0.88273	0.3217	2 618	0	1

	N Non spending (=0)	N Spending (=1)
d_IT_goods	745	2 222
d_IT_serv	174	2 793
d_com_goods	2 068	899
d_com_serv	48	2 919

DENMARK

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	1 785	5.75374128	71.0789005	10 084 188.4	–1.95099	10.10977
ln_IT_serv	2 351	5.99492692	19.148077	14 681 279.3	2.47211	9.1984
ln_com_goods	371	4.00167026	61.2088093	1 495 149.38	–2.0755	8.55514
ln_com_serv	2 384	6.35090735	28.7544919	15 805 295.8	–1.5901	9.42035
ln_eq_inc	2 449	10.59586	28.30042	27 054 646	0	12.84208
d_edu_high	2 378	0.2077	12.88444	498 046	0	1
d_edu_med	2 378	0.47835	15.86597	1 147 051	0	1
d_geo_high	2 449	0.41163	15.89372	1 051 019	0	1
d_geo_med	2 449	0.31017	14.93886	791 961	0	1
d_child	2 449	0.26052	14.17517	665 180	0	1
age	2 449	49.80541	579.37374	127 169 329	17	92
d_cple	2 449	0.46891	16.1167	1 197 284	0	1
d_male	2 449	0.5937	15.86187	1 515 904	0	1
d_comp	2 449	0.73027	14.33351	1 864 621	0	1
d_mob	2 448	0.82725	12.21028	2 111 827	0	1

	N Non spending (=0)	N Spending (=1)
d_IT_goods	663	1 785
d_IT_serv	98	2 351
d_com_goods	2 078	371
d_com_serv	55	2 384

FINLAND

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	2 729	6.26765751	27.9748257	9 709 463.44	1.63594	9.29198
ln_IT_serv	3 864	5.40745698	7.05382313	12 722 004.7	3.93852	8.91027
ln_com_goods	1 525	4.94802058	22.7736373	4 284 957.8	1.59672	8.17597
ln_com_serv	3 961	6.34118089	19.5269664	15 344 423	2.28987	9.03745
ln_eq_inc	4 007	10.40418	18.40887	25 542 269	7.68891	13.17957
d_edu_high	4 007	0.29733	11.31529	729 946	0	1
d_edu_med	4 007	0.40441	12.14941	992 829	0	1
d_geo_high	4 007	0.28849	11.21565	708 234	0	1
d_geo_med	4 007	0.16509	9.19077	405 299	0	1
d_child	4 007	0.25497	10.7895	625 949	0	1
age	4 007	50.79765	436.51845	124 708 234	17	96
d_cple	4 007	0.46375	12.34514	1 138 516	0	1
d_male	4 007	0.6025	12.11484	1 479 135	0	1
d_comp	4 007	0.63672	11.90597	1 563 152	0	1
d_mob	4 007	0.91445	6.92388	2 244 986	0	1

	Non spending (=0)	Spending (=1)
d_IT_goods	1 278	2 729
d_IT_serv	143	3 864
d_com_goods	2 482	1 525
d_com_serv	46	3 961

FRANCE

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	5 031	5.8405576	64.287826	67 241 311.7	–0.7120671	9.87482
ln_IT_serv	7 737	5.35398937	39.4208337	98 541 105.9	2.46599	9.21083
ln_com_goods	2 071	4.3938463	70.9240474	20 933 333.9	–0.0189199	9.34174
ln_com_serv	9 816	6.50230041	40.158912	154 621 466	2.28367	9.86203
ln_eq_inc	10 240	10.42588	34.53472	259 796 116	7.33302	13.67631
d_edu_high	10 240	0.19854	19.67877	4 947 349	0	1
d_edu_med	10 240	0.37579	23.89291	9 364 069	0	1
d_geo_high	10 240	0.40949	24.25862	10 203 785	0	1
d_geo_med	10 240	0.13455	16.83441	3 352 842	0	1
d_child	10 240	0.32009	23.01407	7 976 163	0	1
age	10 240	52.13011	858.636	1 298 998 080	16	98
d_cple	10 240	0.4881	24.65915	12 162 626	0	1
d_male	10 240	0.64746	23.56909	16 133 568	0	1
d_comp	10 240	0.48981	24.66102	12 205 302	0	1
d_mob	10 240	0.48553	24.65581	12 098 623	0	1

	Non spending (=0)	Spending (=1)
d_IT_goods	8 169	2 071
d_IT_serv	2 503	7 737
d_com_goods	8 169	2 071
d_com_serv	423	9 816

GREECE

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	2 071	5.67494524	26.4099317	7 370 206.08	2.37993	9.70445
ln_IT_serv	6 536	4.04174978	19.1268141	16 088 213.3	3.26588	7.73548
ln_com_goods	218	4.76443913	21.1225633	632 675.589	2.33286	6.71894
ln_com_serv	6 525	6.62190962	19.6186673	26 333 108.9	4.4651	9.53101
ln_eq_inc	6 555	10.26827	17.03429	41 000 854	7.09008	12.36437
d_edu_high	6 553	0.16618	9.18815	663 371	0	1
d_edu_med	6 553	0.29584	11.26594	1 180 960	0	1
d_geo_high	6 555	0.4319	12.22638	1 724 545	0	1
d_geo_med	6 555	0.03498	4.53481	139 665	0	1
d_child	6 555	0.31926	11.50685	1 274 783	0	1
age	6 555	53.64969	424.48437	21 4221 321	15	98
d_cple	6 555	0.43096	12.2232	1 720 821	0	1
d_male	6 555	0.74779	10.71924	2 985 909	0	1
d_comp	6 555	0.29478	11.25396	1 177 043	0	1
d_mob	6 555	0.7257	11.01244	2 897 714	0	1

	Non spending (=0)	Spending (=1)
d_IT_goods	4 484	2 071
d_IT_serv	19	6 536
d_com_goods	6 337	218
d_com_serv	30	6 525

HUNGARY

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	3 163	5.01536862	20.750045	6 220 812.46	0.96026	8.34934
ln_IT_serv	5 053	5.04669579	10.9309124	10 519 080.4	1.75877	7.91097
ln_com_goods	1 348	4.30001448	15.1804637	2 102 298.58	–1.26514	6.3286
ln_com_serv	8 590	5.81816795	16.4260223	20 983 571.8	1.34598	8.31933
ln_eq_inc	9 058	9.03015	14.53043	34 649 456	0	12.00055
d_edu_high	9 058	0.19443	8.1459	746 033	0	1
d_edu_med	9 058	0.26628	9.09798	1 021 752	0	1
d_geo_high	9 058	0.36575	9.91359	1 403 415	0	1
d_geo_med	9 058	0.3172	9.57907	1 217 136	0	1
d_child	9 058	0.2787	9.22863	1 069 413	0	1
age	9 058	52.42769	343.02272	201 169 615	18	98
d_cple	9 058	0.32282	9.62366	1 238 679	0	1
d_male	9 058	0.56784	10.19634	2 178 836	0	1
d_comp	9 058	0.34606	9.79157	1 327 850	0	1
d_mob	9 058	0.72994	9.13866	2 800 843	0	1

	Non spending (=0)	Spending (=1)
d_IT_goods	5 895	3 163
d_IT_serv	4 005	5 053
d_com_goods	7 710	1 348
d_com_serv	468	8 590

IRELAND

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	3 246	6.45159309	16.5782844	4 208 929.01	1.66988	11.3014
ln_IT_serv	5 920	5.97216334	10.0241993	7 241 307.77	–0.632703	8.7395
ln_com_goods	2 256	5.02013747	10.9001822	2 255 301.78	1.44674	8.59685
ln_com_serv	6 786	6.86000725	15.9468688	9 751 452.29	0.753592	9.29548
ln_eq_inc	6 884	10.90473	13.68633	15 761 852	0	14.13496
d_edu_high	6 884	0.29089	6.58157	420 461	0	1
d_edu_med	6 884	0.24752	6.25404	357 771	0	1
d_geo_high	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
d_geo_med	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
d_child	6 884	0.34843	6.9047	503 622	0	1
age	6 884	50.5849	237.38918	73 116 119	15	93
d_cple	6 884	0.43472	7.18364	628 356	0	1
d_male	6 884	0.61861	7.03883	894 145	0	1
d_comp	6 884	0.54379	7.21781	786 003	0	1
d_mob	6 884	0.84448	5.25169	1 220 618	0	1

	Non spending (=0)	Spending (=1)
d_IT_goods	3 638	3 246
d_IT_serv	964	5 920
d_com_goods	4 628	2 256
d_com_serv	98	6 786

NETHERLANDS

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	1 328	5.92093095	1.22777775	7 694.77466	2.17478	9.26852
ln_IT_serv	318	4.8079025	0.83791815	1 447.04793	2.72296	7.61367
ln_com_goods	513	4.56684547	0.65736261	2 235.16141	3.23379	6.67969
ln_com_serv	1 543	6.55827623	0.65837806	10 115.4494	3.00663	9.072
ln_eq_inc	1 570	10.21465	0.92248	16 037	0	12.25358
d_edu_high	1 561	0.32204	0.46756	503.03033	0	1
d_edu_med	1 561	0.55821	0.49692	871.94349	0	1
d_geo_high	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
d_geo_med	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
d_child	0	.	.	.	.	.
age	1 570	48.51546	15.31503	76 169	21	80
d_cple	1 570	0.51737	0.49986	812.27593	0	1
d_male	1 570	0.69975	0.45851	1 099	0	1
d_comp	1 570	0.75425	0.43067	1 184	0	1
d_mob	0	.	.	.	.	.

	Non spending (=0)	Spending (=1)
d_IT_goods	242	1 328
d_IT_serv	1 251	318
d_com_goods	1 049	513
d_com_serv	27	1 543

NORWAY

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	2 487	6.54907147	1.61168721	28 559.7803	–2.1749	11.0986
ln_IT_serv	633	6.9330206	1.73479662	7 766.23102	3.48009	10.04464
ln_com_goods	1 239	4.54154347	2.82099959	9 212.83885	–2.16397	9.13012
ln_com_serv	1 343	7.43745396	1.28712027	17 640.5995	2.52827	10.1428
ln_eq_inc	3 376	12.68875	2.25626	82 420	0	15.49446
d_edu_high	3 331	0.34631	0.65925	2 214	0	1
d_edu_med	3 331	0.50447	0.69277	3 225	0	1
d_geo_high	3 376	0.51162	0.69346	3 323	0	1
d_geo_med	3 376	0.16995	0.52106	1 104	0	1
d_child	3 376	0.30364	0.63792	1 972	0	1
age	3 376	46.61104	21.54417	302 762	18	87
d_cple	3 376	0.33327	0.65395	2 165	0	1
d_male	3 376	0.6162	0.67466	4 003	0	1
d_comp	3 376	0.70814	0.63069	4 600	0	1
d_mob	2 245	0.93433	0.34404	4 045	0	1

	Non spending (=0)	Spending (=1)
d_IT_goods	889	2 487
d_IT_serv	2 743	633
d_com_goods	2 137	1 239
d_com_serv	2 033	1 343

SLOVAK REPUBLIC

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	622	4.51625853	32.8215105	112 5671.5	1.30822	9.94398
ln_IT_serv	4 111	3.92592474	12.9247025	6 441 508.78	1.82741	7.23481
ln_com_goods	55	5.73003885	20.7735358	122 683.688	3.84232	7.81888
ln_com_serv	3 958	5.48052073	15.5465995	8 638 517.94	1.0289	8.31205
ln_eq_inc	4 710	9.00302	12.66753	17 108 745	6.1449	11.42473
d_edu_high	4 710	0.14072	6.98538	267 407	0	1
d_edu_med	4 710	0.72977	8.92098	1 386 798	0	1
d_geo_high	4 710	0.26469	8.86242	502 991	0	1
d_geo_med	4 710	0.42047	9.91643	799 026	0	1
d_child	4 710	0.38757	9.78709	736 511	0	1
age	4 710	49.5244	312.7075	94 112 899	17	96
d_cple	4 710	0.39418	9.8168	749 075	0	1
d_male	4 710	0.60131	9.83596	1 142 697	0	1
d_comp	4 710	0.35374	9.605	672 231	0	1
d_mob	4 710	0.79958	8.04185	1 519 460	0	1

	N Non spending (=0)	N Spending (=1)
d_IT_goods	4 088	622
d_IT_serv	599	4 111
d_com_goods	4 655	55
d_com_serv	752	3 958

SPAIN

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	5 129	4.71021544	76.0831709	40 263 742.8	0.0332466	8.6287
ln_IT_serv	3 230	3.98295701	83.9193388	21 169 003.1	0.0332466	7.81564
ln_com_goods	1 401	3.97700598	64.4638812	9 653 785.47	0.0332466	7.28884
ln_com_serv	8 709	6.17725972	31.91825	87 337 783.9	0.0332466	8.45701
ln_eq_inc	8 881	10.17436	23.58545	146 736 676	0	11.92429
d_edu_high	8 881	0.18377	15.60828	2 650 402	0	1
d_edu_med	8 881	0.14138	14.04117	2 039 005	0	1
d_geo_high	8 881	0.50607	20.14872	7 298 682	0	1
d_geo_med	8 881	0.20271	16.2015	2 923 529	0	1
d_child	8 881	0.33994	19.08987	4 902 721	0	1
age	8 881	55.21051	621.22543	796 257 577	18	98
d_cple	8 881	0.4247	19.9204	6 125 150	0	1
d_male	8 881	0.78003	16.69344	11 249 775	0	1
d_comp	8 876	0.46253	20.09516	6 668 082	0	1
d_mob	0	.	.	.	.	.

	N Non spending (=0)	N Spending (=1)
d_IT_goods	3 752	5 129
d_IT_serv	5 651	3 230
d_com_goods	7 480	1 401
d_com_serv	172	8 709

SWEDEN

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	1 362	6.38683327	43.9708966	14 349 225.7	2.63944	9.32393
ln_IT_serv	1 968	5.74755892	24.099841	20 367 229.1	4.02573	8.75936
ln_com_goods	657	5.16436234	33.1084675	5 476 075.82	3.98651	9.62397
ln_com_serv	2 049	6.50898562	29.7385101	24 694 219.7	3.62983	8.68225
ln_eq_inc	2 079	10.31928	38.68247	40 079 169	0	12.80687
d_edu_high	2 079	0.3513	20.63819	1 364 400	0	1
d_edu_med	2 079	0.42152	21.34838	1 637 140	0	1
d_geo_high	2 079	0.25444	18.82984	988 228	0	1
d_geo_med	2 079	0.14737	15.32482	572 368	0	1
d_child	2 079	1	0	3 883 911	1	1
age	2 079	48.97739	687.40494	190 223 821	18	89
d_cple	2 079	0.50594	21.6148	1 965 035	0	1
d_male	2 079	0.61207	21.06637	2 377 216	0	1
d_comp	0	.	.	.	.	.
d_mob	2 079	0.92045	11.69853	3 574 947	0	1

	N Non spending (=0)	N Spending (=1)
d_IT_goods	713	1 362
d_IT_serv	111	1 968
d_com_goods	1 422	657
d_com_serv	30	2 049

SWITZERLAND

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	1 350	4.28461563	43.2403598	5 758 717.35	0.405465	8.966726
ln_IT_serv	2 960	3.79941247	13.8624474	11 609 170.1	1.07044	6.21936
ln_com_goods	157	4.20070947	38.5436374	645 343.25	0	6.68324
ln_com_serv	3 075	4.78139688	24.0237847	15 263 488.3	1.22009	7.58943
ln_inc	3 087	8.92661	19.71894	28 597 926	4.83898	11.55437
d_edu_high	3 087	0.32683	15.11291	1 047 044	0	1
d_edu_med	3 087	0.5535	16.01755	1 773 219	0	1
d_geo_high	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
d_geo_med	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
d_child	3 087	0.32632	15.10684	1 045 412	0	1
age	3 087	50.65929	528.63585	162 295 713	18	96
d_cple	3 087	0.6065	15.74035	1 943 016	0	1
d_male	3 087	0.69342	14.85586	2 221 476	0	1
d_comp	3 087	0.07717	8.59802	247 211	0	1
d_mob	3 087	0.18237	12.44185	584 267	0	1

	N Non spending (=0)	N Spending (=1)
d_IT_goods	1 737	1 350
d_IT_serv	127	2 960
d_com_goods	2 930	157
d_com_serv	12	3 075

UNITED KINGDOM

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ln_IT_goods	2 083	6.50565865	2.4473229	50 597.9517	2.94504	11.1015
ln_IT_serv	5 932	5.83512091	1.51207411	126 124.133	2.02875	8.608
ln_com_goods	207	6.79625834	2.67279074	5 147.12145	3.46383	9.73851
ln_com_serv	6 480	6.43731438	1.5876684	152 612.375	–0.273837	9.48103
ln_eq_inc	6 785	10.22842	1.66001	253 652	0	14.14603
d_edu_high	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
d_edu_med	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
d_geo_high	6 245	0.77083	0.82617	18 597	0	1
d_geo_med	6 245	0.15109	0.70399	3 645	0	1
d_child	6 785	1	0	24 799	1	1
age	6 785	51.89904	32.99095	1 287 034	16	98
d_cple	6 785	0.3261	0.89628	8 087	0	1
d_male	6 785	0.61472	0.93046	15 244	0	1
d_comp	6 785	0.64503	0.91486	15 996	0	1
d_mob	0	.	.	.	.	.

	N Non spending (=0)	N Spending (=1)
d_IT_goods	4 702	2 083
d_IT_serv	853	5 932
d_com_goods	6 578	207
d_com_serv	305	6 480

Source: See Table 1

References

ARCEP. *Autorité de Régulation des Communications électroniques et des Postes – La lettre de l'Autorité* [online]. No. 59, 2008. <http://www.arcep.fr/uploads/tx_gspublication/lettre59.pdf>.

BIGOT R. *La diffusion des technologies de l'information dans la société française — 2006* [online]. Paris: ARCEP, 2006. <http://www.arcep.fr/uploads/tx_gspublication/etude-credoc2006.pdf>.

CENTRE FOR EDUCATIONAL RESEARCH AND DEVELOPMENT (CERI). *Looking for the invisible. Educational performance and technology use.* Paris: OECD, 2009.

CRAGG J. Some Statistical Models for Limited Dependant Variables with Application to the Demand for Durable Goods. In *Econometrica,* 39 (5), pp. 829–844.

CREDOC. La diffusion des technologies de l'information dans la société française — 2007 [online]. Paris: ARCEP, 2007. <http://www.arcep.fr/uploads/tx_gspublication/etude-credoc-2007.pdf>.

EUROSTAT. *Data transmission for the HBS round of the reference year 2005.* Doc. ref. HBS/153E/2003/EN, Luxembourg: Eurostat, 2006.

HECKMAN J. Sample selection bias as a specification error. In *Econometrica,* 1979, 47, pp. 153–61.

HORRIGAN J. *A Typology of Information and Communication Technology Users, PEW Internet and American Life Project* [online]. 2007. <http://www.pewinternet.org/~/media/Files/Reports/2007/PIP_ICT_Typology.pdf.pdf>.

McKEOWN L., NOCE A. and CZERNY P. Factors Associated with Internet Use: Does Rurality Matter? [online]. In *Rural and Small Town Canada Analysis Bulletin,* vol. 7, no. 3, September, 2007. <http://www.statcan.ca/english/freepub/21-006-XIE/21-006-XIE2007003.pdf>.

MONTAGNIER P. and WELSUM VAN D. *ICTs and Gender — Evidence from OECD and Non-OECD Countries* [online]. UNCTAD Expert Meeting, 4–5 December 2006. <http://www.unctad.org/sections/wcmu/docs/c3em29p025_en.pdf>.

OECD. *Broadband access and ICT access and use by households and Individuals* [online]. Document ref. DSTI/ICCP/ /IE(2007)4/FINAL, Paris: OECD, 2007. <http://www.oecd.org/dataoecd/44/11/39869349.pdf>.

STATISTICS CANADA. *SHS Data Dictionary, 2006 Data Model Entity (PUMDF).* 7.9.2008, Statistics Canada, 2008a.

STATISTICS CANADA. *User Guide for the Public-use Microdata File, Survey of Household Spending, 2006.* Catalogue no. 62M0004XCB, Statistics Canada, 2008b.

STATISTICS CANADA. *User Guide for the Survey of Household Spending 2006* [online]. Statistics Canada, 2008c. <http://www.statcan.gc.ca/pub/62f0026m/62f0026m2008001-eng.pdf>.

VEENHOF B., CLERMONT Y. and SCIADAS G. Literacy and digital technologies: Linkages and outcomes [online]. In *Connectedness Series,* no. 12, 2005. <http://www.statcan.ca/english/research/56F0004MIE/56F0004MIE2005012.pdf>.

VEENHOF B. et al. How Canadians' Use of the Internet Affects Social Life and Civic Participation [online]. In *Connectedness Series,* no. 16, 2008. <http://www.statcan.gc.ca/pub/56f0004m/56f0004m2008016-eng.pdf>.

YIN W., DEVANEY S. and STAHURA J. Determinants of Household Expenditure on Computer Hardware and Software. In *Journal of Consumer Affairs, September,* 2005, vol. 39, Issue 2, pp. 237–417.

OECD eXplorer: Making Regional Statistics Come Alive through a Geo-Visual Web-Tool

Monica Brezzi[1] | *Organization for the Economic Cooperation and Development — OECD, Paris, France*
Mikael Jern | *National Center for Visual Analytics, ITN, Linkoping University, Linkoping, Sweden*
Lars Thygesen | *Statistics Denmark, Copenhagen, Denmark*

Abstract

Recent advances in web-enabled graphics technologies have the potential to make a dramatic impact on developing highly interactive Geovisual Analytics applications for the Internet. An emerging and challenging application domain is geovisualization of regional (sub-national) statistics. Higher integration driven by institutional processes and economic globalisation is eroding national borders and creating competition along regional lines in the world market. Sound information at sub-national level and benchmark of regions across borders have gained importance in the policy agenda of many countries. In this paper, we introduce "OECD eXplorer" — an interactive tool for analyzing and communicating gained insights and discoveries about spatial-temporal and multivariate OECD regional data. This database is a potential treasure chest for policy-makers, researchers and citizens to gain a better understanding of a region's structure and performance and to carry out analysis of territorial trends and disparities based on sound information comparable across countries. Many approaches and tools have been developed in spatial-related knowledge discovery but generally they do not scale well with dynamic visualization of larger spatial data on the Internet. In this context, we introduce a web-compliant Geovisual Analytics toolkit that supports a broad collection of functional components for analysis, hypothesis generation and validation. The same tool enables the communication of results on the basis of a snapshot mechanism that captures, re-uses and shares task-related explorative findings. Further developments underway are in the creation of a generic highly interactive web "eXplorer" platform that can be the foundation for easy customization of similar web applications using different geographical boundaries and indicators. Given this global dimension, a "generic eXplorer" will be a powerful tool to explore different territorial dimensions (countries, regions, cities, rural areas) according to the different needs of users.

Keywords

Data visualization, geovisual analytics, regional statistics, measures of quality of life in regions, knowledge base management

JEL code

R and Z

[1] Corresponding author: e-mail: monica.brezzi@oecd.org. This paper was prepared in 2009.

INTRODUCTION

We live in a data-rich world where people have become familiar with notions like GDP and sustainable development; statistics that compare countries' economic performance often hit the news headlines. At the same time, people want to see statistics that describe the places where they live and capture the quality of their own lives, taking into account a broader perspective beyond the economic one. The geographical level to which statistics are referred is, therefore, increasingly important. Even though international comparisons of economies and societies tend to be undertaken at the country level, economic performance and social indicators can vary within countries every bit as much as they do between countries (OECD, 2009a). Comparing a region's competitiveness in the global arena requires sound statistics and data, but such information is often limited and complex to be visualized. How can the existing information be made available in a way to enable decision-makers and citizens to fully explore and make use of it? How can different sources of information be pooled together and transformed into knowledge for decision-making?

This paper reflects a challenging applied research task to stimulate, at global level, an exchange of best practices through collaborative Geovisual Analytics reasoning (Thomas and Cook, 2005). Tools are introduced to help gathering and sharing national and local initiatives aimed at measuring economic, social and environmental developments and to engage policy makers, statisticians and the public in collaborative activities. The global dimension of such a task responds to build a repository of *progress indicators*, where experts and public users can use Geovisual Analytics tools to compare situations for countries, regions or local communities.

While the benefits of Geovisual Analytics tools are many, it remains a challenge to adapt these tools to the Internet and reach a broader user community. Important features include dynamic web-enabled visualization and animation that enables statisticians and regional experts to explore geospatial demographics data from multiple perspectives (Brodbeck and Girardin, 2003, Roberts, 2004, Jern et al., 2005), discover interesting relationships, share their discoveries with colleagues (Wohlfart and Hauser, 2007, Jern et al., 2008) and finally communicate selected relevant knowledge to the public. These discoveries often emerge through the diverse backgrounds and experiences of expert domains and are precious in a creative analytics reasoning process.

In this context, we introduce a dynamic web-enabled demonstrator "OECD eXplorer"[2], a customized tool for interactively analyzing, communicating and collaborating gained insights. It includes a snapshot mechanism that captures, re-uses and shares findings. We customize tailor-made and task-oriented applications based on layered components thinking. The key feature of OECD eXplorer is its deployment to Internet users for dynamic web. A first version of OECD eXplorer was released in November 2008 on the OECD web-site. An extended version, released in March 2009, enhances the possibility to explore trends over time, provides novel functions for presenting stories about the statistics and combines metadata and maps status.

This paper is organized as follows: Section 1 describes the needs for such a tool for the OECD work on regional statistics, followed by related work in other areas (Section 2). The architecture and main components of the GAV Flash framework are described in Section 3, the integrated snapshot mechanism in Section 4 and Adobe Flash-based time animation in Section 5. Section 6 discusses how regional information is presented in OECD eXplorer, drawing examples from the OECD Regional database and recalling how the main issues that emerged during the evaluation phase have been addressed. The last section concludes and sets out future global project developments.

[2] <http://stats.oecd.org/OECDregionalstatistics>.

1 THE OECD USER PERSPECTIVE

OECD countries have experienced a growing interest in regional development in recent years (OECD, 2009a). The performance of regional economies and the effectiveness of regional policy help determine a nation's growth and shape the measure of well-being across countries. For the past years the OECD has been studying regional disparities and regional economic growth in its member countries in order to evaluate innovative strategies for development and spread successful policies.

This interest has generated new demand for sound statistical information at the sub-national level on factors that are important to regional competitiveness. The *OECD Regional database* (OECD, 2009c) is a unique source of statistical information at sub-national level for all the OECD Countries. It contains yearly time-series for around 40 indicators on demography, economic and labour market performance, education, healthcare, environmental outputs and knowledge-based economy. Within each country regions are classified on the basis of two territorial levels: the higher level consists of 335 large regions in the OECD countries, while the lower level comprises 1,681 small regions. OECD has since long felt the need to make regional data much more easily available on the web in an interactive and user-participative way. In particular, to make a more extensive use of dynamic web-enabled maps which can, more effectively than a graph, convey the four dimensions included in the regional database: statistical indicator, time, regional and country value (Figure 1).

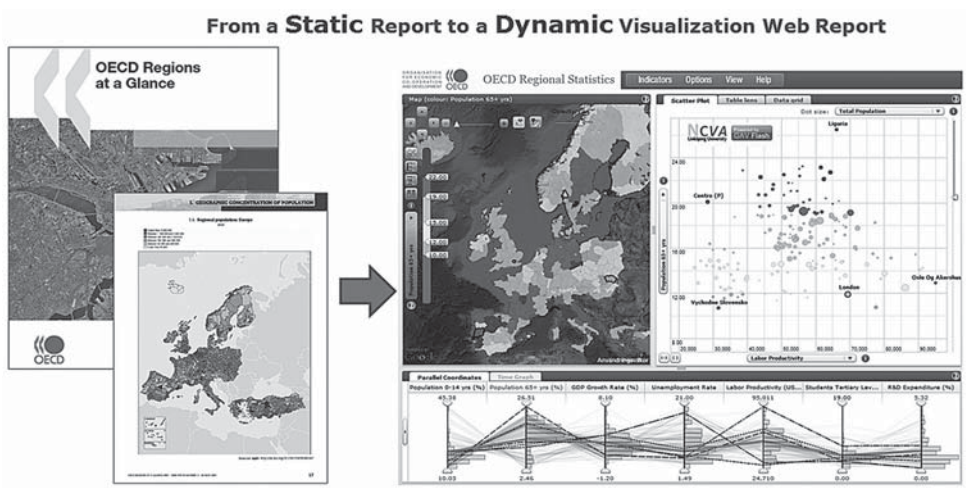
In addition, timely information on the progress of a local community requires crossing different sources of information and new ways to generate and share information for decision-making.

Finally, many analysts and citizens want to create content and express themselves through "user-created knowledge" and a more pro-active, collaborative role in content creation, distribution and shared use. Active users and user-centred innovation are increasingly important and could have a beneficial social impact.

Target groups for such a knowledge-generating, collaborative GeoAnalytics tool are quite diverse. A primary target group is policy-makers in charge of regional development policy, who can make use of this tool in their decision process. As well as statisticians and analysts involved in policy analysis and evaluations. Citizens and the media would also be able to get informed and at the same time participate in increasing the knowledge on how life is lived — and can be improved — from region to region.

Because of the different expertise and needs of the target-groups, the tool should be flexible and adaptable to different audiences.

Figure 1 A dynamic OECD Visualisation Web Report



Source: OECD Regions at a Glance

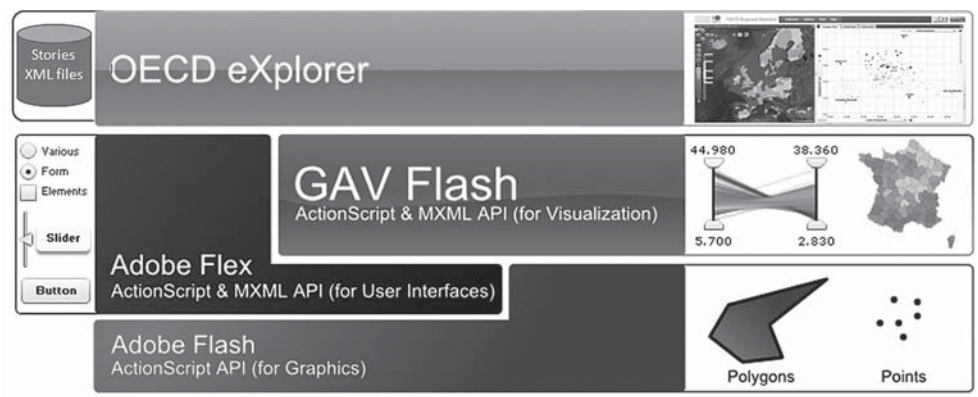
2 RELATED WORK

Massive volumes of geospatial statistical data are today generated all over the world but not used as effectively as one would wish for. Too little focus has been given to making GeoAnalytics technologies useful and accessible to statisticians. A common and functional multivariate geovisualization research method, such as the parallel coordinates plot (PCP) used in CommonGIS (Andrienko et al., 2003), is still unknown to the broad statistics user community. Despite the advances in web graphics technologies, comparatively little research has been focused on more advanced dynamic GeoAnalytics applied to Internet users. The most well-known tool is GeoVISTA Studio, an open source Java-based visual programming environment that is commonly used for developing Internet applications within the research community (Geovista studio, 2006). Rosling has dramatically increased the awareness of geovisualization tools among the world of statisticians by exploring his most efficient Gapminder time-animated bubble plots (Rosling, 2006).

The importance of providing explorative sessions in geovisualization and incorporated features to capture and reuse interactions and integrate them into electronic documents was early demonstrated by MacEachren (MacEachren and Brewer, 2001) and Jern (Jern, 2001). CCMaps presents a conditioned choropleth mapping tool that allows users to save snapshots and reuse them for presentation purpose. More recent efforts were made by Visual Inquiry Toolkit (Guo et al., 2006) that allows users to place pertinent clusters into a "pattern-basket" to be reused in the visualization process. Robinson describes a method they call "Re-Visualization" and a related tool ReVise that captures and re-uses analysis sessions (Robinson, 2006); Keel describes a visual analytics system of computational agents that support the exchange of task-relevant information and incremental discoveries of relationships and knowledge among team members commonly referred to as sense-making (Keel, 2006). Wohlfart describes a storytelling approach combined with interactive volume visualization and an annotated animation (Wohlfart and Hauser, 2007).

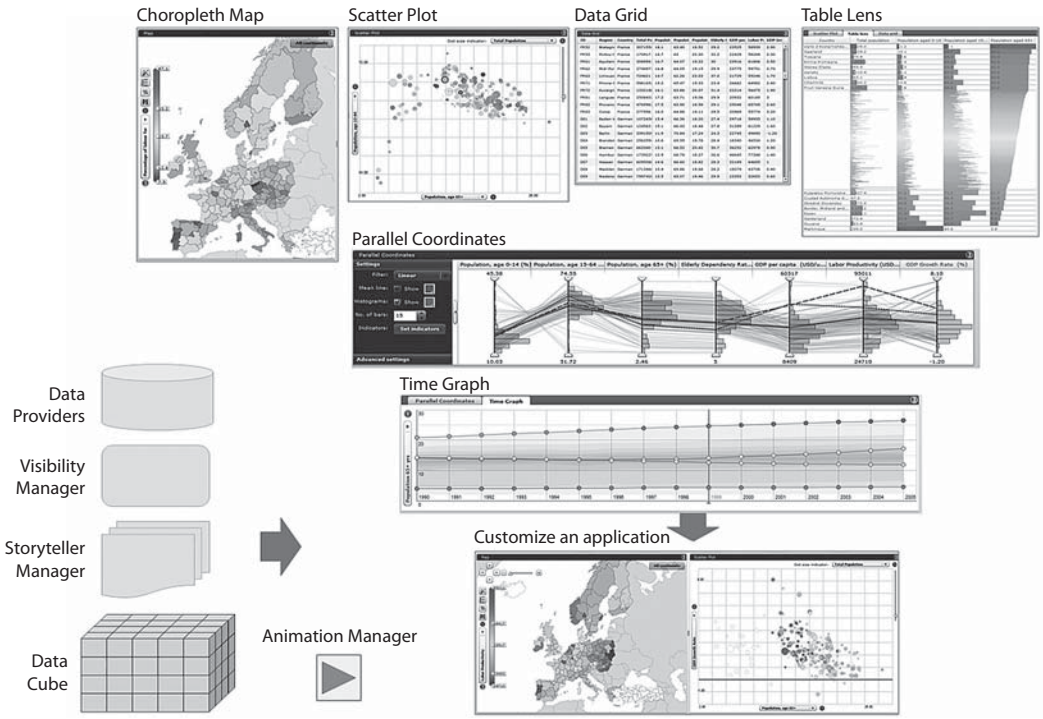
Many capture and reuse approaches are limited to be used within the same application environment that may well require a software license and are not always easily accessible to team members without installing external software. Increased computer security practices tend to limit this possibility. In this context, we introduce a web compliant layered component toolkit facilitating a snapshot mechanism that captures, re-uses and shares active properties for individual functional components. We demonstrate that such an implementation can provide a more open and collaborative GeoAnalytics framework for public use (OECD, 2009b).

Figure 2 GAV Flash uses the Adobe Flash API for graphics and integrates with the Adobe Flex Framework



Source: NComVA

Figure 3 GAV Flash component architecture



Source: NComVA

3 SYSTEM IMPLEMENTATION

OECD eXplorer was developed and customized using NCVA's own GAV Flash framework. A previous version for Microsoft's Net and DirectX is described in Jern and Johansson (Jern and Johansson, 2007). In this new version, our geovisual analytics methods have now been adapted for the Web 2.0 using Adobe's Flash basic graphics and Flex for user interfaces. Programmed in Adobe's object-oriented (O-O) language ActionScript, GAV Flash (Figure 2) and OECD eXplorer facilitate 100 % deployment to the Internet — a requirement from OECD and its associated global statistics community.

GAV Flash includes a collection of common geo-visualization and information visualization components, data analysis algorithms, tools that connect the components to each other and data providers that can load data from various sources. Interactive features that support a spatial analytical reasoning process are introduced such as tooltips, brushing, highlight, visual inquiry and conditioned statistics filter mechanisms that help detecting outliers.

As GAV Flash is built upon Adobe Flex, a developer has access to all Flex user interface functionalities. By combining buttons, panels and sliders with visual representations, applications can be easily customized. The open architecture allows new tools to be incorporated with the already existing components, e.g. statistical analysis tools or visual representations.

All GAV Flash components rest on different basis such as visualizations, managers or data providers and interact with each other, without knowing the exact nature of the others (Figure 3). A visual component needs a data provider to function, but the source of data can be different, from a simple Excel sheet to an extensive database. The data run through the data provider in order to create a common for-

Figure 4 GAV Flash Map Layer Component with three map layers glyphs, shaded map and Google map

Source: NComVA

matting that all the other components can understand. By separating the data structure from the visual representations, applications are created that work regardless of input, so that data can be supplied from the outside and linked into the system with minimal programming skills.

All visual components (Figure 3) use the same data storage class, Data Cube. This class is also used in the analysis to calculate different properties of the indicators, such as percentiles and histograms. Data loading is implemented separately for each data source type. For example, the Excel Reader loads an Excel spreadsheet and creates a data cube containing the data from the spreadsheet and the rest of the framework does not need to know anything about the data source. Data providers can be customized to support many types of sources. Excel is one of the simpler ones, while direct data base connections via a standard such as SDMX (SDMX, 2009) are more advanced and often require advanced technical skills from the user.

3.1 GAV Map Layers

The choropleth map component facilitates a novel layered architecture for simultaneous transparent views of multiple map layers. It means that each class of spatial information is represented by its own layer, e.g. glyph, shaded map and Google map. These layers can then be combined and controlled to be displayed, hidden or transparent depending on the needs of the user (Figure 4).

In a typical GAV Flash application (Figure 1) each visual component is added to a *Selection Manager* and a *Visibility Manager*. The selection manager receives selection events (such as clicking in the Map) from each component and then forwards these events to all other components. The visibility manager also listens for and forwards events, but in this case from filters that hide spatial regions based on some criteria. Both the selection manager and the visibility manager store a state (e.g. selected, visible) of each data record.

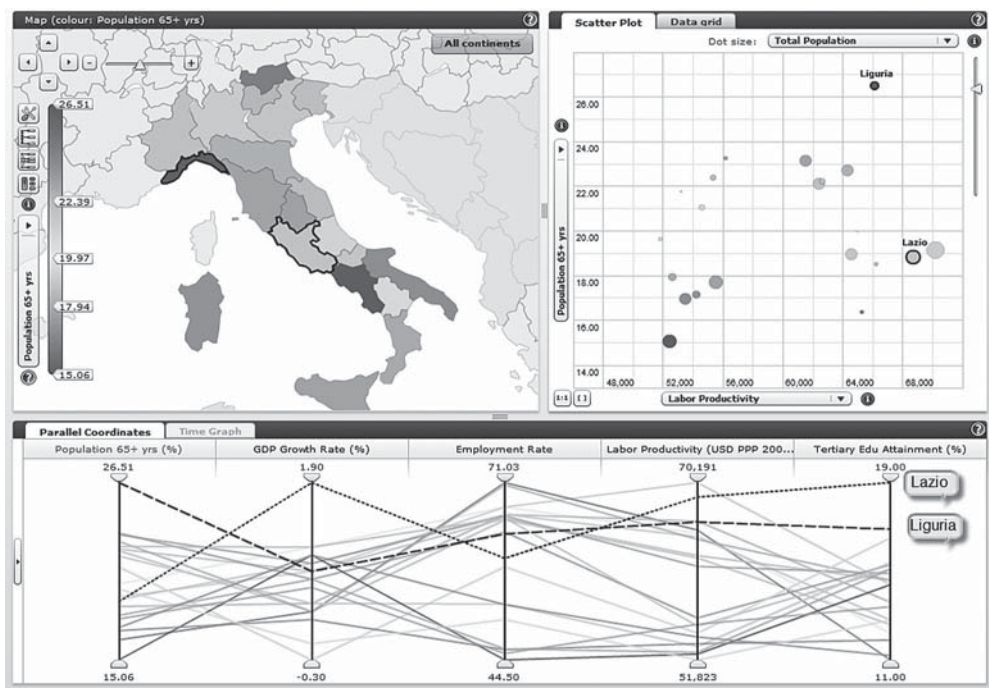
A minimal GAV Flash application contains a data provider and an interactive visualization component such as the *Choropleth Map (or Parallel Coordinates Plot, Scatter Plot etc)*. An application with a single visualization component does not need any selection or visibility managers since they only deal with interaction between two or more components. The visualization uses a *Data Cube*, supplied by the *Data Provider* to calculate and create its own special view of the data, such as a colour scheme, correlation plot or a histogram.

3.2 Dynamically linked views

Spatial and multivariate data are effectively analysed through the use of multiple-linked and coordinated views (Roberts, 2004, Jern et al., 2005). In order to detect complex patterns it is convenient to view data through a number of different visual representations simultaneously, each of which is best suited to highlight different features. GAV Flash facilitates an object-oriented architecture with compulsory components (selection manager and visibility manager etc. — see Figure 3) integrated with Adobe Flex layout management. This integration provides alternative and different views of the statistical data that can help stimulate the analytical visual thinking process characteristic of the GeoAnalytics reasoning. The views are separated by interactive splitters allowing the user to scale the individual views and allocate more space to the visual representation that is most important.

Spatial-temporal and multivariate statistical data are simultaneously explored through the use of dynamically linked multiple views to detect complex patterns and problems as described in more detail below. The views are coordinated using the GAV data linking method based on the data cube model (see Figure 3) and colouring scheme. Any filtering or highlighting made in one of the linked

Figure 5 OECD eXplorer uses dynamic multiple linked views



Source: NComVA

functional components is transmitted to all the others. Figure 5 is an example of an eXplorer linked view scenario (choropleth map, scatter plot and parallel coordinates) applied to the Italian TL2 regions. The chosen colour map indicator is "percentage of population aged 65 and more". Two regions, Lazio and Liguria, are highlighted in all three views for comparison. Each string (line segment) in the PCP (see paragraph 3.3) represents an Italian region, coloured according to a chosen indicator. Each of the five vertical axes represents an indicator (the red one is the one analyzed). Lazio and Liguria are represented with different highlighted dotted profile lines and can be compared for all indicators in the same visual representation.

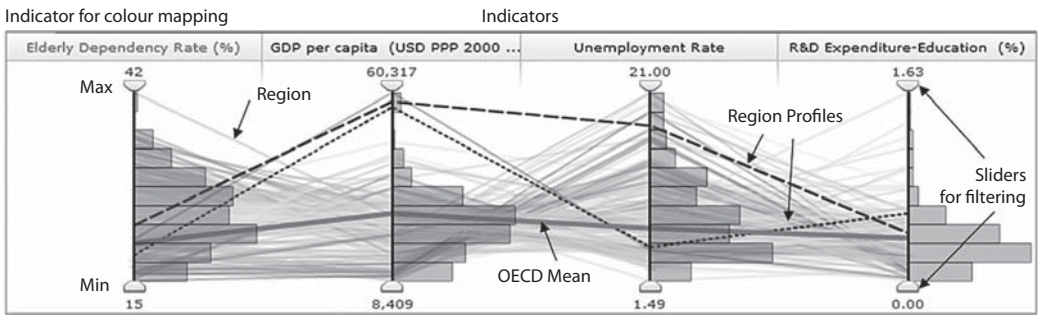
3.3 Parallel Coordinates Plot (PCP) in Adobe Flash

The strength of the PCP has already been demonstrated in many scientific environments (Inselberg, 1985, Andrienko et al., 2006); it represents a proven geovisualization technique that enables visual representation of spatial multivariate data and hence a key explorative mechanism in a Geovisual Analytics application. The technique supports a large number of tasks for analyses of relationships between indicators. In the case of OECD eXplorer, each region is represented by a string passing through the parallel axes. Each axis represents a single indicator in the data cube. A string forms a visual representation of the characteristics of one region. Differences between selected regions can be found by visually comparing the profiles representing them. Dynamic range sliders are attached to each axis and the user can dynamically select or combine filter methods thus altering constraints on indicator values shown in the other views (Figure 6).

This PCP has been extended with special features that are important to statistical exploration, such as histograms and filter operations based on percentile statistics (Jern and Johansson, 2007). Histograms attached to each axis are used to visualize the distribution of indicator data, splitting the axes into a user defined number of equally high rectangular areas (bins). The width of a rectangle indicates the frequency of regions intersecting that bin, the more regions within an area the wider the rectangle (Figure 6). Dynamic Flash animated histograms have proven helpful by OECD for analysing multivariate temporal indicator data.

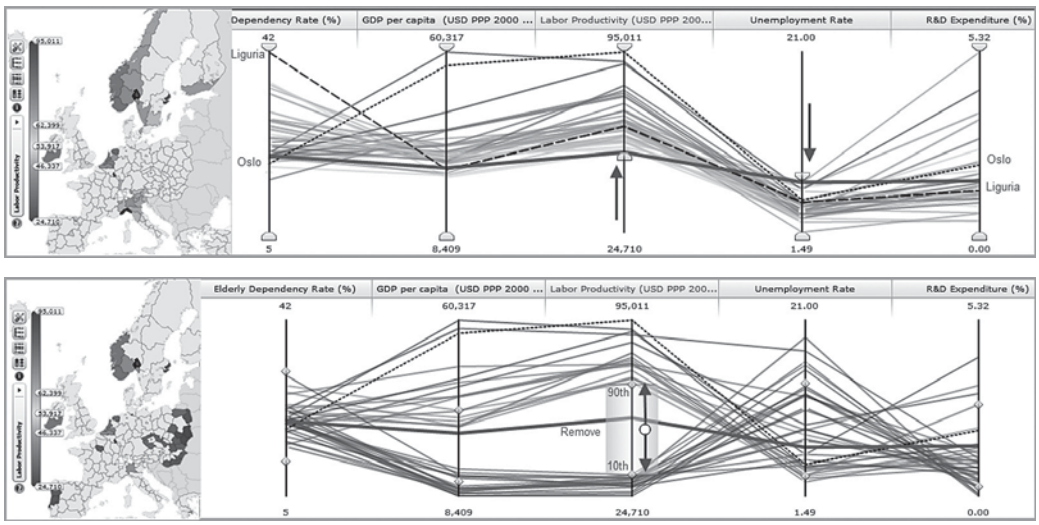
Statistical filter methods based both on ranges chosen by the user and on percentile calculations are embedded in the PCP attached to an indicator. Figure 7 (top) shows only regions which fulfil two conditions controlled by dynamic sliders along the indicators axes: a) Labour productivity must be higher than the OECD mean value and b) unemployment rate must be lower than the mean value. Two regions Oslo and Liguria which comply with these conditions are highlighted. Figure 7 (bottom) shows the regions with the highest (above 90th percentile) and the lowest (below 10th percentile) labour productivity. Oslo's profile is highlighted.

Figure 6 Parallel Coordinate Plot for analysing and comparing regional data



Source: eXplorer component by NcomVA

Figure 7 Parallel Coordinates Plot to filter data

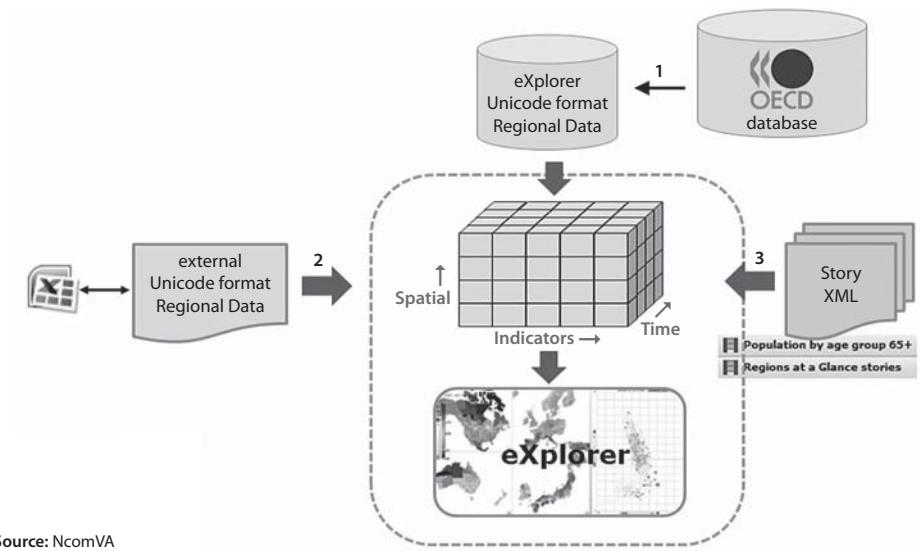


Source: eXplorer component by NcomVA

3.4 Getting data into eXplorer

There are several ways to load data into eXplorer. First, data can be preloaded using a Unicode text file format (UNICODE), which is a simple spreadsheet format generated by Excel. The OECD regional database is preloaded in eXplorer (arrow 1), thus the user has only to retrieve the indicators and years he wants to show by clicking on the name of the indicator. Second, data can be imported by the user from a client local data file using an Excel-like format "Unicode" (arrow 2). With this option the user is able

Figure 8 Different ways to load data into eXplorer



Source: NcomVA

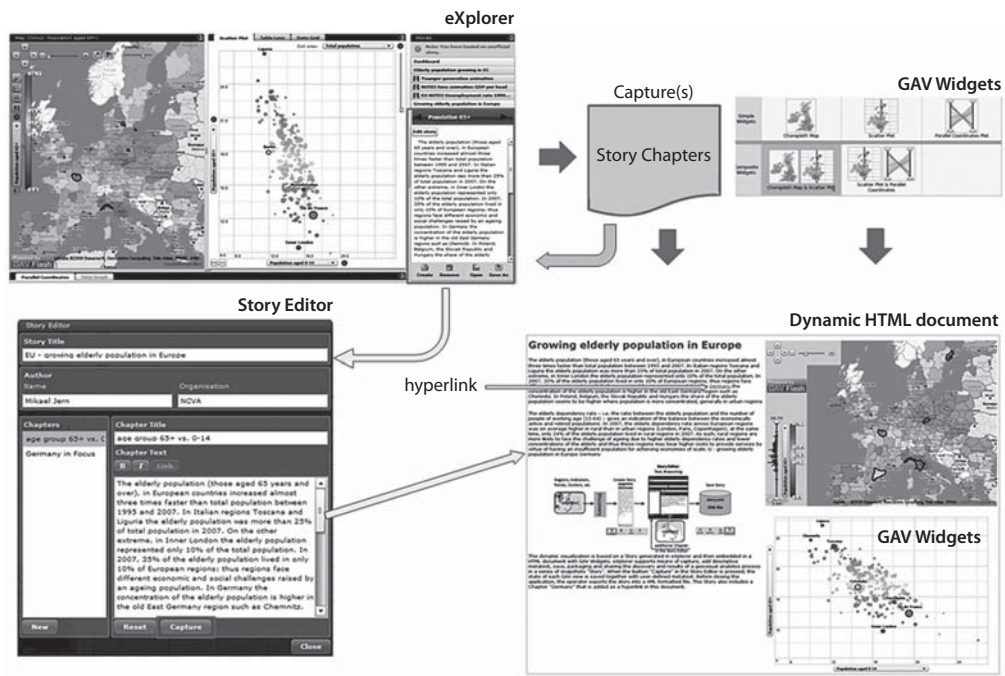
to visualize additional data together with the OECD regional database, achieving a great flexibility. Finally, external data can be included in story xml (arrow 3, see paragraph 4). Stories built from temporal data are marked with a "movie" icon.

4 COLLABORATIVE GEOANALYTICS THROUGH STORYTELLING

Complex and collaborative geovisual analytics tasks require the external representation and visual organization of information. These methods can help users compare, organize, comprehend and reflect on what they know, quickly access specific information when needed, remember relevant thoughts and ideas, as well as exchange knowledge and develop a shared understanding with other people. Computer generated visualizations usually explicitly state relationships among information items, thus allowing for quick and non-ambiguous explorations of an information space. Information constellations generated by human beings are often vague in regards to relationships, thus inviting more creative interpretations of an information space. The GAV Flash Framework integrates tools for both collaborative interactive visualization and sense-making. A story indicates a successful suggestion and subsequently fosters additional suggestions based on similar considerations. This learning mechanism allows our storytelling system to improve the accuracy of its suggestions as well as to dynamically adapt to particular users, tasks and circumstances. Colleagues can review a story arrangement and respond with suggestions and comments, which can subsequently foster additional suggestions based on similar considerations.

OECD eXplorer facilitates the architecture to support means of capture an image, add descriptive text, save, packaging and sharing the results of a geovisual analytics process in a series of snapshots "Story" (Figure 9). When the button "Capture" in the *Story Editor* is pressed, the state of each GAV Flash (i.e. map, PCP, scatter plot or time animation) in OECD eXplorer is saved together with user-defined text. Before

Figure 9 How to capture, save and share a story on Explorer



Source: eXplorer component by NcomVA

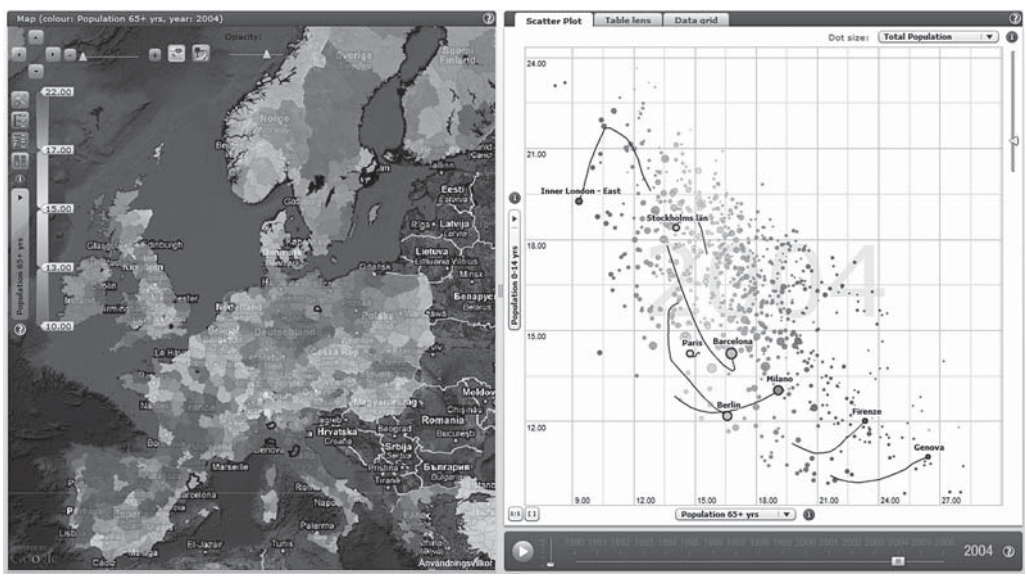
closing the application, the user can export the story into an XML formatted file. Other users can follow the analyst's way of logical reasoning by loading selected stories through descriptive text combined with interactive visualization. At any time a team member can access stories and apply them in OECD eXplorer or any other GAV Flash application assembled from the same components. A comprehensive story in the context of a remote collaborative sense-making activity can thus be created by the analyst through a set of linked snapshots (chapters). Users will discuss relevant issues through storytelling based on solid evidence, thus raising awareness and increasing the common knowledge on a certain phenomenon.

5 APPLYING FLASH-BASED TIME ANIMATION

GAV Flash employs a data model (Figure 3) optimized for handling spatio-temporal and multivariate indicator data sets in a GeoAnalytics context (Franzén and Jern, 2006). This conceptual data model can be seen as a data cube with three dimensions: space, time and indicators. The spatial dimension is represented by the OECD TL2 or TL3 regions, the indicators are the socio-economic indicators (GDP growth, elderly dependency rate, etc) and time are the yearly values of the indicators. The general method for finding a value in the cube is by its position (space, time and indicator). Of course different definitions of space, time and indicators can be applied in eXplorer.

Space-time-indicator awareness means that the data cube can be analysed and visualized across all three dimensions simultaneously. OECD eXplorer performs this task by integrating and time-linking three visual representation views: choropleth map, PCP and scatter plot (bubble). The bubble plot is demonstrated to be an effective Flash-based time animation method (Rosling, 2006), but integrated and linked with a choropleth map and a multivariate indicator frequency histogram embedded in a PCP, makes it even more useful. The GAV open architecture, which can handle large datasets and integrated snapshot mechanisms, is another important extension to this emerging animation technique. The spatial visualization for the time-linked views was evaluated by OECD and Statistics Sweden and found to be both intuitive and innovative. Statisticians discovered interesting trends in all three views respectively.

Figure 10 Time Animation in eXplorer applied to OECD TL3 regions



Source: OECD eXplorer

6 OECD EXPLORER: THE HISTORY AND THE RESULTS

When the OECD and Statistics Sweden organised a seminar on innovative methods of generating knowledge from statistical information in Stockholm in May 2008, OECD and NCVA became partners, and it became apparent that the tools developed by NCVA had a lot of potential to meet the wishes of OECD. This started a cooperation that resulted, a few months later, in the launch of OECD eXplorer on the OECD's web site.

NCVA demonstrated ability and will to adapt the tool to the needs expressed by the OECD, and later to the proposals expressed by pilot key users during a two months trial, thus adjusting eXplorer in an ideal way to user needs. The current version of OECD eXplorer has two identical tools to analyze, respectively, statistics on large regions and small regions. This choice was made to avoid complexity and long downloading time of large geometry files. OECD regions are analysed through four separate application scenarios grouping countries in continents and according to the territorial level chosen. The latest addition, *OECD Factbook eXplorer*, provides all the country statistics contained in the flagship publication *OECD Factbook 2009*.

The choropleth map dynamically linked with a PCP and a data grid gives the users facility for simultaneous analysis of the geospatial data. The use of the filtering function in OECD eXplorer, through the percentile distribution, helps visualizing the regional distribution within national borders and comparing it across countries. The use of lines and mean values in the PCP can help identify those regions that outperform their country or the OECD area as a whole and those that lag behind. Finally the scatter plot can highlight correlation among indicators and patterns of association among different types of regions. And the time animation adds the possibility of viewing the changes in all these patterns.

6.1 Evaluation

The OECD eXplorer development followed a user-centric design approach (Andrienko et al., 2003). A combination of domain experts from OECD, statistical managers and selected users of OECD statistics outside the organisation have been involved in the various stages of the prototype design and implementation, providing user feedback about usability and utility evaluation. The user-centric design process involved public access to beta versions of the web-based tool, allowing focus group discussions. The overall involvement and reactions have been very positive. Many useful suggestions for improving the functionality were made and have been incorporated in successive iterations.

A number of characteristics of the current version of OECD eXplorer derive from comments received during the evaluation phase. Firstly, users asked to have help functions and tutorial features for dynamic web-enabled GeoAnalytics tools, since the targeted audiences is not expert in geo- or information visualization technologies. Secondly, users asked to keep the entire structure sufficiently simple, while maintaining some functions to analyze data and not only visualize them. In this context, for example, the PCP was considered not to be self-evident to traditional users of statistics, as this is a technique that has not previously been employed in the statistics community and is not described in the methodological literature on statistics; therefore it was decided to keep it hidden in the start-up phase; at the same time it was regarded as a valuable addition to the statistical toolbox, especially the possibility of dynamically filtering to discover outliers and use profiles to make comparisons between highlighted regions. Finally, the dynamic links between views (context and focus maps, scatter plot, PCP and table grid) were evaluated as very important.

The test phase also revealed limitations that had to be addressed, e.g. resource demands when accessed on the web, resulting in too long times for data uploading and response. It also became clear, for example, that not all the institutional users had the possibility to have the latest versions of Flash installed on their client pc's.

6.2 Towards an open eXplorer

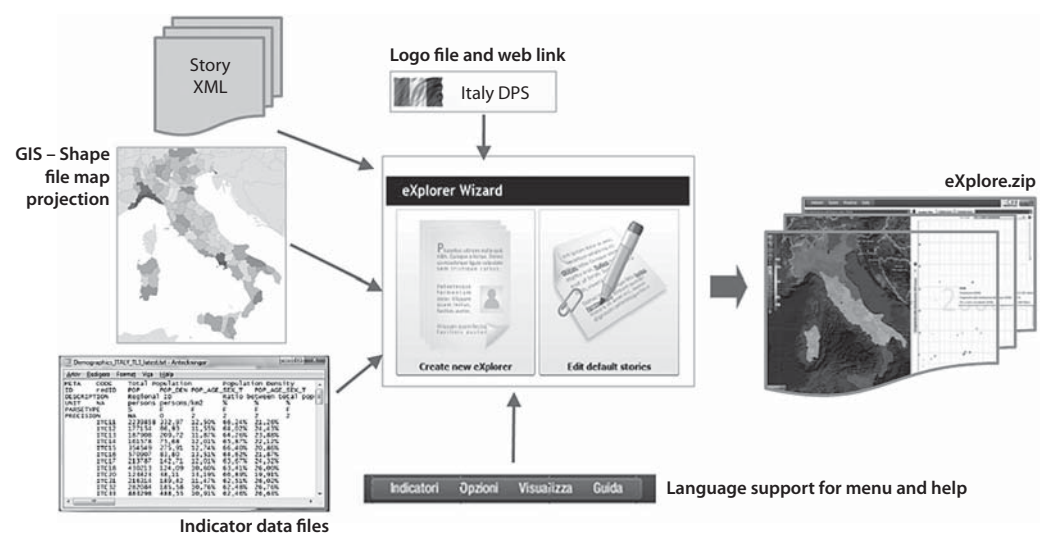
Following the release of OECD eXplorer in March 2009, further developments have been achieved in the direction of customizing the application to better suit different potential users. The aim is to build an "open eXplorer" or a generic platform that can be distributed to institutional partners (National Statistical Offices but also other organizations), and easily customized according to their needs. Explorer would then become the common platform used by different organizations increasing the integration and exchange of information. These recent developments have followed the same user-centric perspective used in the first releases of OECD eXplorer: Potential users have been involved in expressing their specific needs and evaluating the technical solutions adopted by NCVA.

So far the following tasks can be customized by NCVA's *"eXplorer Wizard"* developing a new eXplorer implementation (see Figure 11): Shape files (regional boundaries), Indicator files and map projection, so that different territorial partitions can be visualized (focusing on one country, depicting more detailed geographic boundaries etc.), company logo, language support for the help file, and panel graphical users interface buttons so that the application can be hosted on an institutional website using an associate language. Finally different layouts for maps view are allowed and stories can be preloaded to give even further flexibility of use to eXplorer.

CONCLUSIONS AND FUTURE DEVELOPMENTS

The authors expect that the web-enabled OECD eXplorer (OECD, 2009b) will enhance the use and understanding of OECD regional statistics, thus contributing to build evidence-based decisions. At the same time, it will encourage the practical use of advanced GeoAnalytics science technologies because of its easy accessibility on the Internet. It will enable the analyst to take a more active role in the discovery process of exploring regional indicators, for example, to identify those regions that outperform others, or detect common pattern among regions across countries. The tool will increase the interest in and knowledge of local factors for development among specialist as well as non-specialist users. Feed-back from domain experts and statisticians who have started using the tool shows that a sense of analytical

Figure 11 The program "eXplorer Wizard" can be used by different organizations to customize eXplorer



Source: NcomVA

reasoning and speed-of-thought interaction is achieved through the ability to dynamically link views in OECD eXplorer and thus see the multi-dimensionality of regional development.

Major achievements on the technical side can be summarized as follows:

- Geovisual Analytics framework and layered component architecture have been developed in the O-O language ActionScript with 100 % deployment to Internet users;
- Possibility to capture and save discoveries and knowledge (including time animation) with attached analytical describing text f to be communicated and shared with team members or the public;
- The on-going development of an open "eXplorer" platform that can be the foundation for easy customization of similar dynamic web applications using different geographical boundaries and indicators and be publicly available.

The OECD eXplorer version launched in March 2009 together with the report *OECD Regions at a Glance* (OECD, 2009a) includes three key new features:

- An SDMX interface that allows direct import of data to be presented in the tool. This gives seamless access to the entire OECD Regional database on the data warehouse "OECD.Stat" via an SDMX web service, as well as to any other database supporting SDMX. SDMX is the world standard for exchange and sharing of statistical data and metadata, recommended by the UN Statistical Commission as the preferred standard (see SDMX). It is already implemented in the information systems of many organizations, and many more are coming. This feature enables users to explore data of their own choice and integrate information coming from different sources. This feature is, therefore, of particular interest to regional policy-makers who will be able to drill-down the information on specific topics or on their own country, while using a common benchmark based on the OECD regional database. In the evaluation test preceding the release of OECD eXplorer, many users have welcomed this feature.
- Possibility to explore trends over time (yearly time series) for the indicators in the regional database (Section 5).
- Snapshot mechanism (Section 4) for presenting stories about the statistics embedded with interactive visualization and integrated into, for example, a HTML document structure.

The positive results of this project together with the ambition to develop a generic eXplorer platform have made NCVA a key partner of the OECD Global Project on "Measuring the Progress of Societies". A generic eXplorer platform would serve as the basis for the "Wiki4progress" project, which aims to become the catalyst of existing initiatives around the world on the measurement of progress of societies. Wiki4progress should represent the place where both experts and public could share their analysis and evaluations to raise awareness, inform and get informed through statistical information and storytelling based on solid evidence.

ACKNOWLEDGEMENTS

This applied research case study was carried out by NCVA in close collaboration with OECD who supplied data and comprehensive evaluation of the application. The research is in part supported by funding from the "Visualization Program" coordinated by the Swedish Knowledge Foundation. The authors wish to thank colleagues Tobias Åström and Markus Johnsson at the National Center for Visual Analytics (NCVA) at ITN, Linkoping University. The authors also wish to thank Delegates of the OECD Working Party on Territorial Indicators for valuable feedback on previous versions of the OECD eXplorer.

References

ANDRIENKO V., ANDRIENKO N. and VOSS H. GIS for Everyone: the Common GIS project and beyond. In PETERSON M. (ed.) *Maps and the Internet.* Elsevier Science, 2003, pp. 131–146.

ANDRIENKO G. and ANDRIENKO N. Visual exploration of the spatial distribution of temporal behaviors. In *Proceedings of the International Conference on Information Visualisation IEEE.* London: IEEE Computer Society, 2005, pp. 799–806.

ANDRIENKO G. et al. The parallel coordinate plot in action: design and use for geographic visualization. *International Journal of Geographical Information Science,* 2006a, 20(10).

ANDRIENKO G. et al. Reactions to geovisualization: an experience from a European project. *International Journal of Geographical Information Science,* 2006b, 20(10).

BRODBECK D. and GIRARDIN L. Design study: using multiple coordinated views to analyze geo-referenced high-dimensional datasets. In *Proceedings of the Coordinated and Multiple Views in Exploratory Visualization.* London: IEEE Computer Society, 2003, pp. 104–111.

CARR D., WHITE D. and MacEACHREN A. M. Conditioned choropleth maps and hypothesis generation. *Annals of the Association of American Geographers,* 2005, 95(1), pp. 32–53.

FRANZÉN J. and JERN M. *GeoAnalytics — Exploring spatio-temporal and multivariate data.* Reviewed proceedings IV, London: IEEE Computer Society, 2006.

GUO D. et al. A visualization system for space-time and multivariate patterns (VIS-STAMP). *IEEE Visualization and Computer Graphics,* 2006, 12(6).

INSELBERG A. The plane with parallel coordinates. *The Visual Computer,* 1985, 1(2), pp. 69–92.

JERN M. Smart Documents for Web-Enabled Collaboration. In VINCE J. A. and EARNSHAW R. A. *Digital Content Creation.* Springer Verlag, 2001.

JERN M. et al. *Tailor-made Exploratory Visualization for Statistics Sweden.* London: IEEE Computer Society, 2005.

JERN M. and JOHANSSON S. *GeoAnalytics Visual Inquiry and Filtering Tools in Parallel Coordinates Plot.* Seattle: ACM GIS, 2007.

JERN M. et al. *Visual Analytics presentation tools applied in HTML Documents.* Reviewed proceedings, IV08. London: IEEE Computer Society, 2008.

KEEL P. *Collaborative Visual Analytics: Inferring from the Spatial Organisation and Collaborative use of information, VAST,* 2006, pp. 137–144.

MacEACHREN A. M. and BREWER I. Geovisualization to mediate collaborative work: Tools to support different-place knowledge construction and decision-making. In *20th International cartographic conference, Beijing, China,* 2001.

OECD. *OECD Regions at a Glance.* Paris: OECD, 2009a.

OECD. *OECD eXplorer* [online]. Paris: OECD, 2009b. <http://stats.oecd.org/OECDregionalstatistics>.

OECD. *OECD Regional database* [online]. Paris: OECD, 2009c. <http://stats.oecd.org/WBOS>.

ROBERTS J. Exploratory visualization with multiple linked views. In MacEACHREN A. M., KRAAK M. and DYKES J. *Exploring Geovisualization.* Amsterdam, 2004.

ROBINSON A. *Re-Visualization: Interactive Visualization of the Progress of Visual Analysis, workshop proceedings.* VASDS, 2006.

ROSLING H. *Gapminder* [online]. 2006. <http://www.gapminder.org/world>.

SDMX [online]. 2009. <www.sdmx.org>.

THOMAS J. and COOK K. *Illuminating the Path: The Research and Development Agenda for Visual Analytics.* 2005.

UNICODE. *What is unicode?* [online]. 2009. <http://unicode.org/standard/WhatIsUnicode.html>.

WOHLFART M. and HAUSER H. Story Telling for Presentation. In *Volume Visualization.* EuroVis2007, 2007.

Site-Oriented Statistics and its Geoinformatic Potential

Petr Klauda[1] | *Czech Statistical Office, Prague, Czech Republic*

Abstract

A promising method used by many European statistical institutes, including the Czech Statistical Office, is the consistent orientation towards the addressed model of statistical surveys. This approach can be viewed as a way of standardization of most statistical tasks, which adds quality to the output and makes it more comprehensible and accessible for citizens. The existence of a nationwide spatial identification of addresses and buildings in the Czech Republic bears a significant potential for the application of geoinformatic technologies and their use in various areas of statistical service. The said data basis, which is unique in Europe, opens new qualitative possibilities for the collection, processing, analysis and presentation of statistical data, including the discovering of mutual connections.

Keywords

Address data, spatial data, GRID, spatial analysis, statistical georeports

JEL code

Z19

1 SPATIAL DATA AND STATISTICS

Every statistical task is based on data which can be delimited in terms of subject-matter, time and territory, thus answering three basic questions: What? When? and Where? Also every output from statistical tasks can therefore be delimited in terms of subject-matter, time and territory and the informative value depending on the form of such output is related to the answering of the above questions.

In this respect, spatial data create a certain spatial framework by means of which statistical data may be processed, analyzed and presented in the respective spatial context. This spatial framework can be based on physical real-world objects (e.g. localization of buildings and addresses, delimitation of network of streets) or abstract real-world objects (administrative boundaries, territorial, statistical and other units), or may be created artificially (GRID shapefiles).

2 ADDRESS-ORIENTED STATISTICAL DATA

At present, address-oriented collection and recording of statistical data is done by nationwide population censes, selective household surveys, agronomic censes, statistical researches in building industry and statistical registers. Therefore, the collected data belong to the area of demography, social, economic and agricultural statistics and to the area of recording of economic units, buildings and flats, including their technical and economic characteristics. Direct and exact geographical identification using the address provides for these statistics a detailed spatial dimension which can be further elaborated (Udržalová, 2007).

[1] E-mail: petr.klauda@czso.cz.

Placing statistical data in a spatial framework allows us to apply analytical methods using map algebra and geostatistics, including methods such as the interpolation method ("Kriging"), quadrant method, kernel method, etc. Furthermore, if we relate the data directly to a localized place or building, the results may be even more interesting. Information that is precisely related to a certain location has a number of advantages:

- Possibility to make calculations for any administrative, territorial, residential, statistical and other units or for GRIDs of any size;
- Possibility of a very accurate (detailed) spatial analysis and presentation;
- Possibility to select different regional levels (accuracy) when presenting aggregated data (different GRID sizes);
- Possibility to precisely select the optimum presentation level (spatial detail) in order to maximize the informative capacity of the given indicator (phenomenon) while maintaining the maximum possible degree of accuracy (detail) and ensuring the protection of personal data;
- Possibility to dynamically generate aggregated statistical information for any selection of spatial area (statistical georeports);
- Possibility to create clusters and buffer zones with given characteristics.

However, the address-oriented approach cannot be applied to all statistical data. Several basic conditions must be met, ensuring that their collection in such detailed form is possible and meaningful, including particularly the following:

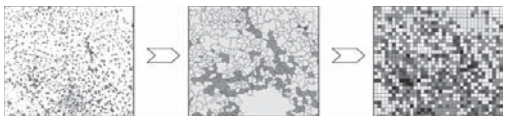
- The source of information must be identifiable in space through its address;
- Area data or sufficiently representative data must be collected (selective surveys do not meet this condition);
- The costs of an address-oriented data collection must be adequate to the value of its output;
- The observed indicator must be meaningfully interpretable also on the micro level (address);
- The use of this method must be based on detailed knowledge of the given data collection methodology;
- The spatial detail (size of the GRID cell) must be selected taking into account the nature of the given indicator, the extent and accuracy of obtained data and the necessity to secure full anonymity.

3 GRID REGULAR POLYGON SHAPEFILES

As it has already been said, when complying with the above conditions, address-oriented data may be very easily presented and analyzed using the GRID regular polygon shapefiles, which is the latest statistical phenomenon gaining its importance mainly in the Europ-wide context, where international comparability of statistical data must be ensured. GRID shapefiles are defined by geographic origin, geometric shape (squares, hexagons), surface area and the size of the elementary cell, which defines mainly the spatial resolution (accuracy) of the presented or analyzed data. The advantages of a regular polygon shapefile are very significant for statistical purposes and can be summarized as follows (Czech Statistical Office and HP s. r. o., 2010):

- Independence of the administrative structure which is changing over time and territory;
- Comparable spatial units of equal size;
- Easy anonymization of presented data;
- Possibility to hierarchize spatial presentations — different GRID sizes (macro, mezzo, micro level, etc.);
- Possibility to apply sophisticated geo-analytical methods (map algebra, geostatistics).

Figure 1 Different forms of presentation of address-oriented data (point density → administrative units → GRID)



Source: Own construction

GRID shapefiles also eliminate the basic problems and disadvantages of administrative and similarly defined spatial units, namely:

- Irregularity;
- Instability over time (changes of boundaries);
- Lack of internal homogeneity (for example, built-up areas, where most socioeconomic phenomena take place, represent only 1.7 % of the total area of the Czech Republic).

4 ADDRESS-ORIENTED DATA COLLECTION

An important prerequisite for the implementation of the above-described concept is the minimization of costs of address-oriented statistical surveys. One of the methods of cost reduction is the optimization of possible field research. The application of geoinformatic techniques and address-oriented geodata may support both logistic and qualitative aspects of these assignments. The solution of geographically oriented optimization tasks aimed at selecting a sufficiently representative sample of respondents, division of territorial districts to individual inquirers or optimization of inquirers' travels to perform field research are some of possible examples of optimization measures. Also, printed or on-line dynamic maps of the researched districts can be provided to inquirers for their better orientation. All these measures can help to achieve a better efficiency and quality of the collection, mainly by reducing administrative work and ensuring maximum productivity of the work performed by individual inquirers.

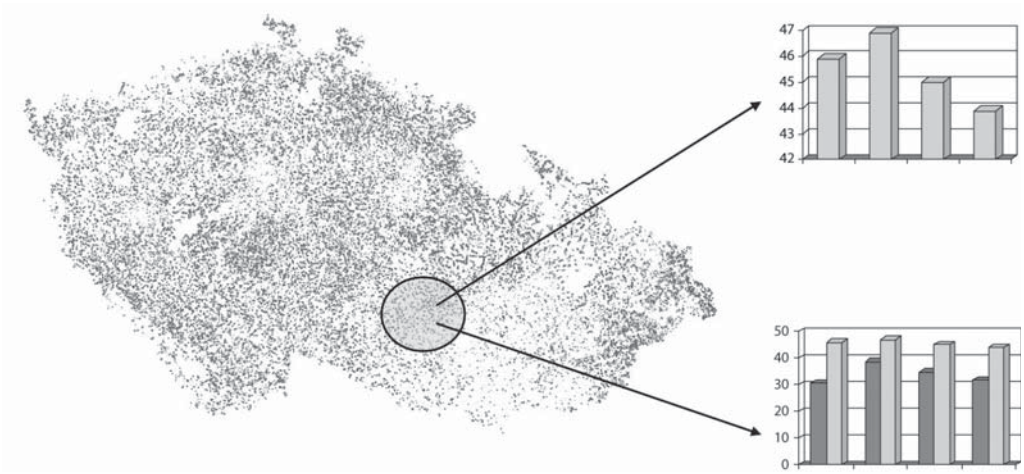
Address-oriented data collection, particularly with nationwide coverage, will always be a very time consuming and costly task. In this respect, statistical registers, being address-oriented and nationwide in their very nature, as well as other administrative data sources including the necessary data are very valuable.

5 USE OF ADDRESS-ORIENTED DATA — STATISTICAL GEOREPORTS

Site-oriented data localized in space can be used in very interesting user-focused services (statistical georeports), which can provide a statistical view on the user delimited area independently on any spatial structures (administrative or other area units etc.). These outputs are applicable in many areas and tasks. Some examples are here (Klauda, 2005):

- Geomarketing (territorial support of advertising campaigns, mapping of customers, customer and retail center network optimization, territorial analysis of the influence of competitors, revealing of geodemographic connections, etc.);
- Zoning (selection of suitable localities for investment and residential development including the analysis of their parameters and properties taking into account the demographic characteristics of the population at the given site);
- State administration and self-administration (optimization of the network of schools, bus and railway transport — analysis of commuter outflows to work and schools, modeling of spatial aspects in the labour market, etc.);
- Crisis management (flood plans, integrated rescue systems);
- Environment (modeling of the influence of emissions and radiation produced by industrial plants on their environment and the population, noise maps, radon risks in residential areas);
- Insurance services (calculation of insured amounts based on spatial aspects — houses in flood areas or highly polluted areas, classification of insured persons according to local conditions, etc.).

Figure 2 Statistical georeports — statistical data aggregations according to spatial selection



Source: Own construction

CONCLUSION

The Czech Statistical Office works with spatial data in selected fields of applied statistics, in databases and statistical registers on a long-term basis. Since 1999 it has been developing its own geographic activities in the field of collection, processing and presentation of data in an efficient cooperation with the Ministry of the Interior and the Czech Geodetic and Cadastral Office. The Czech Statistical Office contributes to the creation and implementation of a national geoinformatic infrastructure, increases awareness and facilitates access to geodata and geoinformation for all uses and gathers feedback.

Thanks to that, it has a wide database of statistical and spatial data, including the basic technological equipment and important know-how, allowing it to manipulate with address-oriented data also in spatial representation. However, it does not have a sufficient geoinformatic infrastructure necessary for standard implementation in statistical assignments and outputs. This situation should change radically if we succeed in creating such infrastructure including a statistical geoportal developed in accordance with Directive 2007/2/EC of the European Parliament and of the Council establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). This project should also fulfill the vision of gradual transformation of the existing forms of statistics to site-oriented statistics (geo-statistics) and the related offer of new user services.

References

CZECH STATISTICAL OFFICE and HP s. r. o. *Studie proveditelnosti: StatGIS, Fáze II: Hlavní studie (Feasibility Study: StatGIS, Phase II: Main Study).* Czech Statistical Office and HP s. r. o.: Prague, 2010.

KLAUDA, P. *Prostorové systémy pro podporu rozhodování* [Diplomová práce] *(Spatial systems to support decision-making,* Master's Thesis*).* Prague, 2005.

UDRŽALOVÁ, Z. Adresní místa. (Address Sites). In *ArcRevue* 2/2007, pp. 5–7.

Papers

The journal of *Statistika* has the following sections:
The *Analyses* section publishes high quality, complex, and advanced analyses based on the official statistics data focused on economic, environmental, and social spheres. Papers shall have up to 12 000 words or up to twenty (20) 1.5-spaced pages.

The *Methodology* section gives space for the discussion on potential approaches to the statistical description of social, economic, and environmental phenomena, development of indicators, estimation issues, etc. Papers shall have up to 12 000 words or up to twenty (20) 1.5-spaced pages.

The *Book Review* section brings reviews of recent books in the field of the official statistics. Reviews shall have up to 600 words or one (1) 1.5-spaced page.

Language

The submission language is English only. Authors are expected to refer to a native language speaker in case they are not sure of language quality of their papers.

Recommended Paper Structure

Title (e.g. On Laconic and Informative Titles) — Authors and Contacts — Abstract (max. 160 words) — Keywords (max. 6 words/phrases) — JEL classification code — Introduction — … — Conclusion — Annex — Acknowledgments — References — Tables and Figures

Authors and Contacts

Rudolf Novak*, Institution Name, Street, City, Country
Jonathan Davis, Institution Name, Street, City, Country
* Corresponding author: e-mail: rudolf.novak@domainname.cz, phone: +420 111 222 333

Main Text Format

Times New Roman 12 (main text), 1.5 spacing between lines. Page numbers in the lower right-hand corner. Italics can be used in the text if necessary. *Do not* use **bold** or underline in the text. Paper parts numbering: 1, 1.1, 1.2, etc.

Headings

1 FIRST-LEVEL HEADING (Times New Roman 12, bold)
1.1 Second-level heading (Times New Roman 12, bold)
1.1.1 Third-level heading (Times New Roman 12, bold italic)

Footnotes

Footnotes should be used sparingly. Do not use endnotes. Do not use footnotes for citing references (except headings).

References in the Text

Place reference in the text enclosing authors' names and the year of the reference, e.g. "White (2009) points out that…", "… recent literature (Atkinson et Black, 2010a, 2010b, 2011, Chase et al., 2011, pp. 12–14) conclude…". Note the use of alphabetical order. Include page numbers if appropriate.

List of References

Arrange list of references alphabetically. Use the following reference styles:
[for a book] HICKS, J. *Value and Capital: An inquiry into some fundamental principles of economic theory.* Oxford: Clarendon Press, 1939.
[for chapter in an edited book] DASGUPTA, P., et al. Intergenerational Equity, Social Discount Rates and Global Warming. In PORTNEY, P. et WEYANT, J., eds. *Discounting and Intergenerational Equity.* Washington, D.C.: Resources for the Future, 1999.
[for a journal] COASE, R.H. The Problem of Social Cost. *Journal of Law and Economics.* 1960, 3 (October), pp. 1–44.
[for an online source] CZECH COAL. *Annual Report and Financial Statement 2007* [online]. Prague: Czech Coal, 2008. [cit. 20.9.2008]. <http://www.czechcoal.cz/cs/ur/zprava/ur2007cz.pdf>.

Tables

Provide each table on a separate page. Indicate position of the table by placing in the text "insert Table 1 about here". Number tables in the order of appearance Table 1, Table 2, etc. Each table should be titled (e.g. Table 1 Self-explanatory title). Refer to tables using their numbers (e.g. see Table 1, Table A1 in the Annex). Try to break one large table into several smaller tables, whenever possible.

Figures

Figure is any graphical object other than table. Attach each figure as a separate file. Indicate position of the figure by placing in the text "insert Figure 1 about here". Number figures in the order of appearance Figure 1, Figure 2, etc. Each figure should be titled (e.g. Figure 1 Self-explanatory title). Refer to figures using their numbers (e.g. see Figure 1, Figure A1 in the Annex).

Figures should be accompanied by the *.xls, *.xlsx table with the source data. Please provide cartograms in the vector format. Other graphic objects should be provided in *.tif, *.jpg, *.eps formats. Do not supply low-resolution files optimized for the screen use.

Formulas

Formulas should be prepared in formula editor in the same text format (Times New Roman 12) as the main text.

Paper Submission

Please email your papers in *.doc, *.docx or *.pdf formats to statistika.journal@czso.cz. All papers are subject to double-blind peer review procedure. You will be informed by our managing editor about all necessary details and terms.

Contacts

Journal of Statistika | Czech Statistical Office
Na padesátém 81 | 100 82 Prague 10 | Czech Republic
e-mail: statistika.journal@czso.cz
web: www.czso.cz/statistika_journal

Contact us

Managing Editor: Jiří Novotný
phone: (+420) 274 052 017
fax: (+420) 274 052 133
e-mail: statistika.journal@czso.cz
web: www.czso.cz/statistika_journal
address: Czech Statistical Office | Na padesátém 81 | 100 82 Prague 10 | Czech Republic

Subscription price (4 issues yearly)

CZK 356 (incl. postage) for the Czech Republic,
EUR 199 (incl. postage) for other countries.

Printed copies can be bought at the Publications Shop
of the Czech Statistical office (CZK 66 per copy).
address: Na padesátém 81 | 100 82 Prague 10 | Czech Republic

Subscriptions and orders

MYRIS TRADE, s. r. o.
P. O. BOX 2 | 142 01 Prague 4 | Czech Republic
phone: (+420) 234 035 200,
fax: (+420) 234 035 207
e-mail: myris@myris.cz

Design and layout: Ondřej Pazdera
Typesetting: Jana Chocholoušová
Print: Czech Statistical Office