STATISTIKA

STATISTICS AND ECONOMY JOURNAL

VOL. **101** (4) 2021

CZECH STATISTICAL OFFICE

CONTENTS

ANALYSES

CONSULTATION

INFORMATION

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 professional double-blind peer reviewed open access journal included in the citation database of peer-reviewed literature **Scopus** (since 2015), in the **Web of Science** *Emerging Sources Citation Index* (since 2016), and also in other international databases of scientific journals. Since 2011, Statistika has been published quarterly in English only.

Publisher

The Czech Statistical Office is an official national statistical institution of the Czech Republic. The Office´s 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

Determinants of Economic Growth in the European Union Countries

Emília Zimková[1] | *Matej Bel University, Banská Bystrica, Slovakia*
Michaela Vidiečanová[2] | *Matej Bel University, Banská Bystrica, Slovakia*
Petra Cisková[3] | *Matej Bel University, Banská Bystrica, Slovakia*

Abstract

This paper reassesses the long-debated relationship between the financial system development and economic growth. We use not only indicators for financial access, efficiency, stability and depth of the bank-oriented financial sector, but we also consider Eurozone membership, corruption perception and competitiveness of countries to examine the determinants of economic growth. We apply a panel data approach to 27 European countries over the 2004–2017 period. By splitting the time span, we examine whether the effect of financial system development, Eurozone membership, corruption perception and competitiveness on economic growth is affected by the occurrence of financial and debt crises. Our results indicate that loans to private sector do not always support economic growth. Our research also reveals that corruption perception has a negative impact on economic growth, and so does membership in Eurozone during a crisis.

Keywords	***DOI***	***JEL code***
Economic growth, financial system, Eurozone membership, corruption perception, competitiveness, panel regression	*https://doi.org/10.54694/stat.2021.16*	*F43, B26, O11, O43, O44*

INTRODUCTION

There is no doubt that the issue of the sources of economic growth is one of the most frequently researched areas in economics. While until the global financial crisis in 2008–2009 reputable studies argued that the financial system development increases the country's performance, improves resources allocation, supports technology development and financial stability, after the global financial crisis the situation changed, and the financial systems became the central subject of criticism when explaining the causes of economic recession. However, economists have still not come to the battle of wills about an impact of financial systems on economic growth. Over the years, we find in economic literature the views

[1] Matej Bel University, Faculty of Economics, Department of Finance and Accounting, Tajovského 10, 975 90 Banská Bystrica, Slovakia. Corresponding author: e-mail: emilia.zimkova@umb.sk, phone: (+421)484466322.
[2] Matej Bel University, Faculty of Economics, Department of Finance and Accounting, Tajovského 10, 975 90 Banská Bystrica, Slovakia. E-mail: michaela.vidiecanova@umb.sk, phone: (+421)484466314.
[3] Matej Bel University, Faculty of Economics, Department of Quantitative Methods and Information Systems, Tajovského 10, 975 90 Banská Bystrica, Slovakia. E-mail: petra.ciskova@umb.sk, phone: (+421)484466617.

of authors, who in their analyses lean to one side or the other. To authors who claim a positive impact of finance on economic development belong Dollar (1992), Ben-David (1993), Wacziarg and Welch (2008), Lucas (2009), Perera and Paudel (2009), Madsen and Ang (2016). On the contrary, to the authors who claim a negative impact of finance on economic growth belong Demirguc-Kunt and Detragiache (1998), Fisher and Chenard (1997), Pinar and Damar (2006), Ayadi et al. (2013), Karagiannis and Kvedaras (2016), and Barradas (2018).

The aim of this study is to analyse the impact of selected financial soundness indicators of the bank-oriented financial systems on economic growth in the European Union countries. To enrich the existing strand of literature, we investigate not only the determinants typical for access, efficiency, stability and depth of the financial system, but also non-financial determinants as the perception of corruption, competitiveness in the economic environment, and the impact of the Eurozone membership. As time perspective, we analyse period from 2004 to 2017, as it enables us to examine an impact of selected determinants on the economic growth in the period before and after the financial crisis.

The analysis confirmed positive impact of credit on economic growth before the financial crises and a negative impact since crisis. This finding is in line with Ayadi, Emrah, Sami and Willem (2013), and Barradas (2018). It is evident that loans injected into economy during and after crises covered only operational needs of companies and were used for household consumption but did not support the economic growth (with the exception of Poland).

Our research also reveals that the membership in Eurozone during crisis disables the use of autonomous monetary and exchange tools and therefore has a negative impact on economic growth. As expected, the corruption perception shows a negative impact on economic growth. It is in line with Wu and Wei (2002).

The paper is structured as follows. The first part contains a brief literature overview, the second part provides description of data and model specifications, the third part reports main research findings that are discussed and in the final one concludes.

1 LITERATURE REVIEW

In their landmark study, Čihák, Demirguc-Kunt, Feyen and Levine (2013) invented several measures to benchmark financial systems: (a) the size of financial institutions and markets (financial depth), (b) the degree to which individuals can and do use financial institutions and markets (access), (c) the efficiency of financial institutions and markets in providing financial services (efficiency), and (d) the stability of financial institutions and markets (stability). These four characteristics were measured both for financial institutions and financial markets (equity and bond markets), thus covered bank-oriented and market-oriented financial systems. Their seminal paper enabled using these measures to characterize and compare financial systems across countries and over time and to assess the relationship between these measures of the financial system and key financial sector policies.

Tadesse (2002) examined whether economic growth was more favourably affected by bank-oriented financial systems or market-oriented financial systems. He found out that bank-oriented financial systems were much better in countries with underdeveloped financial systems and market-oriented systems outperformed bank-oriented systems in countries with developed financial systems. This finding was confirmed by Demirguc-Kunt, Feyen and Levine (2013), and Cambacorta, Yang and Tsatsaronis (2014). Arcand, Berkes and Panizza (2015) found out that the financial depth ceased to have positive impact on the economic growth when private sector credit exceeded 100% of gross domestic product (GDP). Similar research was conducted by Beck, Georgiadis and Straub (2014), who came to a similar result; their threshold was 109% of the share of private credit to GDP. Caporale et al. (2009) examined the interconnection between financial development and economic growth in the new EU member states (Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia) by Granger causality and revealed that the causality came from financial development to economic growth. Georgantopolus et al. (2015) focused on examining the relationship between financial systems and economic growth. They evaluated this relationship for the 28 EU member states divided into two groups; countries that belong to the Eurozone (17 countries) and those, that did not adopt the euro (11 countries). They argued that the adoption of euro required closer and more centralized political, economic, fiscal and financial cooperation between the Eurozone members. The group with countries of the Eurozone showed a significant contribution of the financial sector to the economic growth. On the contrary, the group of countries outside the Eurozone showed a significant negative impact of the banking sector on the economic growth. Armin, Ibrahim and Azman-Saini (2012) examined the relationship between the financial development and economic growth in 15 economically developed countries before and after the euro adoption. The results of panel regression showed that the financial development is important in supporting economic growth in both periods examined. They found out that the impact of the banking sector on economic growth was greater after the euro adoption, while the impact of developments on the market diminished over the examined period. The results of Stolbová, Battiston, Napoletano and Roventini (2017) showed an overall trend in increasing of financing in the Eurozone, as well as in individual countries of the EU, during the examined period. However, the pace of increasing of financing is different. The result of network analysis was the finding that a large part of the assets owned by financial institutions are in fact securities issued by other financial institutions.

As stated in the introduction of this study, within our analysis we focus our attention solely on the bank-oriented financial systems and their interaction with economic growth in the EU countries. Our intention is to analyse the determinants of economic growth according to the definition specified by Čihák et al. (2013). We, therefore, examine the impact of the bank-oriented financial system on economic growth in terms of four main characteristics; in terms of access, efficiency, stability and depth of the financial system. Moreover, our analysis also includes determinants which reflect the membership in the Eurozone, perception of corruption and competitiveness in economics, but also variables, which reflect world financial crisis in years 2008–09 and debt crisis in years 2010–2012. Thus we would like to broader the studies of Čihák (2013), Nyasha and Odhiambo (2017), Chu (2020) and other, financial determinants by other important areas, such as corruption perception, level of competitiveness of countries and the Eurozone membership.

2 DATA AND MODEL SPECIFICATION

2.1 Data specification

In this paper we analyse unbalanced panel data of the 27 EU countries in the period from 2004 to 2017. Researched economic growth is expressed by the gross domestic product per capita. Based on the study of Čihák et al. (2013) in order to measure the impact of bank-oriented financial system on the economic growth, we use (a) financial resources provided to the private sector over GDP as a measure of financial depth, (b) number of commercial bank branches to measure access of financial services, (e) return on equity as a measure of efficiency, and (d) Z-score to reflect stability of a financial system.

In our study we also included corruption perception index (CPI) and global competitiveness index (GCI), which significantly resonate at present times. The corruption perception index reflects the information about perceiving the level of corruption in 180 countries of the world, including all of the EU countries and it is published by Transparency International. The global competitiveness index has been published since 2004 by the World Economic Forum. The index reflects the competitiveness of countries in the world and is composed of more than 100 indicators, which assess different pillars of competitiveness. In addition, the study includes three dummy variables: Eurozone membership, global financial crisis (GFC) and debt crisis (DC). Period of global financial crisis and debt crisis is adopted from OECD Economic Outlook (2021) and it does not consider possible individual differences among analysed countries. Table 1 shows the description, sources of the variables used as well as their expected impact on the economic growth.

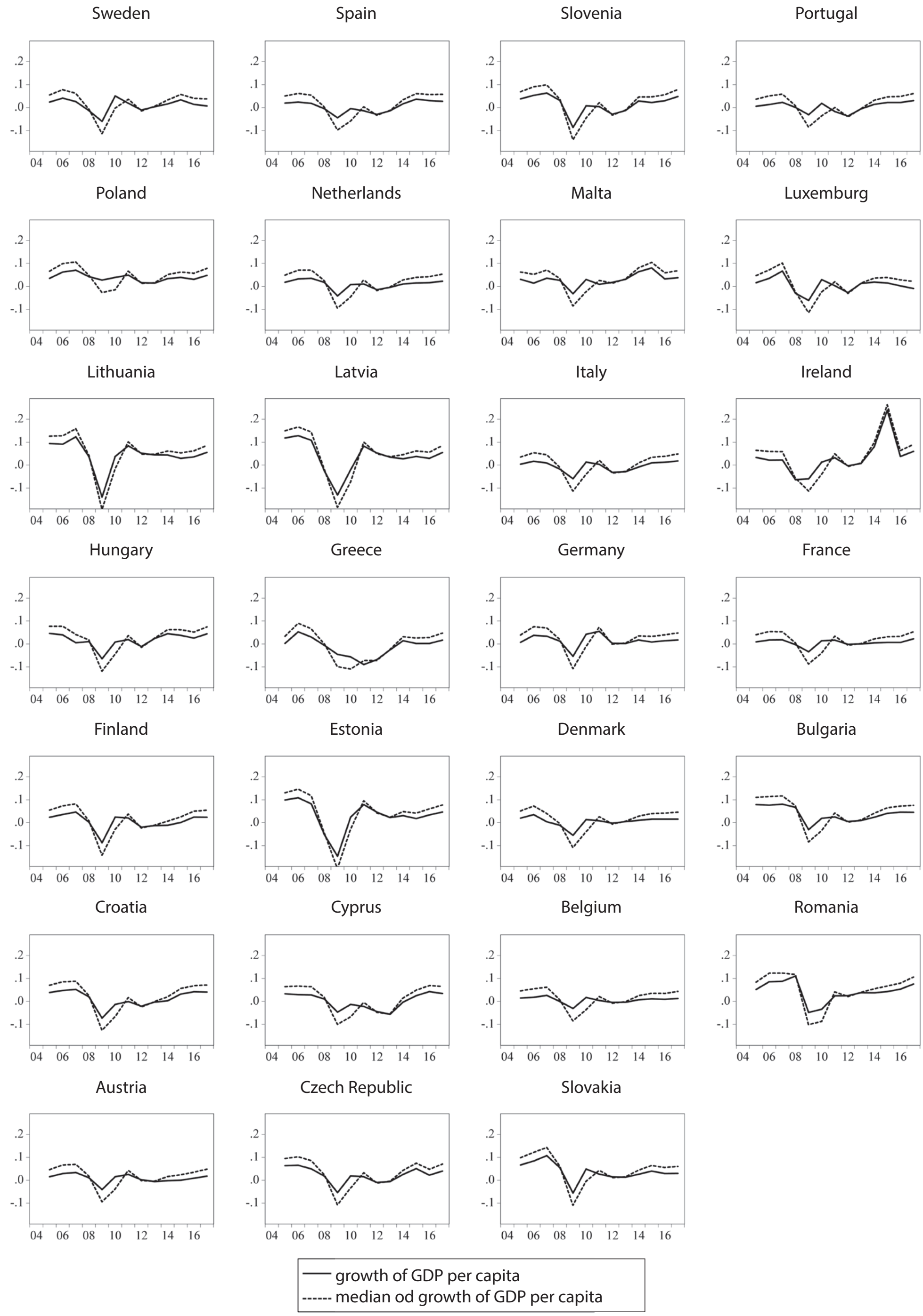
Table 1 Description of variables

Variable	Description	Unit of measure	Source	Expected impact on economic growth
GDP per capita – dependent variable	Gross domestic product divided by midyear population	Year over year percentage change	(World Bank)	
Access	Number of commercial bank branches per 100 000 adults	Year over year percentage change	(World Bank)	+
Efficiency	Bank ROE – commercial banks' pre-tax income to yearly averaged equity	Year over year percentage change	(World Bank)	+
Stability	Bank Z-score – it captures the probability of default of a country's commercial banking system	Year over year percentage change	(World Bank)	+
Depth	Domestic credit to private sector – financial resources provided to the private sector	Year over year percentage change	(World Bank)	-
CPI	Corruption perception index	Year over year percentage change	(Transparency International)	+
GCI	Global competitiveness index	Year over year percentage change	(World Bank)	+
Eurozone	Countries that are members of the Eurozone	Dummy variable (if 1 selected country is a member of Eurozone; 0 otherwise)	-	+
GFC	Global financial crisis in 2008 and 2009	Dummy (1 indicates the global financial crisis in 2008 and 2009; 0 otherwise)	-	-
Debt crisis	Debt crisis from 2010 to 2012	Dummy (1 means the year when the debt crisis has begun in European Union)	-	-

Source: Authors

The economic growth is expressed as gross domestic product per capita in annual percentage change. Its development in individual countries of the EU, together with the median, can be seen in Figure 1. The figure clearly visualizes whether the annual change in economic growth was above or below the median value of the analysed countries. It should be noted that median is the same line in all partial charts in Figure 1. As shown in Figure 1, individual countries of the EU recorded year-over-year growth in the period from 2004 to 2008. In 2008, as a result of the economic crisis, there was an economic downturn in almost all the EU countries. The exception was Poland, whose growth rate is above the median value throughout all analysed period. A significant negative impact on the economy confirm the studies of Checherita-Westphal and Rother (2012), Gómez-Puig and Sosvilla-Rivero (2015), and Pegkas et al. (2020). In 2010, the countries of the EU were hit by the debt crisis, after which GDP fell again, however, to a lesser extent than was in the case of the economic crisis. Cyprus and Portugal were hit hard by both crises; GDP fell even deeper during the debt crisis than during the economic crisis. Most of the EU countries were able to stabilize quickly and initiate the GDP growth, except for Greece. The Greek economy was in recession from 2008 to 2016. Based on the graphical results, we decided in this study to focus more on examining the impact of determinants before and after the economic crisis.

Figure 1 Development of GDP per capita in the countries of the European Union



Source: Authors

Table 2 presents descriptive statistics for the whole data set of the 27 EU countries in the years 2004–2017. The number of observations varies depending on the variable, as the data set is unbalanced. In terms of indicators, we can point out, for example, the negative value of the median of efficiency variable. This indicates that approximately 50% of the EU countries had a negative ROE during the period examined. The values of the access to finance indicator, which are represented by the number of branches per 100 000 inhabitants, indicates a decrease in the number of branches, which coincides with the ever-expanding trend of branchless banking.

Table 2 Descriptive statistics

	Mean	Median	Standard deviation	Minimum	Maximum
GDP per capita	1.768	1.833	3.991	–14.56	23.94
Access	–2.510	–2.278	10.44	–73.65	108.6
Efficiency	18.61	–10.30	700.9	–1684	11944
Stability	135.4	2.103	1163	–99.56	14587
Depth	1.698	1.268	9.387	–33.93	42.05
CPI	0.701	0.000	5.311	–19.15	24.24
GCI	0.224	0.000	3.452	–16.99	25.00
Eurozone	0.587	1.000	0.493	0.000	1.000
GFC	0.143	0.000	0.350	0.000	1.000
Debt crisis	0.214	0.000	0.411	0.000	1.000

Notes: All variables excluding dummy variables are expressed as year over year percentage change; Access – bank branches per 100 000 adults; Efficiency – bank return on equity (ROE in %); Stability – bank Z-score; Depth – domestic credit to private sector (% of GDP); CPI – Corruption perception index; GCI – Global competitiveness index; Eurozone – dummy variable (if 1 – country is a member of Eurozone, 0 – otherwise), GFC – dummy variable (if 1 – global financial crisis in 2008 and 2009, 0 – otherwise); DC – Debt crisis – dummy variable (if 1 – Debt crisis from 2010 to 2012, 0 – otherwise).
Source: Authors

One of the important assumptions when performing this analysis and constructing models, to the extent that we present in this study, is the assumption of uncorrelated independent variables. In Table 3 we present the results of correlation analysis. The results show only a small degree of dependence between the explanatory variables.

Table 3 Correlation matrix

	Access	Efficiency	Stability	Depth	CPI	GCI	Eurozone	GFC	Debt crisis
Access	1								
Efficiency	0.0200	1							
Stability	–0.0097	–0.0189	1						
Depth	0.2501	0.0241	–0.0435	1					
CPI	0.1168	–0.1046	–0.0092	0.0804	1				
GCI	0.0635	0.0230	–0.0144	–0.1574	0.0999	1			
Eurozone	–0.0777	–0.0405	–0.0795	–0.2146	–0.1637	0.0412	1		
GFC	0.0703	0.0330	0.1079	–0.1637	–0.1400	–0.0817	–0.0123	1	
DC	–0.0623	0.0056	0.0745	–0.2180	0.0353	0.0007	0.0302	–0.2132	1

Source: Authors

We subjected the continuous variables to stationarity tests and used unit root tests LLC, IPS, ADF and PP tests. Results indicate that the selected variables are stationary and do not contain a unit root. In case of the depth variable, we recorded the presence of unit root at the significance level of 0.05 by ADF test. Since LLC, IPS and PP tests at the significance level of 0.05 do not statistically confirm the presence of a unit root, we can state that these data are stationary.

Table 4 Unit root test for each continuous variable

	LLC	IPS	ADF	PP	Result
GDP per capita	–10.073 (0.000)	–5.760 (0.000)	125.81 (0.000)	119.492 (0.000)	Stable
Access	–35.279 (0.000)	–6.982 (0.000)	80.061 (0.008)	119.266 (0.000)	Stable
Efficiency	–4.257 (0.000)	–5.135 (0.000)	115.040 (0.000)	240.795 (0.000)	Stable
Stability	–7.155 (0.000)	–5.740 (0.000)	126.443 (0.000)	265.684 (0.000)	Stable
Depth	–6.399 (0.000)	–1.773 (0.038)	69.492 (0.076)	120.093 (0.000)	Stable
CPI	–5.224 (0.000)	–5.391 (0.000)	120.050 (0.000)	208.248 (0.000)	Stable
GCI	–10.906 (0.000)	–8.560 (0.000)	175.035 (0.000)	334.606 (0.000)	Stable

Source: Authors

2.2 Model specification

A substantial part of empirical research in this study focuses on the analysis of panel data, through GLS.

The model used in the research can be written in the following form:

$$EG_{it} = \alpha_i + \beta_j \, FS'_{it} + \gamma_k \, IN'_{it} + \delta_l \, DV'_{it} + \mu_{it} \, , \qquad (1)$$

where:

$$FS'_{it} = (A_{it}, E_{it}, S_{it}, D_{it})',$$
$$IN'_{it} = (CPI_{it}, GCI_{it})', \qquad (2)$$
$$DV'_{it} = (EA_{it}, GFC_t, DC_t)'.$$

The symbol i represents a specific country, t stands for a time, j, k, l correspond to specific type of financial structure variable, index variable and dummy variable. EG_{it} is the explained variable economic growth, FS'_{it} represents vector of financial structure indicators, IN'_{it} is the vector of index variables, DV'_{it} is the vector of dummy variables, A_{it} is the variable reflecting the access to the financial system, reflects the efficiency of the financial system, E_{it} represents the financial system stability, D_{it} is the financial system depth, CPI_{it} is the corupption perception index, GCI_{it} is the global competitiveness index, EA_{it} reflects the membership and entry into the Eurozone, GFC_t is the dummy variable reflecting the financial crisis, DC_t represents the dummy variable of the debt crisis and μ_{it} represents the error term.

Within our analysis we examine five models in three time periods. The first time span is the period from 2004 to 2017. The second time span is the sub-period from 2004 to 2009 and the third time span is the sub-period from 2010 to 2017. In these time periods we gradually examine five different models, which are composed of individual vectors of the hypothetical model (1).

A similar methodology was also addressed by Armin, Ibrahim and Azman-Saini (2012), Přívara and Trnovský (2021), Pegkas et al. (2020), Shittu et al. (2020), Agapova and Vishwasrao (2020), Mazurek (2017), and Pinar and Damar (2006). We contribute to existing literature by including the indicators that are nowadays topical.

3 RESULTS AND DISCUSSION

Table 5 provide results for the panel regression for the whole period from 2004 to 2017, while Table 6 contains results for the period span from 2004 to 2009 and the Table 7 for the period span from 2010 to 2017, respectively. Based on the Hausman test, we use the random effect of GLS, as Zhang, Wang and Ren (2021), Naghshpour (2019), and Prochniak (2011).

The regression results over the 2004–2017 period contain 5 models. Model 1 takes into consideration an impact of main financial indicators to assess access, efficiency, stability and depth of the bank-oriented financial systems on economic growth. Model 2 analyses not only an impact of financial indicators

Table 5 Regression results over the 2004–2017 period

Explanatory variable	Explained variable				
	GDP per capita				
	Model 1	Model 2	Model 3	Model 4	Model 5
	Random effect	Random effect	Random effect	Fixed effect	Random effect
Access	0.049** (0.0208)	0.038* (0.0207)	0.053*** (0.0164)	0.049*** (0.0176)	0.046*** (0.0171)
Efficiency	0.0006** (0.0003)	0.0007** (0.0003)	0.0007*** (0.0002)	0.0008*** (0.0002)	0.0008*** (0.0002)
Stability	–0.0003 (0.0002)	–0.0003 (0.0002)	0.00005 (0.0014)	0.0001 (0.00015)	0.00002 (0.00014)
Depth	–0.024 (0.0241)	–0.016 (0.0243)	–0.013 (0.0208)	0.002 (0.0213)	–0.008 (0.0211)
CPI		0.135*** (0.0409)		0.0584* (0.0348)	0.062* (0.0341)
GCI		0.097 (0.0641)		0.0445 (0.0537)	0.071 (0.0528)
Eurozone			–1.737*** (0.519)		–1.675*** (0.483)
GFC			–5.891*** (0.482)	–5.750*** (0.4944)	–5.706*** (0.4902)
Debt crisis			–2.148*** (0.414)	–2.118*** (0.494)	–2.126*** (0.416)
intercept	1.853*** (0.3147)	1.714*** (0.2673)	4.304*** (0.468)	3.108*** (0.2293)	4.160*** (0.4311)
Rho	0.3193	0.2817	0.240	0.2418	0.2345
Hausman test (p-value)	2.306 (0.680)	10.969 (0.089)	1.414 (0.965)	14.349 (0.0453)	9.3677 (0.3122)
AIC	1 863.03	1 850.60	1 738.22	1 727.97	1 734.86
SIC	1 882.09	1 877.28	1 768.71	1 857.53	1 772.97
Log-likelihood	–926.52	–918.30	–861.11	–829.99	–857.43

Notes: All numerical variables (except dummy variables) – year over year percentage change; Access – bank branches per 100 000 adults; Efficiency – bank return on equity (ROE in %); Stability – bank Z-score; Depth – domestic private credit to GDP (% of GDP); CPI – corruption perception index; GCI – global competitiveness index; Eurozone – dummy variable (if 1 – country is a member of Eurozone, it reflects also the year of joining to Eurozone, 0 – otherwise), GFC – dummy variable (if 1 – global financial crisis in 2008 and 2009, 0 – otherwise); Debt crisis – dummy variable (if 1 – debt crisis from 2010 to 2012, 0 – otherwise); p-values are in parentheses; *, **, *** indicate significant at 10, 5 and 1% level, respectively.
Source: Own estimates

but also an effect of the corruption perception (CPI) and competitiveness in economy (GCI). Model 3 abstracts from the control variables CPI and GCI and examines an impact of the Eurozone membership and it also takes into account an impact of economic and debt crisis. Model 5 examines the impact of all explanatory, control and also dummy variables. Model 4 differs from Model 5 by abstracting the impact of the Eurozone membership.

Except for Model 1, indicators access and efficiency demonstrated a significant positive impact on economic growth within the whole period. Surprisingly, indicators stability and depth are statistically insignificant at all significance levels in all regression models. With the extension of model by the corruption perception index and the global competitiveness index the levels of significance of financial structure indicators did not change. The model supported the statement that diminishing corruption has a positive impact on economic growth. We can also say that the financial structure in terms of efficiency and access has a significant impact on the economic growth even when taking into account the economic and debt crisis. Membership in the Eurozone had a significantly negative impact on all levels of significance

Table 6 Regression results over the 2004–2009 period

Explanatory variable	Explained variable				
	GDP per capita				
	Model 1	Model 2	Model 3	Model 4	Model 5
	Random effect	Random effect	Random effect	Fixed effect	Random effect
Access	0.073* (0.0387)	0.059 (0.0381)	0.038 (0.287)	0.029 (0.0288)	0.029 (0.0286)
Efficiency	0.0003 (0.0002)	0.0005 (0.0003)	0.0006*** (0.0002)	0.0007*** (0.0002)	0.0006*** (0.0002)
Stability	0.00001 (0.0003)	0.00002 (0.0003)	0.0005* (0.0003)	0.0004* (0.0002)	0.0004* (0.0003)
Depth	0.213*** (0.0493)	0.178*** (0.0503)	0.100** (0.0397)	0.110*** (0.0385)	0.091** (0.0399)
CPI		0.237*** (0.0784)		0.145** (0.060)	0.126** (0.0604)
GCI		0.079 (0.0934)		0.071 (0.0702)	0.082 (0.0701)
Eurozone			–1.298* (0.671)		–1.098 (0.6729)
GFC			–6.733*** (0.666)	–6.470*** (0.666)	–6.480*** (0.662)
Debt crisis	-	-	-	-	-
intercept	–0.437 (0.5978)	–0.206 (0.5831)	3.897*** (0.7260)	2.983*** (0.5478)	3.745*** (0.7171)
Rho	0.310	0.254	–0.057	–0.048	–0.047
Hausman test (p-value)	2.849 (0.583)	2.801 (0.833)	2.284 (0.892)	4.249 (0.751)	2.710 (0.951)
AIC	765.92	758.99	688.92	686.82	685.99
SIC	780.18	778.95	708.88	709.64	711.66
Log-likelihood	–377.96	–372.49	–337.46	–335.41	–333.99

Notes: All numerical variables (except dummy variables) – year over year percentage change; Access – bank branches per 100 000 adults; Efficiency – bank return on equity (ROE in %); Stability – bank Z-score; Depth – domestic private credit to GDP (% of GDP); CPI – corruption perception index; GCI – global competitiveness index; Eurozone – dummy variable (if 1 – country is a member of Eurozone, 0 – otherwise), GFC – dummy variable (if 1 – global financial crisis in 2008 and 2009, 0 – otherwise); Debt crisis – dummy variable (if 1 – debt crisis from 2010 to 2012, 0 – otherwise); p-values are in parentheses; *, **, *** indicate significant at 10, 5 and 1% level, respectively.
Source: Own estimates

in the models for the whole period examined. As expected, economic and debt crisis had a significantly negative impact on the economic growth.

In contrast to the whole analysed period (Table 5), the indicator depth demonstrated a significant positive impact on economic growth in the pre-crisis period. It is evident, that the financial depth indicator had a strong positive impact in the pre-crisis period, but after the financial crisis its impact on the economic growth changes to negative one (Table 7). Adding the dummy variable, which includes the economic crisis, the impact of indicators has changed only slightly. The financial structure stability indicator in the extended models 3, 4, and 5 came out at the level α = 0.1 as statistically significant with a positive impact on economic growth, which met our assumptions.

Table 7 Regression results over the 2010–2017 period

Explanatory variable	Explained variable				
	GDP per capita				
	Model 1	Model 2	Model 3	Model 4	Model 5
	Random effect	Random effect	Random effect	Fixed effect	Random effect
Access	0.0168 (0.0168)	0.0198 (0.0169)	0.0223 (0.0164)	0.0247 (0.0165)	0.025 (0.0166)
Efficiency	–0.0011** (0.0005)	–0.0011** (0.0005)	–0.001** (0.0005)	–0.001** (0.0005)	–0.001** (0.00049)
Stability	–0.0004*** (0.0001)	–0.0004*** (0.0001)	–0.0003** (0.0001)	–0.0003** (0.00013)	–0.0003** (0.00013)
Depth	–0.2884*** (0.0295)	–0.2909*** (0.0298)	–0.2781*** (0.0292)	–0.2841*** (0.0293)	–0.2818*** (0.0296)
CPI		–0.0474 (0.0316)		–0.0438 (0.0307)	–0.0442 (0.0308)
GCI		0.1145* (0.0689)		0.0894 (0.0673)	0.0884 (0.0675)
Eurozone			–0.2731 (0.5517)		–0.2928 (0.5434)
GFC	-	-	-	-	-
Debt crisis			–1.0753*** (0.3018)	–0.9992*** (0.2992)	–1.0225*** (0.3032)
intercept	1.1554*** (0.3100)	1.1306*** (0.3038)	1.7927*** (0.5247)	1.5538*** (0.3335)	1.7697*** (0.5200)
Rho	0.1228	0.1060	0.0817	0.0666	0.0682
Hausman test (p-value)	1.4495 (0.8356)	5.292 (0.5068)	5.0211 (0.5411)	5.245 (0.6301)	8.926 (0.3486)
AIC	954.701	954.660	946.787	948.501	947.423
SIC	971.340	977.955	970.083	975.13	977.494
Log-likelihood	–472.350	–470.330	–466.394	–466.252	–464.771

Notes: All numerical variables (except dummy variables) – year over year percentage change; Access – bank branches per 100 000 adults; Efficiency – bank return on equity (ROE in %); Stability – bank Z-score; Depth – domestic private credit to GDP (% of GDP); CPI – corruption perception index; GCI – global competitiveness index; Eurozone – dummy variable (if 1 – country is a member of Eurozone, 0 – otherwise), GFC – dummy variable (if 1 – global financial crisis in 2008 and 2009, 0 – otherwise); Debt Crisis – dummy variable (if 1 – debt crisis from 2010 to 2012, 0 – otherwise); p-values are in parentheses; *, **, *** indicate significant at 10, 5 and 1% level, respectively.
Source: Own estimates

When examining the impact of determinants on economic growth after the economic crisis, several significant changes have occurred. The indicators of efficiency, stability and depth came out with negative impact on economic growth in the given period. The negative impact of the domestic credit to private sector was expected based on the studies dealing with finance-growth nexus Ayadi, Emrah, Sami and Willem (2013), Barradas (2018), Arcand, Berkes and Panizza (2015). They reveal similar outcomes and stress importance of regulation in finance. We assume that from short term perspective credits influence the economic growth positively and their expansion supports the economy, however, from long term perspective, and as the economic cycle accelerates, they can become a burden to economic growth, which is also supported by our results. It should be noticed that our indicator of depth included traditional credits to private sector over gross domestic product provided by regulated banking sector. To support economic growth after a crises more sophisticated forms such as crowdfunding might be used. A surprising result is the negative impact of the remaining two indicators of efficiency and stability, for which, on the contrary, we expected a positive impact. In case of the efficiency indicator, declines in interest margins, which reached negative values in the post-crisis period, have a strong adverse impact on the economic growth. Results of our research reveal that economic growth might be driven by non-banking institutions that are not regulated by the central banks. Our results indicate that the Eurozone membership does not contribute to the economic growth during the crises times. This result is not unlikely since individual countries joined in monetary union might miss autonomous monetary and exchange policies to fix their specific problems. The study results suggest a major role for the governments of the Eurozone member countries in designing appropriate macroeconomic policies. They highlight importance of the balanced budget rules recently adopted in the Eurozone. The fiscal rule framework needs to be more effective in reducing high levels of indebtedness in some member countries. Reducing the public debt across euro area countries would allow to set up a common macroeconomic stabilisation function and this, in turn, would help to overcome deep economic crises.

CONCLUSION

This paper contributes to the existing literature about relationship between financial system development and economic growth in European Union countries with new results based on a larger data set and the longer time span 2004–2017. Among determinants of economic growth in the European Union countries it incorporates not only indicators for financial access, efficiency, stability, and depth of the bank-oriented financial sector but it also includes additional topical control variables: Eurozone membership, corruption perception and competitiveness of countries. Splitting the time span allows us to examine whether the chosen determinants of economic growth are affected by the occurrence of financial and debt crises. We apply a panel data GLS approach and estimate results by five models: Model 1 takes into consideration an impact of main financial indicators to assess access, efficiency, stability, and depth of the bank-oriented financial systems on economic growth. Model 2 analyses not only an impact of financial indicators but also an effect of the corruption perception (CPI) and competitiveness in economy (GCI). Model 3 abstracts from the control variables CPI and GCI and examines an impact of the Eurozone membership and economic and debt crisis. Model 5 examines the impact of all variables on economic growth. Model 4 differs from Model 5 by abstracting the impact of the Eurozone membership.

As stated in the paper, over the period 2004–2017 indicators access and efficiency demonstrated a significant positive impact on economic growth while indicators stability and depth are statistically insignificant at all significance levels in all regression models. In contrast to the whole analysed period, the indicator depth revealed a significant positive impact on economic growth in the pre-crisis period and a strong negative impact after financial crises and during debt crises.

Our results also indicate that the Eurozone membership does not contribute to the economic growth during the crises times. This result is not unlikely since individual countries joined in monetary union might miss autonomous monetary and exchange policies to fix their specific problems. The study results suggest a major role for the governments of the Eurozone member countries in designing appropriate macroeconomic policies. They highlight importance of the balanced budget rules recently adopted in the Eurozone. The fiscal rule framework needs to be more effective in reducing high levels of indebtedness in some member countries. Reducing the public debt across euro area countries would allow to set up a common macroeconomic stabilisation function and this, in turn, would help to overcome deep economic crises. The framework should benefit from the long-lasting experience of the fiscal rule frameworks adopted in Switzerland or USA though there are important differences in all three areas. The reflection of the fiscal compact in national rules should ultimately help to increase the resilience of the Eurozone.

Our future research will be in line with the European Green Deal adopted by the European Commission as a set of proposals to make the EU's climate, energy, transport, and taxation policies fit for reducing net greenhouse gas emissions by at least 55% by 2030, compared to 1990 levels. We plan to extend the financial sector determinants by a wider range of control variables, use static and dynamic panel data and employ advanced estimation techniques.

ACKNOWLEDGEMENT

The research done in this paper is associated with the grant scheme VEGA 1/0608/19 The finance-growth nexus in the Visegrad Group countries and VEGA 1/0785/19 Construction and analysis of dynamic nonlinear models of macroeconomic processes.

References

AGAPOVA, A., VISHWASRAO, S. (2020). Financial sector foreign aid and financial intermediation. *International Review of Financial Analysis*, 72: 1–18. <https://doi.org./ 10.1016/j.irfa.2020.101589>.

ARCAND, J. L., BERKES, E., PANIZZA, U. (2015). Too much finance? *Journal of Economic Growth*, 20: 105–148. <https://doi.org./ 10.1007/s10887-015-9115-2>.

ARMIN, S., IBRAHIM, S., AZMAN-SAINI, W. N. W. (2012). The Impact of European Monetary Union on Finance-Growth Nexus. *Transition Studied Review*, 19: 347–356. <https://doi.org./ 10.1007/s11300-012-0249-z>.

AYADI, R., ARBAK, E., BEN-NACEUR, S., GROEN, P. W. (2013). Financial Development, Bank Efficiency and Economic Growth across the Mediterranean. *Mediterranean Prospects*, 30 (March): 1–17. <https://doi.org./10.1007/978-3-319-11122-3_14>.

BARRADAS, R. (2018). The Finance-Growth Nexus in the Age of Financialisation: An Empirical Reassessment for the European Union Countries [online]. *Working Paper*. [cit. 12.6.2021]. <https://repositorio.iscte-iul.pt/bitstream/10071/16985/1/DINAMIA_WP_2018-07.pdf>.

BECK, R., GEORGIADIS, G., STRAUB, R. (2014). The finance and growth nexus revisited. *Economics Letters*, 124: 382–385. <https://doi.org./ 10.1016/j.econlet.2014.06.024>.

BEN-DAVID, D. (1993). Equalizing Exchange: Trade Liberalization and Income Convergence. *Quarterly Journal of Economics*, 108: 653–679. <https://doi.org/10.2307/2118404>.

CAMBACORTA, L., YANG, J., TSATSARONIS, K. (2014). Financial structure and growth. *BIS Quarterly Review*, 1: 21–35.

CAPORALE, G. M., RAULT, CH., SOVA, R., SOVA, A. (2009). *Financial Development and Economic Growth: Evidence from Ten New EU Members* [online]. Discussion Papers, Berlin. [cit. 14.6.2021]. <https://ideas.repec.org/p/diw/diwwpp/dp940.html>.

CHECHERITA-WESTPHAL, C., ROTHER, P. (2012). The impact of high government debt on economic growth and its channels: an empirical investigation for the euro area. *European Economic Review*, 56: 1392–1405. <https://doi.org/10.1016/j.euroecorev.2012.06.007>.

CHU, K. L. (2020). Financial structure and economic growth nexus revisited. *Borsa Istanbul Review*, 20: 24–36. <https://doi.org/10.1016/j.bir.2019.08.003>.

ČIHÁK, M., DEMIRGÜÇ-KUNT, A., FEYEN, E., LEVINE, R. (2013). Benchmarking Financial Systems around the world. *Policy Research Working Paper*, 6175: 2–56.

DEMIRGUÇ-KUNT. A., DETRAGIACHE, E. (1998). Financial Liberalization and Financial Fragility. *World Bank and Research Department of IMF*, 83 (June): 3–36.

DEMIRGÜÇ-KUNT, A., FEYEN, E., LEVINE, R. (2013). The evolving importance of banks and securities markets. *World Bank Economic Review*, 27: 476–490. <https://doi.org/10.2139/ssrn.2083622>.

DOLLAR, D. (1992). Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976–85. *Economic Development and Cultural Change*, 40: 523–544. <https://doi.org/10.1086/451959>.
FISCHER, K. P., CHENARD, M. (1997). Financial Liberalization Causes Banking System Fragility. *Working Paper*, 97 (June): 1–32.
GEORGANTOPOLOUS, A., TSAMIS, A., AGORAKI, K., M. (2015). The Euro-adoption effect on the bank, market, and growth nexus: New evidence from EU panels. *The Journal of Economic Asymmetries*, 12: 41–51. <https://doi.org/10.1016/j.jeca.2015.01.002>.
GÓMEZ-PUIG, M., SOSVILLA-RIVERO, S. (2015). The causal relationship between debt and growth in EMU countries. *Journal of Policy Modeling*, 37: 974–989. <https://doi.org/10.1016/j.jpolmod.2015.09.004>.
KARAGIANNIS, S., KVEDARAS, V. (2016). *Financial development and economic growth: a European perspective.* Publications Office of the European Union: 2–67. <https://doi.org/ 10.2791/53648>.
LUCAS, R. E. (2009). Ideas and Growth. *Economica*, 76: 1–19. <https://doi.org/10.1111/j.1468-0335.2008.00748.x>.
MADSEN, J., ANG, J. (2016). Finance-led growth in the OECD since the nineteenth century: How does financial development transmit to growth? *Review of Economics and Statistics*, 98: 552–572. <https://doi.org/10.1162/REST_a_00513>.
MAZUREK, J. (2017). On determinants of the economic growth of European countries during 2005–2015. *Comparative Economic Research*, 20: 21–33. <https://doi.org/10.1515/cer-2017-0010>.
NAGHSHPOUR, S. (2019). Determinants of economic growth rates of the newly formed countries of the former Soviet Union and Eastern Bloc. *Economic Policy in Emerging Economies*, 12: 425–442. <https://doi.org/10.1504/IJEPEE.2019.104638>.
NYASHA, S., ODHIAMBO, N. M. (2017). Financial Development and Economic Growth Nexus: a Revisionist Approach. *Economic Notes: Review of Banking, Finance and Monetary Economics*, 47: 223–229. <https://doi.org/10.1111/ecno.12101>.
OECD Economic Outlook. (2021). Paris: OECD Publishing, 1(109). ISBN 978-92-64-81691-6.
PEGKAS, P., STAIKOURAS, CH., TSAMADIAS, C. (2020). On the determinants of economic growth: Empirical evidence from the Eurozone countries. *International Area Studies Review*, 23: 210–229. <https://doi.org/10.1177/2233865920912588>.
PERERA, N., PAUDEL, R. (2009). Financial development and economic growth in Sri Lanka. *Applied Econometrics and International Development*, 9: 157–164.
PINAR, A., DAMAR, E. (2006). Financial Sector Deepening and Economic Growth. *Working Paper*, 19.
PŘÍVARA, A., TRNOVSKÝ, K. (2021). The impact of remittances on household savings in the Baltics. *Investment Management and Financial Innovations*, 18: 335–345. <http://dx.doi.org/10.21511/imfi.18(1).2021.27>.
PROCHNIAK, M. (2011). Determinants of economic growth in Central and Eastern Europe: the global crisis perspective. *Post-Communist Economies*, 23: 449–468. <https://doi.org/10.1080/14631377.2011.622566>.
SHITTU, O. W., ISMAIL, N. A., LATIFF, A. R. A. (2020). Modeling external debt – growth nexus: how relevance is governance? *Journal of financial crime*, 27: 1323–1340. <https://doi.org/10.1108/AJEMS-07-2017-0171>.
STOLBOVÁ, V., BATTISTON, S., NAPOLETANO, M., ROVENTINI, A. (2017). Financialization of Europe: a comparative perspective. *ISI Growth Working Paper*, 22.
TADESSE, S. (2002). Financial architecture and economic performance: International evidence. *Journal of Financial Intermediation*, 11: 429–454. <https://doi.org/10.1006/jfin.2002.0352>.
ZHANG, K., WANG, Y., REN, S. (2021). An Empirical Cross Country Analysis on the Determinants of Economic Growth. *Design Engineering*, 1: 38–52.
WACZIARG. R., WELCH, K. (2008). Trade Liberalization and Growth: New Evidence. *World Bank Economic Review*, 22: 187–231. <https://doi.org/10.1093/wber/lhn007>.
WU, S. J., WEI, Y. (2002). *The Life and Death Implications of Globalizations.* [online]. Washington, D. C.: International Monetary Fund. [cit. 12.6.2021]. <https://users.nber.org/~confer/2002/iasef02/wei.pdf>.

A Critical View on Pension Savings in Slovakia

Jana Špirková[1] | *Matej Bel University, Banská Bystrica, Slovakia*
Martin Dobrovič | *Matej Bel University, Banská Bystrica, Slovakia*
Miroslava Vinczeová | *Matej Bel University, Banská Bystrica, Slovakia*

In almost all countries around the world, pension systems are based on several pillars. This is also the case of Slovakia with its three-pillar pension system. The paper presents a case study underlying the risks that can seriously affect the amount of future pensions. The case study clearly indicates that current pensions in Slovakia paid under all three pillars do not correspond with the expectations from the implementation of the three-pillar pension system. The aim of the paper is to the risks that can seriously affect the amount of future pensions. Our own contribution is the determination of the amount of pension for a specific pensioner specified in the presented case study. Within the saving phase of pension contributions the development of investment fund returns, the amount of future pensioner´s contributions, as well as administrative costs are analyzed on a monthly basis. The payout phase is modelled using actuarial functions applying the mortality tables of Slovakia.

Keywords	DOI	JEL code
Pension annuity, modelling, life expectancy, yield, pillars of pension savings	*https://doi.org/10.54694/stat.2021.11*	*G12, G22, H55, H75*

INTRODUCTION

Since almost all economically active people are interested in their future retirement, a great number of economists and mathematicians are engaged in research into the pension systems around the world. Over the years, there has been general recognition that serious economic problems have plagued and continue to plague many aged persons in almost all countries. We will mention at least some of the contributions we worked with in our study. For example, the paper Schulz and Carrin (1972) examines the personal savings rates required to provide an average worker with adequate retirement funds, the influence of various definitions of adequacy, inflation, and economic growth on the magnitude of the retirement preparation.

After 2000, most defined benefit pension plans saw a decline in funding ratios, mainly due to lower asset prices. Weller and Wenger (2009) identify four indicators for the reckless investment behavior of pension plans: no portfolio balance, conflicts of interest for employers, conflicts of interest for managers, and failure to implement best investment practices. Rydqvist et al. (2014) make a very interesting finding which contributes to policy debates on effective taxation and to financial economics research on the long-term effects of taxation on corporate finance and asset prices. Konicz and Mulvey (2016)

[1] Matej Bel University, Faculty of Economics, Tajovského 10, 975 90 Banská Bystrica, Slovakia. Corresponding author: e-mail: jana.spirkova@umb.sk, phone: (+421)908901464.

provide some guidelines to individuals with defined contribution pension plans on how to manage pension savings both before and after retirement. Albrecher et al. (2016) offer a summary of the main topics and findings from the Swiss Risk and Insurance Forum 2015. Dolls et al. (2018) also suggest how can retirement savings be increased, and they emphasize that the increase in tax deductible items increases the increase in savings on pension accounts.

Kieren and Weber (2021) suggest that policy makers should consider offering combined sequential and annuity solutions. Beneficiaries who refuse a full annuity could retain some of their accumulated wealth while gaining protection against the risk of longevity.

Modelling of pension saving and insurance in Slovakia is presented in several papers. The most significant of them include the works Melicherčík et al. (2015a, 2015b). Moreover, we were inspired by psychological assessment of this issue in Bačová and Kostovičová (2018).

The pension system in Slovakia is based on the three pillars. The first one is a pay-as-you-go pillar, it is a defined benefit scheme regulated by the Act no. 461/2003 Coll. on Social Insurance.[2] Compulsory pension insurance in the Slovak Republic is administered by the Social Insurance Agency. Persons insured only in the first pillar pay contributions only to the Social Insurance Agency and, at old age, they will be provided with the income only by the Social Insurance Agency in accordance with the number of years of service, income during their working life and current pension amount in a given year. The second pillar is partially voluntary and is a funded scheme representing appreciation in the funds of the pension management companies under the Act no. 43/2004 Coll. on Old-age Pension Savings. The third pillar is entirely established on a voluntary basis and the terms and conditions of its operation are laid down by the Act no. 650/2004 Coll. on Supplementary Pension Savings. Its main advantage is obtaining of the supplementary pension for an employee and the tax benefit granted to the employer who remits the contributions for an employee in a certain amount.

In the Slovak Republic, the second pension pillar was established on 1 January 2005. Its aim was not to ruin the public finances, ensure a reasonable pension amount and appreciate savers´ money. The main background to its establishment particularly included the unfavourable demographic trends and jeopardised future solvency of the Social Insurance Agency. This pillar is regulated by the Act no. 43/2004 Coll. on Old-age Pension Savings and have been made an integral part of the pension system in the Slovak Republic.

The amount of social security contributions is 18% of the assessment base, i.e., the average monthly wage. Since its beginning, contributions to the second pillar were 9%, however in 2012, the premium rate for the second pillar was decreased by the amended act to 4% and this was applied until 2016. In the following years, the premium rate was increasing by 0.25 percentage points until 2020, since it has remained at the level of 5% of the assessment base.

In Slovakia, savers´ funds in the second pillar are managed by the pension management companies. Employees choose a management company according to their interest and preferences and they entrust it with appreciation of their funds. According to the Act no. 43/2004 Coll., the pension management companies are obliged to administer one guaranteed bond fund and one non-guaranteed equity fund.

Appreciation obviously does not only depend on the choice of the management company but also on the investment strategy of the fund chosen by the client.

Only insurance or reinsurance companies meeting the requirements set for such activities according to the Act no. 39/2015 Coll. on Insurance and amending and supplementing certain acts can perform insurance and reinsurance activities. In this case, the pension management companies but also supplementary pension insurance companies have to follow this act. They also secure payments of (provisional) pensions resulting from the character of their activities. Pursuant to Section 23 of this act,

2 Links of the Acts are in the References.

the insurance and reinsurance companies are required to manage funds of their clients with prudence. This fact affects the total amount of pension in the payout phase of the pension insurance.

In the Slovak Republic, the third pillar of the pension system was established by the Act no. 650/2004 Coll. on Supplementary Pension Saving and amending and supplementing certain acts of 1 January 2005. This pillar is managed by the supplementary pension companies. Supplementary pension saving is voluntary and the employees themselves decide on the entry into this scheme. The main advantage of the conclusion of the supplementary pension saving contract is a contribution provided by an employer. Information concerning supplementary pension saving is published in the collective agreements of companies as a result of bargaining between the trade unions and employers or an employer agrees on them with the authorised representatives of the employees. The voluntary nature of the employer´s contributions, however, excludes risk at work. In such cases, the employer is obliged to pay contributions to the employees of at least 2% of their assessment base. An employee who performs hazardous work may or may not save in the third pillar.

The third pillar is a defined contribution and funded scheme as well as the second one and, therefore, its aim is not only to save but also appreciate cash resources accumulated in the funds.

According to the Income Tax Act no. 595/2003 Coll., an employee can reduce his/her tax base by EUR 180 which represents a maximum tax allowance related to the contributions paid to the third pillar. The amount of this type of the tax allowance is derived from the amount an employee provably has paid. At the same time, a taxpayer may deduct this tax allowance only when he/she concluded a contract after 31 December 2013 and paid supplementary pension contributions based on the participation agreement and, moreover, he/she has not concluded any other contract in accordance with the Act no. 650/2004 Coll.

This advantage of the tax base reduction may also be used by the employer in relation to contributions to the employees. It differs in particular in terms of the amount of the claimed tax allowance. The employer may contribute up to 6% of the employee´s gross wage and these contributions are exempted from the income tax and social security and healthcare insurance contributions.

This paper presents a case study in which the future amount of pension is modelled for a pensioner born on 1 May 1958 who retired on 1 January 2021 and, obviously, was paying social insurance required by law. He entered the second pillar on 1 January 2005 at the age of 46 and 8 months and on 1 July 2009 he also started to contribute to the third pillar of pension saving. Our future pensioner joined the second pillar at a relatively late age and therefore, the contribution period and opportunities to appreciate his savings were not very favourable. This has translated into the amount of his pension. In the study, we lay emphasis on real appreciation of savings and careful determination of the pension amount which a pensioner will receive.

Our own contribution is the determination of the amount of pension for a particular pensioner specified in the presented case study. Based on the Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services, we apply the unisex life tables of the Slovak Republic of 2019 to model a survival function.

With regards to appreciation of financial resources during the payout phase under the second and third pillars, we use yields modelled by means of the Smith-Wilson yield curve published on the website of the European Insurance and Occupational Pensions Authority, 2021 and AAA-rated bond yields which are traded in the euro area, modelled by means of the Svensson yield curve published on the European Central Bank´s website (European Central Bank, 2021), both of 30 November 2020.

Based on the above assumptions, we set the following hypotheses at the beginning of our work:

H1: The life expectancy of a 62-year-old pensioner is (for 2020) 19.14 years.

If a 62-year-old pensioner specified in our study lives further 19 years, the entire amount saved in the second and third pillars will be returned to him in the form of monthly pension payments.

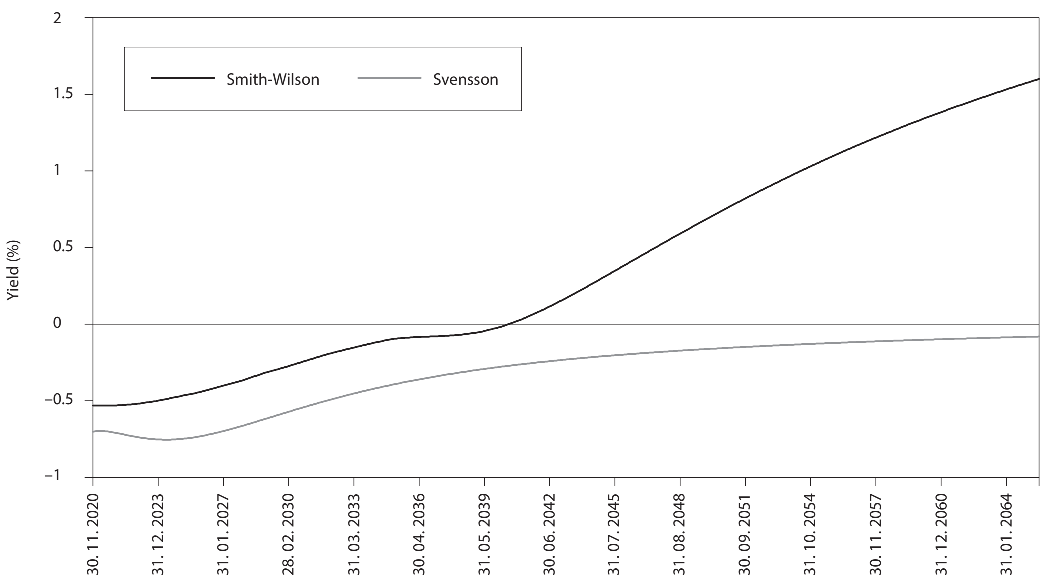
H2: If a 62-year-old pensioner specified in our study receives a pension from all three pillars, his monthly pension will be at least 30% higher than a pension that he would receive only from the first pillar.

The paper is organized as follows. Section 2 – Preliminaries, gives information about basic data related with yields of risk-free bonds traded in the euro area and life tables of the Slovak Republic. It also gives basic actuarial notation of the survival function and probability of death. Section 3 – Savings phase of old-age and supplementary pension savings, describes the development of appreciation in individual funds in more detail. Section 4 – Payout phase of pension savings, offers basic models for determining the amount of monthly pensions from all three pillars of pension savings. Section 5 – Conclusion, offers our thoughts and ideas on how to more accurately model the amount of pensions, but also our attitude as such to the idea of long-term savings in such funds.

1 PRELIMINARIES

The basic building blocks in our modelling are the above-mentioned Smith-Wilson and Svensson yield curves which are shown in Figure 1. Technical notes relating to both yield curves published on the websites of mentioned institutions and make an integral part of our modelling (Technical notes, 2020; Technical documentation – EN, 2021).

Figure 1 Smith-Wilson and Svensson yield curves, business day – November 30, 2020



Source: Own construction using the European Insurance and Occupational Pensions Authority and the European Central Bank (2021)

Since the survival and mortality probabilities represent the basic building block in actuarial modelling, we will remember them:

${}_tp_x$ – the probability that individual at age x survives at least to age $x + t$,

${}_tq_x$ – the probability that individual at age x dies before age $x + t$,

${}_{r|t}q_x$ – the probability that individual at age x survives r years, and then dies in the subsequent t years, that is, between ages $x + r$ and $x + r + t$.

The probability ${}_{r|t}q_x$ can be calculated by formula (Dickson et al., 2009):

$$_{r|t}q_x = {}_rp_x - {}_{r+t}p_x\,. \qquad (1)$$

As our model assumes monthly pension payments while the pensioner is alive to determine monthly mortality probabilities as main variables, we apply the so-called fractional age assumption – for integer x, provided the uniform distribution of deaths in every age interval $[x, x + 1]$, and for $0 \leq s < 1$, assume that $_sq_x = s \times q_x$. For more information see, for example Dickson et al. (2009).

2 SAVINGS PHASE OF OLD-AGE AND SUPPLEMENTARY PENSION SAVINGS

We suppose in our case study that the saver was born on 1 May 1958 and is a man. We also assume that he entered the second pillar on 1 May 2005 at the age of 46 years and 8 months and was saving until 31 December 2020. He retired on 1 January 2021 at the age of 62 years and 8 months. This assumption is determined in accordance with the Act no. 461/2003 Coll. and the Annex 3a prescribing the retirement age. On 1 January 2021, the saver´s period of service is 40 years and 8 months. To determine the amount of contributions paid to the Social Insurance Agency and pension management companies, we build our model on the development of the average monthly salary since 2005.

We assume that the saver also makes contributions to the third pillar and suppose that his employer contributes 2% of the assessment base. This fact is mostly conditional upon the employer contributing the same amount.

Modelling of the respective performances of both funds is set on the first day of the month for the second and third pillars.

2.1 Valuation in funds and investment strategy of the second pillar

The level of appreciation and saved amount depend on the development of the pension unit value in the fund. Our model includes funds appreciation for the period whose performance is already known. The Act no. 43/2004 Coll. also determines an investment strategy regulating a transfer of net assets from another fund to the bond fund. Section 92 of the Act provides for the saver´s net assets transfer according to his age.

Appreciation of savings depends on the pension unit value and time of purchase. Due to the availability of data on the websites of pension management companies, our model shows saving in the equity and bond funds of VÚB Generali d. s. s., a. s.

To determine fund performance, we use the indicator of the cumulative fund return. This indicator reflects the case of a full amount of one-off investment at the beginning of the period. However, saving in the second pillar is regular, thus the contributions are credited continually which also affects appreciation of the funds.

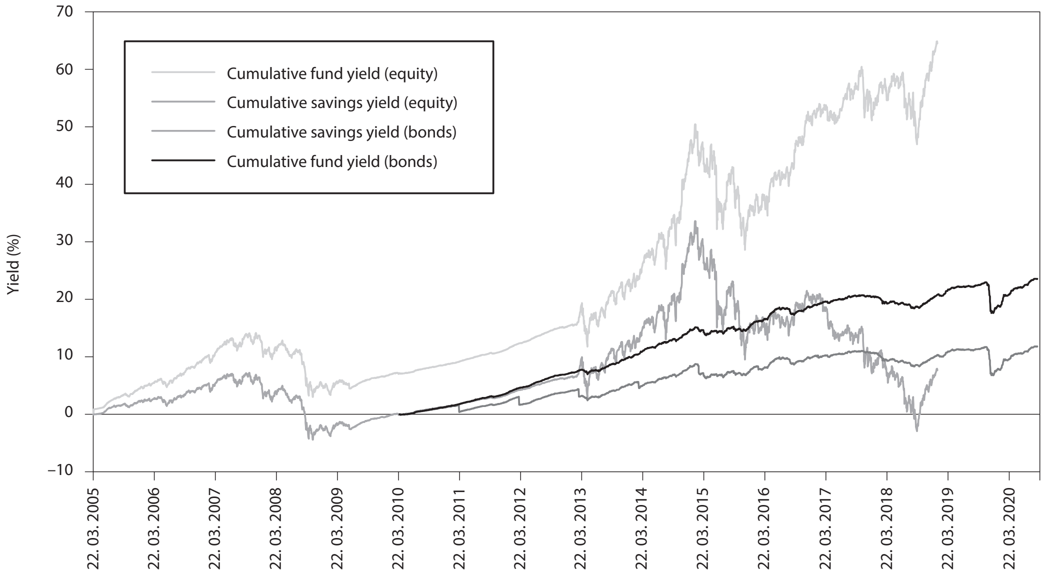
Overall yield of savings is significantly affected mainly by fluctuations of the fund appreciation. Compared to the cumulative fund return, the real savings appreciation is substantially lower, in some phases of saving it is even negative. This situation is illustrated in Figure 2 defining appreciation of both funds.

The cumulative savings return in the equity fund was 7.65% on 1 May 2019 but as a result of the assets transfer, this figure does not reflect real appreciation. If the charges are also taken into consideration, the cumulative fund return of 9.66% would be achieved.

The cumulative fund return represents appreciation based on the pension unit value in the fund, while the cumulative savings return also takes the amount of paid contributions and the saver´s account value into account. The data are modelled for the period of saving in the bond fund from 1 May 2010 to 31 December 2020 without accounting for charges.

As Figure 2 shows, the return of the fund is higher than that one of savings which mainly results from the fund investment strategy as individual contributions are paid regularly. It also implies that if the total amount of contributions was paid as a one-off payment, the saver would achieve appreciation

Figure 2 Equity and bonds fund cumulative yields in the second pillar (free of charges)



Source: Own construction using data from the web page of the Generali, d.s.s. (2020)

equalling to the cumulative fund return of 23.38%, without taking account of charges. The return of the bond fund in our model, in the case of regular contributions was 11.29% on 31 December 2020 and if its value also took charges into account, the final cumulative savings return would only be 9.12%, see Figure 2. The cumulative savings return is accompanied by larger fluctuations compared to the cumulative fund return since always on 1 May in a particular year, a part of the account value of the equity fund is deposited to the fund following the set investment strategy. Charges mentioned in connection with funds are stipulated in Section 63 of the Act no. 43/2004 Coll. Pursuant to this Act, the pension management companies usually invoice maximum charges.

The total contributions to the second pillar made by our saver bring the total to EUR 9 969.22. If he saved in the funds of VÚB Generali d. s. s., the resulting amount saved by means of an investment strategy would be EUR 11 300.58 after their appreciation in the funds and including all charges.

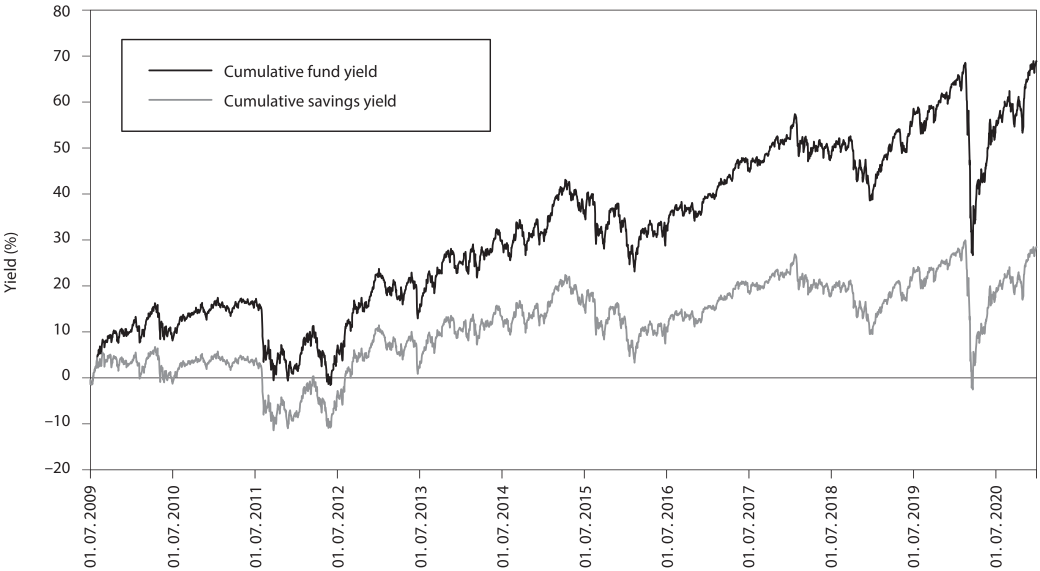
2.2 Valuation in funds and investment strategy of the third pillarr

In this part of the paper we focus on appreciation in the funds of the company NN Tatry – Sympatia, d. d. s., a. s. which has been operating in the Slovak market since 2006 (Company NN Tatry – Sympatia, 2020). Cumulative appreciation of the growth contribution pension fund amounted up to 68.92% from 2009 to the end of 2020. This situation is illustrated in Figure 3.

The growth fund was repeatedly achieving significantly higher appreciation compared to the savings in the fund. This is primarily due to regular investments and fluctuations in financial markets. The cumulative fund return reached up to 28.70%, excluding charges in the fund.

The charges in the contribution fund influenced the level of savings appreciation more than in the case of the second pillar. It is mainly due to higher rates which are the basis for the calculation of those charges. The saver´s return achieved after including the charges is 11.53% representing substantially lower savings appreciation than the fund potential. Our saver was saving in the third pillar from 2009

Figure 3 Equity fund cumulative yields in the third pillar (free of charges)



Source: Own construction using data from the web page of the NN Tatry – Sympatia, d.s.s. (2020)

to 2020. As a result of the amended Act no. 650/2004 Coll., the amount of a charge for fund management and appreciation was changed several times during this period. All amendments of charging policy are reflected in our model. The amount of the charge for the whole saving period was EUR 526.61 for fund management and EUR 202.94 for appreciation.

Thus, if we saved in the growth fund of NN Tatry – Sympatia, d. d. s., the resulting saved sum would amount to EUR 5 548.35 after appreciation in the fund including all charges.

3 PAYOUT PHASE OF PENSION SAVINGS

We recall that in our case study, we consider the saver who retired on 1 January 2021 at the age of 62 years and 8 months. He was pension insured in the first pillar for 40 years and 8 months. He started to save in the second pillar at the age of 46 years and 8 months and was contributing to it for 15 years. He was contributing to the third pillar since 2009, for 10 years and 6 months. The payout phase of our saver´s pension insurance consists of the payout either only from the first pillar or from the first as well as second pillars or also from all three pillars.

In this part we also use the term insurance as we model the amount of pensions in the second and third pillars by means of the saved amount representing a single premium for a lifetime monthly arrear pension.

3.1 Permanent pension from the first pillar

The amount of the permanent[3] monthly pension MP_1 paid from the first pillar is determined in accordance with the Act no. 461/2003 Coll. based on the relation:

[3] Life time or whole life pension.

$$MP_1 = APSP \times PIP \times CPV\,, \qquad (2)$$

where:

$APSP$ – average personal salary point,

PIP – pension insurance period,

CPV – current pension value, which is determined based on the growth of the average wage.

The average personal wage point $APSP$ is determined as a proportion of the amount of personal wage points calculated over particular calendar years for the period of reference and the number of years of the pension insurance period. The personal wage point is a proportion of a personal assessment base and general assessment base. The general assessment base is 12 times the average monthly wage in the Slovak economy for a given year determined by the Statistical Office of the Slovak Republic. Since we assume that our pensioner was receiving the average wage throughout his active working life, the average personal point is $APSP = 1$.

The current pension value CPV applicable as of 31 December of a calendar year is changed as of 1 January of the following year by means of the index representing a proportion of the average wage determined for the third quarter of the preceding calendar year and average wage determined for the third quarter of the calendar year two years preceding the calendar year since which the pension value has been adjusted. The sum of the pension value identified in this way is always applied from 1 January to 31 December of a calendar year. The current pension value applied to calculate the pension benefit is a pension value valid in time of entitlement to draw a pension. It reflects a monetary value of the average personal wage point. In 2021, its value is $CPV = 14.2107$ EUR. For more information, see Social Insurance Agency, 2021.

The pension insurance period PIP is 40 years and 8 months, i.e., $PIP = 40.6667$ years.

Based on Formula (2), our saver´s pension amount paid from the first pillar is $MP_1 = 573.51$ EUR.

3.2 Permanent pension from the second pillar

As we model the permanent pension paid from the second pillar (but also from the third pillar) by means of the actuarial functions, we first recall basic symbols and formulas:

- S – accumulated sum in the Old-Age Pension Saving Scheme, gross single premium;
- $R_{t^*}(z)$ – Smith-Wilson yield curve or Svensson yield curve, respectively; continuous compounding yield in % p.a. from business day t^*;
- $P(z) = \exp\left(-\frac{R_{t^*}(z)}{100\%} \times z\right)$ – discounting factor with continuous compounding yield $R_{t^*}(z)$;
- x – retirement age;
- ω – maximum age to which a person can live to see (regarding used life tables $\omega = 105$ years;
- t – number of months in the model, i.e., $t = 1, 2, \dots, \omega - x + 1$;
- α – initial costs as an absolute amount in monetary units independent on an accumulated sum;
- β – administrative costs as a % p.a. from yearly annuity;
- γ – collection costs as a % p.a. from yearly annuity.

This product contains permanent monthly annuity and payment of a lump sum equal to not yet paid monthly annuities in the case of the beneficiary death during the period of the first seven years of pension payment (Section 32, Paragraph 2 of the Act no. 43/2004 Coll.).

Monthly pension annuity MP_2 of this product is given by formula:

$$MP_2 = \frac{S \times \left(1 - \frac{\alpha}{100\%}\right)}{12 \times \left(a_x^{(12)} \times \left(1 + \frac{\beta}{100\%} + \frac{\gamma}{100\%}\right) + (MA)_{x:7}^{(12)}\right)}, \qquad (3)$$

where:

$$a_x^{(12)} = \sum_{t=1}^{12\times(\omega-x+1)} \frac{1}{12} \times {}_{\frac{t}{12}}p_x \times P\left(\frac{t}{12}\right) \tag{4}$$

is expected present value of whole life benefits in advance in the amount of 1⁄12 of monetary units (m.u.), 12-times per year, conditional upon the beneficiary life, and:

$$(MA)_{x:7}^{(12)} = \sum_{t=1}^{83} \frac{84-t}{12} \times {}_{\frac{t}{12}|\frac{1}{12}}q_x \times P\left(\frac{t+1}{12}\right) \tag{5}$$

is the expected present value of the sum of not yet paid monthly annuities in the case of the beneficiary death during the period of the first seven years of pension payment. For more information, see Špirková, Szűcs and Kollár (2019).

One-off costs of the insurance company – α in the amount of EUR 200, administration costs β as 1% of the yearly pension and collection costs γ in the amount of 0.2% of the yearly pension are also embedded in Formula (3).

Provided that our future pensioner chooses the funds of VÚB Generali d. d. s., a. s., he would receive the pension from the second pillar amounting to EUR 44.63 when calculating the returns by means of the Swensson yield curve. When applying the Smith-Wilson curve, this amount would be EUR 46.57.

3.3 Permanent pension from the third pillar

The so-called monthly benefit paid from the third pillar of pension saving will be calculated on the basis of the relation:

$$MP_3 = \frac{S \times \left(1-\frac{\alpha}{100\%}\right)}{12 \times a_x^{(12)} \times \left(1+\frac{\beta}{100\%}+\frac{\gamma}{100\%}\right)}, \tag{6}$$

where relations and variables have the same meaning as in the relation for the calculation of the pension paid from the second pillar.

However, here it is a monthly permanent arrear pension the saver will be paid while he is alive. After his death, the survivors will receive neither any pension nor other benefits.

When calculating the same costs as in the calculation of the pension paid from the second pillar and using the Svensson curve, our pensioner would receive the pension from the third pillar amounting to EUR 21.91 and, when applying the Smith-Wilson curve, the pension amount would rise to EUR 22.88. For comparison, see Tables 1 and 2.

3.4 Pension form the first and second pillars simultaneously

If our saver received the pension from the first pillar as well as the second one, he would be paid the decreased pension from the first pillar as his pension insurance period will be reduced and recalculated according to the formula:

$$PIP^* = 365 \text{ days} \times \sum_{i=1}^{16} \left(1 - \frac{c_i\%}{22.75\ \%}\right), \tag{7}$$

where c_i is an amount of contributions paid to the second pillar (in %) in particular years and the value (18.00 + 4.75)% = 22.75% represents the total contributions to the Social Insurance Agency and contributions to the Reserve Fund of Solidarity in %.

Our saver´s pension insurance period was reduced by 4.6081 years and thus, his pension from the first pillar (when also receiving the pension from the second pillar) will be EUR 512.42.

Tables 1 and 2 clearly demonstrate the pension our pensioner may expect.

Table 1 Monthly pensions from individual pillars using the Svensson yield curve, business day – November 30, 2020

	Monthly pension annuity (EUR)
Only the first pillar	573.51
First and second pillars	512.42 + 44.63 = 557.05
First, second and third pillars	512.42 + 44.63 + 21.91 = 578.96

Source: Own construction

Table 2 Monthly pensions from individual pillars using the Smith-Wilson yield curve, business day – November 30, 2020

	Monthly pension annuity (EUR)
Only the first pillar	573.51
First and second pillars	512.42 + 46.57 = 558.99
First, second and third pillars	512.42 + 46.57 + 28.88 = 581.87

Source: Own construction

CONCLUSION

The saving as well as payout phase of pension insurance is affected by many factors. If our pensioner decided not to join the second pillar in the past, he would currently receive his monthly pension only from the first pillar in the amount of EUR 573.51. However, if he also joined the second pillar, as assumed in our case study, his so-called pension insurance period in the first pillar would be reduced by more than 4 years and 7 months which resulted in the decline of his pension from the first pillar to EUR 512.42. It would be reasonable when the monthly pension paid from the second pillar would be high enough, when added to the pension paid from the first pillar, to sufficiently exceed the monthly pension only paid from the first pillar. Our pensioner saved EUR 11 300.58 in the second pillar which is substantially less than the overall fund potential. As results from our model, his total pension paid from both pillars would only amount to EUR 558.99. Admittedly, this situation is because our pensioner joined the second pillar at a relatively late age when it is generally no longer even recommended to join it. Furthermore, regular investments of low sums and the situation in financial markets could not guarantee high appreciation.

In our case study, we also assume that our future pensioner was saving in the voluntary third pillar 2% of his assessment base and the same amount was also contributed by his employer. Therefore, if he earned the average wage and saved in this pillar in the equity fund of the NN Tatry – Sympatia d. d. s., a. s., his savings appreciation would reach 11.53%, counting for EUR 5 548.35. In view of the relatively high charges, appreciation of this investment is significantly lower as the fund potential reached 68.92%. To be closer to the fund potential, at least the mixed investment strategy would have to be chosen. The problem of one-off and regular investments could be solved for example by state interference when the saver would be provided with a repayable grant at the beginning of saving which would be repaid during a half of the pension insurance period and he would regularly save during the second half as up to now.

Thus, the pension paid from the third pillar could, to a certain extent, match the pension to the amount of EUR 581.87. We can see that voluntary contributions made by the employee and employer are a part of this adjustment representing the opportunity cost. Both parties could invest this money in a much more efficient way compared with the voluntary savings.

Assessment of formulated hypotheses

To H1: Our pensioner saved in the second and third pillars sum of EUR 16 848.93. We calculated the amount of the monthly pension at EUR 66.54. This means that the amount saved would be exhausted after more than 21 years. Therefore, we reject hypothesis H1.

To H2: As can be seen from Tables 1 and 2, our pensioner will not receive the pension he would probably expect. In our modelling, the fact that if economically active people save in all three pillars, their pension will be "decent", has not been proven. Our quite high expected percentage, set out in hypothesis H2, has not been demonstrated. Therefore, we reject hypothesis H2.

In this paper, we have used the Life tables of the Slovak Republic, 2019, published on the website of the Statistical Office of the Slovak Republic, however, the insurance companies model pensions by means of adjusted tables since – based on their expertise – people who are healthier, have better genetic background, are relatively richer join the second and third pillars. This results in a significantly higher probability of their older-age survival (Kainhofer, Predota and Schmock, 2021). The effect of such modelling will be that the pension amounts can even be substantially lower than we present.

It is recognised that profit testing of such products including individual financial flows, not only those modelled in this paper, is also an integral part of modelling. Modelling of financial flows including inflation, international accounting standards and other items entering the accounts of insurance companies is also essential to determine the pension amount correctly. This is modelling we want to address in the near future. In addition to actuarial modelling, the idea of a broader discussion on pension savings on social networks within participatory governance is very interesting, as mentioned by Murray Svidroňová et al. (2018).

ACKNOWLEDGEMENT

The work of Jana Špirková has been supported by the Slovak Scientific Grant Agency VEGA no. 1/0150/21 – Profit testing of pension insurance products.

References

ACTUARIAL ASSOCIATION OF EUROPE. (2020). [online]. Brussels, Belgium: Actuarial Association of Europe. [cit. 14.2.2021]. <https://actuary.eu/?s=PEPP>.

ACT No. 39/2015 of 3 February 2015 on insurance and amending certain laws. (2015). [online]. Bratislava: National Council of the Slovak Republic. [cit. 14.2.2021]. <https://www.nbs.sk/_img/Documents/_Legislativa/_BasicActs/A39-2015.pdf>.

ACT No. 43/2004 Coll. of 20 January 2004 on the old-age pension scheme. (2004). [online]. Bratislava: National Council of the Slovak Republic. [cit. 14.2.2021]. <https://www.nbs.sk/_img/Documents/_Legislativa/_BasicActs/A43-2004.pdf>.

ACT No. 650/2004 Coll., dated 26 October 2004, on supplementary pension saving. (2004). [online]. Bratislava: National Council of the Slovak Republic. [cit. 14.2.2021]. <http://www.nbs.sk/_img/Documents/_Legislativa/_BasicActs/A650-04.pdf>.

ACT No. 461/2003 on Social Insurance. (2003). [online]. Bratislava: National Council of the Slovak Republic. [cit. 10.1.2021]. <http://www.ilo.org/dyn/natlex/natlex4.detail?p_lang=en&p_isn=67708>.

ACT No. 595/2003 Coll. of 04 December 2003 on income tax. (2003). [online]. Bratislava: National Council of the Slovak Republic. [cit. 14.2.2021]. <https://www.mfsr.sk/files/archiv/63/ActNo_595_2003_Coll_on_Income_Taxver_2020.pdf>.

ALBRECHER, H., EMBRECHTS, P., FILIPOVIĆ, D., HARRISON, G. W., KOCH, P., LOISEL, S., VANINI, P., WAGNER, J. (2016). Old-age Provision: Past, Present, Future. *Eur. Actuar. J.*, 6(2): 287–306. <https://doi.org/10.1007/s13385-016-0136-9>.

ASSOCIATION OF PENSION MANAGEMENT COMPANIES (in Slovak: Asociácia dôchodkových správcovských spoločností). (2004). [online]. Bratislava: Asociácia dôchodkových správcovských spoločností. [cit. 10.1.2021]. <https://www.adss.sk>.

BAČOVÁ, V., KOSTOVIČOVÁ, L. (2018). Too Far Away to Care about? Predicting Psychological Preparedness for Retirement Financial Planning among Young Employed Adults [online]. *Ekonomický časopis/Journal of Economics*, 66(1): 46–63. [cit. 10.12.2018]. <https://www.sav.sk/journals/uploads/0125125101\%2018\%20Bacova-Kostovic%20+%20RS.pdf>.

BILÍKOVÁ, M., MIHALECHOVÁ, J., SAKÁLOVÁ, K. (2016). Adequacy test for longevity risk reserves for a portfolio of model annuities from Pillar II of the Slovak pension system. *8^{th} International Scientific Conference Managing and Modelling of Financial Risks*, Ostrava, Czech Republic, 5^{th}–6^{th} September: 61–68.

COMPANY NN TATRY – SYMPATIA PENSION MANAGEMENT COMPANY, JSC (in Slovak: Spoločnosť NN Tatry – Sympatia d. d. s., a. s.). (2005). [online]. Bratislava: Spoločnosť NN Tatry – Sympatia d. d. s., a. s. [cit. 10.12.2020]. <https://tatrysympatia.nn.sk>.

COMPANY VÚB GENERALI PENSION MANAGEMENT COMPANY, JSC. (in Slovak: Spoločnosť VÚB Generali d.s.s., a.s.). (2004). [online]. Bratislava: VÚB Generali dôchodková správcovská spoločnosť, a.s. [cit. 10.12.2020]. <https://www.vubgenerali.sk>.

Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services. (2004). [online]. Brussels, Belgium: Official Journal of the European Union L 373/37. [cit. 5.12.2020]. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32004L0113\&from=EN>.

DICKSON, D., HARDY, M., WATERS, H. (2020). *Actuarial Mathematics for Life Contingent Risks.* Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108784184>.

DOLLS, M., DOERRENBERG, P., PEICHL, A., STICHNOTH, H. (2018). Do retirement savings increase in response to information about retirement and expected pensions? *Journal of Public Economics*, 158: 168–179. <https://doi.org/10.1016/j.jpubeco.2017.12.014>.

Current Pension Unit (in Slovak: Aktuálna dôchodková jednotka). (2004). [online]. Bratislava: Sociálna poisťovňa. [cit. 14.1.2021]. <https://www.socpoist.sk/dochodkova-hodnota-na-rok-2021/69137s>.

Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II). (2009). [online]. Brussels, Belgium: Official Journal of the European Union L 335/1. [cit. 14.11.2020]. <https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:335:0001:0155:en:PDF>.

EUROPEAN INSURANCE AND OCCUPATIONAL PENSIONS AUTHORITY. (2011). *Risk-free interest rate term structures.* [online]. Frankfurt am Main, Germany: European Insurance and Occupational Pensions Authority. [cit. 7.5.2021]. <https://www.eiopa.europa.eu/tools-and-data/risk-free-interest-rate-term-structures_en>.

EUROPEAN CENTRAL BANK. (2020). *Euro area yield curves* [online]. Frankfurt am Main, Germany: European Central Bank. [cit. 7.5.2021]. https://www.ecb.europa.eu/stats/financial_markets_and_interest_rates/euro_area_yield_curves/html/index.en.html.

KAINHOFER, R., PREDOTA, M., SCHMOCK, U. (2005).*The New Austrian Annuity Valuation Table AVO 2005R.* [online]. Vienna: Announcements from the Austrian Association of Actuaries (Mitteilungen der Aktuarvereinigung Osterreichs). [cit. 14.2.2021]. <http://www.avoe.at/pdf/mitteilungen/H13_w3.pdf>.

KIEREN, P., WEBER, M. (2021). When saving is not enough – wealth decumulation in retirement. *Journal of Pension Economics and Finance:* 1–28. <https://doi.org/10.1017/S1474747221000196>.

KONICZ, A. K., MULVEY, J. M. (2015). Optimal savings management for individuals with defined contribution pension plans. *European Journal of Operational Research*, 243(1): 233–247. <https://doi.org/10.1016/j.ejor.2014.11.016>.

Life Tables of the Slovak Republic. (2019). [online]. Bratislava, Slovakia: Statistical Office of the Slovak Republic. [cit. 10.12.2020]. <https://slovak.statistics.sk>.

MELICHERČÍK, I., SZŰCS, G., VILČEK, I. (2015a). Investment Strategies in the Funded Pillar of the Slovak Pension System [online]. *Ekonomický časopis/Journal of Economics*, 63(2): 133–151. <https://www.proquest.com/scholarly-journals/investment-strategies-funded-pillar-slovak/docview/1681852796/se-2?accountid=17223>.

MELICHERČÍK, I., SZŰCS, G., VILČEK, I. (2015b). Investment strategies in defined-contribution pension schemes [online]. *Acta Mathematica Universitatis Comenianae*, 84(2): 191–204. <http://www.iam.fmph.uniba.sk/amuc/ojs/index.php/amuc/article/view/139/186>.

MINISTRY OF LABOUR, SOCIAL AFFAIRS AND FAMILY OF THE SLOVAK REPUBLIC. (2020). *Labour and Employment* [online]. Bratislava, Slovakia: Ministry of Labour, Social Affairs and Family of the Slovak Republic. [cit. 10.12.2020]. <https://www.employment.gov.sk/en>.

MURRAY SVIDROŇOVÁ, M., KAŠČÁKOVÁ, A., VRBIČANOVÁ, V. (2018). Can Social Media be a Tool for Participatory Governance in Slovak Municipalities? *The NISPAcee Journal of Public Administration and Policy*, 11(2): 81–101. <http://dx.doi.org/10.2478/nispa-2018-0014>.

NATIONAL BANK OF SLOVAKIA. (2020). *Financial market supervision* [online]. Bratislava Slovakia: National Bank of Slovakia. [cit. 10.12.2020]. <https://www.nbs.sk/en/home>.

RYDQVIST, K., SPIZMAN, J., STREBULAEV, I. (2014). Government policy and ownership of equity securities. *Journal of Financial Economics*, 111(1): 70–85. <https://doi.org/10.1016/j.jfineco.2013.09.001>.

SOCIAL INSURANCE AGENCY. (2020). *Social Insurance System in Slovakia* [online]. Bratislava, Slovakia: Social insurance agency. [cit. 18.1.2021]. <https://www.socpoist.sk/social-insurance-system-in-slovakia/24533s>.

SCHULZ, J., CARRIN, G. (1972). The Role of Savings and Pension Systems in Maintaining Living Standards in Retirement. *The Journal of Human Resources*, 7(3): 343–365. <https://doi.org/10.2307/145088>.

ŠPIRKOVÁ, J., SZŰCS, G., KOLLÁR, I. (2019). Detailed View of a Payout Product of the Old-Age Pension Saving Scheme in Slovakia [online]. *Ekonomický časopis/Journal of Economics*, 67(7): 761–777. <http://hdl.handle.net/11159/4233>.

Technical Documentation. (2011). [online]. Frankfurt am Main, Germany: European Insurance and Occupational Pensions Authority. [cit. 10.2.2021]. <https://www.eiopa.europa.eu/tools-and-data/risk-free-interest-rate-term-structures_en>.

Technical Notes [online]. Frankfurt am Main: European Central Bank. [cit. 17.12.2020]. <https://www.ecb.europa.eu/stats/financial_markets_and_interest_rates/euro_area_yield_curves/html/technical_notes.pdf?update=20160513>.

WELLER, C., WENGER, J. (2009). Prudent investors: The asset allocation of public pension plans. *Journal of Pension Economics and Finance*, 8(4): 501–525. <https://doi.org/10.1017/S147474720800396X>.

Analysis of the Impact of Expenditures on Education and R&D on GDP in Central European Countries

Vasily Derbentsev[1] | *Kyiv National Economic University named after Vadym Hetman, Kyiv, Ukraine*
Yurii Pasichnyk[2] | *Institute of Agrarian Economics, Kyiv, Ukraine*
Leonid Tulush[3] | *Institute of Agrarian Economics, Kyiv, Ukraine*
Ievgeniya Lomako[4] | *Cherkasy State Technological University, Cherkasy, Ukraine*

Abstract

The main purpose of this paper is to identify the impact of public spending on education and research and development (R&D) on the formation of gross domestic product (GDP) in nine Central European countries, which are divided into two clusters – "old" and "new" EU members.

The study took into account official Eurostat data of both the EU and national statistical organizations for the period 2010–2019. The analysis of this impact was carried out using a system approach, statistics and econometric framework including panel data regression, Wald, Breusch-Pagan, Hausman tests. The main finding of the present study is the identification of additional income in terms of GDP in Euro per capita for selected countries, which is obtained from adequately spent public funds for education and R&D. Our results showed that the strongest influence of these expenditures for the "old" members was in Germany and Austria, and for the "new" – in Slovenia and Czechia. It is proved that this impact is different in individual countries and is determined by the public financial policy of national governments.

Keywords	***DOI***	***JEL code***
Government spending, economic growth, econometric modeling, panel data regression	*https://doi.org/10.54694/stat.2021.20*	*H52, I25, O32, C23*

INTRODUCTION

In today's globalized world, EU countries are trying to ensure dynamic economic growth and improve the welfare of the population. According to many economists in the realities of the XXI century, the main

[1] Kyiv National Economic University named after Vadym Hetman, Kyiv, Institute of Information Technologies in Economics, Informatics and System Research Department, Prospect Peremogy 54/1, 03057, Kyiv, Ukraine. E-mail: derbv@kneu.edu.ua.

[2] The National Scientific Centre Institute of Agrarian Economics, Str.Heroiiv Oborony, 10, 03127, Kyiv, Ukraine. Corresponding author: e-mail: y.v.pasich@gmail.com.

[3] The National Scientific Centre Institute of Agrarian Economics, Str.Heroiiv Oborony, 10, 03127, Kyiv, Ukraine. E-mail: tulush@ukr.net.

[4] Cherkasy State Technological University, Shevchenko Boulevard, 460, 18006, Cherkasy, Ukraine. E-mail: jane.pozhar@gmail.com.

driving force of social development is digitalization, introduction of new technologies, intellectualization of production processes.

It should be noted that in this context the fundamental basis for using these driving forces is to ensure the implementation of the following algorithm: *education → research → innovation*.

The study focuses on the impact on GDP of only two factors – education and R&D, as the most relevant at present. It should be noted that in some EU countries there may also be a significant impact on GDP and other factors, such as spending on economic programs, the investment climate, etc. However, a significant increase in the number of factors markedly expands the scope of research and such an approach is appropriate in fundamental research.

Education provides an opportunity for intellectual development of certain social strata, which then transform the acquired knowledge into scientific products with a corresponding innovative focus on real or virtual goods and services.

Such economically powerful EU countries as Germany, Great Britain, France attach great importance to these social levers, because they determine economic growth, which in turn provides the welfare of the population in the medium and long term.

Basic modern economic theories that study social development (a new Theory and Narrative of Economic Growth, Grounded in Innovation, Behavioral Economics, Experimental Economics, Ethical Economic Theory, Economic Theory of Happiness) in their methodological tools have appropriate provisions for education, science, and innovation.

Now experts characterize the modern economy as a knowledge economy, which is based on the relevant actions of the authorities. Public funding for education and R&D in the EU is constantly growing. Agreeing with this approach, we note that the modern economy also has other characteris.

Thus, if the total public spending on education in the EU-27 in 2016 amounted to 265.8 billion Euro, then in 2019 – 307.8 billion Euro, which is 15% growth over four years (Eurostat, 2021a). Accordingly, R&D expenditures for the same period amounted to 590.0 billion Euro and 656.3 billion Euro, an increase of 11.3% (Eurostat, 2021b).

These levers of economic growth in the EU are given significant importance. Therefore, an increasing spending on education and R&D is an important component of the fiscal policy of governments that adjust these expenditures according to public demand and real opportunities.

1 LITERATURE REVIEW

Given that the EU pays close attention to education and science at both the interstate and national levels, it should take these levels into account when analyzing strategic legal acts and the point of view of scientists.

At the EU level, several strategic documents have been approved, which define the directions of development. One of the defining perspective documents of Project Europe 2030 (European Commission, 2010) postulates that intelligence; innovation and creativity have now become the relevant standards of social development.

Another strategic document "EU Delivering on the UN 2030 Agenda" (European Commission, 2017) states that sustainable development is at the heart of the development of the European Union, and each initiative aims to improve the life of citizens.

These documents also stipulate that the key factors contributing to the transition to sustainability are education, science, technology, R&D, and innovation.

Considering the important role of universities in the development of both individual countries and the EU as a whole, the European University Association with more than 100 representatives proposed the Europe's Universities 2030 program, which defines the role of universities in shaping sustainable development (European University Association, 2021).

Bernadette et al. (2011) conducted a study aiming to map out the policies and strategies that exist in Europe to improve and promote the teaching of subjects academic research.

An original study of the impact of education quality on GDP was carried out by Hanushek and Woessmann (2020) who argued that increasing student achievement by 25 PISA scores across the EU would add 71 trillion Euros in reduced value to EU GDP compared to the current state of affairs.

A group of researchers (Pastor et al., 2018) assessed the contribution of universities to economic growth and GDP per capita in the EU during 2000–2015. They identified the impact of universities on supply in their national economy, especially the R&D universities contribution in technological assets and found out that GDP per capita was now more than 11% higher than in the scenario without universities contribution.

In the paper of Dima et al. (2018), the Global Competitiveness Index (GCI) for R&D spending as a percentage of GDP was analyzed by determining the impact of various indicators related to the knowledge economy using a regression model and panel data. The authors identified the crucial role of both innovation and education as determinants of EU competitiveness.

Taking into account unforeseen events in the organization of the educational process, in particular, the COVID pandemic, government spending on education in their EU countries has not decreased significantly, although face-to-face communication between faculty and students has been limited. To maintain appropriate contacts, a key concept of learning status and learning space was developed, which is based on prior learning and further online learning using appropriate algorithms and neural networks (Bukovsky et al., 2020).

Let's analyze the empirical studies in terms of individual EU countries. The reduction in public spending due to the economy has led to negative consequences of the reduction of public research and university budgets in Italy, so it is proposed to increase public R&D and increase universities funding (Nascia and Pianta, 2018).

Examining the problems of public debt in Czechia during the decline in GDP, Chekina and Vorhach (2020) found that the general government sector is in deep deficit, and debt is growing rapidly. They proposed a solution that should include increasing revenue over costs and stimulating economic growth, and as an exception to the concept, reducing the deficit which is necessary to stimulate investment activities, in particular, focused on new technologies, science, R&D and debt reduction in the long term.

The analysis of the impact of higher education funding on GDP growth in a number of countries, including Poland, Estonia, Czechia, Slovakia, Latvia, Lithuania, and Slovenia, reveals a trend when higher education expenditures correspond to higher qualifications, which contributes to higher GDP.

At the same time, a significant dependence of education expenditures on GDP growth in interstate comparisons has not been established, which is explained by the historical features of economic development of individual countries, the specifics of the national economy, labor markets and others (Chekina and Vorhach, 2020; UNESCO, 2010; UNESCO, 2012).

A comprehensive study by Anderson and Odei (2018), Hronova (2019) on the importance of research and development funding as part of ancillary government policy to coordinate interactive relations between firms and research institutions in Czechia, Slovakia, Hungary and Romania with the analysis of research papers revealed trends formation of the dynamics of economic growth.

In the context of economic growth at the micro level, the importance of research and development costs for the competitiveness of manufacturing enterprises in Poland has been proven Grzelak et al. (2018), and Roszko-Wójtowicz et al. (2019).

The study of current issues on the impact on GDP of the economic sentiment indicator (ECI), which includes five indicators, was conducted using a panel co-integration analysis for EU countries during 2000–2018, which explained the current and future values of relevant macroeconomic parameters performed by Tomic et al. (2020).

A fundamental study of the impact of R&D on long-term economics using the Bayesian model averaging (BMA) on the example of Czechia was conducted by Horváth (2011). These studies proposed to capture the R&D intensity by the number of Nobel prizes in science. Using this indicator, obtained estimates show that R&D exerts a positive effect on long-term growth.

2 MATERIALS AND METHODS

According to the purpose of this study, approaches to the classification of Central European countries were used, data for analysis were prepared, methods were selected and appropriate models were substantiated.

2.1 Selection of a statistical base and modeling tools for research

Regarding the choice of Central European countries, individual researchers, depending on the purpose of the study, include in this group different countries. Thus, Dlouhá et al. (2016) when considering the problems of higher education in Central Europe include in this group such "new" countries as Czechia, Hungary, Poland, Serbia, Slovakia, and Slovenia. In this context, the "new" are the countries that joined the EU after 2014.

Römisch (2020) analyzing medical processes, respectively distinguishes Italy, Croatia, Slovakia, Czechia, Austria, Slovenia, Hungary, Germany, Poland. The Czech Geographical Directory (Svobody, 2007) includes Poland, Czechia, Slovenia, Hungary, Austria, Switzerland, Liechtenstein, and Germany.

Thus, this group of countries is variable, but since this study uses Eurostat data, we turn to the grouping, which is officially recognized in statistics. European Commission in the allocation of funds in the field of innovation, low-carbon strategies, natural and cultural resources for sustainable growth, as well as in the field of transport, which were then accounted for by Eurostat, identified the countries of Central Europe as follows: Czechia, Croatia, Germany, Hungary, Italy, Poland. Slovenia, Slovakia (European Commission, 2014).

This is the list of Central European countries that we will serve a basis for our research, and we will choose these identified nine European countries for the specification and modeling. To compare the results of the study among these countries, we distinguish two clusters – "old" EU countries – Austria, Germany, Italy and "new" – Czechia, Croatia, Hungary, Poland, Slovenia, and Slovakia.

It should be noted, that economic growth is complex multidimensional process, which is influenced directly and indirectly by many factors: capital, innovation, entrepreneurship, human capital, education, demographic processes, labor productivity, the level of development of technologies, political and social institutions, etc.

In our research, we focused on two main aspects related to Government incentives (funding) of economic growth in:

(i) innovative technologies (R&D expenditures);

(ii) education, which is one of the aspects of improving the quality of human capital.

Moreover, using more explanatory factors can lead to a multicollinearity problem.

Thus, for our analysis we used annual data of GDP (y, target variable), domestic R&D expenditures (X_1) and expenditures on education (X_2) as independent variables according to Eurostat (Eurostat, 2021; Annex: Tables A1, A2, A4, A6). All data were taken in Euro per capita in order to make the sample comparable, leveling the difference in GDP and population between countries (Annex: Tables A3, A5, A7).

As time period for our analysis, it was selected period from 2010 to 2019. We did not use earlier data because the period before the global crisis of 2008–2009 was characterized by other trend of economic development. Therefore there is no reason to believe that the data belong to one statistical sample, or generated by the same dynamic system. Data for 2020 are also not representative due to the global economic recession caused by the COVID 2019 pandemic.

As was noted in the Introduction, the purpose of present study is to assess the impact of public expenditures on Research and Development (R&D), and Education on economic growth in several European countries. Therefore, we used panel (longitudinal) regression as a modeling tool. Panel data combines both cross-sectional and time-series data (Baltagi, 2021; Croissant and Millo, 2008). At each time point there are spatial data for economic "units" (in our case, GDP per capita in several countries), and for each such unit are available the corresponding data form one or more time series (for each country there are time series for expenditures on R&D and education).

As software tool for simulation panel data models in our research were used STATA.

2.2 Panel data model notation

Let is the dependent variable (GDP per capita) for the i-th country at time (year) t; x_{it} – set of explanation (independent) variables (k-dimension vector); ε_{it} – regression error vector; $i = 1, 2, \dots, n$; $t = 1, 2, \dots, T$.

Let's introduce the following notation for each i-th country:

$$y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix}, X_i^{(j)} = \begin{bmatrix} x'_{i1} \\ \vdots \\ x'_{iT} \end{bmatrix}, \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}. \quad (1)$$

In the case of this study the number of independent variables $k = 2$, so in Formula (1) the superscript in parentheses (j) denotes the ordinal number of the explanatory variable ($j = 1, 2$): X_1 – is domestic R&D expenditures in Euros for each country; X_2 – expenditures in Euros for education in each country. In Formula (1) superscript in middle term (such as x'_{i1}) denotes transpose operator.

Typically, majority panel data applications use a one-component random error composition model (Baltagi, 2021; Croissant and Millo, 2008):

$$\varepsilon_i = u_i + \mu_{it}, \quad (2)$$

where u_i characterize individual effects that are unobservable and constant over time (individual heterogeneity); μ_{it} residual error.

Let's also defined "pooled" variables:

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \quad (3)$$

where y, ε – are $nT \times 1$– dimension vectors, X is $nT \times k$ matrix.

$$y = X\beta + \varepsilon. \quad (4)$$

Formula (2) assumes that all μ_{it} errors are uncorrelated with each other in both i (for different units) and t (for different time periods), and are uncorrelated with all explanatory variables x_{it}.

Consider the most common panel regression model specifications: Pooled Regression model, Fixed Effect model, and Random Effect model.

The Pooled Regression model (PR) specifies constant coefficients, the usual assumptions for cross-sectional analysis.

$$y_{it} = \alpha + x'_{it}\beta + \varepsilon_{it}. \quad (5)$$

This is the most restrictive panel data model. PR is usually applied in the absence of significant differences (heterogeneity) between the sampled units.

Panel data allows taking into account differences in economic units. Let's write one of the possible implementations in the following form:

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it},\ i = 1, 2, \dots, n;\ t = 1, 2, \dots, T. \tag{6}$$

In Formula (6) expresses the individual effect of the unit i, which does not depend on time while the regressors do not contain the constant. In this notation it is generally accepted assumption that errors ε_{it} are (Baltagi, 2021; Croissant and Millo, 2008):

(i) ε_{it} are uncorrelated with each other by i and t: $E(\varepsilon_{it}) = 0$, $Var(\varepsilon_{it}) = \sigma_\varepsilon^2$,

(ii) ε_{it} are uncorrelated with repressors for all i and t.

These assumptions guarantee unbiased and consistency OLS parameter estimates.

Intra-group ("within") and inter-group ("between") estimates are often used to find the parameters of a panel data model. "Within" estimates can be obtained by building a model for deviations from group means, and between estimates by building a model for group means.

Regression "between" is the original model rewritten in terms of time-averaged values of variables:

$$\overline{y}_i = \alpha_i + \overline{x}'_i\beta + \overline{\varepsilon}_i, \tag{7}$$

where:

$$\overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it},\ \overline{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it},\ \overline{\varepsilon}_i = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{it},\ i = 1, 2, \dots, n. \tag{8}$$

The "within" regression or Fixed Effect (FE) models is the original regression model (6) which is rewritten in terms of the deviations from the time mean values of the variables. If subtract (8) from (6) term by term ("within" transform), than it's possible to get:

$$(y_{it} - \overline{y}_i) = \alpha_i + (x'_{it} - \overline{x}_i)\,\beta + (\varepsilon_{it} - \overline{\varepsilon}_i). \tag{9}$$

In Formula (9) α_i also expresses the individual effect of the unit i, which does not depend on time. Then the FE estimator is equivalent to the "within" estimator and can be written in this form:

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)'\right)^{-1}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(y_{it} - \overline{y}_i)'\right). \tag{10}$$

If unobservable factors do not correlate with regressors, to obtain more efficient estimates, it is possible to consider a panel data model with Random Effects (RE): it is assumed that the missing variables are one of the components of errors.

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it} + u_t,\ i = 1, 2, \dots, n;\ t = 1, 2, \dots, T. \tag{11}$$

In Formula (11) u_t characterize within-units errors, invariant in time for each economic unit; ε_{it} – between-units errors. In other words, RE models assume that individual differences are random.

In addition to the assumptions (i) and (ii) about errors for the FE model, we will also assume that u_t are $IID(0, \sigma_u^2)$:

(iii) errors u_i are uncorrelated, $E(u_i) = 0$, $Var(u_i) = \sigma_\varepsilon^2$,

(iv) errors u_i are uncorrelated with repressors for all i and t.

2.3 Choosing the best model specification

It should be noted that for panel data models of regions and countries, the model with FE is most often used, since each of the objects (unit) of such a sample has its own individual characteristics, and the purpose of building a model is, in particular, to obtain a forecast for a specific sample object.

There are several effective statistical approaches for choosing the most adequate panel data model (Baltagi, 2021).

To test the model Pooled Regression vs. Fixed Effects and select the best specification, the Wald test (based on Fisher's F-test) and the Likelihood-Ratio test (LR test) are the most frequently used. Wald's test checks the hypothesis that all individual effects are equal to zero in the FE model. The null hypothesis (H0) is that all in (6) $\alpha_i = 0$ (Baltagi, 2021; Croissant and Millo, 2008).

To select RE model vs. FE model, as a rule, it is usually applied the Hausman test. The RE model takes place only when the random effects are uncorrelated with regressors. So in this test the null hypothesis (H0) is that $cov(\alpha_i, X_{it}) = 0$ and model is correctly specified. Thus, to accept the FE model, the null hypothesis must be rejected both in the F-test (there is a panel structure) and in the Hausman test (only the estimates of the model with FE are consistent, and the estimates of the model with RE are inconsistent).

The Breusch-Pagan test is a test for the presence of a random individual effect and tests a hypothesis $Var(u_i) = 0$.

3 EMPIRICAL RESULTS

3.1 Correlation analysis

To test our hypothesis about the dominant influence of spending on education and research on economic growth (and as a consequence population income), we conducted a correlation analysis.

The average values of the selected factors, grouped by country are shown in Table 1. All data are presented in Euro per capita.

Table 1 Average values of the selected factors, grouped by countries

	Y	X1	X2
2010	18 830.0	339.4	918.3
2011	19 250.0	367.1	923.4
2012	19 098.9	395.8	915.4
2013	19 034.4	404.7	928.1
2014	19 302.2	420.5	930.4
2015	19 687.8	354.5	983.7
2016	20 067.8	435.6	981.3
2017	20 688.9	366.3	1 013.4
2018	21 241.1	499.1	1 079.7
2019	21 614.4	528.7	1 131.2

Source: Authors calculations based on Eurostat (2021a, 2021b)

The correlation coefficients between the target variable (y) and the corresponding factors are given in the last row. As one can see, there is a high statistical relationship between the factors and the target variable: the correlation between GDP per capita and expenditure on R&D and education is 0.76 and 0.98, respectively. At the same time, there is also a significant correlation between the explanatory variables (0.78), which can lead to instability of the OLS estimates and their variances.

It was also calculated correlation coefficients for "between" regression, in which factors were grouped by the time (Formula 8). Results are presented in Table 2.

Table 2 Average values of the selected factors, grouped by the time

Country	GDP per capita	R&D expenditures	Education expenditures
I	$\bar{y}$	$\bar{x}_1$	$\bar{x}_2$
Austria	36 592	1 199.5	1 948.0
Czechia	16 291	299.6	751.3
Croatia	10 979	95.3	541.4
Germany	34 135	1 064.8	1 554.6
Hungary	11 195	154.1	585.2
Italy	26 279	370.6	1 118.3
Poland	10 911	111.8	585.3
Slovenia	18 465	424.1	1 139.2
Slovakia	14 087	121.6	601.3
	Correl.	0.95	0.97

Source: Authors calculations based on Eurostat (2021a, 2021b, 2021c)

Results of Table 2 also show a high relationship between the target variable and factors for the case of inter-group inter-group ("between") estimates.

Thus, the results of the correlation analysis confirm our hypothesis regarding the main drivers of economic growth.

3.2 Models specification

Parameter estimates for various model specifications presented in Section 2.2, as well as the results of statistical tests, are given in the Annex. Summarized estimations for panel data models are shown in Table 3.

As one can see (final estimation results are summarized in the Annex: Tables A8–A12), all model specifications showed good fit accuracy by R-squared criterion: within 0.93–0.94.

The best results were obtained with «between» regression (7). In this case, the R-squared between value reflects the quality of the regression fit and is large enough (0.94), i.e. the change in the average over time factors for each country has a more significant impact on each variable than the temporal fluctuations of these factors relative to the average. This is an additional argument for the need to take into account individual effects against the pool model. But the coefficients of this model (in particular, β_1) turned out to be insignificant according to the *t*-criterion (Table A9 in the Annex).

To select the most adequate model of the dependence of the level of economic growth on expenditures on education and R&D, we except R-squared criterion applied the following statistical tests:

a) Wald test: for comparison the FE (9) model versus the PR model (5).

b) Breusch-Pagan test: for comparison the RE (11) model versus the PR model (5).

c) Hausman test: for comparison the RE (11) model versus the FE regression model (9).

Wald's test checks the null hypothesis that all individual effects are equal to zero. STATA automatically tests this hypothesis at the same time as evaluating the FE model and displays the result in the last row of Table A10 in the Annex. In this case the null hypothesis is rejected at any level of significance. Thus, the FE regression model is better fit data than the PR model.

Table 3 Summarized estimations for panel data models

		Pooled regression model, PR	Between	Fixed effect (within), FE	First differences	Random Effect, RE
R&D	X1	7.12	6.48	3.31	4.85	4.23
Education	X2	12.23	14.08	6.74	2.55	7.50
Constant		3 853.90	3 307.64	11 851.77	-	10 716.04
R2		0.93			0.35	
R2-within			0.61	0.61		0.61
R2-between			0.94	0.94		0.94
R2-overall			0.93	0.93		0.93

Source: Authors calculations by using STATA software

The Breusch-Pagan test checks the presence of a random individual effect and tests the following null hypotheses: $Var(u_i) = 0$. As one can see in Table A12 in the Annex, the null hypothesis is rejected and so model RE is preferred over model PR.

Finally, the Hausman test was used to choose between the FE and RE model specification, which tests for correlation between random effects with regressors. The null hypothesis is $H0$: $corr(u_i, x_{it}) = 0$. The test results not allow reject or assume the null hypothesis, because model fitted on these data fails to meet the asymptotic assumptions of the Hausman test (Table A13 in the Annex).

3.3 Analysis of empirical result

Our results show that, other things being equal, higher level of economic development is associated with (caused by) higher values of expenditures on education and R&D. In all models, we obtained the expected positive signs at the coefficients of the regressors.

At the same time, an interesting fact is that additional spending on education, on average, has almost twice the effect on economic growth than spending on R&D. So, according to the RE model, each additional Euros invested in education leads to an increase in GDP per capita by 7.5 Euro, and additional Euros invested in R&D – about 4.23 Euro.

In this specification constant (10 716.04 Euro) characterizes the impact (or contribution, share) on economic growth (GDP per capita) of unobservable variables (individual characteristics of countries (averaged). Moreover, in this model, it is postulated that the differences are of a random nature and, on average, are leveled. Thus, this value can be interpreted as the average level of GDP per capita across countries, independent of spending on education and research. In particular, this level may be due to factors of the neoclassical production function of the Cobb-Douglas type: labor and capital.

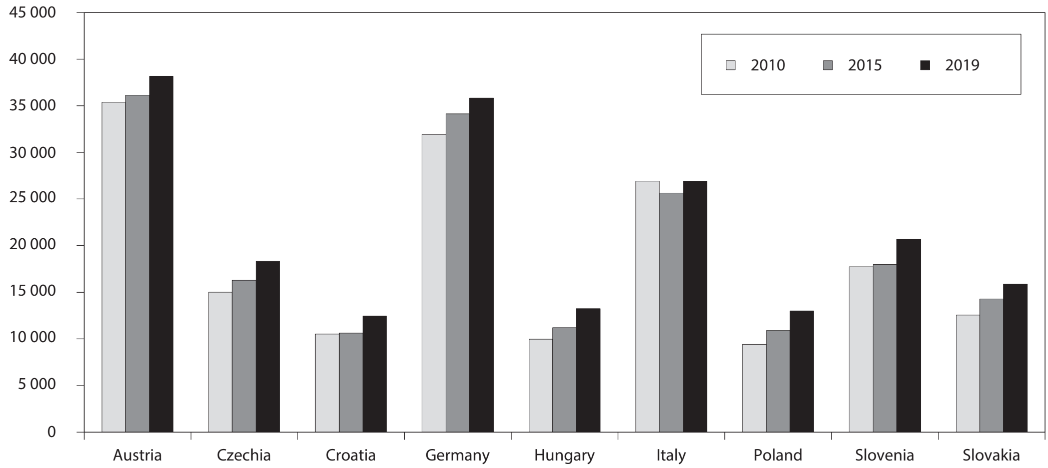
Smaller values are given by the FE model: 6.74 and 3.31 Euro, respectively. Constant (11 851.77) also characterizes the impact on economic growth of missed and unobservable variables (individual characteristics of countries that do not change over time).

Value of by R-squared criterion allows us to conclude that the variation of the factors (expenditure on R&D and education per capita) explains 93–94% of variations in the dynamics of GDP per capita.

This fact confirms our hypothesis that the main drivers of economic development in the medium term are human capital (which we approximated by Education expenditures) and innovative technologies (in our models, this factor is described by R&D expenditures).

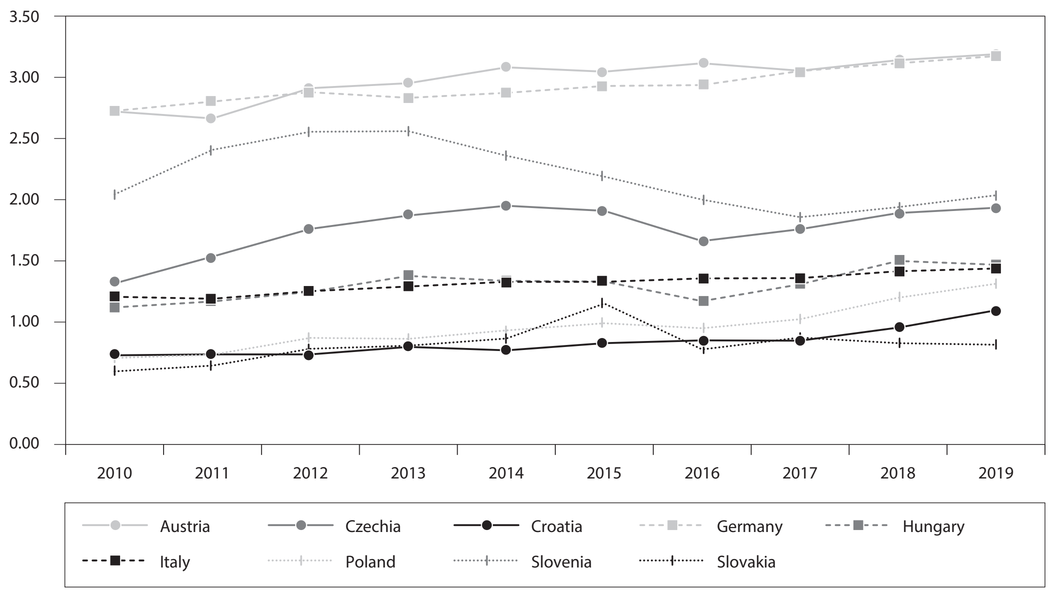
An additional analysis was also carried out for two clusters: for "old" EU members (Germany, Italy, and Austria – the 1st cluster) and countries that joined the EU after 2000 (the 2nd cluster). As we can

Figure 1 Real GDP per capita for selected counties in 2010, 2015, and 2019 (Euro)



Source: Authors calculations based on Eurostat (2021c)

Figure 2 Expenditure on R&D (% GDP)



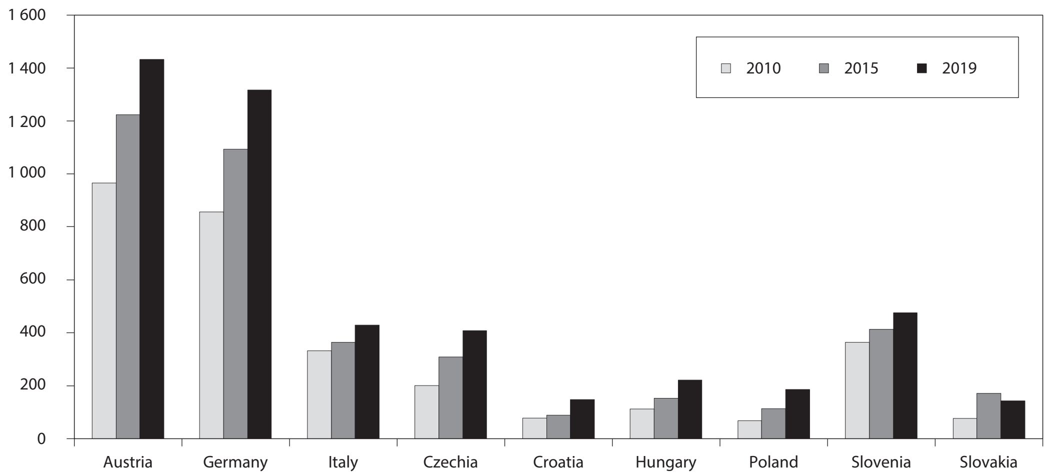
Source: Authors calculations based on Eurostat (2021b, 2021c) see (Figure 1), the old members of the EU both led in terms of living standards in 2010 and continue to outstrip other countries now.

At the same time, it can be noted that for Germany and Austria the living standards (in Euro per capita) is gradually increasing, for Italy it has remained at the same level for last 10 years. This may also be due to the fact that Italy spends much less on R&D in both absolute (% of GDP, Figure 2) and relative terms (Euro per capita, Figure 3).

As for the countries of the 2nd cluster, the greatest absolute increase in the living standard during the study period was observed in Slovenia and Czechia.

Moreover, in relative terms for R&D expenditures, Slovenia has already surpassed Italy, and Czechia has practically caught up with it.

Figure 3 Expenditure on R&D per capita in 2010, 2015, and 2019 (Euro)



Source: Authors calculations based on Eurostat (2021b)

With regard to spending on education, the situation is not so clear here. In absolute terms, the countries of the 1st cluster are not leaders, but at the same time they maintain high standards: about 4–5% of GDP (Figure 4). As to the countries of the 2nd cluster, then for the period under study, expenditures on education were within the same limits: at the level of 4–5% of GDP.

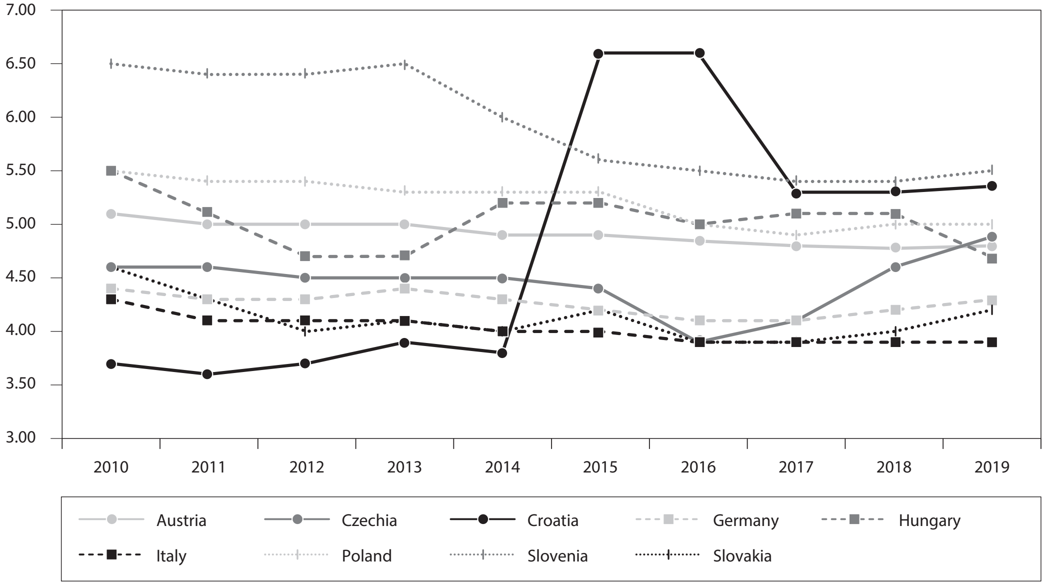
At the same time, unlike other countries, the data for Croatia are somewhat anomalous: in 2015–2016, spending on education was almost doubled (from 3.5 to 6.5%), and then decreased to 5.4% (Figure 4). But spending on education per capita in 2015 and 2019 was almost the same (Figure 3). At the same time, the standard of living (GDP per capita) during this period increased by 17% (1 800 Euros per capita). This fact requires additional analysis.

In relative terms, the situation is similar to R&D expenditure: Germany and Austria are the leaders during the selected period. And Slovenia by 2019 outstripped Italy in this indicator, and the Czechia came close to it (Figure 5).

These empirical facts allow us to make a preliminary conclusion that, all other things being equal, if the existing trends continue for 5 years, Slovenia will catch up with Italy in terms of living standards, and Czechia will come close to it.

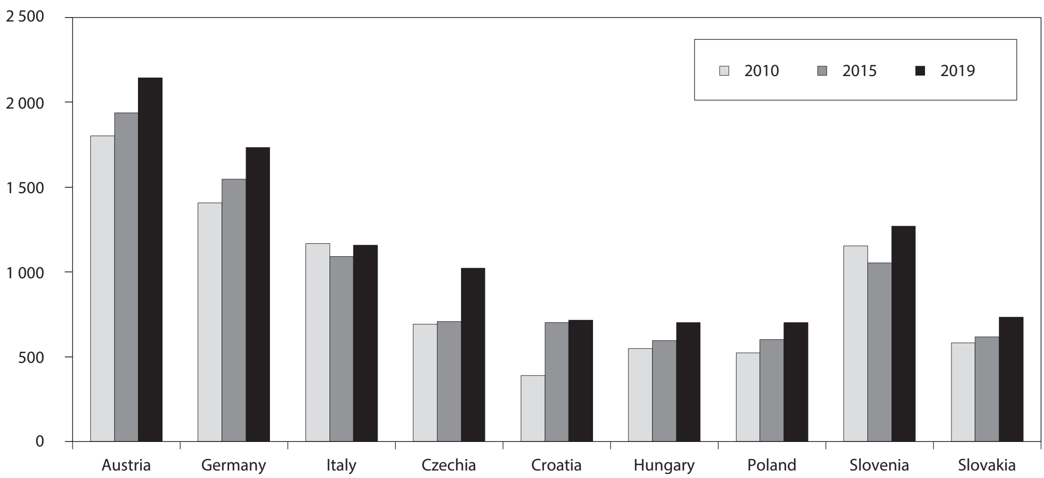
To test our hypotheses, we also performed panel data regression analyzes for selected clusters. Final results for model selection which based on the Hausman test are presented in the Annex (Table A14

Figure 4 Expenditure on education (% GDP)



Source: Authors calculations based on Eurostat (2021a, 2021c)

Figure 5 Expenditure on education per capita in 2010, 2015 and 2019 (Euro)



Source: Authors calculations based on Eurostat (2021a)

for the 1st cluster and Table A15 for the 2nd) and summarized estimations for panel data models for the 2nd cluster are shown in Table 4.

The results of the Hausman test do not allow for rejecting the main hypothesis about the presence of random effects for the 2nd cluster (Table A 15 in the Annex). However, the coefficients of models 1 and 2 are practically the same. Models are also adequate in terms of multiple determination R-squared.

Table 4 Summarized estimations for panel data models for the 2nd cluster

		Pooled regression model, PR	Between	Fixed effect (within), FE	Random Effect, RE
R&D	X1	17.27	19.63	15.25	15.36
Education	X2	3.04	1.61	4.11	4.07
Constant		8 051.50	8 575.29	7 707.08	7 710.37
R2		0.82			
R2-within			0.64	0.64	0.64
R2-between			0.85	0.85	0.85
R2-overall			0.82	0.82	0.82

Source: Authors calculations by using STATA software

For both specification (FE and RE) one can see that the importance (elasticity) of the research expenditure factor (X_1) has increased significantly: for FE almost 5 times (from 3.31 to 15.25). At the same time, the significance of the factor of spending on education has slightly decreased (from 6.74 to 4.11 for FE specification).

From our point of view, this can be explained by the fact that in terms of education expenditures, the countries of the 2nd cluster in relative terms are not so far behind the leaders.

Thus, the results of our modeling show that for the countries of the 2nd cluster (which became EU members after 2 000), an additional 1 Euro (per capita) invested in research brings an effect of 15 Euro (per capita), and 1 Euro invested in education brings an effect of 4 Euro (per capita).

The analysis shows that among the selected countries the most positive dynamics are shown by Slovenia and Czechia.

4 DISCUSSION

Public spending on education and R&D in the context of economic growth is an important aspect of government policy, and in Central Europe this issue is also discussed by scholars who study spending, impact factors, households' response, and so on.

Note that in our research, as a strong argument was used economic and mathematical methods, based on econometrics framework using real Eurostat data

Evaluating the impact of educational changes in primary, secondary and higher education using the following decomposition on the effects of level and substitution between 2005 and 2017 in Hungary, Poland, Slovakia, and Czechia (Fischer et al., 2020) identified the effect of differences in education structure of households whose incomes significantly affect economic growth.

However, there are other studies (Delgado et al., 2014) where using nonparametric local-linear regression estimation and test for the relevance of nonparametric transformations to conduct a nationwide and systematic search for the significance of average learning duration while studying the most complete databases on education, it was found that the average duration of schooling is not a statistically significant variable in economic growth regression.

At the same time, empirical studies often show that the scope and importance of training depends on the choice of observations or model specifications.

Earlier it was hypothesized that education has a greater effect in those countries where economic change is faster. Its examination in a wider set of 32 countries (Hanushek et al., 2015) found that the range of differences in skill levels in different countries is even greater than previously thought.

The main observed pattern was economic growth, which corresponds to the ability of skilled workers to adapt more easily to economic change. This refined hypothesis, of course, is subject to questions of causality, but taking into account a number of alternative effects does not change the general trend.

Several studies have identified the link between education, research and innovation, and substantiated the positive impact on the economy (Power, 2015; WEF, 2016–2017; Grant, 2017; Rundle, 2021), but in practice governments in some countries will send different amounts funding for these public goods depending on the defined state policy.

Our study presents examples of different levels of funding for education and research in nine Central European countries where, in general, the relationship between these government expenditures and the level of GDP in the compared units of measurement is confirmed (Figure 1, Figure 5).

Sometimes research focuses on situations in which investment in education and human capital development affects the economic growth of middle-income countries. The problem is that the return on accumulated physical capital is declining, and the rate of productivity growth and technological innovation depend to a large extent on skilled human capital.

Therefore, no matter what policy options a country chooses, it still faces the challenge of implementing them because over time they have different opportunities than those that originally led them to middle-income status (Larson et al., 2016).

The calculations of the experts of the International organizations convince of the importance of acquiring the appropriate level of education in school, which is vital for reducing the level of unemployment, inequality and poverty and promoting growth. Thus, for every US dollar spent on education, there can be 10 to 15 US dollars of economic growth.

Also, if 75% of 15-year-olds in the world's 46 poorest countries could reach the lowest level in mathematics in OECD countries, economic growth would increase by 2.1% from baseline, and 104 million people could be brought out of extreme poverty (UNESCO, 2012).

At the same time, according to our calculations for Central European countries, the additional costs of education on average have almost twice the effect of economic growth than the costs of R&D. Thus, according to the RE model, each additional Euro invested in education leads to an increase in GDP per capita of 7.5 Euros, and each Euro invested in R&D leads to an increase of 4.23 Euros. Fewer values are given by the FE model: 6.74 and 3.31 Euros, respectively, due to different model specifications.

Comparing the views of experts from international organizations and scientists highlighted different approaches to the role and importance of public spending on education in R&D in ensuring economic growth. These differences are due to the objectives of the study, the applied scientific apparatus, in particular analytical, the amount of data analyzed, and so on.

CONCLUSION

Ensuring dynamic economic growth is an important societal challenge in any country. That is why governments use certain expenditures, in particular on education and R&D, as important levers to achieve this goal. Achieving the highest possible rate of economic development is especially important for EU countries, where living standards in most countries are quite high compared to other regions of the world and governments are constantly trying to prevent its decline.

In this context, scientists are involved in solving this problem. They publish relevant scientific developments that analyze the situation with the dynamics of economic growth, the factors influencing these processes, propose levers, forms, methods that should promote the GDP growth and increase incomes with the solution of related life problems.

Our study confirmed the importance of investing in these expenditures, as education provides from 6.7 to 7.5 Euros, and R&D – from 3.3 to 4.3 Euros of growth (GDP per capita). At the same time, despite compelling arguments, not all of the nine countries identified are following the growing dynamics of these costs.

Thus, during 2010–2019, in the "old" EU members, expenditures on education in absolute terms increased the most in Germany – 31.3%, and among the "new" – in Croatia 64.6%. R&D is given more importance, and in the "old" EU member states these expenditures increased the most in Austria – 1.57 times, and in the "new" ones in Poland – 2.64 times. These data confirm the existence of different government approaches in policy-making with respect to the priority of the use of certain levers of economic growth.

It should also be noted that the “Europe 2030 Strategy” emphasizes the importance of the role of education, science, technology, research and innovation in ensuring the sustainable development of the European Union and recommends that national governments take action to improve the living standards of citizens.

The study compares the approaches of scientists to the role and importance of the impact of public spending on education and R&D in Central Europe. It is revealed that various approaches are used to perform the research by scientists from different countries, which include algorithms, research methods, mathematical apparatus, economic theories, a set of statistical data.

The analysis of scientific achievements revealed some relevant aspects – the importance of obtaining quality knowledge in schools and universities, the positive increase in the number of young people covered by training in relevant institutions at the pace of development, employment prospects after graduation, levers of economic growth, the feasibility of increasing education and R&D.

Despite the postulates of economic theory on the importance of these expenditures on the dynamics of economic growth revealed, only their absolute increase will not lead to simultaneous GDP growth, because it is necessary to take into account the quality of education, long-term economic growth programs, and initiative in the government policy.

The peculiarity of this study is that, despite the diversity of scientific research on this issue, the relationship between such expenditures as education and R&D in the context of the overall impact on GDP is insufficiently covered. Based on the results of the correlation analysis, two models were formed, which made it possible to identify the significance of the impact of these expenditures on GDP separately for the “old” and “new” EU members.

The use of the obtained results will help to guide the governments of Central European countries in the formation of public policies, given that the study was performed with appropriate scientific justification. In addition, in the process of making government decisions to increase GDP, it is advisable to adjust them with other levers of influence.

As for the direction of further research, it seems promising to us to test other classes of econometric models, which allows incorporate time lag effect between factors and target variable. In particular, there are Vector Autoregressive Models (VAR) and Models with Distributed Lags (DLM).

VARs take into account the fact that, on the one hand, innovations stimulate economic growth, and on the other hand, funding for innovation and R&D depends on GDP. This creates a multiplier effect.

DLM have such advantage that they take into account the presence of a time lag between economic growth and an increase in labor productivity and competitiveness due to an improvement in the quality of human capital and the level of technology.

It also seems promising to test a longer time period to obtain more reliable estimates of the model parameters. In this case, however, it is necessary to conduct additional research on the statistical properties of the model factors. It is need to define whether the time series before 2008 describe the same statistical process, or after 2008 there have been significant structural changes that do not allow considering the dynamics of economic growth in different sub-periods as a single process.

Moreover, it would be advisable to expand the set of explanatory variables that characterize the contribution of human capital and innovations to economic growth.

References

ANDERSON, H. J., ODEI, S. A. (2018). The Influence of Public Support on University-Industry. Government Collaboration: the Case of the Czech Republic, Slovakia, Hungary and Romania. *Statistika: Statistics and Economy Journal*, 98(4): 352–361.

BALTAGI, B. H. (2021). Econometric Analysis of Panel Data. *Springer Texts in Business and Economics*, Springer, Ed. 6, No. 978-3-030-53953-5 (September). Handle: RePEc: spr:sptbec:978-3-030-53953-5. <http://doi.org/ 10.1007/978-3-030-53953-5>.

BERNADETTE, F., HORVATH, A., COUTINHO, A., SOFIA, A. (2011). *Science education in Europe. National policies, practices and research.* <https://doi.org/10.2797/7170>.

BUKOVSKY, I., STEINBAUER, P., BUDIK, O., ICHIJI, K., NORIYASU, H. (2020). Learning Entropy of Adaptive Filters via Clustering Techniques. *Sensor Signal Processing for Defense Conference (SSPD):* 66–70. <https://doi.org/10.3390/e21020166>.

CHEKINA, V. D., VORHACH, O. A. (2020). The impact of education expenditures on economic growth: empirical estimation. *Econony of Industry*, 3(91). <http://doi.org/10.15407/econindustry 2020.03.096>.

CROISSANT, Y., MILLO, G. (2008). Panel data econometrics in R: the plm package. *Journal of Statistical Software*, 27(2): 1–43. <https://doi.org/10.18637/jss.v027.i02>.

DELGADO, M. S., HENDERSON, D. J, PARMETER, C. F. (2014). Does Education Matter for Economic Growth? *Oxford Bulletin of Economics and Statistics*, 76(3): 334–359. <https://doi.org/10.1111/obes.12025>.

DIMA, A. M., BEGU, L., VASILESCU, V. D., MAASSEN, M. A. (2018). The Relationship between the Knowledge Economy and Global Competitiveness in the European Union [online]. *Sustainability*, 10(6): 1706. <https://www.mdpi.com/2071-50/10/6/1706>. <https://doi.org/10.3390/su10061706>.

DLOUHÁ, J., GLAVIČ, P, BARTON, A. (2017). Higher education in Central European countries – Critical factors for sustainability transition [online]. *Journal of Cleaner Production*, 151: 670-684. <https://doi.org/10.1016/j.jclepro.2016.08.022>.

EUROPEAN COMMISSION. (2014). *Central Europe. Interreg. Programme description. Regions* [online]. <https://ec.europa.eu/regional_policy/en/atlas/programmes/2014-2020/austria/2014tc16rftn003>.

EUROPEAN COMMISSION. (2017). *EU delivering on the un 2030 agenda Sustainable development in Europe and the World* [online]. <https://ec.europa.eu/info/sites/info/files/factsheet-eu-delivering-2030-agenda-sustainable-development_en.pdf>.

EUROPEAN COMMISSION. (2010). *Project Europe 2030. Challenges and Opportunities. A report to the European Council by the Reflection Group on the Future of the EU 2030* [online]. <https://www.consilium.europa.eu/media/ 30776/qc3210249enc.pdf>.

EUROPEAN UNIVERSITY ASSOCIATION. (2021). *Europe's Universities 2030* [online]. <https://eua.eu/101-projects/784-europe%E2%80%99s-universities-2030.html>.

EUROSTAT. (2021a). *General government expenditure by function* [online]. <https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>.

EUROSTAT. (2021b). *GERD by sector of performance* [online]. <https://ec.europa.eu/ eurostat/databrowser/view/rd_e_gerdtot/default/table?lang=en>.

EUROSTAT. (2021c). *GDP and main components (output, expenditure and income)* [online]. <https://ec.europa.eu/eurostat/databrowser/view/nama_10_gdp/default/ table?lang=en>.

FISCHER, J., FLUSKOVÁ, H., VLTAVSKÁ, K. (2020). At-Risk-of-Poverty Rate or Social Exclusion in Visegrad Countries 2005–2017: Impact of Changes in Households' Structure. *Statistika: Statistics and Economy Journal*, 100(4): 351–364.

GRANT, C. (2017). *The contribution of education to economic growth* [online]. Knowledge, evidence and learning for development K4D, Institute of Development Studies. <https://assets.publishing.service.gov.uk/media/5b9b87f340f0b67896977bae/K4D_HDR_The_Contribution_of_Education_to_Economic_Growth_Final.pdf>.

GRZELAK, M., LASKOWSKA, I., ROSZKO-WÓJTOWICZ, E. (2018).The Importance of Research and Development Expenditure for the Competitiveness of Manufacturing Enterprises in Poland. The Application of Panel Models. *Econometrics*, 22(3): 9–22. <https://doi.org/10.15611/eada.2018.3.01>.

HANUSHEK, E. A., SCHWERDT, G., WIEDERHOLD, S., WOESSMANN, L. (2017). Coping with change: International differences in the returns to skills. *Economic Letters*, 153(April): 15–19. <http://dx.doi.org/10.1016/j.econlet.2017.01.007>.

HANUSHEK, E. A., WOESSMANN, L. A. (2020). Quantitative look at the economic impact of the European Union's educational goals. *Education Economics*, 28(3): 225–244. <https://doi.org/10.1080/09645292.2020.1719980>.

HANUSHEK, E. A., SCHWERDT, G., WOESSMANN, L., ZHANG, L. (2017). General education, vocational education, and labor-market outcomes over the life-cycle. *Journal of Human Resources*, 52(1): 48–87. <https://doi.org/10.3368/jhr.52.1.0415-7074R>.

HANUSHEK, E., SCHWERDT, G., WIEDERHOLD, S., WOESSMANN, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73: 103–130. <http://dx.doi.org/10.1016/j.euroecorev.2014.10.006 0014-2921>.

HORVÁTH, R. (2011). Research & Development and LongTerm Economic Growth: a Bayesian Model Averaging Analysis [online]. <https://core.ac.uk/download/pdf/6480625.pdf>. doi:10.1016/j.econmod.2011.08.007
HRONOVÁ, S. (2019). Economic Behaviour of the General Government and Sustainability of Public Finances – Comparative Analysis of the Czech Republic and Selected EU Countries. *Statistika: Statistics and Economy Journal*, 99(4): 350–368.
LARSON, G., LOAYZA, N., WOOLCOCK, M. (2016). *The Middle-Income Trap: Myth or Reality? Research Policy Briefs From the World Bank Malaysia Hub* [online]. <http://documents.worldbank.org/curated/en/9655114.68194956837/pdf/104230-VISEDRevised-RPB-1-Middle-Income-Trap.pdf>.
NASCIA, L., PIANTA, M. (2018). Research and innovation policy in Italy. Munich Personal Archive. RePEc [online]. *MPRA Paper*, No. 89510, posted 24 Oct. <https://mpra.ub.uni-muenchen.de/89510/1/MPRA_paper_89510.pdf>.
PASTOR, J. M., PERAITA, C., SERRANO, L., SOLER, A. (2018). Higher education institutions, economic growth and GDP per capita in European Union countries. *European Planning Studies*, 26(8): 1616–1637. <https://doi.org/10.1080/09654313.2018.1480707>.
POWER, L., KERRY, A. (2015). Millington and Stephanie Bengtsson [online]. *Building Capacity in Higher Education Topic Guide, HEART.* <http://www.heart-resources.org/wpcontent/uploads/2015/09/Capacity-Building-in-Higher-Education-Topic-Guide.pdf?x30250>.
RÖMISCH, R. (2020). *Covid-19 effects on Central Europe September* [online]. Interred Central Europe. <https://www.interreg-central.eu/Content. Node/discover/Covid-19-effects-on-Central-Europe.pdf>.
ROSZKO-WÓJTOWICZ, E., GRZELAK, M., LASKOWSKA, I. (2019). The Impact of R&D Expenditure on Productivity in the Manufacturing Industry in Poland. *Econometrics*, 23(4): 112–126. <https://doi.org/10.15611/eada.2019.4.08>.
RUNDLE, E. (2021). *How Education and Training Affect the Economy* [online]. Press release content from KISSP. <https://www.investopedia.com/articles/economics/09/education-training-advantages.asp>.
SVOBODY, D. (2007). *Výuka zeměpisu* [online]. <https://ostrava.educanet.cz/www/zemepis/index.htm>.
TOMIĆ, D., ŠIMURINA, J., JOVANOV, L. (2020). The Nexus between Economic Sentiment Indicator and Gross Domestic Product: a Panel Cointegration Analysis. *Zagreb International Review of Economics and Business*, 23(1): 121–140. <https://doi.org/10.2478/zireb-2020-0008>.
UNESCO. (2010). *The Central Role of Education in the Millennium Development Goals* [online]. MDG Summit High-Level Round Table, Secondary Education, New York: United Nations. <http://www.unesco.org/fileadmin/multimedia/hq/ed/ed_new/images/education_for_all_international_coordination_new/PDF/analyticalnote.pdf>.
UNESCO. (2012). *Global Education Monitoring Report* [online]. <http://unesdoc.unesco.org/images/0021/002180/218003e.pdf>.
WORLD ECONOMIC FORUM. *Annual Report 2015–2016* [online]. <https://www.weforum.org/reports/annual-report-2015-2016>.
WORLD ECONOMIC FORUM. *The Global Competitiveness Report 2016–2017.* Insight Report, 400 p.

ANNEX

Table A1 Real GDP (million Euro), National accounts indicator (ESA 2010)

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Austria	295 896.6	310 128.7	318 653.0	323 910.2	333 146.1	344 269.2	357 608.0	369 341.3	385 361.9	397 575.3
Czechia	157 920.8	165 202.2	162 587.5	159 461.5	157 821.3	169 558.2	177 438.5	194 132.9	210 927.8	223 950.3
Croatia	45 195.1	44 924.6	44 007.9	43 806.3	43 398.6	44 612.0	46 619.3	49 238.5	51 950.1	54 237.9
Germany	2 564 400.0	2 693 560.0	2 745 310.0	2 811 350.0	2 927 430.0	3 026 180.0	3 134 740.0	3 259 860.0	3 356 410.0	3 449 050.0
Hungary	99 576.3	102 020.6	99 984.0	102 034.3	106 061.3	112 701.0	116 129.8	126 891.0	135 931.0	146 061.8
Italy	1 611 279.4	1 648 755	1 624 358.7	1 612 751.3	1 627 405.6	1 655 355.0	1 695 786	1 736 593	1 771 063	1 789 747.0
Poland	362 190.9	379 860.0	387 947.0	392 310.7	408 967.8	430 465.8	427 091.8	467 426.6	497 842.3	532 329.2
Slovenia	36 363.9	37 058.6	36 253.3	36 454.3	37 634.3	38 852.6	40 443.2	43 009.1	45 862.6	48 392.6
Slovakia	68 188.7	71 304.5	73 575.8	74 448.8	76 269.8	79 767.6	81 051.5	84 532.2	89 505.5	93 865.2

Source: Eurostat (2021a), <*https://appsso.eurostat.ec.europa.eu/nui/show.do*>

Table A2 Population on 1 January (persons)

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Austria	8 375 164	8408121	8 451 860	8 507 786	8 584 926	8 700 471	8 772 865	8 822 267	8 858 775	8 901 064
Czechia	10 486 731	10505445	10 516 125	10 512 419	10 538 275	10 553 843	10 578 820	10 610 055	10 649 800	10 693 939
Croatia	4 289 857	4275984	4 262 140	4 246 809	4 225 316	4 190 669	4 154 213	4 105 493	4 076 246	4 058 165
Germany	80 222 065	80327900	80 523 746	80 767 463	81 19 7537	82 175 684	82 521 653	82 792 351	83 019 213	83 166 711
Hungary	9 985 722	9931925	9 908 798	9 877 365	9 855 571	9 830 485	9 797 561	9 778 371	9 772 756	9 769 526
Italy	59 364 690	59394207	59 685 227	60 782 668	60 795 612	60 665 551	60 589 445	60 483 973	60 359 546	60 244 639
Poland	38 062 718	38063792	38 062 535	38 017 856	38 005 614	37 967 209	37 972 964	37 976 687	37 972 812	37 958 138
Slovenia	2 050 189	2055496	2 058 821	2 061 085	2 062 874	2 064 188	2 065 895	2 066 880	2 080 908	2 095 861
Slovakia	5 392 446	5404322	5 410 836	5 415 949	5 421 349	5 426 252	5 435 343	5 443 120	5 450 421	5 457 873

Source: Eurostat (2021), *<https://ec.europa.eu/eurostat/databrowser/view/tps00001/default/table?lang=en>*

Table A3 Real GDP per capita (Euro)

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Austria	35 390	36 300	36 390	36 180	36 130	36 140	36 390	37 030	37 800	38 170
Czechia	15 020	15 310	15 170	15 160	15 480	16 290	16 670	17 490	17 990	18 330
Croatia	10 520	10 530	10 310	10300	10 310	10 630	11 100	11 600	12 040	12 450
Germany	31 940	33 200	33 280	33 330	33 920	34 130	34 610	35 380	35 720	35 840
Hungary	9 960	10 180	10 090	10 310	10 770	11 210	11 480	12 010	12 680	13 260
Italy	26 930	27 020	26 090	25 480	25 420	25 640	26 020	26 490	26 780	26 920
Poland	9 400	9 850	9 980	10 100	10 440	10 890	11 240	11 790	12 420	13 000
Slovenia	17 750	17 870	17 360	17 160	17 620	17 990	18 550	19 430	20 220	20 700
Slovakia	12 560	12 990	13 220	13 290	13 630	14 270	14 550	14 980	15 520	15 860

Source: Eurostat (2021c), *<https://appsso.eurostat.ec.europa.eu/nui/show.do>*

Table A4 Total expenditure on education (million Euro)

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Austria	15 090.7	15 506.4	15 932.6	16 195.5	16 324.2	16 869.2	17 322.7	17 728.3	18 401.3	19 083.6
Czechia	7 264.4	7 599.3	7 316.4	7 175.8	7 102.0	7 460.6	6 920.1	7 959.4	9 702.7	10 937.6
Croatia	1 672.2	1 617.3	1 628.3	1 708.4	1 649.1	2 944.4	3 076.8	2 609.6	2 753.3	2 904.4
Germany	112 833.6	115 823.1	118 048.3	123 694.4	125 879.5	127 099.6	128 524.3	133 654.3	140 969.2	148 180.2
Hungary	5 476.7	5 203.0	4 699.2	4 795.6	5 515.2	5 860.5	5 806.5	6 471.4	6 932.5	6 855.5
Italy	69 285.0	67 599.0	66 599.0	66 122.2	65 096.2	66 214.2	66 135.7	67 727.1	69 071.5	69 800.0
Poland	19 920.5	20 512.4	20 949.1	20 792.5	21 675.3	22 814.7	21 354.6	22 903.9	24 892.1	26 619.6
Slovenia	2 363.6	2 371.7	2 320.3	2 369.5	2 258.0	2 175.7	2 224.4	2 322.5	2 476.6	2 661.6
Slovakia	3 136.7	3 066	2 943	3 052.4	3 050.1	3 350.2	3 161.0	3 296.8	3 580.2	3 942.3

Source: Eurostat (2021a)

Table A5 Education expenditure (Euro per capita)

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Austria	1 801.8	1 844.2	1 885	1 903.6	1 901.5	1 938.8	1 974.5	2 009.5	2 077.2	2 144
Czechia	692.7	723.4	694.7	682.6	673.9	706.9	654.6	750.2	911.1	1 022.8
Croatia	389.8	379.5	382.0	402.2	390.3	702.6	740.6	635.6	675.4	715.6
Germany	1 406.5	1 441.9	1 466.0	1 531.5	1 550.3	1 546.7	1 557.5	1 614.3	1 698	1 733.6
Hungary	548.5	523.9	474.2	485.5	559.6	594.6	592.6	661.8	709.4	701.7
Italy	1 167.1	1 138.1	1 115.8	1 087.8	1 070.7	1 091.4	1 091.5	1 117.9	1 144.3	1 158.6
Poland	523.3	538.9	550.4	546.8	570.3	600.9	562.4	603.1	655.2	701.2
Slovenia	1 152.9	1 153.8	1 127.0	1 149.6	1 094.6	1 054.0	1 076.7	1 123.7	1 190.2	1 269.9
Slovakia	581.7	567.3	543.9	563.6	562.6	617.4	581.6	604.9	656.9	733.3

Source: Authors calculations based on Eurostat data

Table A6 Internal expenditure on R&D (million Euro)

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Austria	8 066.4	8 276.3	9 287.8	9 571.3	10 275.2	10 499.1	11 145.0	11 289.8	12 110.2	12 688.8
Czechia	2 095.1	2 552.0	2 877.3	2 996.7	3 090.7	3 250.2	2 963.3	3 433.3	4 006.5	4 348.3
Croatia	335.1	336.4	330.0	354.7	339.9	374.8	402.4	423.5	501.8	600.8
Germany	70 014.2	75 569.1	79 110.4	79 729.5	84 246.8	88 781.8	92 173.6	99 553.6	104 669.0	109 544.4
Hungary	1 126.1	1 204.6	1 257.3	1 415.1	1 428.8	1 510.9	1 371.7	1 672.9	2 051.4	2 158.6
Italy	19 624.9	19 810.6	20 502.5	20 983.1	21 781.3	22 157.0	23 171.6	23 793.7	25 232.2	25 909.6
Poland	2 607.5	2 836.2	3 429.9	3 436.3	3 864.0	4 316.5	4 112.3	4 834.0	6018.5	7 046.9
Slovenia	745.9	894.2	928.3	935.0	890.2	853.1	812.0	802.3	892.7	989.3
Slovakia	416.4	468.4	585.2	610.9	669.6	927.3	640.8	749.0	750.9	776.6

Source: Authors calculations based on Eurostat data

Table A7 Internal expenditure on R&D (Euro per capita)

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Austria	965.9	988.2	1 104.6	1 132.4	1 207.7	1 223	1 281	1 286.9	1 372.7	1 432.3
Czechia	200.3	243.4	273.9	285.0	294.0	308.4	280.8	324.5	377.6	408.3
Croatia	77.9	78.4	77.2	83.2	80.0	88.7	96.0	101.9	122.2	147.4
Germany	855.9	942.0	984.8	990.1	1 043.1	1 093.4	1 121.7	1 206.4	1 264.2	1 316.8
Hungary	112.4	120.6	126.6	142.8	144.7	153.3	139.5	170.8	209.8	220.9
Italy	331.6	333.7	345.2	351.6	358.3	364.5	382	392.7	417.2	429.3
Poland	68.6	74.5	90.1	90.3	101.6	113.6	108.3	127.3	158.5	185.6
Slovenia	364.4	436.2	451.6	454.1	431.9	413.5	393.4	388.4	431.9	475.4
Slovakia	77.2	86.9	108.3	112.9	123.6	171.0	118.1	137.8	138.0	142.5

Source: Eurostat (2021c), <*https://appsso.eurostat.ec.europa.eu/nui/show.do*>

Table A8 Pooled Regression Model estimates

```
. * Pooled OLS estimator

.
. reg $ylist $xlist
```

Source	SS	df	MS
Model	7.5623e+09	2	3.7811e+09
Residual	601505049	87	6913851.14
Total	8.1638e+09	89	91728049.2

Number of obs	=	90
F(2, 87)	=	546.89
Prob > F	=	0.0000
R-squared	=	0.9263
Adj R-squared	=	0.9246
Root MSE	=	2629.4

Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
X1	7.12156	2.418518	2.94	0.004	2.314494	11.92863
X2	13.23238	2.032005	6.51	0.000	9.193551	17.27121
_cons	3853.903	1076.697	3.58	0.001	1713.851	5993.955

Source: Authors' calculations by using STATA

Table A9 Between Regression Model estimates

```
Between regression (regression on group means)  Number of obs     =         90
Group variable: id                              Number of groups  =          9

R-sq:                                           Obs per group:
     within  = 0.6090                                         min =         10
     between = 0.9376                                         avg =       10.0
     overall = 0.9262                                         max =         10

                                                F(2,6)            =      45.05
sd(u_i + avg(e_i.))=  2897.838                  Prob > F          =     0.0002
```

Y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
X1	6.4787	9.777829	0.66	0.532	-17.44679	30.40419
X2	14.07058	8.220808	1.71	0.138	-6.045009	34.18618
_cons	3307.643	4244.991	0.78	0.465	-7079.475	13694.76

Source: Authors' calculations by using STATA

Table A10 Fixed Effects Regression Model estimates

```
Fixed-effects (within) regression               Number of obs     =         90
Group variable: id                              Number of groups  =          9

R-sq:                                           Obs per group:
     within  = 0.6091                                         min =         10
     between = 0.9376                                         avg =       10.0
     overall = 0.9263                                         max =         10

                                                F(2,79)           =      61.55
corr(u_i, Xb)  = 0.8791                         Prob > F          =     0.0000

------------------------------------------------------------------------------
           Y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          X1 |   3.308877   1.334922     2.48   0.015     .6517805    5.965974
          X2 |   6.742611   1.100305     6.13   0.000     4.552509    8.932713
       _cons |   11851.77   803.4874    14.75   0.000     10252.47    13451.07
-------------+----------------------------------------------------------------
     sigma_u |  5576.6761
     sigma_e |  681.79508
         rho |  .98527303   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(8, 79) = 151.87                     Prob > F = 0.0000
```

Source: Authors' calculations by using STATA

Table A11 Random Effects Regression Model estimates

```
Random-effects GLS regression                   Number of obs     =         90
Group variable: id                              Number of groups  =          9

R-sq:                                           Obs per group:
     within  = 0.6087                                         min =         10
     between = 0.9375                                         avg =       10.0
     overall = 0.9263                                         max =         10

                                                Wald chi2(2)      =     160.03
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000
theta          = .92559883

------------------------------------------------------------------------------
           Y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          X1 |   4.234197   1.437408     2.95   0.003      1.41693    7.051464
          X2 |   7.496329   1.187942     6.31   0.000     5.168007    9.824652
       _cons |   10716.04   1348.892     7.94   0.000     8072.257    13359.82
-------------+----------------------------------------------------------------
     sigma_u |  2889.8064
     sigma_e |  681.79508
         rho |  .94727158   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Source: Authors' calculations by using STATA

Table A12 Breusch and Pagan test results

```
Breusch and Pagan Lagrangian multiplier test for random effects

        Y[id,t] = Xb + u[id] + e[id,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                       Y |   9.17e+07       9577.476
                       e |   464844.5       681.7951
                       u |    8350981       2889.806

        Test:   Var(u) = 0
                             chibar2(01) =   274.96
                          Prob > chibar2 =   0.0000
```

Source: Authors' calculations by using STATA

Table A13 Hausman test results

```
. hausman fixed random

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
          X1 |    3.308877     4.234197       -.9253197                .
          X2 |    6.742611     7.496329       -.7537182                .
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =  -138.87    chi2<0 ==> model fitted on these
                                        data fails to meet the asymptotic
                                        assumptions of the Hausman test;
                                        see suest for a generalized test
```

Source: Authors' calculations by using STATA

Table A14 Hausman test results for the 1st cluster

```
. hausman fixed random

                 ——— Coefficients ———
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random        Difference          S.E.
-------------+----------------------------------------------------------------
          X1 |    3.308877     4.234197       -.9253197               .
          X2 |    6.742611     7.496329       -.7537182               .
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =  -138.87    chi2<0 ==> model fitted on these
                                        data fails to meet the asymptotic
                                        assumptions of the Hausman test;
                                        see suest for a generalized test
```

Source: Authors' calculations by using STATA

Table A15 Hausman test results for the 2nd cluster

```
. hausman fixed random

                 ——— Coefficients ———
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random        Difference          S.E.
-------------+----------------------------------------------------------------
          X1 |    15.25227     15.36214       -.1098746        1.256156
          X2 |    4.111231     4.074996        .0362358        .2873838
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        0.02
                Prob>chi2 =      0.9916
```

Source: Authors' calculations by using STATA

Approximate Valuation of Life Insurance Portfolio with the Cluster Analysis: Trade-Off Between Computation Time and Precision

Jan Fojtík[1] | *Prague University of Economics and Business, Prague, Czech Republic*
Jiří Procházka[2] | *Prague University of Economics and Business, Prague, Czech Republic*
Pavel Zimmermann[3] | *Prague University of Economics and Business, Prague, Czech Republic*

Abstract

Valuation of the insurance portfolio is one of the essential actuarial tasks. Life insurance valuation is usually based on a projection of cash flows for each policy which is demanding computation time. Furthermore, modern financial management requires multiple valuations under different scenarios or input parameters. A method to reduce computation time while preserving as much accuracy as possible based on cluster analysis is presented. The basic idea of the method is to replace the original portfolio by a smaller representative portfolio based on clusters with some weights that would ensure the similarity of the valuation results to the original portfolio. Valuation is then significantly faster but requires initial time for clustering and the results are only approximate – different from the original results. The difference is studied for a different number of clusters and the trade-off between the approximation error and calculation time is evaluated.

Keywords	DOI	JEL code
Life insurance, portfolio valuation, cash flow projection, cluster analyses	*https://doi.org/10.54694/stat.2021.23*	*C38, C63, G22*

INTRODUCTION

The proper valuation of the life insurance portfolio is one of the essential actuarial tasks. In general, there are several metrics describing the portfolio value such as the liability value, profit or loss or distributive

[1] Faculty of Informatics and Statistics, Prague University of Economics and Business, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: xfojj00@vse.cz.
[2] Faculty of Informatics and Statistics, Prague University of Economics and Business, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: xproj16@vse.cz.
[3] Faculty of Informatics and Statistics, Prague University of Economics and Business, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: zimmerp@vse.cz.

earnings. These metrics will be referred to as the economic metrics. Calculation of these *economic metrics* will be referred to as the *portfolio valuation*. Proper calculation of these metrics is important for reporting or accounting purposes, especially with new legislation of IFRS 17 or Solvency II. Actual and accurate valuation of life insurance portfolio has also a great importance in management decision or creating future business plans.

Ordinary methods for valuating life insurance products are usually not complicated or mathematically sophisticated. The valuation is based on per-policy projection – projecting the expected development and economic metrics of each policy (Selimovic, 2010) which is computationally demanding. Furthermore, actuaries usually need to valuate the portfolio multiple times under different assumptions of expected reality – e.g. valuation on different interest, mortality or lapse rates. For this purpose, the portfolio needs to be valuated under a wide range of different scenarios (Giamouridis et al., 2016; or Kaucic and Daris, 2015). Even with the newest technologies such as optimized actuarial software and powerful hardware, the results are derived with significant delay. For example, for a portfolio of a smaller insurance company with 300 000 policies, projecting 1 000 scenarios for 50 years (600 months), valuation may, depending on the performance of hardware and software, last even a month. The valuation time increases with the portfolio size or the number of required scenarios. The results derived with such a delay may be outdated or not satisfying actual market assumptions. Therefore, new opportunities of faster valuation present an active area of research.

Several researches suggest that data-mining methods can be used to solve this task (see Mohammed et al., 2016; or Devale and Kulkarni, 2012). A recent study (Janecek, 2017) presents two solutions for faster liability modeling based on mathematical proxy-functions and interpolations among different interest rate scenarios. Authors Freedman and Reynold (2008), and Fojtik et al. (2017) suggest cluster analysis as a good alternative way to accelerate valuation of life insurance portfolio. Presently, the application of cluster analysis has been in life insurance used mainly for the client segmentation purposes (Jandaghi et al., 2015), or experience analysis and assumption-setting process (Purushotham, 2016). Another approach for faster life insurance portfolio valuation is based on least squares Monte Carlo techniques. These techniques are studied in Turnbull (2014), Nteukam et al. (2014), and Krah and Nikolič (2018).

The approach researched in this paper is based on cluster analysis and will be referred to as *clustering approach*. This approach can be described in the following steps:

1. Calculate the economic metrics for each policy of the original portfolio based on a small number of initial runs – e.g. the basic (best estimate) scenario and few stress scenarios for the most important inputs.
2. Perform cluster analysis on economic metrics and potentially even on the policy parameters for a fixed, reasonably chosen number of clusters. Find a suitable representative (one policy) for each cluster. The portfolio of representatives will be referred to as the *reference portfolio*.
3. Calculate the projections of the economic metrics for all further scenarios only for the representatives of the clusters.
4. Find suitable weights of the representatives ensuring that the projection of economic metrics of the weighted reference portfolio will be as close as possible to the projection of the economic metrics of the original portfolio.

The more clusters (and hence representatives) are used, the slower the calculation and more precise the approximation. This constitutes a trade-off between computation time and accuracy. Analysis of this trade-off is the main contribution of this article.

In the first chapter the most typical method for life-insurance portfolio valuation is introduced including detail description of each component and calculation. The second chapter presents the alternative approach for faster valuation of life-insurance portfolio based on clustering analysis including basic parametrization, advantages and settings. The third chapter presents the main analysis of this paper.

In this section the alternative valuation approach is applied on artificial portfolio and its time and accuracy performance is discussed.

1 PROJECTING CASH FLOW PER EACH POLICY

The most common methods for life-insurance portfolio valuation are based on the cash flow projection. Typically, these models are a projection of future cash in-flows and out-flows for each policy or an aggregate of policies. Modelling all policies one by one will be referred to as *per-policy projection*. A detailed description of per-policy projection can be found in Dickson and Hardy (2013).

1.1 Components of cash flow model

In this section, we present a cash flow model that can be considered as a typical valuation approach for life insurance policies. Specific companies of course have specific variations of these models. However, it is not essential (nor possible) for this analysis to cover all particular deviations of valuation models. The typical components of the cash flow model are:

In-flows:

- Premium.

Out-flows:

- Claims,
- Surrenders,
- Maturities,
- Commissions,
- Expenses.

The cash flow is calculated as the difference between the expected in-flows and the expected out-flows. The cash flow $CF_{i,t}$ formula for the i^{th} policy at the time t is:

$$CF_{i,t} = EPrem_{i,t} - EClaims_{i,t} - ESurr_{i,t} - EMat_{i,t} - EComms_{i,t} - EExpens_{i,t}\,, \tag{1}$$

where the $EPrem_{i,t}$ is the expected value of premium paid at the beginning of the period t. The $EClaims_{i,t}$ is the expected value of coverage paid for loss or policy event that occurs during period t. The $ESurr_{i,t}$ is the expected value of surrenders paid for early contract cancellation at the end of the period t before the end of the policy period. The $EMat_{i,t}$ is the expected value of maturities paid at the end of the policy. The ECommsi,t is the value of commissions and $EExpens_{i,t}$ is the value of expenses expected to be paid at the beginning of the time period t. Cash flows and its components are projected in discrete time intervals indexed by t (months or years). The expected values of each component are the nominal values adjusted by the probability that the component will be paid.

The probability of death and the probability of surrender (pre-mature end of contract) are involved in the model. The notation q_x refers to the probability of death for a person in age x. The s_t denotes the probability of surrender in policy period t.

1.1.1 Expected premium

The premium $Prem_{i,t}$ is paid at the begin of the period t only if the insured person is alive and the contract has not been cancelled before. The expected value of premium for the i^{th} contract and period t is:

$$EPrem_{i,t} = \begin{cases} Prem_{i,t} & t = 1 \\ Prem_{i,t} \prod_{j=1}^{t-1} (1 - q_{x+j-1})(1 - s_j) & t > 1 \end{cases}, \tag{2}$$

where x is the age of insured person at the beginning of the valuation. The premium is paid at the beginning of the projected period therefore, the first payment in time $t = 0$ is not adjusted by any probability.

1.1.2 Expected claim

The claim $Claim_{i,t}$ is paid only if the accident happened in period t, the insured person was alive before the accident occurred and the contract has not been cancelled before the accident. The claim expected value for the i^{th} contract in time t is:

$$EClaim_{i,t} = \begin{cases} Claim_{i,t}(1 - s_t) & t = 1 \\ Claim_{i,t}\, q_{x+t-1}(1 - s_t) \prod_{j=1}^{t-1}(1 - q_{x+j-1})(1 - s_j) & t > 1 \end{cases}, \tag{3}$$

where x is the age of the insured person at the beginning of the valuation. The claim is assumed to be paid at the end of projected period t.

1.1.3 Expected surrender

The surrender value $Surr_{i,t}$ is paid only if the contract is cancelled before the end of the policy period and the insured person is still alive. The expected surrender value for i^{th} contract in time t is:

$$ESurr_{i,t} = \begin{cases} Surr_{i,t}(1 - q_x)s_t & t = 1 \\ Surr_{i,t}\, q_{x+t-1}(1 - q_{x+t-1})s_t \prod_{j=1}^{t-1}(1 - q_{x+j-1})(1 - s_j) & t > 1 \end{cases}, \tag{4}$$

where x is the age of the insured person at the beginning of the contract. The surrender is assumed to be paid at the end of projected period t.

1.1.4 Expected maturity

The maturity value $Mat_{i,t}$ is paid only if the contract has not cancelled and the insured person is alive at the end of the contract in time T. As the maturity is paid only at the end of the contract the expected maturity value before the end of contract is equal to zero. For the i^{th} contract and time equal to end policy period T the expected maturity is:

$$EMat_{i,t} = \begin{cases} 0 & t < T \\ Mat_{i,t} \prod_{j=1}^{t}(1 - q_{x+j-1})(1 - s_j) & t = T \end{cases}, \tag{5}$$

1.1.5 Expected commissions

The commission $Comms_{i,t}$ is paid at the beginning of the period t only if the insured person is alive and the contract has not been cancelled before. The expected commission value for the i^{th} contract in time t is:

$$EComms_{i,t} = \begin{cases} Comms_{i,t} & t = 1 \\ Comms_{i,t} \prod_{j=1}^{t-1}(1 - q_{x+j-1})(1 - s_j) & t > 1 \end{cases}, \tag{6}$$

where x is the age of the insured person at the beginning of the contract. The commission is assumed to be paid at the beginning of the projected period therefore, the first commissions in time $t = 0$ is not adjusted by probabilities.

1.1.6 Expected expenses

The expenses $Expens_{i,t}$ are paid in period t only if the insured person is alive and the contract has not been cancelled before. The expected value of expenses for the i^{th} contract in time t is:

$$EExpens_{i,t} = \begin{cases} Expens_{i,t} & t = 1 \\ Expens_{i,t} \prod_{j=1}^{t-1} (1 - q_{x+j-1})(1 - s_j) & t > 1 \end{cases}, \tag{7}$$

where x is the age of the insured person at the beginning of the contract. The expenses are assumed to be paid at the beginning of the projected period therefore, the first expenses in time $t = 0$ are not adjusted by probabilities.

1.2 Economic metrics

In this section the calculation of basic economic metrics is presented.

1.2.1 Present value of liability

The portfolio value is usually calculated as the present value of future cash flows *PVCF*. The present value of future cash flows for i^{th} policy is calculated as discounted sum of cash flows:

$$PVCF_i = \Sigma_t CF_{i,t} v^t , \tag{8}$$

where v^t is the flat discount factor at projection time t. For simplicity PVFC will be referred to the value of liability. The total value of the liability for the whole portfolio is sum of *PVCF* for each policy.

1.2.2 Present value of profit and loss

The profit of the i^{th} policy in projection time t is calculated as:

$$Profit_{i,t} = CF_{i,t} + \Delta Reserve_{i,t} + InvIncome_{i,t} , \tag{9}$$

where the $\Delta Reserve_{i,t}$ is the change of reserve (change of fund value between times $t - 1$ and t) and $InvIncome_{i,t}$ is the investment income realized between times $t - 1$ and t. The present value of profit for i^{th} policy is calculated as discounted sum of individual profits as:

$$PVprof_i = \Sigma_t Profit_{i,t} v^t, \tag{10}$$

where v^t is the flat discount factor at projection time t. The portfolio present value of profit is calculated as sum of present values of profits for all policies.

1.2.3 Present value of premium

The present value of premium for i^{th} policy in time t is calculated as discounted sum of individual expected premium payments:

$$PVPrem_i = \Sigma_t Eprem_{i,t} v^t, \tag{11}$$

where v^t is the flat discount factor at projection time t. The portfolio present value of premium is calculated as sum of present values of premium for all policies.

2 CLUSTER ANALYSIS AND PARAMETER SETTING

As stated above, some acceleration techniques focus on replacing the original portfolio by representatives of clusters. The cash flow projection is then performed for these representatives since there is potentially a major reduction of computation time. This is however connected with certain inaccuracy. Inaccuracy in this context means the difference between the projection results of the original portfolio and the weighted projection results of cluster representatives for a given scenario.

There are two essential decisions within the clustering approach to make. The number of clusters required and the set of clustering variables to be used. The number of clusters as well as the number of clustering variables increases computational time on one hand but may increase the accuracy on the other hand. A reasonably selected set of clustering variables may also increase the accuracy. Therefore, both should be optimized at least to some extent. The number of clusters represents the size of the reference portfolio which should be manually pre-selected. In the following analysis, the different number of clusters will be analysed to present its accuracy and speed trade-off.

2.1 Clustering variables

There are two distinct types of clustering variables available in this task. The first type is the basic policy characteristics such as age, gender, premium, sum assured or policy reserve. The second type is the economic metrics such as profits, premiums, claims or even cash flows in individual projected periods or intervals and its sensitivities to certain stress scenarios. The character of both types of clustering variables is very different. The economic metrics describe more the dynamics of the policy rather than position as the basic characteristics.

It has been confirmed in Fojtik (2017), that for the purposes of scenario valuation using the economic metrics as clustering variables leads to significantly better results than using the policy characteristics. Clustering variables assumed in this study are:

- the present values of future cash flows,
- the present value of profit or loss,
- the present value of premium,
- the sum of cash flows in the first five years,
- the sum of profits in the first five years,
- the sum of claims in the first five years,
- the sum of expenses and commissions in the first five years,
- the sum of premium in the first five years.

The selected clustering variables represent the common metrics describing the insurance portfolio from economical perspective. When using clustering variables, the problem of the different scale of nominal values may arise. To avoid this problem, it is advised to standardize the data before clustering. The standardization process consists of subtracting the mean and dividing it by the standard deviation.

2.2 Clustering algorithm

For the purposes of this paper, the non-hierarchical medoid based algorithm CLARA (Clustering LARge Application) was applied (Ng and Han, 2002). The CLARA algorithm is suitable for handling large datasets such as a life insurance portfolio (Hebak, 2013). In general, clusters are created by grouping similar observations. In the case of life insurance portfolio, clusters of policies with similar economic variables selected as clustering variables are created. The dissimilarity between the two policies is defined by the distance measure. In this paper, the dissimilarities are measured by the Euclidean distance between policies. The Euclidean distance is defined as the sum of squared differences between the i^{th} and the j^{th} policy:

$$d_{i,j} = \sqrt{\sum_{m=1}^{M} (Z_{m,i} - Z_{m,i})^2}\ , \qquad (12)$$

where $Z_{m,i}$ and $Z_{m,j}$ are the standardized values of the m^{th} clustering variable of the i^{th} respectively j^{th} policy. The setting of R function is as follows: clara (x = Clustering_data, k = K, samples = 200, rngR = TRUE, stand = TRUE, correct.d = TRUE, metric = "Euclidean", pamLike = TRUE). The data object Clustering_data is insurance portfolio with clustering variables only and K stands for the number of clusters (size of reference portfolio). For more information about additional parameters see the documentation of the clustering function in the package *cluster* (Maechler et al., 2021).

2.3 System of weights

A system of weights must be assigned to the reference portfolio in order to replicate the projection of the original portfolio. Authors Freedman and Reynold (2008) suggest using the number of policies in each cluster as weight. This ensures that the number of reference policies matches the size of the original portfolio. Another option of the weighting system is scaling by some financial variable. In this paper, we present a weight based on the ratio of the present values of cash flows between the original portfolio and the reference portfolio. The weights are calculated for each cluster individually on the basic (best estimate) scenario. The weight of the kth cluster is given by:

$$w_k = \frac{PVCF_k^{Orig}}{PVCF_k}, \tag{13}$$

where the $PVCF_k^{Orig}$ represents the total present value of cash flows of policies from the original portfolio belonging to the k^{th} cluster and the $PVCF_k$ is the present value of cash flows of the k^{th} representative. This ensures that the reference portfolio will, in the basic scenario, replicate the present value of cash flow of the original portfolio exactly.

The approximate total value of the m_{th} projected variable $\tilde{X}$ is for each scenario calculated as a weighted sum of the m_{th} projected variable $X_{m,k}$ of the representatives from the reference portfolio:

$$\tilde{X}_m = \sum_{k=1}^{K} w_k X_{m,k}. \tag{14}$$

The symbol K stands for the number of clusters given by the size of the reference portfolio which is set manually before the clustering. Note that the weights are built on the best estimate scenario but can be used for projecting other stress scenarios.

2.4 Error measure

As we are trying to replicate the results of the per-policy projection of the original portfolio, the error in this context is the relative difference between the approximate value calculated by Formula (14) of the m_{th} projected variable and the corresponding variable of the original portfolio.

The error measure of the m_{th} variable is given by the following:

$$e_m = \frac{\tilde{X}_m}{X_m} - 1, \tag{15}$$

where the $\tilde{X}$ is obtained by Formula (14) and X_m is the total value of the m_{th} projected variable of the original portfolio. The total error of the reference portfolio is then defined as the average square root sum of squares over all selected variables as:

$$e = \frac{\sqrt{\sum_m e_m^2}}{M}. \tag{16}$$

It is advisable to measure the error only for the important variables in terms of actuarial modelling. In this paper, the error is measured on the clustering variable from section Clustering variables.

2.5 Computation time

The main goal of the clustering approach is to reduce the number of projected policies in order to speed up life insurance portfolio valuation with an acceptable level of inaccuracy. The total *computation time* consists of two components – *clustering time* and *valuation time*. The valuation time is required for projecting the policies and the clustering time is required for reducing the size of the original portfolio and building the reference portfolio.

Let's assume that the one scenario valuation of one policy by classical per-policy cash flow model takes in average time T_{avg}. The valuation time of $N_{scenarios}$ on the whole original portfolio of size $N_{policies}$ then lasts approximately:

$$T_{avg}\, N_{policies}\, N_{scenarios}\,. \tag{17}$$

In the case of the original portfolio, the total computation time is equal to the valuation time because no clustering is performed. But in the case of the reference portfolio, the total computation time is given by the sum of clustering and valuation time of the reference portfolio of size $N_{reference}$ as:

$$T_{clustering} + T_{avg} N_{reference}\, N_{scenarios}\,. \tag{18}$$

where the first component $T_{clustering}$ is the time required for clustering. For simplicity, let's assume that the average valuation time of one policy will remain approximately the same after the reduction. The acceleration by clustering approach is then:

$$\frac{T_{avg} N_{policies} N_{scenarios}}{T_{clustering} + T_{avg} N_{reference} N_{scenarios}}\,. \tag{19}$$

The section Analysis confirms that the size of the reference portfolio (number of clusters) increases both – the clustering time as well as the valuation time. The significant time saving is evident especially when testing more scenarios.

3 ANALYSIS

The goal of this analysis is to present an approach how to select the suitable number of clusters for the specific portfolio and the number of scenarios that preserves the high accuracy and significantly speeds up the portfolio valuation.

There are three essential aspects that need to be considered before selecting the number of clusters, namely:

- accuracy of the clustering approach,
- clustering time,
- total acceleration.

In this part, we present the relation between accuracy and clustering time with respect to the different number of clusters and the total acceleration for the different number of scenarios.

3.1 Experimental artificial portfolio

The analysis of the clustering approach is performed on an artificial life insurance portfolio that consists of universal-life insurance policies. The portfolio includes 100 000 policies. The 8 different policy products are ensuring a reasonable level of heterogeneity that may be observed in real portfolios. Each product has 12 500 policies. The products differ in the premium frequency, length of policy period or the system of benefit payments. The basic parameters of the portfolio are presented in Table 1. The artificial portfolio includes the basic policy characteristics and the metrics of economic profit based on best estimate assumptions.

In the first step of making the artificial portfolio, the basic policy characteristics were generated for each product individually.

Table 1 Basic overview of the artificial portfolio

Product	**A**	**B**	**C**	**D**
Average age	25	25	25	25
Average policy period	30	30	5	5
Max age	80	80	50	50
Term coefficient	1	1	0.25	0.25
Min policy period	10	10	5	5
Policy duration	10	10	1	1
Sum assured	500 000	500 000	500 000	500 000
Premium frequency	Regular	Single	Regular	Single
Benefit type	SA	SA	SA	SA
Product	**E**	**F**	**G**	**D**
Average age	30	30	30	30
Average policy period	30	30	5	5
Max age	80	80	50	50
Term coefficient	1	1	0,25	0,25
Min policy period	10	10	5	5
Policy duration	10	10	1	1
Sum assured	1 000 000	1 000 000	1 000 000	1 000 000
Premium frequency	Regular	Single	Regular	Single
Benefit type	SA+CV	SA+CV	SA+CV	SA+CV

Source: Own construction

The policy characteristics were generated as follows:

- The age of the client at the start of the valuation is generated from Poisson distribution with the specific mean for each product presented in Table 1.
- The policy period was calculated as follows:
    1. Firstly, the maximal length of the policy period *h* is calculated from the age of the client obtained from the previous step to a maximum possible age considered. The maximal possible age is presented in Table 1.

2. Secondly, the maximal period *h* is multiplied by a random variable generated from uniform distribution with minimum set to 0.1 and maximum to the term coefficient presented in Table 1.
3. In the last step, the minimal length of the policy period for each product is ensured by parameter *Min policy period* from Table 1.

- The policy duration is given by the policy period obtained from the previous step multiplied by a specific coefficient l generated from uniform distribution with minimum set to 0 and maximum 1. The l coefficient ensures that the policy duration is lower than the policy period for each contract.
- Sum assured was generated from normal distribution with the same mean and standard deviation parameter (Sum assured). This parameter can be seen in Table 1. To eliminate negative or very low values, the lower bound of the sum assured was set as 10 000.
- The premium was calculated by the deterministic pricing formulas for premium (see Cipra, 2014), for two specific types of benefit payments:
  – SA: the benefit of sum assured is paid only in case of death. The premium is calculated as for the term insurance products.

Figure 1 Examples of Cash-Flow projection in artificial portfolio

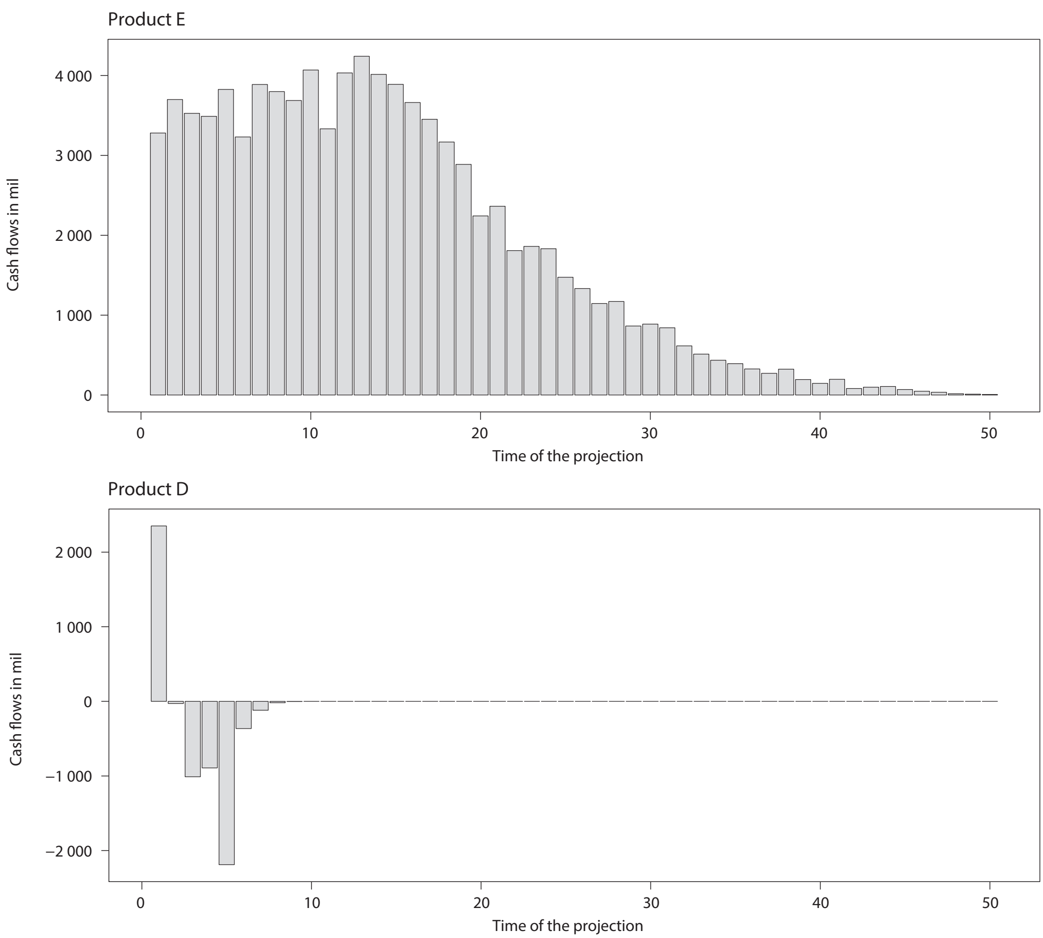
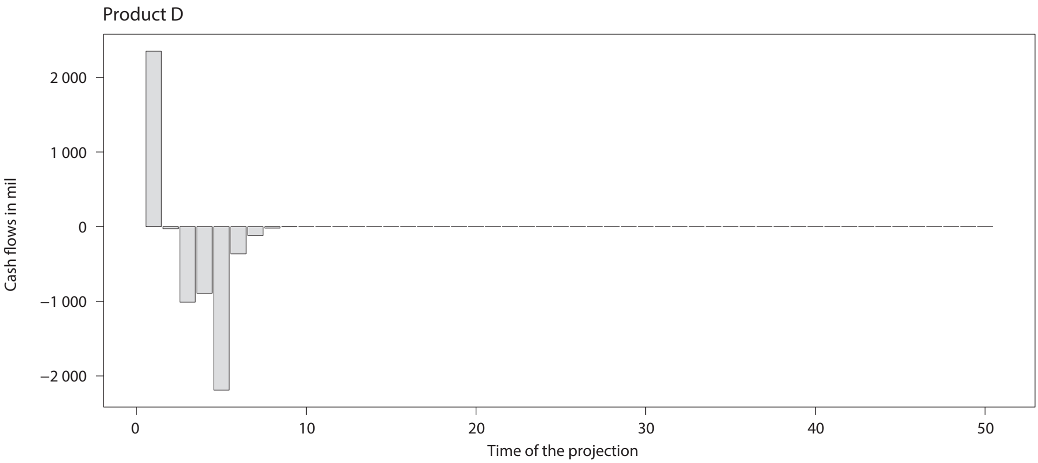
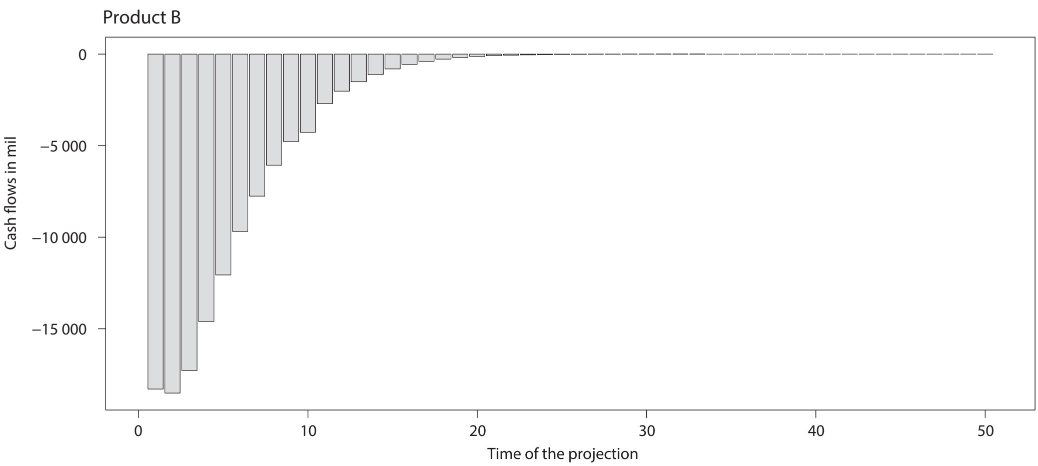


Figure 1 (continuation)





Source: Own construction

- SA + CV: the benefit is paid in two cases – death and maturity (surviving to the end of the policy period). The premium is calculated as for the endowment insurance product where the death benefit is a sum assured and the survival benefit is the value of the fund at the end of the policy period.

- The fund value was calculated as a difference between premium paid over the policy duration with interest minus the expenses paid over the policy duration.

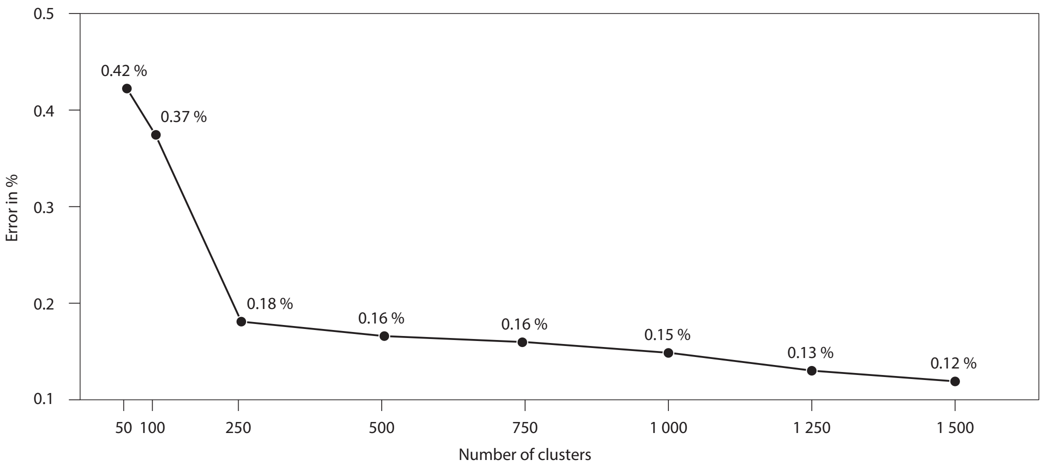
After generating the policy characteristics, the economic metrics are calculated by per-policy projection described in the section Components of cash flow model.

The examples of cash flows can be seen in Figure 1 on three selected products.

3.2 Number of clusters and the accuracy

Figure 2 presents the accuracy of the approximation for a different number of clusters used. The analysis is provided on the original portfolio designed in section Artificial portfolio. As stated previously, accuracy increases (error decreases) with the number of clusters. At first, the error decreases very fast. Somewhere around 250 clusters, the decrease of the error slows down significantly and continues steadily.

Figure 2 Relation between the accuracy and the number of clusters

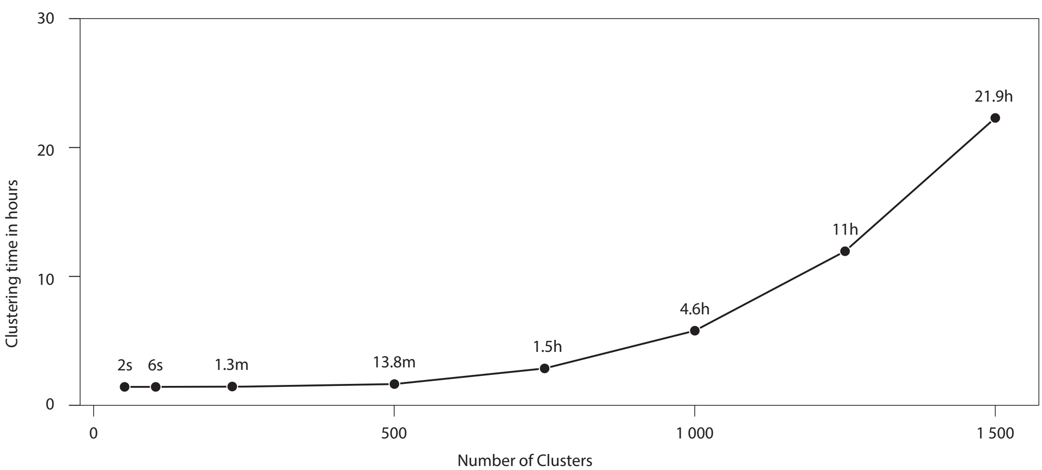


Source: Own construction

3.3 Number of clusters and the clustering time

Figure 3 presents the clustering time for the different number of clusters. The clustering time increases with the number of clusters. The increase is not linear but significantly faster. Therefore, the results for a high number of clusters (more than 10 000) may not be achieved in real time.

Figure 3 Relation between the clustering time and the number of clusters

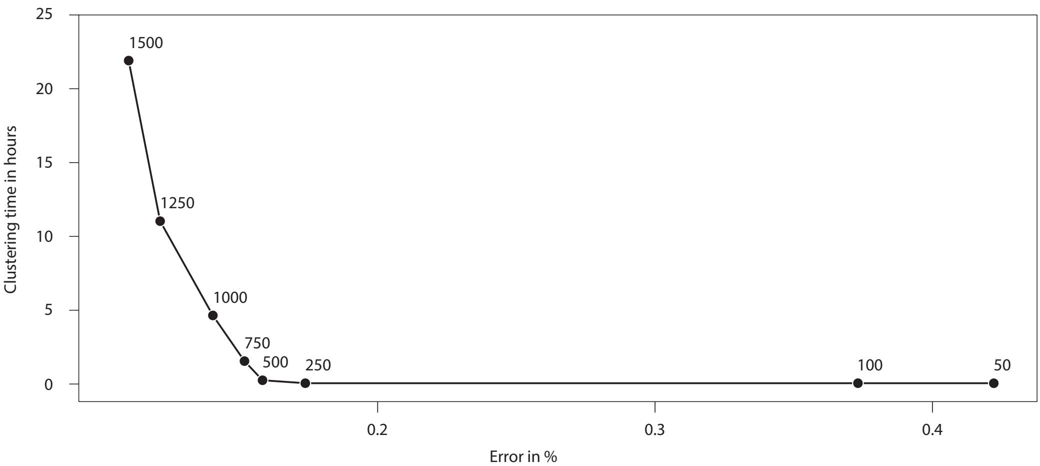


Source: Own construction

3.4 Accuracy and clustering time trade-off

Figure 4 puts the previous analysis together and presents the relation between the clustering time and accuracy achieved for the different number of clusters. The label of the line represents the number of clusters. The trade-off between the accuracy and clustering time, in this case, suggests that the reasonable number of clusters is somewhere between 250 and 500 where the additional increase in clustering time is not compensated by the significant increase in the corresponding accuracy.

Figure 4 Trade-off between the accuracy and clustering time



Source: Own construction

3.5 Acceleration of clustering approach

Table 2 present the acceleration calculated by formula 19 for the different number of clusters and one scenario. Acceleration naturally decreases towards 0 with the increasing number of clusters. For example, using the reference portfolio of 500 policies defined by the clustering approach seems to be beneficial because the whole calculation is 21 times faster already for one scenario. But using the reference portfolio of 1 500 policies is 4 times slower for one scenario as the valuation time and especially the clustering increase materially.

Table 3 presents the acceleration of the clustering approach for the different number of scenarios and the different number of clusters. The acceleration may differ for the different number of scenarios

Table 2 Reference portfolio acceleration for one scenario

Number of clusters	50	100	250	500	750	1 000	1 250	1 500
Valuation time	0.17	0.33	0.83	1.67	2.5	3.33	4.17	5
Clustering time	0.03	0.1	1.29	13.83	90.24	274	662	1 314
Calculation time	0.2	0.44	2.12	15.5	92.74	277	667	1 319
Acceleration	1 709	762	156	21	3.59	1.2	0.5	0.25

Source: Own construction

but usually raises with the number of scenarios. For the high number of scenarios, clustering time is negligible, and the acceleration is proportional to the number of clusters.

The results from Table 3 suggest that using 1 500 clusters for modelling only one scenario does not save any time but modelling 10 scenarios would be 2.44 times faster. The boundary $N^{+}_{Scenario}$ defines the minimal number of scenarios, where acceleration is higher than 1 (the clustering approach is beneficial). This boundary has the following:

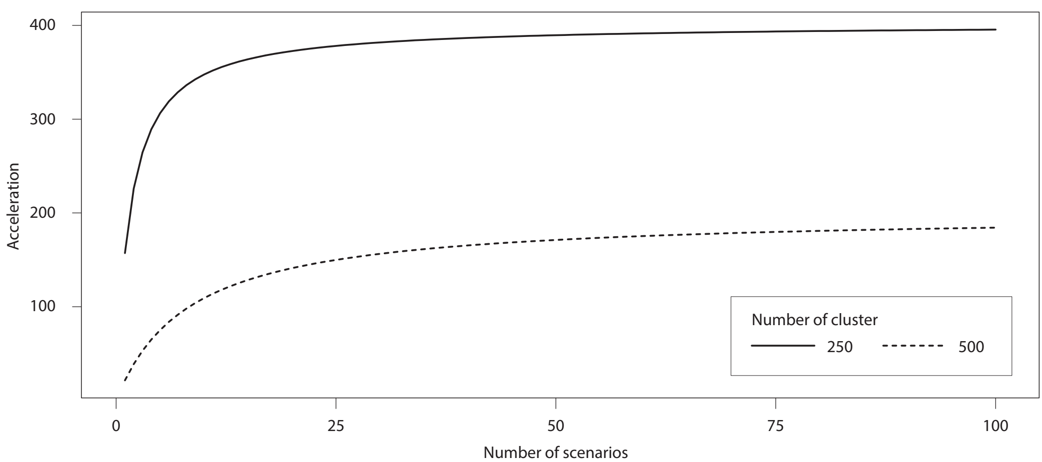
$$N^{+}_{Scenario} > \left[\frac{T_{Clustering}}{T_{avg}\,(N_{Orig} - N_{Refer})} \right], \qquad (20)$$

Table 3 Reference portfolio acceleration for one scenario

Number of clusters	Number of scenarios				
	1	10	50	100	1000
50	1 709.04	1 966.52	1 993.21	1 996.6	1 999.66
100	762.77	969.84	993.82	996.9	999.69
250	156.96	346.37	387.99	393.9	399.38
500	21.51	109.29	171.53	184.67	198.35
750	3.59	28.92	77.43	97.97	128.69
1 000	1.20	10.83	37.78	54.84	92.39
1 250	0.50	4.73	19.13	30.87	69.02
1 500	0.25	2.44	10.65	18.37	52.79

Source: Own construction

Figure 5 Comparision of acceleration for two clustering settings



Source: Own construction

For 1 500 clusters the clustering approach seems to be faster for at least 5 scenarios.

In section Accuracy and clustering time trade-off, it has been mentioned that a reasonable number of clusters based on error criteria should be between 250 or 500 clusters. Using the 250 clusters to replicate the original portfolio on a high number of scenarios (100 and more) is almost 400 times faster than modelling the original portfolio. Using the 500 clusters on the same number of scenarios is almost 200 times faster. This means that using the reference portfolio of 250 respectively 500 policies the analysts may test 400 respectively 200 times more scenarios in the same amount of time with a very high level of accuracy.

Figure 5 compares the acceleration between 250 and 500 clusters. The acceleration for 250 clusters dramatically increases when modelling a low number of scenarios and the growth slowly stabilizes after 20 scenarios. Using 500 clusters the acceleration grows slowly and doesa not stabilize so fast as using a lower number of clusters.

This task may be posted as an optimization task, where we search for maximum accuracy given the computation time available or minimum computation time for a given acceptable accuracy.

CONCLUSION

The proper valuation of the life insurance portfolio is one of the essential actuarial tasks. Traditionally used valuation techniques are based on modelling all policies of the portfolio which is time demanding. This takes effect, especially when valuating a high number of scenarios. Reducing the portfolio size in terms of the number of policies seems to be a good approach to speed up the computation time of the valuation.

Cluster analysis is one of the tools that can be applied to accelerate multiple scenario valuation of life insurance portfolio by reducing the size of the original portfolio into smaller reference portfolio. Results are on one hand obtained much faster as the per-policy projection is performed only for the reference portfolio. On the other hand, certain inaccuracy occurs as there is a difference between the projection results of the reference and the original portfolio.

The proper application of clustering approach requires the setting of several parameters such as selection of clustering variables and the suitable size of the reference portfolio determined by the number of clusters. The selection of clustering variables may increase the precision of the clustering approach. It can be advised to select clustering variables as the variables that the model should reproduce with the highest accuracy. The higher number of clusters may increase accuracy but also increase the computation time. When comparing the computation time, one has to include also the clustering time. The accuracy of the approximation is driven by the number of clusters used. An increasing number of clusters, on the other hand, increases both the clustering time as well as the valuation time of the reference portfolio. From our experiment, we may conclude:

1. The general level of error is relatively low.
2. An error of the approximation decreases with the increasing number of clusters at first relatively fast. At some point, the decrease slows down and continues steadily further at a slower rate.
3. For the high number of scenarios, the clustering time tends to be negligible, and the acceleration is proportional to the ratio of the size of the reference portfolio to the size of the original portfolio.
4. This means that the reasonable number of clusters is for our experiment somewhere between 250 and 500 as for the higher number of clusters, error rate decreases only slowly while as computation time increases rather fast and for the lower number of clusters the situation is opposite.

ACKNOWLEDGEMENT

The support of grant scheme Methods for fast estimation of life insurance liabilities with respect to different investment strategies IG410017 is gladly acknowledged.

References

CIPRA, T. (2014). *Financial and insurance formulas.* Physica Springer.

DEVALE, A. B., KULKARNI, R. V. (2012). Applications of data mining techniques in life insurance. *International Journal of Data Mining and Knowledge Management Process* (IJDKP), 2(4). <https://doi.org/10.5121/ijdkp.2012.2404>.

DICKSON, D. C., HARDY, M. R. (2013). *Actuarial Mathematics for life contingent risks.* 2nd Ed. Cambridge University Press, 616 p.

FOJTÍK, J. PROCHÁZKA, J. ZIMMERMANN, P. MACKOVÁ, S. ŠVEHLAKOVÁ, M. (2019). Alternative approach for fast estimation of life insurance liabilities. *20th Application of Mathematics and Statistics in Economics*, Sklarska Poreba, Poland. <https://doi.org/10.15611/amse.2017.20.11>.

FREEDMAN, A. REYNOLD, C., W. (2008). *Cluster analysis: A spatial approach to actuarial modeling* [online]. Milliman research report. [cit. 13.9.2021]. <https://bit.ly/2Uq48Kr>.

GIAMOURIDIS, D. SAKKAS, A., TESSAROMATIS, N. (2016). Dynamic Asset Allocation with Liabilities. *European Financial Management*, 23(2): 254–291. <https://doi.org/10.1111/eufm.12097>.

HEBAK, P. (2013). *Statistické myšlení a nástroje analýzy dat.* Informatorium.

JANDAGHI, G., MOAZZEZ, H., MORADPOUR, Z. (2015). *Life Insurance Customers segmentation using fuzzy clustering* [online]. Tehran: Faculty of Management and Accounting, Farabi College, University of Tehran. [cit. 13.9.2021]. <https://bit.ly/2Vyrga2>.

JANEČEK, M. (2017). Acceleration Techniques for Life Cash Flow Projection Based on Many Interest Scenarios – Cash Flow Proxy Functions [online]. *Czech Actuarial Society*. <https://bit.ly/2MBPtcb>.

KAUCIC, M., DARIS, R. (2015). Multi-Objective Stochastic Optimization Programs for a Non-Life Insurance Company under Solvency Constraints. *Risks*, 3(3): 390–419. <https://doi.org/10.3390/risks3030390>.

KRAH, A. S., NIKOLIČ, Z., KORN, R. (2018). A Least-Squares Monte Carlo Framework in Proxy Modeling of Life Insurance Companies. *Risks*, 6(2): 62. <https://doi.org/10.3390/risks6020062>.

MAECHLER, M., ROUSSEEUW, P., STRUYF, A., HUBERT, M., HORNIK, K. (2021). *Cluster: Cluster Analysis Basics and Extensions* [online]. R package version 2.1.2. <https://CRAN.R-project.org/package=cluster>.

MOHAMMED, M., YOUSSEF, B., TAOUFIQ, G. (2016). Time-Saving Approach for Optimal Mining of Association Rules. *International Journal of Advanced Computer Science and Applications*, 7(10). <https://doi.org/10.14569/IJACSA.2016.071031>.

NG, R. T., HAN, J. (2002). CLARANS: A method for clustering objects for spatial data mining. *IEEE transactions on knowledge and data engineering*, 14(5): 1003–1016. <https://doi.org/10.1109/TKDE.2002.1033770>.

NTEUKAM, O., JIAEN, R., PLANCHET, F. (2014). Internal model in life insurance: Application of least squares Monte Carlo in risk assessment. *Ekonomia. Rynek, Gospodarka, Społeczeństwo*, 41: 81–93. <http://dx.doi.org/10.17451/eko/41/2015/93>.

PURUSHOTHAM, M. (2016). Cluster analysis: Applications in experience analysis and assumption-setting [online]. *The Actuary.* <https://theactuarymagazine.org/cluster-analysis>.

SELIMOVIC, J. (2010). Actuarial Estimation of Technical Provisions' Adequacy in Life Insurance Companies. *Interdisciplinary Management Research*, 6: 523–533.

TURNBULL, C. (2014). Implementation of Least-Squares Monte Carlo (LSMC) in a Life Insurance Context – a Case Study [online]. *Moody's analytics risk perspectives, managing insurance risk*, 3. [cit. 13.9.2021]. <https://www.moodysanalytics.com/risk-perspectives-magazine/managing-insurance-risk/principles-and-practices/implementation-of-least-squares-monte-carlo-in-a-life-insurance-context>.

Social Progress Index for Urban and Rural Areas of a Region: Evidence from Peru

Maylee Inga-Hancco[1] | *Universidad Continental, Huancayo, Peru*
Adamari Indigoyen-Porras[2] | *Universidad Continental, Huancayo, Peru*
Sergio Parra-Alarcón[3] | *Universidad Continental, Huancayo, Peru*
Juan Cerrón-Aliaga[4] | *Universidad Continental, Huancayo, Peru*
Wagner Vicente-Ramos[5] | *Universidad Continental, Huancayo, Peru*

Abstract

The present study describes the methodological process proposed by the Social Progress Imperative Global Organization to calculate the Social Progress Index in urban and rural areas of the province of Huancayo, Peru, in 2020. The survey was based on 229 observations regarding basic human needs, foundations of well-being and opportunities. The result produced an index of 56.04 for urban areas and 53.98 for rural areas; results that are in the low and low middle range respectively, identifying deficiencies in the quality of economic policies, with respect to the sanitation service, where more than 30% do not have access to drinking water, and others. It was concluded that the index showed no improvement with respect to 2019, likewise the social gaps still persist and the well-being of the aforementioned population was not increased.

Keywords	***DOI***	***JEL code***
Social progress, social gaps, basic human needs, foundations of well-being, opportunities	*https://doi.org/10.54694/stat.2021.7*	*I31, I25, I28*

INTRODUCTION

The social progress of a nation leads to economic development and this favors economic growth, however, the opposite does not always happen. For this reason, some people consider "happiness" as a very relevant term when talking about social well-being, which in turn means not only satisfying the Basic Human

1 Professional Academic School of Economics, Faculty of Business Sciences, Universidad Continental, Av. San Carlos 1980, Huancayo, Perú. E-mail: 75801931@continental.edu.pe.
2 Professional Academic School of Economics, Faculty of Business Sciences, Universidad Continental, Av. San Carlos 1980, Huancayo, Perú. E-mail: 77354404@continental.edu.pe.
3 Professional Academic School of Economics, Faculty of Business Sciences, Universidad Continental, Av. San Carlos 1980, Huancayo, Perú. E-mail: 70305829@continental.edu.pe.
4 Professional Academic School of Economics, Faculty of Business Sciences, Universidad Continental, Av. San Carlos 1980, Huancayo, Perú. E-mail: jcerrona@continental.edu.pe.
5 Research coordination, Faculty of Business Sciences, Universidad Continental, Av. San Carlos 1980, Huancayo, Perú. E-mail: wvicente@continental.edu.pe.

Needs (BHN) of the population of a certain territory. A few years ago it was believed that the Gross Domestic Product (GDP) was an indicator that could measure the well-being of individuals, as the years went by they realized that when they talked about GDP, they only talked in monetary terms and I do not know how to evaluate other aspects that generate well-being for people. According to Greve (2016), the use of indices related to well-being and happiness could improve decision-making by including elements having also another type of value besides monetary measures.

Likewise, it should be considered that within a territory there are two important areas for the development of an individual, that is, in almost all the countries of the world there are rural and urban areas, of which several studies show that there exist significant differences to develop and achieve social well-being within the group designated at birth. The United Nations (2010) report states that the differences between these areas may vary according to the characteristics and needs that each country has, therefore, it would be a mistake to generalize the differences between the two. However, there are similarities: the standard of living and population density is higher in urban areas, but the most significant criteria to determine the difference between the two is the percentage of the economically active population in the agricultural sector, for a lifetime dignified with access to water and sewage, electricity, access to health and education. In Malaysia, Mansor et al. (2013) find that people born in urban areas have a better quality of life compared to people born in rural areas, in aspects such as education, health, communication, security and transportation. In the report developed by Arroyo et al (2018), it is mentioned that territorial inequality still persists in rural and urban areas in Latin America, where they reaffirm that the rural poor are in a worse situation than those in urban areas.

In Peru, there is sufficient information related to economic and social issues; both at the country level and at the regional level, but not at the provincial level given that within the Peruvian territory there is a diversity of territories and also a sociocultural diversity, this information that is available is not enough to be able to understand the growth or the development of the economic that exists.

Several provinces explain much of the departmental economic and social development results; it can be said that there is a bias between the areas within a department which may lead to overestimating or underestimating the results obtained, which are explained by the fact that cities show an accelerated population growth, but not an integration of the population at the same rate and that is why the level of poverty increases within the urban area, the number of informal settlements increases and informal employment also grows consequently it has a higher level of pollution, citizen security is reduced and there is a lower tax collection. Therefore, measuring the Social Progress Index at the provincial level and taking into account the differences that exist between urban and rural areas can reveal truly accurate information.

As Estes (2019) pointed out, in a review of the social progress of the continents of the world, where the continents with improvements in social development are Europe and North America, where there is a high commitment to the environment; the least developed regions were Africa and Asia, but now we see improvements in financial investments, where high advances in technology, development and innovation were visualized; in Latin America development is being seen with constant growth, however, there are weak points that remain to be improved, such as problems in the health system, high rates of population growth, government instability, civil conflict, among others; then it was determined that most countries can satisfy the BHN, however there are still great challenges to optimal well-being.

Several countries in the world develop and measure their social progress through various methods and indices, so that they can obtain a significant result enabling to apply better policies focused on these results. A study conducted by Bilan et al. (2019), in Ukraine showed that one of the main factors for having macro-economic stability was to have a good Social Progress Index (SPI). Clark et al. (2020), made a purchase of Ireland with the most countries of the European Union, and they considered that social statistics are important for a better approach to public policies. Another study conducted by Cárcaba et al. (2017), of social progress in Spain, where it was observed that the municipalities of the south have a positive social trend, but in Mediterranean areas a social regression is shown.

In Peru, SPI measurements were made at the level of its main regions, according to Collazos et al. (2018), obtained a SPI of 71.23 on a scale from 0 to 100, a result that means that they have a medium-high level of social progress in the district of San Miguel, department of Lima, Peru. It should be shown that the most developed dimension has been that of BHN (85.70). Arias et al. (2020), conducted a study in the city of Puno, Puno department, Peru; which has characteristics similar to Huancayo province, obtaining an SPI of 57.95 which means that it has a low average level. Given the academic and economic progress that has brought with it the passage of time in Peru, in the case of Huancayo Province there is still no detailed information on the standard of living of the population.

Therefore, this research took Huancayo into account because it is a participant in the Emerging and Sustainable Cities Program (ESC) according to the Inter-American Development Bank (2017); that is, Huancayo, like 50 other cities in Latin America and the Caribbean, has some important characteristics, which make it eligible for the ESC; these are: Intermediate city according to the total population of the country, has an accelerated growth both demographically and economically, shows a good environment of governance and social stability, is a reference at the regional level of quality of life and has a high potential to integrate with the world.

For all the aforementioned, the purpose of this study was to calculate, analyze and compare each dimension that makes up the SPI in urban and rural areas of the Huancayo province, for the period 2020, using the methodology of the Global Organization of Social Progress Imperative, and thus identify the most dimension, to be able to improve it through policy recommendations.

1 LITERATURE REVIEW

1.1 Social progress

The concept of social progress has emerged a few years ago, with the initiative of looking for an indicator that better describes economic development, according to Montuschi (2013), this term is related to economic growth, well-being and happiness, that is to say, it is the process that a society undergoes so that it can improve the quality of life of its inhabitants. In addition, it consists of an idea that society has to improve in many aspects, for example, social, environmental, political, and beyond the economic point of view. Many countries also regard prosperity as part of social progress, so they seek better social development for people living within a territory, so this concept optimally encompasses development in societies. In fact, Stern et al. (2020), define Social Progress as the ability of a society to meet the basic human needs of citizens, so that in this way they establish the basic components that allow citizens and communities to improve and sustain the quality of their lives and create the conditions for all people to reach their full potential. In the same way, Stiglitz et al. (2008), determine that it is important to improve the indicators that adequately reflect the structural changes of modern societies, which is why they propose that they should focus on measuring and analyzing the well-being of the population and not economic production.

1.2 Social Progress Index

The SPI is an indicator that helps countries to evaluate the level of social progress that they have according to their components, in this way it is possible to know what size of a country is in a deficient state and thus apply policies that can improve these aspects. A very important concept is offered by Porter et al. (2014), where they define the SPI as a tool that allows to capture a comprehensive set of social outcome measures in a transparent way, also allows countries to identify specific areas of strength or weakness in terms of social progress, as well as to compare with peer countries at both the level of individual and global indicators. After many years Stern et al. (2020), updated their research work on the SPI and describe it as the tool to be able to identify social aspects but not economic, that is to say, collect information so that based on the results policy decisions can be made, that is to say to be able to evaluate the input results, which allow there to be an improvement in the final results, this in turn makes it possible to make a comparison of different places with reference to their level of progress. In this way this index is essential to see the contrast of public policies through basic social indicators such as; access to electricity, clean water, education, and others. However, Montuschi (2017), refers that SPI is related to the happiness of individuals, that is, if a person lives in a place with greater social progress, this will be much happier because these people will be living in a better society, their quality of life will be more optimal. In the case of Desai (1998), he criticizes GDP as an indicator of human development, since it does not take into account what, how and for whom it is produced, so he recommends the SPI as an indicator that complements GDP and thus evaluate well-being optimally.

The Social Progress Index is directly related to general social welfare, thus evaluating an approximation of the optimal conditions of an individual in a given territory. According to Collazos et al. (2018), these indicators have elements of different natures, in turn with different approaches such as the degree of happiness, social health, ecology, human development or social well-being in general. So, this indicator is optimal to see the results of the policies already implemented, and see what progress has been in society. The structure of SPI consists of 12 components of which are grouped into three dimensions according to the information they collect, the ones are; Basic Human Needs, Fundamentals of Well-being and Opportunities.

1.3 Social Progress Index dimensions

According to Porter et al. (2017) to evaluate the SPI 4 essential principles are used for the design, which are the following: social and environmental indicators, performances, non-efforts, comprehensive and outstanding for all countries and applicable. They were divided into three dimensions for a better assessment of social progress. The SPI considers three fundamental dimensions that are: Basic Human Needs (BHN), Fundamentals of well-being (FW) and Opportunities (OP) that allow to understand the level of real well-being of the population (Marquina and Del Carpio, 2017).

a) *Basic Human Needs.* This dimension includes the fundamental needs that individuals require to develop effectively at physiological levels, that is, goods and services that can cover the existence of the subjects within a delimited territory. Likewise, according to Porter et al. (2014), the first dimension captures the degree to which the most essential conditions for survival are met. These essential needs must be met to create the minimum standards for further progress. In turn they have 4 very important components that help evaluate this dimension, Nutrition and basic medical care, Water and Sanitation, Housing and Personal Security.

b) *Fundamentals of Well-being.* This dimension according to Stern et al. (2020), selected items such as the benefit of a modern health system, available information, receiving basic education and communicating freely, allows the citizen to have an environment conducive to a better life. So, this dimension has 4 components that are relevant when estimating the SPI, then it is detailed what they are: Access to Basic Knowledge, Access to Information and Communications, Health and Well-being, Environmental Quality.

c) *Opportunities.* This dimension is understood by the freedom of expression on the part of individuals, that is to say the level that they can develop freely through their ideas and expressions. In addition, these apply to all those infrastructures that allow individuals to develop fully. This dimension according to Stern et al. (2020), is perhaps the most controversial and the most difficult to measure.

2 MOTIVATION AND HYPOTHESES

The report of the SPI of the regions of Peru carried out by The Business School of Centrum Pontificia Universidad Catolica del Peru (Centrum PUCP) in 2019 was used like a reference to realize this study.

Consequently, this study made a report with a different approach, studying the rural and urban areas to show the gaps that exist between them within Huancayo province that itself explains much of the regional economy of Junín. A provincial and differentiated analysis between urban and rural areas will capture results according to the reality Huancayo province, otherwise (regional analysis) could overestimate or underestimate the social welfare of the same. This application will serve for further analysis of the index with the approach given in this study, that is to say, for those cities that show an accelerated demographic growth, but not an integration of the population at the same pace and that as a consequence increases the poverty rate, increases the number of informal settlements and informal employment, high level of pollution, citizen insecurity and low tax collection.

2.1 General hypothesis

Taking into account the objectives already set for the present study, the general hypothesis was taken; the SPI of the urban area is greater than the SPI of the rural area of the Huancayo Province for the year 2020.

2.2 Specific assumptions

The specific hypotheses that will provide the necessary support for the general hypothesis described above are as follows:

- H1: The basic human needs index (INHB) in the urban area is higher than the index of basic human needs in the rural area in the Huancayo Province for the year 2020.
- H2: The index of fundamentals of well-being (IFB) in the urban area is higher than the index of fundamentals of well-being in the rural area in the Huancayo Province for the year 2020.
- H3: The index of opportunities (IO) in the urban area is higher than the index of opportunities in the rural area in the Huancayo Province for the year 2020.

3 METHODOLOGY

3.1 Data

The SPI report of the regions of our country made in 2019 by the Centrum of the Catholic University of Peru was taken as a reference. For the SPI 2020 calculation, it's considered a descriptive study, with quantitative, non-experimental and cross-sectional approach. The segment of the population was made up of inhabitants aged 14 to 64 years of the Huancayo population. The urban population within the selected age range was 343 022 and the rural area was 25 823 (INEI, 2018). The sample was determined taking into account the formula of a finite population, where a margin of error equal to 10% and a confidence level of 90% are taken, as a result a sample of 115 and 114 was obtained, for the urban and rural areas, respectively.

3.2 Missing values

It identifies relevant data on the variables, if it does not exist a data instrument is built, such is the case of the present research. Table A1 in the Appendix shows 56 questions, organized into 12 components and 3 dimensions based on the data collection instrument standardized by Centrum Católica.

3.3 Standardization

It converts the indicators to the same scale in three steps:

1. First, it establishes scenarios as favorable and non-favorable to determine the concrete limits of the scale that are based on theoretical or historical values.
2. Second, the indicators will be reversed when the increasing values show the lowest SPI values
3. Finally, the indicators are standardized into scores before the calculation of the principal component analysis (PCA).

3.4 Development of the PCA

One of the important points that you should consider before continuing the third step of the standardization, is to perform a factor analysis. De la Fuente (2011) argues that this analysis is mainly based on reducing the set of variables, and, therefore, it is grouped according to its variances and the information provided by each indicator. To perform this analysis, it is necessary to calculate Kaiser, Meyer and Olkin (KMO) statistics, that allows to analyze that this score obtained which is above 0.5, similar studies conducted by Stern et al. (2020), and Garcia and Jimenez (2015) consider that the scores obtained should be close to the unit or be greater than 0.5, so this result allows you to correctly calculate the principal components.

Therefore, once the KMO test has been performed, we proceed to continue with the calculation of the Principal Component Analysis. It could be calculated by Stata, specifically with the *predict factor* command to find significant indicators. To continue, the sequence of steps for the SPI calculation, is described.

3.5 Component scores

It is necessary to consider the formula that reflects the sum of indicators in a main component, where c = Social Component of the progress index and i = respondent.

$$Component\ Value_c = \textstyle\sum_i (w_i * indicator_i).$$

To convert each major component into a component score on a scale from 0 to 100, we use a simple minimum-maximum formula, where X = component value and i = respondent.

$$Component\ Score_c = \frac{(X_J - Worst\ cases)}{(Best\ cases - Worst\ cases)} * 100.$$

3.6 Dimension scores

For each dimension, the arithmetic average of each of its components that form the dimension will be taken. Be the case that the province under study does not have a score on the components of a dimension, therefore, an SPI score, the following formula is used to calculate a dimension score, where d = dimension and c = component.

$$Dimension_d = \frac{1}{4} \textstyle\sum_d Component\ Score_d.$$

3.7 Index scores

For the overall SPI score, the arithmetic mean of the 3 dimensions is calculated and the following formula is used, where d = dimension.

$$SPI\ score = \frac{1}{3} \textstyle\sum_d Dimensions_d.$$

For a proper comparison of the data obtained, Table 2 is used as a reference of the report, which shows that the score of 100 is the highest value that can be obtained and the score with value 0 is the minimum value (Marquina and Del Carpio, 2017).

Table 1 Division of the level of social progress according to the score obtained

SPI score	Level of social progress
Of [85, 100]	Very high
Of [75, 85]	High
Of [65, 75]	Medium high
Of [55, 65]	Medium low
Of [45, 55]	Low
Of [35, 45]	Very low
From [0, 35]	Low end

Source: Retrieved from Regional Social Progress Index (Marquina and Del Carpio, 2017)

4 RESULTS

In the present research, the KMO statistical l score for the urban and rural areas shown in Table 3 has a result greater than 0.5.

Table 2 Keyser-Meyer-Olkin urban and rural

Dimension	Component	KMO urban	KMO rural
Basic human needs	Nutrition and basic medical care	0.615	0.500
	Water and sanitation	0.452	0.506
	Housing	0.564	0.579
	Personal safety	0.564	0.504
Fundamentals of well-being	Access to basic knowledge	0.556	0.500
	Access to information and communications	0.578	0.490
	Health and well-being	0.553	0.544
	Ecosystem sustainability	0.554	0.495
Opportunities	Personal rights	0.568	0.409
	Personal freedom and freedom of choice	0.663	0.470
	Tolerance and inclusion	0.670	0.462
	Access to higher education	0.551	0.517

Source: Stern, Krylova and Harmacek (2020)

In the absence of information at the provincial level, a virtual questionnaire was conducted for a sample of the total population of the Huancayo Province in urban and rural areas, taking into account the three indicators that make up the Social Progress Index. When the main components analysis was carried out, 21 main components were obtained in the urban area and 19 main components in the rural area. These results allowed us to identify which indicators within the dimension components turned out to have more information of the total of the 56 questions asked.

In addition, both in urban and rural areas the standardization was carried out where we had some indicators that by its very nature allowed us to give it a score scale according to well-being, such as the indicator that measures the security that people feel had a criterion from 1 to 5, where 1 is not at all safe and 5 very safe. In the score of the components in both urban and rural areas, the maximum amount of variations in the data was obtained and, therefore, only groups the indicators that are really significant to be able to describe the SPI in the Huancayo Province corresponding to each zone. Finally, the result of the SPI of Huancayo in urban area, can be reviewed in Table 3. Likewise, the SPI of Huancayo in rural area in Table 4.

The SPI in urban areas is divided according to the 3 dimensions, BHN dimension had a score of 62.20. This dimension is the one that scored the best. This can be explained because 78% of persons surveyed have the façade of their home brick, the floors of their homes in 35% are made of tiles and 33% are cement, on the other hand, there is a low percentage of overcrowding, only 21% of people share a room with some other family member. A very important and very relevant fact is that 98% of the respondents have electricity through the public network, in turn 61% of these people have a kitchen, refrigerator and washing machine, this means that the people surveyed do not yet have essential devices in the housing indicator.

On the other hand, the percentage that is worrisome is that there is still a gap of 17% of persons without health insurance. However, it is still close to average. The dimension of FW has a score of 52.49, which is explained by some more representative indicators for example 47% of people having internet, but only 25% of these have a good connection, so there is a gap in quality of technological services.

Table 3 Sequence of steps for the urban SPI Huancayo

Dimension	Main factors by dimension	Component score	Dimension scores	SPI (by dimension)	TOTAL SPI
Basic human needs	Factor 3 Factor 9 Factor 10 Factor 18 Factor 13 Factor 19	Normalized factor 3 = $\frac{(\text{factor 3} - \text{mini 3})}{(\text{max 3} - \text{mini 3})}$ (for each corresponding factor)	D1 = (nfactor3 + nfactor9 + nfactor10 + nfactor18 + nfactor13 + factor 19)/6	D1 = 62.20	56.04
Fundamentals of well-being	Factor 5 Factor 7 Factor 12 Factor 16 Factor 1 Factor11 Factor 21 Factor 15	Normalized factor 5 = $\frac{(\text{factor 5} - \text{mini 5})}{(\text{max 5} - \text{mini 5})}$ (for each corresponding factor)	D2 = (nfactor5 + nfactor7 + nfactor12 + nfactor16 + nfactor1 + factor 11 + factor 21 + factor 15)/8	D2 = 52.49	
Opportunities	Factor 2 Factor 4 Factor 8 Factor 17 Factor 6 Factor 14 Factor 20	Normalized factor 2 = $\frac{(\text{factor 2} - \text{mini 2})}{(\text{max 2} - \text{mini 2})}$ (for each corresponding factor)	D3 = nfactor2 + nfactor4 + nfactor8 + nfactor17 + nfactor6 + nfactor 14 + nfactor20)/7	D3 = 53.43	

Source: Own construction

The dimension of OP has a score in the urban area 53.43, this index is reflected by some indicators; for example, 89% of the people surveyed do not agree at all with corruption, 4% tolerate it. 31% of people at some point in their lives have felt discriminated. They are considered within the low mid-range being not favorable when it comes to measuring the social progress of the urban area of the province of Huancayo. The SPI evaluated in the Huancayo Province presented a score of 56.04 points for the urban area, which is located at a medium low level of progress.

The results of the SPI of the Rural Zone in the Huancayo Province for the size of BHN obtained as a score 51.80, it should be considered that the predominant material in the rural area is brick or cement block that represents 49% of the sample, while 38% of the houses are built of adobe or wall, on the other hand 64% of the rural sample has access to drinking water service while 36% do not have a drinking water connection in their homes, which is a major problem for the essential needs to be met, in the personal security component it means that 54% of the sample under study was a victim of theft while 21% have not been victims of any crime, the percentage of theft presents to be more than 50%.

On the side of the dimension of FW the score was 45.55 which is below average, in the component of health care, 72% attend 1 to 3 times a year a health center, 17% do not attend and 3% of respondents attend more than 6 times a year, with regard to access to information and telecommunications internet service in rural areas, 75% of respondents consider that the coverage is regular, 28% that it is bad and only 3% indicate that it is good for the dimension of FW 45.55 being the lowest score because it is below 0.5.

Table 4 Sequence of steps for the rural SPI Huancayo

Dimension	Main factors by dimension	Component score	Dimension scores	SPI (by dimension)	TOTAL SPI
Basic human needs	Factor 1 Factor 2 Factor 6 Factor 9 Factor 15 Factor 16 Factor 17	Normalized factor 1 = $\frac{(\text{factor } 1 - \text{mini } 1)}{(\text{max } 1 - \text{mini } 1)}$ (for each corresponding factor)	D1 = (nfactor1 + nfactor2 + nfactor6 + nfactor9 + nfactor15 + factor 16 + nfactor17)/7	D1 = 51.80	53.98
Fundamentals of well-being	Factor 3 Factor 5 Factor 8 Factor 11 Factor 13 Factor 19	Normalized factor 3 = $\frac{(\text{factor } 3 - \text{mini } 3)}{(\text{max } 3 - \text{mini } 3)}$ (for each corresponding factor)	D2 = (nfactor3 + nfactor5 + nfactor8 + nfactor11 + nfactor13 + factor 19)/6	D2 = 45.55	
Opportunities	Factor 4 Factor 7 Factor 10 Factor 12 Factor 14 Factor 18	Normalized factor 4 = $\frac{(\text{factor } 4 - \text{mini } 4)}{(\text{max } 4 - \text{mini } 4)}$ (for each corresponding factor)	D3 = (nfactor4 + nfactor7 + nfactor10 + nfactor12 + nfactor14 + nfactor 18)/6	D3 = 64.57	

Source: Own construction

Finally, in the third dimension of OP has the highest score of 64.57%, in the component of freedom of expression that 81% of the rural population does not agree in relation to corruption and a reduced percentage of 5% accepts it, then in the component of tolerance and inclusion 31% of the population has been discriminated against at some point either for reasons of place of birth, for some disability, socioeconomic level, among others. It also shows that 24% of the population has suffered some form of psychological, physical or sexual violence during the last 12 months. Therefore, it was concluded that in rural areas it presented a score of 53.98, placing it at a low level of medium social progress.

DISCUSSION AND CONCLUSION

The SPI allows a comparison in relation to other realities, so it allows us to know the situation in which a certain territory is located through the score obtained and identified in the division table of social progress, but even more importantly, it allows us to know and obtain information on aspects that are relevant within the three dimensions such as the BHN, FW and OP, which are being set aside, nor are the relevant measures being taken to solve the problem that raises a reality within a territory, limiting sustainable social development.

If you compare the SPI of the province of Huancayo, both rural and urban areas, both are in the lower middle level, however, according to the organization Social Progress Imperative, the SPI of Peru is 74.22 placing in this way in the 59^{th} place of 163 countries. Now, this first look and comparison gives us to understand that the Huancayo Province still has many dimensions to improve, not to mention that there are deficiencies in each of them.

Noting that Huancayo is below Peru's score; we can compare ourselves with the rest of the economies. According to the World Bank, Peru is classified as a low- and middle-income country; within this classification of the World Bank is Argentina, a country that is in economic crisis and that is still similar to Peru within this classification. The SPI for Argentina is 80.66, placing it on 41^{st} place out of 163 countries; superior to that obtained by Peru and of course to Huancayo. We can also make the comparison with a European economy such as Albania with the population smaller than that of Peru and with a geographical extension also lower than that of Peru; however, the SPI of Albania is 75.41, although it is true that it is not far from the result obtained for Peru according to the organization Social Progress Imperative, the geographical and demographic differences are very significant. Peru being a country with a greater territorial extension and a greater amount of productive force as well as the advantages in terms of natural wealth, it was to be expected that the results will be the opposite. This has only shown that the national shortcoming also affects the departments and especially the provinces of Peru, as is the case of Huancayo.

The social progress index for the urban area of the Huancayo Province had a rating of 56.04 points (low average). While in the rural area a score was obtained not far from that of the urban area, of 53.98, (Low), affirming the general hypothesis raised. As for the Dimension of Basic Human Needs, in the Rural area it was registered at the low level with 51.8 points, and in the Urban area a score of 62.20 considered as a level of social progress Medium Low. In this case, a gap of 10.4 is observed between zones. It is therefore recommended that in areas such as health, an adjustment and budget be made in first-level care establishments, so that it can encourage the reduction of social gaps, as well as promote advertising for access to social insurance, because 24% of the population in rural areas does not have social insurance, and in the case of urban areas the percentage of the population that does not access social security is 17%; that is, this deficiency is found in both the rural and urban areas. To itself, more than 30% in urban area and 18% in rural area do not have the availability of water service within 24 hours, the institution in charge jointly with the corresponding municipality should carry out supervisions and controls for the adequate distribution of water, by which they should be responsible for developing projects to cover this need, it is also known that the amount of investment in water and sanitation is very low, therefore, the increase in budget for water and sanitation is necessary.

In the dimension of FW, the Urban and Rural areas obtained scores of 52.49 (Low) and 45.5 (Low), respectively. Although both qualify as low social progress, there is a gap is the result of a higher percentage of people who have access to higher education in the Urban Zone (78%). In other words, programs aimed at students who are finishing high school must be improved. In addition, awareness about the importance of pursuing higher education (Institutes and Universities) should be raised. Whereas, we have the component of Access to Information and Communication; of the 54% of the rural population and 69% of the Urban population who have access to the Internet, 78% of the sample in the Urban area and 68% of the sample of the Rural Zone, qualify the internet service as regular. Therefore, prioritizing broadband installation projects for comprehensive connectivity will help reduce the gap and increase SPI in both areas. As well, the corresponding municipalities could agree and request support from the regional government for the coordination of recreational activities, that is, to increase citizen participation in sports activities such as athletics, football, volleyball, intern and others, and give incentives, awards, such as educational scholarships for being an outstanding athlete.

This could help to improve the health of citizens and avoid diseases typical of sedentary lifestyle.

Additionally, the Urban area obtained 53.43 points (Low) and the Rural area 64.57 (Medium Low) in the dimension of Opportunities. This is due to the fact that a higher proportion of the population of rural area are beneficiaries of social programs. In addition, 43% and 65% of the urban and rural areas, respectively, live on land by possession, thus increasing informal settlements and overcrowding for housing. For this reason, programs that provide housing benefits and opportunities must facilitate access and reduce the requirements for those who apply to the social program, thus having the formal registration of housing, since in rural areas more than 30% do not have a registered title deed.

The aforementioned recommendations cover the 3 dimensions of the Social Progress Index, as they are important for the population to develop progressively over the years finding its well-being and a sustainable development, so the corresponding institutions should intervene to comply with these.

Finally, it is concluded that, through empirical evidence, there is a gap between both areas, this evidence indicates that the quality of life is higher in the urban area, and it differs in a better quality of education, health, access to entertainment, among others, but they are not as notable as before due to migration between these areas. Since now it is not only the fact of having access to public services that matters, nor the right that exists within a territory, now the level of quality that has been implemented in each of these services prevails.

The information obtained when using the SPI of rural and urban areas is useful to manage the reduction of gaps with specific steps taken by the pertinent authority and around the deficiencies that each area has, in addition to a continuous monitoring of applied policies, carry out an impact evaluation and finally being able to narrow the gap. However, the differences in SPI that exist between rural and urban areas in other countries may differ due to the policies previously implemented in each of them.

References

ARIAS, D., MOSCAIRO, M., PARIAPAZA, C., VALENCIA, A. (2020). *Índice de progreso social de la ciudad de Puno.* Pontificia Universidad Católica del Perú.

ARROYO, J., LEYTON, C., DELAJARA, M., SERDÁN, A., MOLINA, C. (2018). *Presentación de Informe Latinoamericano sobre Pobreza y Desigualdad 2017 de RIMSIP* [online]. <http://bibliodigitalibd.senado.gob.mx/handle/123456789/4049>.

BANCO INTERAMERICANO DE DESARROLLO (2017). *Huancayo: Hacia la sostenibilidad metropolitana bajo el liderazgo de un gobierno local moderno* [online]. <https://www.iadb.org/es/desarrollo-urbano-y-vivienda/programa-ciudades-emergentes-y-sostenibles>.

BILAN, Y. V., VASYLIEVA, T. A., LIULOV, O. V., PIMONENKO, T. V. (2019). EU vector of Ukraine development: Linking between macroeconomic stability and social progress. *International Journal of Business and Society*, 20(2): 433–450.

CÁRCABA, A., GONZÁLEZ, E., VENTURA, J. (2017). Social progress in Spanish municipalities (2001–2011). *Applied research in quality of life*, 12(4): 997–1019. <http://doi.org/10.1007/s11482-016-9502-7>.

CLARK, C. M., KAVANAGH, C., LENIHAN, N. (2020). *Measuring progress: The sustainable progress index 2020.* Social Justice Ireland: Dublin, Ireland.

COLLAZOS, G., JULCAMORO, W., RAMÍREZ, A., SAKIHAMA, L. (2018). *Índice de progreso social del distrito de San Miguel.* Pontificia Universidad Católica del Perú.

ESTES, R. J. (2019). The social progress of nations revisited. *Social Indicators Research*, 144(2): 539–574. <http://doi.org/10.1007/s11205-018-02058-9>.

GREVE, B. (2017). How to measure social progress? *Social Policy & Administration*, 51(7): 1002–1022.

INEI. (2018). *Junín. Resultados Definitivos Lima 2018* [online]. <https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1576/12TOMO_01.pdf>.

MANSOR, N., ZAKARIA, Z., DAUD, C. H. R. (2013). Quality of life in the 21th century: Narrowing the gap between rural and urban area. *International Journal of Business and Social Science*, 4(5).

MARQUINA, P., DEL CARPIO, L., FAJARDO, V. (2019). *Índice del Progreso Social Regional del Perú 2019.* Pontificia Universidad Católica del Perú.

MONTUSCHI, L. (2013). *Social Progress: Growth and Well-being*. Working Papers Series, No. 533.

MONTUSCHI, L. (2017). *Economic growth, social progress and happiness.* Working Papers Series, No. 620.

NACIONES UNIDAS. (2010). *Principios y recomendaciones para los censos de población y habitación* [online]. Nueva York. [cit. 12.11.2021]. <https://unstats.un.org/unsd/publication/seriesm/seriesm_67rev2s.pdf>.

PORTER, M. E., STERN, S., GREEN, M. (2014). *Social progress index 2014.* Washington, DC: Social Progress Imperative.

STERN, S., KRYLOVA, P., HARMACEK, J. (2020) *Social Progress Index: Methodology Summary.* Washington, DC: Social Progress Imperative.

STIGLITZ, J. E., SEN, A., FITOUSSI, J. P. (2008). *Report by the Commission on the Measurement of Economic Performance and Social Progress* [online]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.215.58&rep=rep1&type=pdf>.

APPENDIX

Table A1 Survey questions for each component

Dimension	Component	Questions (factors)
Basic human needs	Nutrition and basic medical care	1. The bathroom or hygienic service of your home, is connected to 2. On average, how many times a year do you attend a health center? 3. What type of insurance do you have?
	Water and sanitation	4. Do you have the supply of drinking water in your home all day? 5. The water supply in your home comes from: 5.1. Public network outside the home, but inside the building. 5.2. Public network inside the home. 5.3. Tanker truck or other similar. 5.4. Pylon for public use. 5.5. River, acequia, spring or similar. 5.6. Water well.
	Housing	6. The predominant material in the facade of the house is: 6.1. Stone or ashlar with lime or cement. 6.2. Brick or concrete block. 6.3. Quincha (cane with mud). 6.4. Adobe or wall. 6.5. Stone with mud. 6.6. Mat. 6.7. Straw, palm leaves. 7. The predominant material in the floors of the house is: 7.1. Asphalt, vinyl or similar sheets. 7.2. Parquet or polished wood. 7.3. Wood (decking). 7.4. Tiles, terraces or similar. 7.5. Land. 7.6. Cement. 7.7. Dirt floor. 8. The predominant material in the roofs of the house is: 8.1. Wood.

Table A1 (continuation)

Dimension	Component	Questions (factors)
Basic human needs	Housing	8.2. Reinforced. 8.3. concrete. 8.4. Calamine. 8.5. Fiber cement or similar sheets mat. 9. Do you share a room with a member of your family other than your partner? 10. If the above answer is yes How many people do you share the room with? 11. Does the house have electricity through the public grid? 12. What fuel do you use in your home to cook your food? 13. Do you have the following appliances?
	Personal safety	14. Currently, with regard to violence and crime, how safe do you generally feel in the district of Huancayo? 15. Do you consider there is drug sales in your neighborhood or area where you live? 16. Do you consider that prostitution exists in your neighborhood or area? 17. Which of the following crimes have you been the victim of? You can dial more than one.
Fundamentals of well-being	Access to basic knowledge	18. What is your level of education? 19. Do you have other studies?
	Access to information and communications	21. Which of the following services do you have? You can dial more than one. 21.1. Fixed telephony. 21.2. Mobile line post payment. 21.3. Prepaid mobile line. 21.4. Internet. 21.5. Cable. 21.6. None. 22. Is the internet service? 23. What method of communication do you use to inform yourself? You can dial more than one.
	Health and well-being	24. Do you think mental health is important? 25. Have you ever been treated for any mental health issues? 26. Do you have a family member who suffers from any type of mental illness? 27. In the last 12 months, has anyone in your family had any chronic illnesses? 28. In the last 5 years, have any members of your family died of cancer? 29. In the last 5 years, have any members of your family died of cardiovascular disease? 30. Currently, have any members of your family died of covid-19? 31. Do you frequently do any sports? 32. Do you suffer from any addictions?
	Ecosystem sustainability	33. How do you consider the noise level in your area? 33.1. Transport. 33.2. Industry. 33.3. Bars and nightclubs. 34. In your opinion, which of these activities generates the most noise? 35. Do you consider that there is air pollution in your area?
Opportunities	Personal rights	36. Do you believe that there is respect for freedom of expression? 37. Is the house you occupy? 38. Is your home registered in public records? 39. Did you participate in the last district, regional and national elections? 40. Do you benefit from any social programs provided by the state? 41. What programs in your district do you participate in? 41.1. Participatory budgeting. 41.2. Neighborhood councils (vote). 41.3. Neighborhood hearings. 41.4. Neighborhood citizen safety boards. 41.5. I don't participate.
	Personal freedom and freedom of choice	42. In relation to corruption: 42.1. Not at all agree. 42.2. It doesn't matter to me. 42.3. The tolero. 42.4. I accept it.

Table A1 (continuation)

Dimension	**Component**	**Questions (factors)**
Opportunities	Personal freedom and freedom of choice	43. Do you have a public service near your home? 44. Do you think that the sidewalks and tracks are in optimal condition and in good condition? 45. What is your rating in satisfaction with the quality and quantity of cultural grade activities in your district? 46. Do you use birth control? 47. What contraceptive methods do you know? You can dial more than one. 48. In the last 3 years, has any member of your household, being. a teenage woman been pregnant or was a mother.
	Tolerance and inclusion	49. Have you ever been discriminated against? 50. Do you consider older adults to be respected in your district? 51. What actions do you consider to be linked to physical and psychological family violence? 52. Have you suffered from any kind of psychological, physical, or sexual violence during the past twelve months? 53. Did any members of your household suffer family violence?
	Access to higher education	54. Where did you become professional? 55. Have you ever been a recipient of a State Scholarship? 56. The professional program where you studied is located.

Source: Collazos, Julcamoro, Ramírez and Sakihama (2018)

Food Consumption and Availability in Czechia in the Years 1993–2019

Marcela Mácová[1] | *Czech Statistical Office, Prague, Czech Republic*

Abstract

The article deals with comparison of consumption of food and beverages in the Czech Republic in the years 1993 and 2019 as well as of two main sources of food and food materials, their domestic production and cross border movements.[2]

Keywords	**DOI**	**JEL code**
Food and beverages, consumption, agricultural production, self-sufficiency, cross border movements	*https://doi.org/10.54694/stat.2021.14*	*Q02, Q10, Q18*

INTRODUCTION

The nutrition of the population belongs to basic indicators of the standard of living of our entire society. The Food and Agriculture Organization of the United Nations (hereinafter referred to as FAO) and the World Health Organization are the main bodies aiming to address global nutrition issues. Globally, it is estimated that almost 750 million, i.e. nearly 10 percent of the world population, were exposed to severe levels of food insecurity in 2019. Moreover, this number shows an increasing trend in recent years. On the contrary, especially high-income countries are currently facing an increasing obesity rate (FAO et al., 2020) and food waste. As for the latter phenomenon, Kubíčková (2021) estimates that the amount of wasted food in the Czech Republic reaches annually 830 thous. tonnes (81 kg per person per year), of which approximately 254 thous. tonnes (25 kg per person per year) of food waste is generated in households. To successfully address the problems of diet quality and rationalisation, we need reliable knowledge of food production, trade, and consumption.

The consumption of food and beverages in the Czech Republic had an increasing trend during the observed period. It was connected with favourable socio-economic conditions of recent years before the SARS-CoV-2 pandemic situation, which were represented by low unemployment rate and household income growing faster than their expenditures. While the average gross monthly wage increased since 1993 to 2019 by 478.0% to CZK 34.1 thous. (CZSO, 2021a), expenditures on food and non-alcoholic beverages rose by 224.3% to CZK 39.9 thous. per person per year and those on alcoholic beverages and tobacco by 277.0% to CZK 20.2 thous. per person per year (CZSO, 2021b).

[1] Czech Statistical Office, Na padesátém 81, 100 82 Prague 10, Czech Republic. E-mail: marcela.macova@czso.cz, phone: (+420)274054088.

[2] This article is based on Mácová and Klémová (2021a, 2021b, and 2021c), and Mácová and Vodičková (2021).

Czechia has never been fully self-sufficient in all kinds of food, especially in vegetables and fruit (Mácová and Klémová, 2021b). Cross border movements of food and food materials include not only food commodities that cannot be produced in situ due to climatic or other conditions; a significant part of our domestic agricultural production is sold abroad and, on the contrary, the same commodities are imported to us (Mácová and Klémová, 2021c).

1 FOOD CONSUMPTION

The consumption of food and beverages has been surveyed in the Czech Republic (or former Czechoslovakia) since 1948 nearly in the current extent, with some statistical data being recorded even earlier, since 1918 (CZSO, 2014).

1.1 Methodology

The Czech Statistical Office uses a balance method for computation of consumption of food and beverages (CZSO, 2020). The input data are as follows:

- Animal production statistics in a respective year;
- Final harvest figures in a respective year;
- Production of selected industrial products in a respective year;
- Initial and final stocks in agricultural enterprises;
- Initial and final stocks in food processing enterprises;
- Imports and exports of food products based on statistics of cross border movements of the Czech Statistical Office in a respective year;
- Self-supply of food products.

Further data sources for the calculation are the Ministry of Agriculture, the Institute of Agricultural Economics and Information, food producers' unions and other organizations.

The consumption of particular items is calculated as average per capita figures. The number of inhabitants is given as the mid-year population as at 1 July of a respective year.

The input data do not include those food commodities that do not serve for human consumption, e.g. sugar for feeding bees or for production of ethanol, rapeseed oil for the production of FAME (fatty acid methyl esters), animal offal used in pharmaceutical production, milk for feeding or casein production, eggs and yeast for chemistry, seeds, etc.

Since 2002, the structure of food consumption items published follows the CZ-COICOP classification (the Czech version of the international Classification of Individual Consumption According to Purpose standard; designed to enable international comparisons and used in the Household Budget Surveys statistics), namely its first two divisions: 01 Food and non-alcoholic beverages and 02 Alcoholic beverages, tobacco and narcotics.

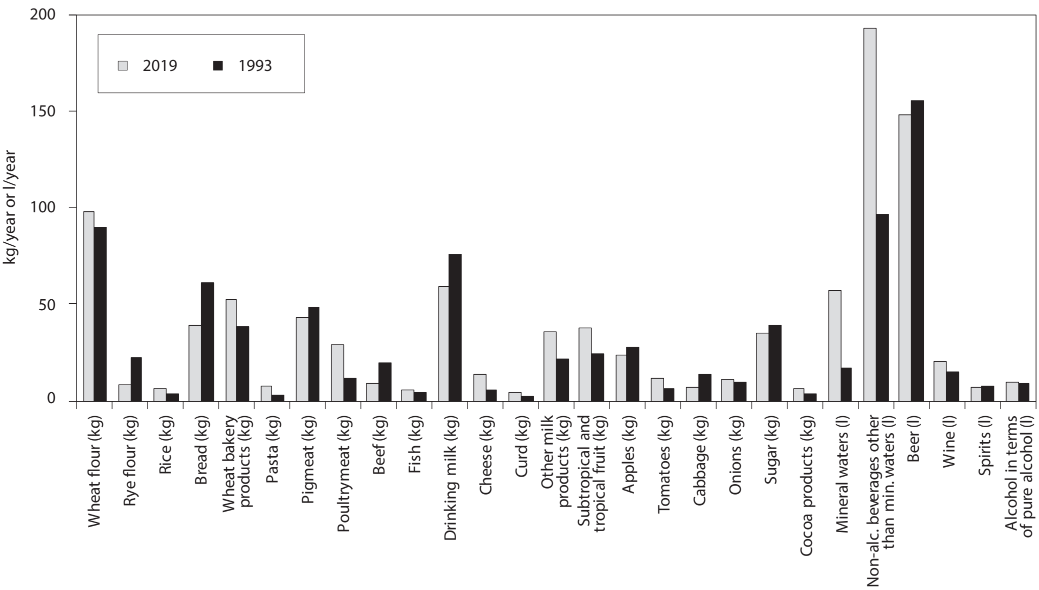
1.2. Main findings

According to the latest data for the year 2019, the average per capita food consumption in Czechia was 796.5 kg/year. It increased by 65.5 kg/year (8.9%) since 1993. Various trends were recorded for particular commodities or their groups: for example, increases were recorded for milk and milk products (+31.0%; to 249.0 kg), fruit (+19.0%; to 86.5 kg) and vegetables (+17.3%; to 87.0 kg); on the other hand, consumption of cereals in terms of flour (−2.2%; to 115.6 kg) or of potatoes (−17.3%; to 69.5 kg) decreased. Meat consumption almost did not change between 1993 and 2019 (−1.3%; 83.2 kg).

Figure 1 shows that there exist contradicting trends in consumption among particular commodities or their groups. As for cereals, the consumption rose for wheat flour and rice but declined for rye flour; among cereal products, it went down for bread but up for wheat bakery products and pasta. Pigmeat remained the most favourite meat type; however, poultrymeat replaced beef at the second position.

The consumption of fish slightly rose. We have consumed less drinking milk but more cheese, curd, and other milk products. Among fruit, the proportion of consumed subtropical and tropical fruit rose, although apples remained to be the most favourite ones. As for vegetables, tomatoes replaced cabbage at the leading position and onions remained to be the second most important item. The consumption of sugar declined but that of cocoa products, especially chocolate, rose. Mineral waters as well as other non-alcoholic beverages recorded a significant increase. Compared to 1993, we have consumed less beer and spirits but more wine in 2019; the consumption of alcoholic beverages in terms of pure alcohol went up.

Figure 1 Per capita consumption of selected food items



Source: Mácová and Klémová (2021a)

There have been some changes observed in the structure of consumption of food consumption during the observed period. The share of milk and milk products in the total volume of food consumed increased from 26.0% to 31.3%, the share of fruit from 9.9% to 10.9%, and the share of vegetables from 10.1% to 10.9%. The share of cereals in terms of flour fell from 16.2% to 14.5 %, the share of meat and fish from 12.1% to 11.2%, and the share of potatoes from 11.5 % to 8.7 %. These changes were caused by a variety of factors: some people responded to the recommendations of health nutritionists (which have changed over the course of the considered period), or to various fashion trends, or the price level of food was a determining factor for others. Last but not least, the food supply itself changed depending on the volume of domestic production and imports of food or material for its production.

1.3 Long-term trends since 1950

Some of the changes in amount and structure of consumed food have started long before the aforementioned period. As describes Vodičková (2017), three main groups of foods can be distinguished according their consumption trends. The first group includes foods whose consumption reached its maximum in the 1950s and has been declining since then. These are especially staple food items as for instance rye flour (decline from 55.5 kg per capita in 1950 to 8.6 kg in 2019), cows' drinking milk (from 164.8 kg to 58.7 kg), bread (from 87.4 kg to 39.0 kg), or potatoes (from 145.9 kg to 69.5 kg). The second group represents food items having their peak consumption in late 1980s and early 1990s; it consists mainly of animal food products as beef (30.7 kg in 1987), milk and milk products (259.6 kg in 1989), pigmeat (50.0 kg in 1990), and also sugar (44.0 kg in 1990). The third group comprises foods whose consumption has been on the rise during recent years, frequently these highly processed or imported: for example subtropical and tropical fruit (increase from 2.2 kg in 1950 to 37.5 kg in 2019), poultrymeat (from 2.4 kg to 29.0 kg), cheese (from 1.6 kg to 13.8 kg), rice (from 0.9 kg to 6.7 kg), pasta (from 2.8 kg to 8.1 kg), or wheat bakery products (from 26.2 kg to 51.7 kg). An immense increase was observed in the consumption of mineral waters and other non-alcoholic beverages (from 12.8 l to 246.8 l); the consumption of alcoholic beverages rose as well (from 4.1 l to 10.0 l in terms of pure alcohol).

1.4 International comparison

There is no binding classification for food consumption statistics either in Czechia or at an international level. The FAO publishes its own calculations in its database (FAO, 2021). Their methodology differs from that of the CZSO, especially at the level of classification of individual items: for example, consumption of vegetables is differentiated into only a few items (tomatoes and its products, onions, other vegetables), whereas consumption of freshwater fish and several groups of marine fish and other aquatic animals is monitored separately. In addition to official statistics on food production and trade, FAO uses expert estimates, both its own and those provided by various specialised institutions.

The most recent data are currently available for 2018; for the sake of clarity, we present a comparison limited to the European Union only. For example, per capita consumption in Czechia compared to the EU28 average was lower for cereals and their products (CZ 102.1 kg; EU 129.9 kg), milk and dairy products (CZ 147.5 kg; EU 187.6 kg), vegetables (CZ 76.7 kg; EU 106.2 kg), fruit (CZ 61.9 kg; EU 87.8 kg) or fish (CZ 9.3 kg; EU 23.1 kg). Consumption of meat (CZ 83.7 kg; EU 80.0 kg) or potatoes (CZ 64.0 kg; EU 61.3 kg) was similar to the EU28 average. We have reached the leading position in the consumption of alcoholic beverages (CZ 175.7 kg; EU 97.8 kg).

2 DOMESTIC PRODUCTION AND FOOD SELF-SUFFICIENCY

Food self-sufficiency belongs to hot topics discussed by politicians and public in many countries, including the Czech Republic.

Since 1993, we have been steadily more than self-sufficient in beer production. Self-sufficiency in milk production and in beef production stayed above the 100% level despite a sharp reduction. Self-sufficiency dropped in the production of butter, cheese and curd, apples, or pigmeat from values of above 100% in 1993 to below this level in 2019. For other commodities, self-sufficiency in 1993 has already been below 100% and since then has even decreased; this applies for example to poultrymeat, eggs, carrots, wheat flour, onions, cabbage, or tomatoes. A reduction of domestic production together with an increase in consumption was recorded for wheat flour, milk, butter, tomatoes and onions. For poultrymeat and cheese and curd, there was an increase in domestic production, but their consumption has increased more significantly. Consumption of pigmeat, beef, eggs, apples, carrots, and cabbage decreased, but their production fell even more. Table 1 shows consumption, domestic production, and self-sufficiency for the abovementioned food commodities in 1993 and 2019.

The range and structure of production in the agricultural sector have changed since the beginning of the 1990s. Between 1993 and 2019, there has been a sharp decline in livestock numbers: by 43.5% for cattle, by 66.4% for pigs, and by 18.6% for poultry. The area under crops decreased as well, for example for cereals (by 15.8%), potatoes (by 78.2%), sugar beet (by 44.8%) or various types of vegetables (by 69.7% for vegetables, total). However, at the same time, yields of many of these crops have increased (by 42.5% for cereals, by 16.7% for potatoes and by 53.0% for sugar beet) so the losses of production area have been compensated to some extent. Similarly, in livestock production, average milk yield per cow increased (by 121.6%) as well as average annual egg yield per hen (by 26.5%), the number of calves reared annually per 100 cows (by 5.7%), and the number of piglets reared annually per one sow (by 70.5%).

Table 1 Consumption, domestic production, and self-sufficiency for selected food commodities

	Self-sufficiency (%)		Domestic production			Consumption as food			Measuring unit
	1993	2019	1993	2019	Index (%)	1993	2019	Index (%)	
Beer	109.5	121.9	168.1	178.0	105.9	153.6	146.0	95.1	l
Milk	175.7	119.1	324.6	288.0	88.7	184.6	241.8	131.0	l
Beef	113.7	104.5	22.9	9.5	41.5	19.8	9.1	46.0	kg
Butter	152.0	65.8	8.1	3.6	44.4	5.3	5.4	101.9	kg
Cheese and curd	123.4	88.3	10.5	16.3	155.2	8.5	18.5	217.6	kg
Apples	108.5	75.3	29.8	18.0	60.4	27.5	23.9	86.9	kg
Pigmeat	104.3	43.4	50.2	18.7	37.3	48.1	43.0	89.4	kg
Poultrymeat	94.9	63.2	11.1	18.3	164.9	11.7	29.0	247.9	kg
Eggs	94.4	84.8	300.2	221.3	73.7	318.0	261.0	82.1	pcs.
Carrots	89.2	48.1	7.3	3.4	46.6	8.2	7.0	85.4	kg
Wheat flour	84.5	69.7	74.8	67.6	90.4	88.5	97.0	109.6	kg
Onions	84.2	43.2	8.5	4.8	56.5	10.1	11.1	109.9	kg
Cabbage	83.3	51.7	11.5	3.5	30.4	13.8	6.9	50.0	kg
Tomatoes	54.4	19.3	3.5	2.3	65.7	6.5	12.0	184.6	kg

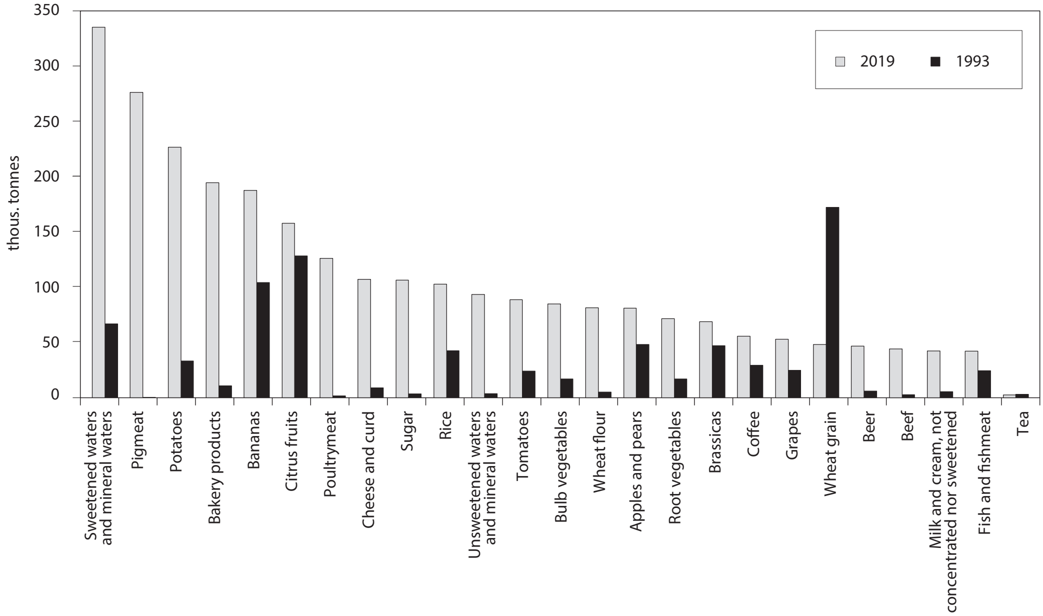
Source: Mácová and Klémová (2021a, 2021b) and author's own computation

The structure of agricultural production has also been adversely affected by the decline in the labour force due to a low attractiveness of working in the agricultural sector. According to the Annual National Accounts, 236.5 thousand persons were employed in the category Crop and animal production, hunting and related service activities in 1993 but only 138.9 thousand persons in 2019, i.e. by 41.3% less. The share of this category in the total employment fell from 4.7% to 2.6%. Last but not least, economic considerations are an important factor for primary producers in deciding which type of agricultural production to pursue. Farmers have an opportunity to benefit from various targeted subsidy programmes (currently, for example, it applies to potatoes, hops, fruit and vegetable production with very high or high intensity, rearing of beef calves, rearing of cows in a market milk production system, etc.). The prospect of selling their products, whether on domestic market or abroad, also plays an important role.

3 CROSS BORDER MOVEMENTS OF FOOD AND FOOD MATERIALS

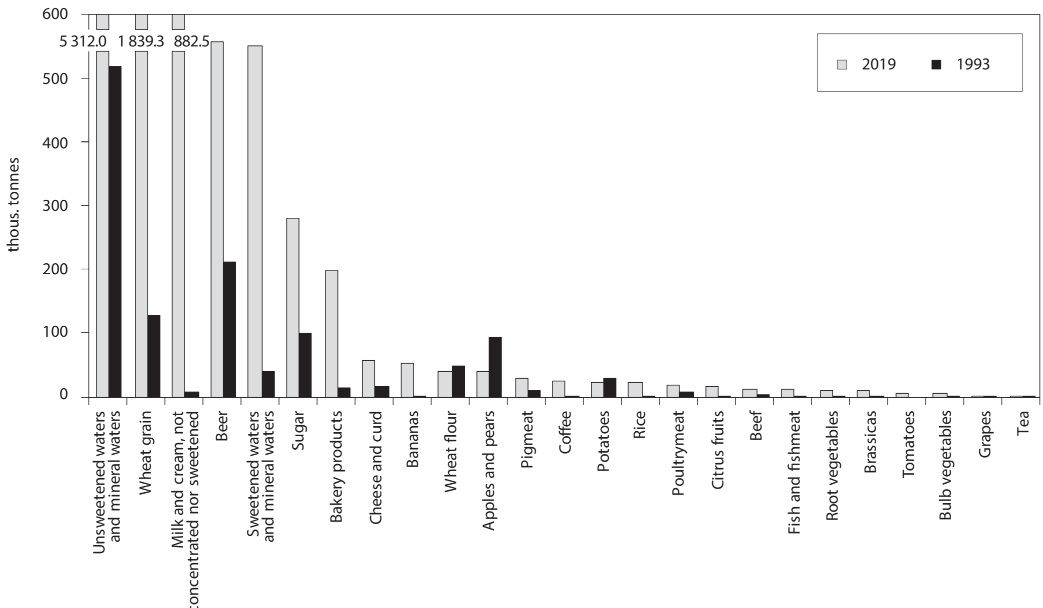
From 1993 to 2019, cross border trade relations with specific countries have evolved differently for various food commodities. The most important partner was and is unsurprisingly Slovakia, followed by Poland and Germany, and then other members of the European Union, but also Russia, for example. Conversely, some goods were and still are imported from countries very far away. At the same time, Czechia plays the role of a transit country, so that commodities that do not originate here also appear on the cross border export side. In the considered period, the re-export of foodstuffs brought to Czechia but not consumed here increased significantly.

Figure 2 Imports of selected food commodities



Source: Mácová and Klémová (2021c)

Figure 3 Exports of selected food commodities



Source: Mácová and Klémová (2021c)

A significant increase in the cross border movement of goods (hereinafter referred to as CBmG) occurred after 2004, when the Czech Republic joined the European Union and the single European market under the EU legislation removing tariff barriers (customs duties etc.) and quantitative restrictions for all types of goods.

Figures 2 and 3 show the amounts in tonnes of food commodities imported and exported, respectively, mentioned in the following text.

All three main meat types – beef, pigmeat, and poultrymeat – have recorded a positive balance of CBmG in 1993, while until 2019 it turned to negative. However, a substantial share of all types of meat consumed domestically comes from animals slaughtered in the country and for relevant live animals – cattle, pigs and poultry – the balance of the CBmG was positive both in 1993 and in 2019 with

Table 2 Main partners of Czechia in cross border movements of food commodities

	Imports		Exports	
	1993	**2019**	**1993**	**2019**
Beef	SK	PL/DE	RU	SK
Pigmeat	DK	DE/ES	RU	SK
Poultrymeat	SK	PL	DE	SK
Fish and fish meat	DE/NL	NO	SK	SK
Milk and cream, not concentrated nor sweetened	SK	SK	SK	DE
Cheese and curd	DE	DE	LB	IT
Potatoes	NL	DE	SK	SK
Tomatoes	SK	NL/ES	SK	SK
Bulb vegetables	NL	NL	SK	SK
Root vegetables	NL	NL	SK	SK
Brassicas	PL	PL	SK	SK
Bananas	CO	CO	SK	SK
Citrus fruits	ES	ES	SK	SK
Grapes	IT	IT	SK	SK
Apples and pears	NL	PL	AT	DE
Coffee	CO/HN	DE	SK	SK
Tea	IN/CN	PL/DE	SK	SK
Sugar	DE	FR	RU	AT
Wheat grain	DE	SK	RU	DE
Wheat flour	SK	SK	PL	SK
Rice	US	IT	SK	SK
Bakery products	SK	PL	SK	SK
Unsweetened waters and mineral waters	SK	SK	AT	PL
Sweetened waters and mineral waters	AT	PL	SK	SK
Beer	SK	PL	SK	SK

Source: Mácová and Klémová (2021c)

an increasing trend. In contrast, fish (other than live) and fish meat came from imports already in 1993 and this has not changed by 2019.

For the commodity milk and cream, not concentrated nor sweetened, we were a net exporter in 1993 and the surplus of CBmG had even significantly increased by 2019. Although the quantity imported increased, exports multiplied many more. On the other hand, the balance of the CBmG for cheese and curd turned from positive to negative.

The CBmG balance for various vegetables and fruits and also for potatoes was already negative in 1993 and has deepened by 2019. Imports of tomatoes, bulb vegetables, root vegetables, brassicas, bananas, citrus fruits, and grapes increased between 1993 and 2019 as well as their exports. Bananas represent an exotic fruit, which have been still imported mainly from their country of origin. The exception among fruits is the commodity apples and pears, whose balance was positive in 1993 but has fallen below zero by 2019.

Not surprisingly, the balance of CBmG for coffee and tea commodities is stably negative. Both imports and re-exports of coffee rose during the observed period, while imports and re-exports of tea remained almost the same. It should be mentioned that imports of coffee and tea underwent an interesting redirection from the countries of their origin to European countries dealing with their re-exports and/or processing.

Despite the decline in domestic sugar production, the balance of the CBmG remained positive between 1993 and 2019, although its quantity increased on both sides.

The CBmG balances for the commodities wheat and wheat flour showed opposite trends: In 1993, we have imported more wheat grain and less flour than in 2019. The amount of exported grain multiplied by more than 14 times until 2019, while the amount of exported flour recorded a slight decrease. The rice traded in Czechia, of course, also comes entirely from imports, which rose from 1993 to 2019 as well as its re-exports. The balance of CBmG for the commodity bakery goods remained positive in the period 1993 to 2019.

Czechia was a net exporter of unsweetened waters and mineral waters both in 1993 and in 2019. Their imports have risen more than 20 times in the period 1993–2019, but exports have increased even more than 100 times. In contrast, the balance of CBmG for sweetened waters and mineral waters was negative in 1993 but it had turned positive by 2019. Unsurprisingly, the CBmG balance for beer is positive in the long term and both imports and exports have increased from 1993 to 2019: imports almost 8 times and exports 2.6 times.

Table 2 shows the main partners of the Czech Republic in cross border movements in 1993 and 2019 for the abovementioned commodities.

CONCLUSIONS

The consumption of food and beverages in Czechia in the period 1993–2019 rose. While the domestic production and the food self-sufficiency in Czechia has declined for all surveyed commodities except beer, the role of cross border movements of food and food materials has significantly increased. Further data collections will show how these trends will be influenced by the SARS-CoV-2 pandemic situation.

ACKNOWLEDGEMENT

The author would like to thank to Lea Klémová for her work on data processing and to Renata Vodičková, Markéta Fiedlerová, Adéla Höffnerová, and an anonymous reviewer for their valuable comments to the manuscript.

References

CZSO. (2014). *Spotřeba potravin – 1948–2010* [online]. Prague: Czech Statistical Office. <https://www.czso.cz/csu/czso/spotreba-potravin-1948-az-2012-n-hjw8eg93rj>.

CZSO. (2020). *Food consumption 2019* [online]. Prague: Czech Statistical Office. <https://www.czso.cz/csu/czso/food-consumption-2019>.

CZSO. (2021a). Labour Statistics: *Time Series of Basic Indicators – January 2021* [online]. Prague: Czech Statistical Office. <https://www.czso.cz/csu/czso/labour-statistics-time-series-of-basic-indicators-january-2021>.

CZSO. (2021b). *Database of National Accounts: GDP by the expenditure approach* [online]. Prague: Czech Statistical Office. [cit. 12.3.2021]. <https://apl.czso.cz/pll/rocenka/rocenkavyber.makroek_vydaj_en>.

FAO. (2021). *New Food Balances* [online]. Rome: FAO. [cit. 19.5.2021]. <http://www.fao.org/faostat/en/#data/FBS>.

FAO, IFAD, UNICEF, WFP, WHO. (2020). *The State of Food Security and Nutrition in the World 2020. Transforming food systems for affordable healthy diets* [online]. Rome: FAO. <https://www.fao.org/3/ca9692en/ca9692en.pdf>.

KUBÍČKOVÁ, L. (2021). Plýtvání potravinami v českých domácnostech. *Statistika&My*, 5 (May): 28–29.

MÁCOVÁ, M., KLÉMOVÁ, L. (2021a). Spotřeba potravin a změny ve struktuře jídelníčku. *Statistika&My*, 5 (May): 21–23.

MÁCOVÁ, M., KLÉMOVÁ, L. (2021b). Kolik potravin si vyrobíme sami? *Statistika&My*, 5 (May): 24–25.

MÁCOVÁ, M., KLÉMOVÁ, L. (2021c). Potraviny cestují přes hranice. *Statistika&My*, 5 (May): 33–35.

MÁCOVÁ, M., VODIČKOVÁ, R. (2021). V Česku stoupá životní úroveň i spotřeba potravin. *Statistika&My*, 5 (May): 20.

VODIČKOVÁ, R. (2017). Spotřebu potravin sledují statistici téměř sto let. *Statistika&My*, 2 (February): 30–31.

Use of Administrative Data for Waste Statistics in the Czech Republic

Marie Boušková[1] | *Czech Statistical Office, Prague, Czech Republic*
Tomáš Harák | *Czech Statistical Office, Prague, Czech Republic*

Abstract

The Czech Statistical Office (CZSO) changed the calculation of the total volume of waste and also changed the definition of municipal waste compared to the previous methodology. This was made possible by the wider use of the existing administrative data source, the Integrated Environmental Reporting System (ISPOP). The change in the definition of municipal waste was a response to recent Eurostat activities, which led to a more precise definition. The original method no longer meets this definition.

Keywords	***DOI***	***JEL code***
Waste, municipal waste, Eurostat, Ministry of the Environment, ISPOP, administrative data	*https://doi.org/10.54694/stat.2021.21*	*Q53*

INTRODUCTION

Economies all over the world are facing new challenges associated with global warming, dwindling raw material resources, and population growth. Efforts to promote the principles of an emission-free industry, circular economy, more environmentally friendly approach to nature and environmental protection, and Green Deal policy in general must be necessarily reflected in reducing the amount of waste produced and striving for its efficient use.

1 EUROPEAN COMMISSION LEGISLATION

The Regulation (EC) No 2150/2002 of the European Parliament and of the Council on waste statistics was adopted on 25 November 2002 by the European Commission. Basic concepts, methods of waste management as well as rules for waste treatment are defined by the Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008, which was supplemented by the Directive (EU) 2018/851 of the European Parliament and of the Council of 30 May 2018. Reporting obligations on waste in the Czech Republic are divided between the Czech Statistical Office (CZSO) and the Ministry of the Environment.

The CZSO is responsible for fulfilling the reporting obligations arising from the Regulation (EC) No 2150/2002 – Regulation on Waste Statistics. The Ministry of the Environment fulfils the obligations

[1] Czech Statistical Office, Na padesátém 81, 100 82 Prague 10, Czech Republic. Corresponding author: e-mail: marie.bouskova@czso.cz.

arising from the Directive 2008/98/EC, the Directive (EU) 2018/851, and other international agreements. From 2020 onwards, the CZSO – beyond the scope of the Regulation (EC) No 2150/2002 – is also obliged to report data on municipal waste. Until that time, data on municipal waste production and management were sent according to the Joint Questionnaire of Eurostat and OECD. Reporting by the Joint Questionnaire was voluntary. While the Czech Statistical Office used information obtained from the ODP 5-01 sample statistical survey (Annual report on waste and secondary raw materials), including data on municipal waste, to fulfil its obligations, the Ministry fulfils its obligations on the basis of information obtained from the Integrated Environmental Reporting System (administrative system of waste records in Czech abbreviated as ISPOP).

2 COOPERATION BETWEEN THE CZSO AND THE MINISTRY OF THE ENVIRONMENT

The issue of different data on waste production published by the CZSO and the Ministry of the Environment has been addressed since 2009. At the same time, duplicate monitoring of identical indicators on the statistical report and within the data collected by the Ministry of the Environment was also criticised. A significant impetus for resolving different data on waste was the signing of the so-called Agreement between the CZSO, the Ministry of the Environment, and the Office of the Government of the Czech Republic on reporting municipal waste production in the Czech Republic from 20 January 2016. Another impetus was the fact that the EU focused more and more on information on municipal waste in connection with setting requirements to limit the dumping of municipal waste in landfills. The result was the acceptance of the obligation to report data on the production and management of municipal waste annually.

The CZSO focused on a gradual transfer of data from the ISPOP system, which is operated by the Ministry of the Environment in cooperation with the Cenia agency. The CZSO first focused on an analysis of the consistency of reported data in both systems for the segment of municipalities. A high degree of compliance of the reported data between the two systems was found. Therefore, the CZSO started to replace the data from the ODP 5-01 report from municipalities with data from the ISPOP system for the reference year 2018. This was followed by the replacement of data from business entities from the ODP 5-01 by the data of the Ministry of the Environment.

3 METHODOLOGICAL ISSUES

Prior to taking over the ISPOP data, it was necessary to convert the classifications used in the ISPOP to the classifications required by the Regulation (EC) No 2150/2002 on waste statistics. In particular, the conversion of waste codes according to the List of Waste into the European Waste Classification for Statistics (EWC-STAT), the determination of the origin of waste, and the method of waste management. Despite initial problems, after consultation with Eurostat, the converters were set up.

The European Waste Classification for Statistics (EWC-STAT) strictly (if the nature of the waste allows) respects the material from which the waste originates, in contrast to the List of Waste (LoW), which also takes into account the sector (activity) from which the waste originates. For example, glass is included in the List of Waste into groups as follows:

10 11 – Wastes from glass and glass products production,
15 01 – Packaging,
16 01 – End-of-life vehicles and waste from the dismantling of these vehicles,
17 02 – Wood, glass, plastics (from construction and demolition),
19 12 – Wastes from waste treatment,
20 01 – Materials from separate collection,

while in the EWC-STAT classification all waste glass belongs to the group 07.1 – Glass waste. The same goes for metals, paper, plastics. Both used data sources – ISPOP and ODP 5-01 – monitor waste according to the List of Waste codes. The conversion to the EWC-STAT classification must therefore be performed regardless of which data source we use. As part of the processing in the CZSO data market, a converter from the LoW classification to the EWC-STAT is stored; the opposite conversion is not possible or unambiguous. Both ISPOP and ODP 5-01 use the Classification of Economic Activities (CZ-NACE), which corresponds to the European standard of the Statistical Classification of Economic Activities in the European Community, Rev. 2 (NACE Rev. 2) to determine the sector of origin of waste.

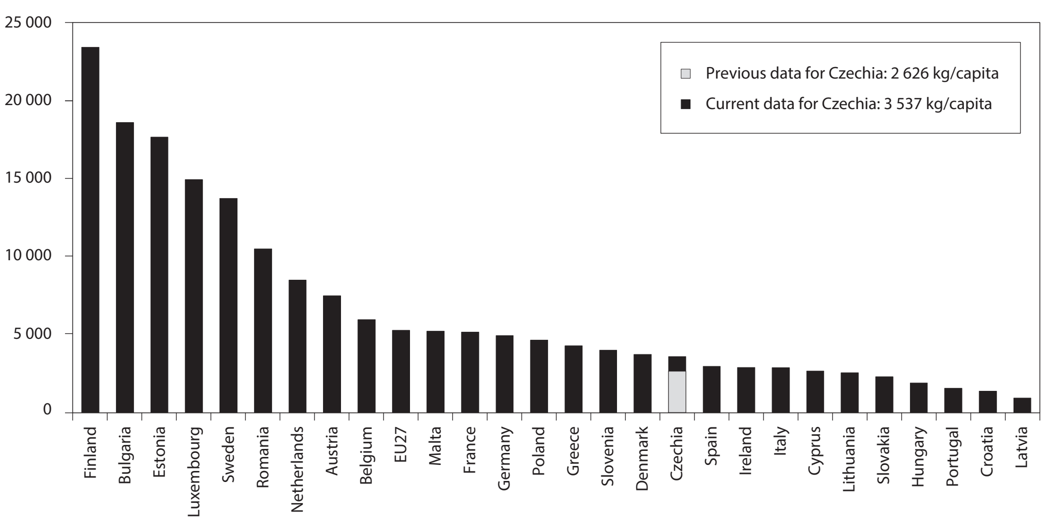
Waste management methods – here the classification used in ISPOP is broader and more detailed than in the Regulation (EC) No 2150/2002 on waste statistics. According to this regulation, only final waste management codes (R – recovery or D – disposal) have to be reported. The Czech classification also uses so-called N (national) codes. The assignment of Czech disposal codes to the European system was consulted with Eurostat and is described in the material "Mathematical expression of the calculation of the system of waste economy indicators" prepared by the Cenia agency.

Earlier requirements of the CZSO for the transmission of data sets and their form also changed. With the gradual expansion and change of the database environment at the CZSO, conditions were created for the transfer of complete data from the ISPOP system. The CZSO was thus able to make full use of data from the ISPOP system for its needs.

4 USE OF ADMINISTRATIVE DATA

The decision on the wider use of administrative data is the result of an analytical evaluation of the compatibility of the data obtained by the statistical report ODP 5-01 and the data reported within the Integrated Environmental Reporting System (ISPOP). It has been repeatedly proven that in a comparable set of respondents there is a high degree of compliance between the two data sources. The scope of the new data source, which also covers enterprises below the threshold for inclusion in statistical monitoring, improved the coverage to almost the entire population, and thus enhances the quality of the presented results. Another advantage of the ISPOP is the possibility of obtaining data according to local units, which will enable the improvement of information on the regional distribution of waste production within the Czech Republic. There is a unique link in the system between the local unit code and the identification

Figure 1 International comparison of waste generated in kilograms per capita (2018)



Source: Eurostat

number of the company/institution/organisation, so it is possible to exclude internal consumption and construct the same data as if reported by the company as a whole.

In the ISPOP system, there is an obligation to report waste data for all entities producing waste above a specified limit (100 tonnes of total waste or 100 kg of hazardous waste).

According to updated results, a total of 37.0 million tonnes of waste was produced in the Czech Republic in 2019, by 634 thousand tonnes less than in 2018. In 2019, the most waste was produced in the Středočeský Region and in Prague, where the volume exceeded 5 million tonnes. The least waste was produced in the Karlovarský Region, in which just over 680 thousand tonnes of waste was generated. 2.6 million tonnes of waste were imported to the Czech Republic, exports amounted to 3.4 million tonnes.

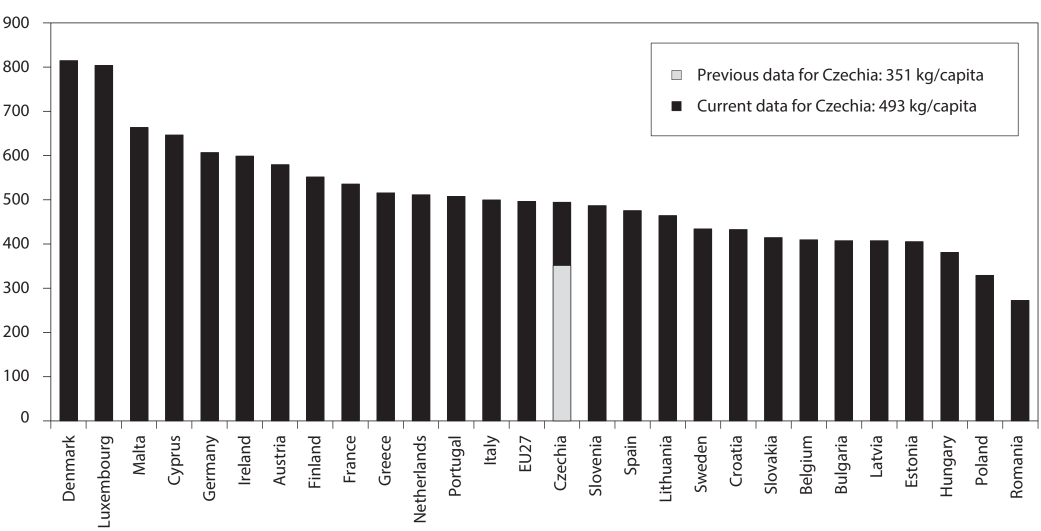
CONCLUSION – CHANGING THE DEFINITION OF MUNICIPAL WASTE

The methodology of compiling the indicator of municipal waste production also underwent changes. The definition of the term used for the Joint Questionnaire of Eurostat and OECD allowed different interpretations; however, for mandatory reporting since 2020, the interpretation of the term municipal waste has been clarified.

Waste produced by households and waste that is comparable in nature and composition to household waste, with the exception of waste from production and waste from agriculture and forestry is considered to be municipal waste. Newly, this indicator also includes waste from enterprises collected outside the municipal collection system, which is similar in nature and composition to household waste and does not come from production. It also includes waste from citizens that is not collected through municipalities. The new methodology reflects the forthcoming changes at the Eurostat level. The methodology for creating the indicator of municipal waste production was widely consulted with Eurostat staff, and external experts from the domestic academic sphere that took part in the audit group.

5.3 million tonnes of municipal waste were generated in 2019, which was by 1.7% more than in the previous year. Municipal waste accounts for about 14% of total waste production. In 2019, there was a total of 499 kg of municipal waste per capita, 6 kg more than a year earlier.

Figure 2 International comparison of municipal waste generated in kilograms per capita (2018)



Source: Eurostat

The use of administrative data for the creation of statistical information has several advantages. The positive effect is undoubtedly a reduction in the administrative burden on respondents. The number of respondents completing the statistical report ODP 5-01 decreased first in 2018 in the segment of municipalities and later in the segment of enterprises. The number of respondents could thus decrease from the original roughly 8 thousand to 1.2 thousand in 2020 while maintaining the quality of data.[2]

References

CZSO. (2019). *Generation, Recovery and Disposal of Waste* [online]. Prague: Czech Statistical Office. <https://www.czso.cz/csu/czso/generation-recovery-and-disposal-of-waste-2019>.

CENIA. (2020). *Mathematical expression of the calculation of the system of waste economy indicators in compliance with the Decree No 383/2001 Sb, on Details of Waste Management.* Prague.

EUROSTAT. (2002). *Regulation (EC) No 2150/2002 on EU waste statistics.* Luxembourg.

EUROSTAT. (2020). *Guidance for the compilation and reporting of data on municipal waste according to Commission Implementing Decisions 2019/1004/EC and 2019/1885/EC, and the Joint Questionnaire of Eurostat and OECD.* Luxembourg.

[2] A new publication *Generation, Recovery and Disposal of Waste – 2020*, published by the CZSO on 30.11.2021, contents the latest results of waste statistics. Available online at: *<https://www.czso.cz/csu/czso/generation-recovery-and-disposal-of-waste-2020>*.

International Conference *Interdisciplinary Information Management Talks* (IDIMT 2021)

Petr Doucek[1] | *Prague University of Economics and Business, Prague, Czech Republic*
Lea Nedomová[2] | *Prague University of Economics and Business, Prague, Czech Republic*

The Interdisciplinary Information Management Talks (IDIMT)[3] conference is among the conferences traditionally organized by the Department of Systems Analysis of the Faculty of Informatics and Statistics at the Prague University of Economics and Business, in cooperation with Johannes Kepler University Linz.

Every year, the IDIMT conference is devoted to a specific topic related to information technologies and their widespread use in society, as well as to management issues and, last but not least, to the education of IT and management professionals. This year's conference called "Pandemics: Impacts, Strategies and Responses" was inspired by the impact of the Covid-19 pandemic on these industries.

The conference's reputation is to be in close touch with current trends and a good indicator of future developments. The conference attracted papers from a total of 133 authors, with 46 submitted papers being accepted together with 11 invited papers. The authors have come from nine different countries: Austria, Czech Republic, Greece, Russia, Slovakia, Uganda, Ukraine, United Kingdom and USA. It was held in early September (1^{st}–3^{rd}) in the historic town of Kutná Hora, in the pleasant environment of a former Jesuit campus that today houses the Central Bohemian Gallery (GASK).

Due to the pandemic, this year's conference was a hybrid conference - it means that some participants, especially those from the Czech Republic, met at the Central Bohemian Gallery, while participants from Austria met at the AIT (Austrian Institute of Technology GmbH) in Vienna. The 'satellite conference site' in Vienna offers to the Austrian participants met, achieving some feeling of a face-to-face conference. These locations were connected into a conference network, to which other conference participants from other countries connected. The connection architecture was very stable, and it turned out that this conference concept with two main centers could be very well used for hybrid conferences.

The conference program was kicked off in the GASK by Petr Doucek, ordinary professor of the Prague University of Economics and Business. After this official opening, Dr. Christian W. Loesch took the floor to acknowledge Professor Gerhard Chroust's 80^{th} birthday and to thank him for his work in the committees of different conferences, including the IDIMT conference. After that, Professor Gerhard Chroust took the floor, professor emeritus at the University of Linz, who recalled the conference's long tradition and the changes over the twenty-nine years of its existence. The next invited speaker, Christian W. Loesch – IBM's former director for Central and Eastern Europe – brought the conference from this general level down into individual topics, with a presentation entitled "Mutual Impact of Covid and ICT".

---

[1] Faculty of Informatics and Statistics, Prague University of Economics and Business, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: doucek@vse.cz.
[2] Faculty of Informatics and Statistics, Prague University of Economics and Business, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: nedomova@vse.cz.
[3] More at: <*https://idimt.org*>.

The main content of his presentation was a highly expert explanation of how the Covid-19 pandemic kept changing and has changed the world of information technologies. His perspective of the impact of the Covid-19 pandemic on ICT starts with the business scenarios of IT companies and their changes over the past year. Another perspective includes the analysis of applications that, due to the pandemic, have spread into cyberspace, such as digital health and robotics applications. In other parts of his presentation, he focused on new materials in technical components, changes in the architecture of both processors and entire information systems. Christian W. Loesch ended his presentation with his vision of the further development of information technologies and their integration into the business environment. The last presentation in the plenary part of the conference was the presentation of Mr. Richard Antonín Novák who introduced the new scientific project "Prague Data Ethics Lab" at the Faculty of Informatics and Statistics. It is an interdisciplinary laboratory at the Department of Systems Analysis of the Prague University of Economics and Business (VŠE), which examines the impact of Big Data/Artificial Intelligence technologies on society. The laboratory's goal is to interconnect different disciplines and workplaces within and outside the Prague University of Economics and Business, and we were thus pleased by the interest of other workplaces, namely the Department of Philosophy of the Prague University of Economics and Business, the Brno University of Technology and the Technical University of Liberec, with which we plan to establish further cooperation. His presentation closed the introductory plenary block. It continued after a break with presentations in the conference's individual sessions.

Among the conference's most-visited sessions were "Digitalization and Industry 4.0 – Changes and Concepts Caused by Covid 19", "Innovations and Strategies in a Pandemic Era", and "Social Media, Fake News, Myths and Ethics". The conference's dominant focus will likely be clear from their names. The "Digitalization and Industry 4.0 – Changes nad Concepts Caused by Covid 19" session presented the results of research on the digitalization of the economy in various countries and on the gradual transition towards a digital economy, primarily in industry, including the Internet of Things. The section also analyzed digitization trends brought about by the Covid-19 pandemic. This session was well complemented by the visionary views in the "Trust in Smart Robotics and Autonomous Systems – Resilient Technology, Economy and Society" and traditional security session „Cyber Security in a Digital World". The "Innovations and Strategies in a Pandemic Era" session contained papers that presented specific impacts of Covid-19 on business models that are newly appearing in Economics. This session also included papers on student start-ups and university activities aiming to support student business projects. The "Social Media, Fake News, Myths and Ethics" session covered trending topics connected with the use of social media and its influence on human society under the impact of Covid-19. Especially important topics of this session were fake news and ethics – both categories very sharp related to information boom during the Covid-19 period.

The remaining sessions primarily covered trends in information science. The "Smart Supply Chain" session covered the creation and presentation of a trust mechanism in the digital economic environment. The "Cybersecurity in a Digital World" session, by now a tradition at the conference, was dedicated to cybersecurity and information security – two highly relevant topics today. The theme of security appeared in other sessions as well, amounting to a thread that interwove the entire conference. The "Performance and Sustainability Management and Corporate Social Responsibility" session was a good complement to the conference's reflections on information science. It brought models for the deployment of new technologies back to the real world of costs and benefits.

As usual, the members of the "Digital Transformation in Crisis Management" section were high-quality experts. The section provided considerations and recommendations that should be implemented during a crisis. We could see a significant change brought about by the Covid-19 pandemic since it was possible to verify various proposed measures and recommendations based on the experience from the last year's pandemic. Interesting was the "COVID-19's Impact on Enterprise Software Development" section that focused on the impact of the pandemic on the work of work teams that develop software. A completely new contribution of the conference was the "COVID-19's Influence on Learning and Teaching" section that discussed the effects of the epidemic on the processes in teaching ICT professionals, presented the experience with used online teaching methods and evaluated the advantages and weaknesses of these methods.

This conference was partially funded through project IGA 409021 of the Faculty of Informatics and Statistics, Prague University of Economics and Business, Česká spořitelna, a.s., and Johannes Kepler University Linz, Austria.

15th Year of the *International Days of Statistics and Economics* (MSED 2021)

Tomáš Löster[1] | *Prague University of Economics and Business, Prague, Czech Republic*
Jakub Danko[2] | *Prague University of Economics and Business, Prague, Czech Republic*

From 9th to 11th September 2021, a worldwide conference of the International Days of Statistics and Economics (MSED) took place at the Prague University of Economics and Business.[3] The conference belongs to traditional professional events; this year, the fifteenth year of this event was held. Prague University of Economics and Business (the Department of Statistics and Probability and the Department of Managerial Economics) was the main organizer, as usual; and was helped by the Faculty of Economics, the Technical University of Košice, and the Ton Duc Thang University, as co-organizers. The conference ranks among important statistical and economic conferences, which can be proved by the fact that Online Conference Proceedings were included in the Conference Proceedings Citation Index (CPCI), which has been integrated within the Web of Science, Clarivate Analytics since 2011.

The traditional goal of this international scientific conference was a presentation of the contributions of individual authors and a discussion of current issues in the field of statistics, demography, economics, and management and their interconnection.

Due to the continuing global situation in connection with COVID 19, this year's conference and presentation were again in a hybrid form (online and real presentation at the university), which caused the participation of foreign nationals to be active.

The online implementation of the conference took place in individual channels of the conference teams in MS Teams (according to partial sections). The number of registered conference participants was a total of 190, of which 102 were foreign, e.g. Russia (63), Slovakia (9), Vietnam (5), Lithuania (4), etc. Other countries from which the participants come include Poland, Switzerland, etc. Among the conference participants were 22 doctoral students.

The received papers were first evaluated in terms of scientific content and suitability of the topic concerning the focus of the conference. After the exclusion of unsatisfactory abstracts, a double independent anonymous review procedure took place in the spring of this year.

We would also like to invite researchers, doctoral students, and the wide professional public to the sixteenth International Days of Statistics and Economics, which will take place at the Prague University of Economics and Business traditionally in early September 2022.

---

[1] Faculty of Informatics and Statistics, Prague University of Economics and Business, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: tomas.loster@vse.cz.

[2] Faculty of Informatics and Statistics, Prague University of Economics and Business, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: jakub.danko@vse.cz.

[3] More at: <*http://msed.vse.cz*>.

Applications of Mathematics and Statistics in Economics (23rd AMSE 2021) International Conference

Stanislava Hronová[1] | *Prague University of Economics and Business, Prague, Czech Republic*

After a year-long pause caused by the Covid-19 epidemic, an international conference called *Applications of Mathematics and Statistics in Economy* (AMSE) took place again this year. Since possibilities to travel were uncertain, the conference was limited to one day (2 September 2021) and an online form. This year's conference was organised by the Department of Statistics from the Faculty of Management and Computer Science, Wroclaw University of Economics. Over 40 experts from the Czech Republic, Slovakia, Poland, and other countries representing the Prague University of Economics and Business in Prague, the Matej Bel University in Banska Bystrica, the Wroclaw University of Economics, the University of Pardubice, the Czech Statistical Office, the Comenius University in Bratislava, the University of California in Davis, and the School of Petroleum Management in Gandhinagar participated in the conference.

It is characteristic for this international conference that knowledge and experience are exchanged, the latest results of research are presented, and new procedures and methods are discussed there. Usual working meetings of representatives of cooperating workplaces from Prague, Wroclaw, and Banská Bystrica as to further heading of scientific and pedagogical cooperation took place in the evening before the conference, also in an online form.

Meetings of the conference took place in five sections as follows: Macroeconomic issues, Microeconomic and financial issues, Old age issues and pension market, Application of statistical methods in economics, Application and history of statistics. It is very difficult to highlight the most interesting papers; therefore let me point out just some of them, which I consider to be of good quality, interesting, and innovative in terms of methodology. Even bigger attention should be paid to those papers the authors of which are young researcher workers.

In the *Macroeconomic issues* section, the biggest attention was caught by the paper called *Aligning the loan-to-deposit ratio with the needs of macroprudential regulation* (by authors Zimková, E. and Boďa, M.). The authors proposed to expand the traditional loan-to-deposit (LTD) ratio as a metric of liquidity and fund towards measuring attainment in financial interme-diation. The proposed metric, financial intermediation measure (FIM), is a normative indicator and amends the descriptive nature of the LTD ratio, and is thus more suited to the needs of macroprudential regulation of financial systems. The metric is grounded in ideas of data envel-opment analysis and is developed in conjunction with a data set on structural indicators provid-ed by the World Bank. Utilization of the FIM was demonstrated

[1] Faculty of Informatics and Statistics, Department of Economic Statistics, Prague University of Economics and Business, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: hronova@vse.cz.

for 48 financial systems of the world using data from 1993 to 2016 under different classifications applied to financial systems. The demonstration shows that the FIM is useful also for performance assessments of market-based financial systems without distinction.

The section of *Microeconomic and financial issues* offered four papers; among them, the biggest attention undoubtedly belongs to a paper of young authors (Fischerová, J., Vincenc, J., Ptáčková, V.) called *Using the business survey and other relevant sources to improve the estimation of financial margins on buying and selling transactions in Czechia.* The authors established the aim of their research work as follows: to find suitable data sources in the area of official statistics that should help solve the question of appropriate capturing of financial margins in the Czech national accounts according to requirements of European institutions (especially in the relation to non-residents). In order to achieve that, they will go through the available data sources about the financial institutions and analyse the current situation in Czechia. Collected information will be helpful for the literature review about the mentioned issue. After that they will evaluate the opportunity to use the Czech *Business and Consumer Survey* (BCS) as the background data source for the additional voluntary questionnaire focused on institutions which filled in the BCS corresponding data. Investment questionnaire under the BCS in the selected services with the edited form will be released for the first time during autumn 2021. The data where respondents evaluate and estimate their investments based on a qualitative scale can be considered as a suitable additional source that will quickly show economic development in the field of financial margins.

Papers in the third section (*Old age issues and pension market*) were devoted to issues related to the population ageing and pension insurance. The paper called *Marriage and individual contracts combining reverse annuity with critical health insurance as a tool for managing the pensioners' budget* (by authors Marciniuk, A. and Zmyslona, B.) was devoted to possibilities of how to interconnect two variants of contracts, marriage and individual. The results of the analy-sis suggested that the choice of the contract version affects the amount of financial flows related to the reverse annuity and the health insurance benefit, depending on the material needs, health condition, and age of the spouses. Financial benefits of the two variants of the contract are dif-ferent.

In the section of *Application of statistical methods in economics*, Jakub Kalenský dealt with issues of real estate prices in various regions of the Czech Republic. In his paper called *Analysis of developments in prices of real estate and construction of predictive models of selected regions of the Czech Republic* he presented the results of the analysis of the development of housing prices for selected regions of the Czech Republic, the design and application of a pre-dictive regional regression model and time series models period of 2000 to 2018. The presented results can be used for the development of management research in the field of banking institu-tions and for top managers in firms.

This year, a usual section called *Application and history of statistics* brought a topic related to the region, in which the conference takes place. A paper called *Reginald Kneifl and the beginnings of the statistical description of "Austrian" Silesia* (by authors Závodský, P. and Šimpach, O.) commemorated an important person, Reginald Kneifl (1761–1826). In the years 1804–1806, Kneifl published a four-volume work (a total of 1 137 pages) *"Topographie des k. K. An-theils von Schlesien,"* in which he described the part of Silesia left after the Austrian-Prussian wars.

A full programme of the AMSE in 2021, including abstracts of all presented papers (from which also a brief description of the contents and aims of all the aforementioned papers follows) can be found at: <*http://www.amse-conference.eu*>. You can also find there information about the history of the AMSE and links to preceding AMSE international conferences.

Papers presented this year at the 23rd AMSE conference will not be published in the proceedings from the conference; nevertheless, they can be offered for a review process in the following journals – Statistika: Statistics and Economy Journal, Silesian Statistical Review, and Forum Statisticum Slovacum.

The proceedings of the past five AMSE conferences (i.e. AMSE 2014, 2015, 2016, 2017, and 2018) have been successfully indexed and are available in the Web of Science database.

The tradition of alternating organisation (Slovakia – Poland – the Czech Republic) fur-ther continues and the 24th AMSE conference (to be organised by colleagues from the departments of statistics of the Prague University of Economics and Business) will take place in the Czech Republic at the turn of August and September 2022.

Mathematical Methods in Economics (MME 2021) International Conference

Petra Zýková[1] | *Prague University of Economics and Business, Prague, Czech Republic*
Josef Jablonský[2] | *Prague University of Economics and Business, Prague, Czech Republic*

The Mathematical Methods in Economics (MME) conference has a very long history and tradition. It is one of the most important scientific events organised in the Czech Republic in the field of operational research, econometrics, mathematical economics, and related research areas. In 2021, the 39th International Conference on Mathematical Methods in Economics was organised in the city of Prague from 8 to 10 September.[3] In addition to the local organiser (the Faculty of Economics and Management, Czech University of Life Sciences, Prague), leading organisers of the MME conference are the Czech Society for Operations Research (CSOR) and the Czech Econometric Society.

The total number of participants in this year's MME conference was almost 120. Due to the Covid-19 pandemic, several participants chose virtual attendance. Participants came from the Czech Republic, United Kingdom of Great Britain and Northern Ireland, China, Poland, the Netherlands, Mexico, Canada, and Slovakia. The programme started with an opening ceremony, where the Chair of the Organising Committee, Robert Hlavatý, introduced the main programme and all of the facilities. After that, the plenary session started with two exciting lectures. The first one, titled *Specification and Estimation of the Effect of Climate on Crop Supply Using Spatial Panel Data*, was presented online by Professor Stephen Clark from the Dalhousie University in Canada. Professor Alessio Ishizaka from the Neoma Business School in France delivered online the second plenary talk about *Visual management tool for multi-criteria decision methods*. After the plenary session, the conference was divided into four parallel sessions. The total number of presentations was more than 80. All accepted papers are published in the *Proceedings of the MME 2021*. As in previous years, they have been submitted for indexing in the Web of Science CPCI database.

It has been a long tradition for PhD students to compete for the best paper during MME conferences. The competition is organised and honoured by the CSOR. All submitted papers were peer-reviewed and the Programme Committee further evaluated the papers with positive referee reports. Eight best-selected papers were presented at the conference in two special sessions and the evaluation committee decided on the winners. The six best papers were awarded after a conference dinner at the Galaxie hotel. Jan Vávra (Charles University, Prague, Czech Republic), with his paper 'GLMM Based Segmentation of Czech Households Using the EU-SILC Database', was the winner of the competition. Karel Kozmík (Charles University, Prague, Czech Republic), with his paper 'Robust First Order Stochastic Dominance

1 Prague University of Economics and Business, Faculty of Informatics and Statistics, Department of Econometrics, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: petra.zykova@vse.cz.
2 Prague University of Economics and Business, Faculty of Informatics and Statistics, Department of Econometrics, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic. E-mail: jablon@vse.cz.
3 More at: <*https://mme2021.czu.cz/en*>.

in Portfolio Optimization', ranked second. Petr Vejmělka (Charles University, Prague, Czech Republic) ranked third with his paper 'State-space modeling of claims reserves in non-life insurance'. Anlan Wang (Technical University of Ostrava, Czech Republic), David Neděla (Technical University of Ostrava, Czech Republic), and Xiaoshan Feng (Technical University of Ostrava, Czech Republic) won the remaining three awards.

The conference was organised at a very good level. All sessions took place in the Faculty of Economics and Management. The welcome evening took place in the hall of the Faculty of Economics and Management.

An essential part of all conferences is a social programme that offers many opportunities to discuss various problems in an informal environment. The organisers have prepared three possible trips. A trip to the Mayrau Mining Museum in Vinařice, which is situated nearby Kladno. A trip to the Old Sewage Treatment Plant in Prague near Stromovka Royal Game Reserve, where participants saw the original technology of wastewater treatment. And a trip to Únětice brewery with beer tasting. The conference dinner took place at the Galaxie hotel near the university campus of the Czech University of Life Sciences, with a barbecue.

This year's annual meeting of the CSOR decided that the 40th MME conference will be organised in the city of Jihlava by the College of Polytechnics Jihlava, on 7–9 September 2022.

Contents 2021 (No. 1–4, Vol. 101)

Statistika: Statistics and Economy Journal, Vol. 101 (1) 2021

ANALYSES

Statistika: Statistics and Economy Journal, Vol. 101 (2) 2021

ANALYSES

INFORMATION

Statistika: Statistics and Economy Journal, Vol. 101 (3) 2021

ANALYSES

METHODOLOGY

Statistika: Statistics and Economy Journal, Vol. 101 (4) 2021

ANALYSES

CONSULTATION

INFORMATION

Papers
We publish articles focused at theoretical and applied statistics, mathematical and statistical methods, conception of official (state) statistics, statistical education, applied economics and econometrics, economic, social and environmental analyses, economic indicators, social and environmental issues in terms of statistics or economics, and regional development issues.

The journal of Statistika has the following **sections**:

The ***Analyses*** section publishes complex and advanced analyses based on the official statistics data focused on economic, environmental, social and other topics. Papers shall have up to 12 000 words or up to 20 1.5-spaced pages.

Discussion brings the opportunity to openly discuss the current or more general statistical or economic issues, in short what the authors would like to contribute to the scientific debate. Contribution shall have up to 6 000 words or up to 10 1.5-spaced pages.

In the ***Methodology*** section we publish articles dealing with possible approaches and methods of researching and exploring social, economic, environmental and other phenomena or indicators. Articles shall have up to 12 000 words or up to 20 1.5-spaced pages.

Consultation contains papers focused primarily on new perspectives or innovative approaches in statistics or economics about which the authors would like to inform the professional public. Consultation shall have up to 6 000 words or up to 10 1.5-spaced pages.

Book Review evaluates selected titles of recent books from the official statistics field (published in the Czech Republic or abroad). Reviews shall have up to 600 words or 1–2 1.5-spaced pages.

The ***Information*** section includes informative (descriptive) texts, information on latest publications (issued not only by the Czech Statistical Office), or recent and upcoming scientific conferences. Recommended range of information is 6 000 words or up to 10 1.5-spaced pages.

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 — Authors and Contacts — Abstract (max. 160 words) — Keywords (max. 6 words / phrases) — Introduction — 1 Literature survey — 2 Methods — 3 Results — 4 Discussion — Conclusion — Acknowledgments — References — Annex (Appendix) — Tables and Figures (for print at the end of the paper; for the review process shall be placed in the text).

Authors and Contacts
Rudolf Novak,1 Institution Name, City, Country
Jonathan Davis, Institution Name, City, Country
1 Address. Corresponding author: e-mail: rudolf.novak @domainname.cz, phone: (+420)111222333.

Main Text Format
Times 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 <u>underline</u> 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.

References in the Text
Place references in the text enclosing authors' names and the year of the reference, e.g. „…White (2009) points out that…", „…recent literature (Atkinson and Black, 2010a, 2010b, 2011; Chase et al., 2011: 12–14) conclude…". Note the use of alphabetical order. Between the names of two authors please insert „and", for more authors we recommend to put „et al.". Include page numbers if appropriate.

List of References
Arrange list of references alphabetically. Use the following reference styles: [book] HICKS, J. (1939). *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory.* 1st Ed. Oxford: Clarendon Press. [chapter in an edited book] DASGUPTA, P. et al. (1999). Intergenerational Equity, Social Discount Rates and Global Warming. In: PORTNEY, P., WEYANT, J. (eds.) *Discounting and Intergenerational Equity.* Washington, D.C.: Resources for the Future. [on-line source] CZECH COAL. (2008). *Annual Report and Financial Statement 2007* [online]. Prague: Czech Coal. [cit. 20.9.2008]. <http://www.czechcoal.cz/cs/ur/zprava/ur2007cz.pdf>. [article in a journal] HRONOVÁ, S., HINDLS, R., ČABLA, A. (2011). Conjunctural Evolution of the Czech Economy. *Statistika: Statistics and Economy Journal*, 91(3): 4–17. [article in a journal with DOI]: Stewart, M. B. (2004). The Employment Effects of the National Minimum Wage [online]. *The Economic Journal*, 114(494): 110–116. <http://doi.org/10.1111/j.0013-0133.2003.0020.x>.
Please **add DOI numbers** to all articles where appropriate (prescribed format = link, see above).

Tables
Provide each table on a separate page. Indicate position of the table by placing in the text "<u>insert Table 1 about here</u>". 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. Separate thousands with a space (e.g. 1 528 000) and decimal points with a dot (e.g. 1.0). Specify the data source below the tables.

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 "<u>insert Figure 1 about here</u>". 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. Specify the source below the figures.

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

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. Articles for the review process are accepted continuously and may contain tables and figures in the text (for final graphical typesetting must be supplied separately as specified in the instructions above). Please be informed about our Publication Ethics rules (i.e. authors responsibilities) published at: http://www.czso.cz/statistika_journal.

Managing Editor: Jiří Novotný
Phone: (+420) 274 054 299 | **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 66 per copy + postage.
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

Czech Statistical Office | Information Services
Na padesátém 81 | 100 82 Prague 10 | Czech Republic
Phone: (+420) 274 052 733, (+420) 274 052 783
E-mail: objednavky@czso.cz

Design: Toman Design
Layout: Ondřej Pazdera
Typesetting: Družstvo TISKOGRAF, David Hošek
Print: Czech Statistical Office

101st year of the series of professional statistics and economy journals of the State Statistical Service in the Czech Republic: ***Statistika*** (since 1964), ***Statistika a kontrola*** (1962–1963), ***Statistický obzor*** (1931–1961) and ***Československý statistický věstník*** (1920–1930).

Published by the Czech Statistical Office
ISSN 1804-8765 (Online)
ISSN 0322-788X (Print)
Reg. MK CR E 4684