

LABOUR MARKET CRISIS MANAGEMENT AFTER CRISIS OF 2008 – INTERVENTION EXPENDITURE AND “EUROPE 2020” INDICATORS

Tímea GYÖRI^a

^a Hungarian University of Agricultural and Life Sciences, Doctoral School of Economics and Regional Sciences H-2100 Gödöllő, Páter Károly street 1. gyoritimi05@gmail.com

Cite this article: Györi, T. (2021). Labour market crisis management after crisis of 2008 – Intervention expenditure and “Europe 2020” indicators. *Deturope*. 13(3), 4-31.

Abstract

The primary aim of this study is to explore how the Member States of the European Union have responded to the crisis, what labour market interventions were preferred, how the structure of labour market expenditures changed between 2008 and 2018. On the other hand, examines the connections between the indicators of the Europe 2020 strategy, as well as the possibilities of condensing the indicators into principal components. Along the dimensions of the created main components, the Member States were grouped by K-mean cluster analysis. The paper also analyses the relationship between the established clusters and the labour market expenditures of each Member State.

Keywords: labour market, Europe 2020, correlation analysis, principal component analysis, cluster analysis

INTRODUCTION

The financial and economic crisis that took place at the end of the first decade of the 2000s also spread to the labour market, as a result of which the European Union's previous employment growth trend was, interrupted (European Commission, 2010). According to Eurostat Labour force surveys in 2009, the total number of unemployed in the 27 Member States increased by 4 million. The unemployment rate (percentage of active population) of the 20–64 age groups exceeded the 9% by the last quarter of 2009 (Fig. 2). By the first quarter of 2010, the employment rate of the priority age group had fallen to 68% of the total population and to 75.5% of the active age population (European Commission, 2013).

As a result of the recession, the EU's economic growth and employment rate have also lagged behind those in the rest of the world. For a sustainable future, the European Commission has set out its Europe 2020 strategy for smart, sustainable and inclusive growth for 2010–2020. The Union's budget for 2014–2020 has been set in line with the objectives of the Europe 2020 strategy.

The main challenge for the Europe 2020 strategy is to be able to prevent an instinctive return to the pre-crisis situation. The crisis has also highlighted the interdependence of our economies

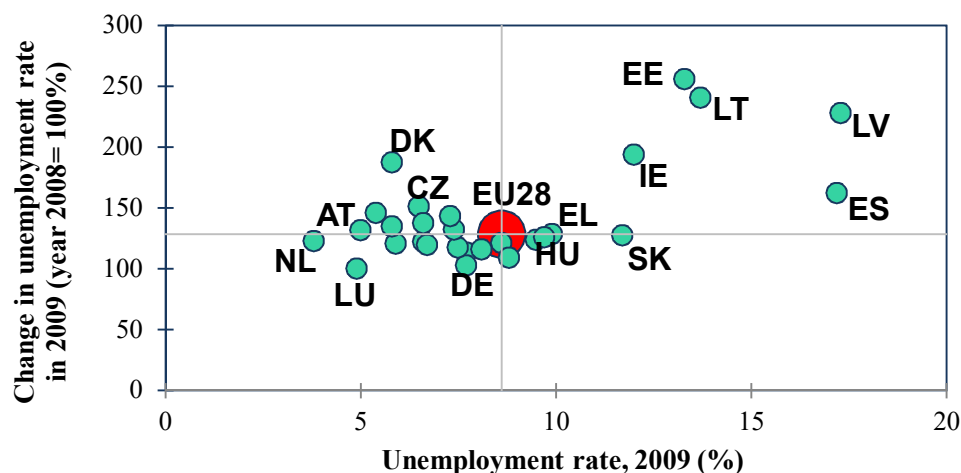
(reforms implemented in one country also affect the performance of others); and our reflection on the crisis that we are much more effective together. It follows from the above that social and territorial cohesion is the basis for achieving the set goals at both Member State and regional level (European Commission, 2010). All EU Member States are committed to achieving the objectives of the Europe 2020 strategy.

Each country has developed a differentiated set of tools to deal with the general labour market crisis. Following the increase in unemployment, social spending increased in almost all Member States (Kálmán, 2015). Many tools for job creation are used in the countries, which involve significant expenditures (Ékes, 2011).

THEORETICAL BACKGROUND

The effects of the economic crisis, which began with the collapse of the US real estate market, have been felt to varying degrees in EU Member States. On the labour market effects at the national level, Fig. 1 provides information. The data refer to the 20–64 age group within the active population. The *x*-axis of the figure shows the 2009 level of the unemployment rate, while the *y*-axis shows the change compared to the previous year (year 2008= 100%). The two linear lines represent EU averages.

Figure 1 Unemployment rate – 2009 level versus changes from during the crisis level



Source: Own calculation and edition base on Eurostat data

Declining labour demand due to the economic downturn between 2008 and 2009 hit the Baltic countries, Spain and Ireland hardest. The highest unemployment rates were recorded in Lithuania in 2009 (17.3%), followed by Spain (17.2%), Latvia (13.7%), Estonia (13.3%) and Ireland (12.0%). Unemployment did not increase in Luxembourg, in Germany also only

minimally. In half of the Member States, the unemployment rate remained below the EU average. The number of unemployed in Denmark has almost doubled, but did not exceed the EU average.

Among the theories dealing with the economic growth and development of regions, endogenous growth theories emphasize the importance of local conditions. Models consider knowledge to be endogenous within an area (Romer, 1994; Capello, 2007; Lengyel, 2010). According to Rehnitz, Barsi, Szabó & Németh (2003), the knowledge base of the workforce is high in successful regions. Living in a lower socio-economic area (SES) has also been linked to poor physical and mental health outcomes, as SES indicators (e.g. income, poverty, education) are major predictors of health and health inequalities around the world (Diez Roux, Borrell, Haan, Jackson & Schultz, 2004; Robinette, Charles & Gruenewald, 2017).

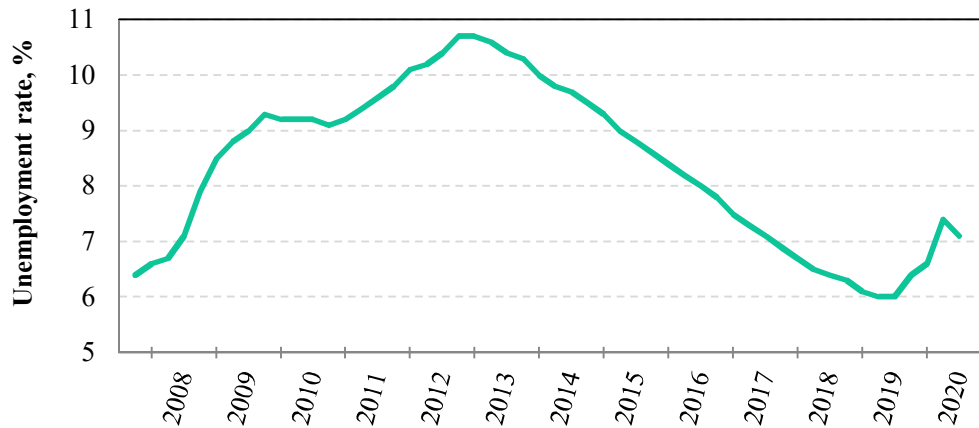
Numerous empirical studies demonstrate that the human capital variable can explain a very large proportion of the variance in per capita GDP between countries (Mankiw, Romer & Weil, 1992). According to experts, the most important element of the long-term solution of economic problems is the drastic increase of the education and knowledge of the population (Jankó, 2010; Hajdú, 2020). A higher level of education can provide better employment opportunities, and by increasing the employment rate, poverty can decrease and well-being increases (Egri, Törőcsik & Tanczos, 2009; Egri, 2017).

The objectives of the Europe 2020 strategy have been formulated taking into account the factors influencing economic growth. Objectives of inclusive growth include raising the employment rate of the 20–64 age groups to 75% and raising the share of tertiary graduates to 40% in the 30–34 age group. Furthermore, one of the EU's education priorities is to reduce the number of early school leavers across the EU to below 10% by 2020. One of the objectives of the Strategic Framework for European Cooperation and Training (Education and Training 2020) is that, on average, at least 15% of adults at European level should participate in lifelong learning by 2020.

The reform efforts introduced in the EU have contributed to job-creating economic growth. Unemployment rates for the 20–64 age group returned to pre-crisis levels in 2018 (Fig. 2), but remained high in several Member States, for example Greece, Spain, Italy, Croatia and Cyprus (European Commission, 2018/a). The relatively short relief period was ended by an economic recession in the context of the coronavirus epidemic at the end of 2019. The Covid19 crisis has had a severe impact on EU labour market outcomes and is expected to shape them in the future. In the current economic climate, financial support has been provided to Member States by the

new SURE resource on the one hand and the Recovery and Resilience Building instrument at the centre of the "Next Generation EU" program on the other (European Commission, 2020).

Figure 2 Unemployment rate for 20–64 age group in EU28 (2008–2020)



Source: Own edition base on Eurostat data

At European level, the Labour Market Policy (LMP) statistics available at national level are considered directly relevant and are used to monitor and evaluate the employment guidelines. LMP's services cover all services and activities of the Public Employment Services as well as other publicly funded services provided to jobseekers (European Commission, 2018/b). Labour market policies vary from one Member State to another, depending on national needs and priorities. National LMP interventions are classified according to the type of activities according to the rules set out in the LMP methodology. The interventions can be grouped as follows: *labour market services, training, employment incentives, supported employment and rehabilitation, direct job creation, start-up incentives, out-of-work income maintenance and support, early retirement*. The data is fairly complete, but efforts are being made to further improve coverage. LMP data become available on average 18 months after the end of the reference year due to the complexity and voluntary nature of data collection.

Objective of research

The aim of the study is, on the one hand, to reveal how the Member States of the European Union have reacted to the crisis; what labour market interventions were preferred, how the structure of labour market (LMP) expenditures changed between 2008 and 2018. On the other hand, it examines between 2008 and 2018 the evolution and context of the five selected indicators of the Europe 2020 strategy (*intramural R&D expenditure % of GDP, employment*

rate at 20–64 aged, tertiary educational attainment at 30–34 aged, early leavers from education and training, people at risk of poverty or social exclusion).

As part of the empirical research, I used the data from 2018 to examine how effectively the selected indicators of the Europe 2020 strategy can be condensed into a principal component. Furthermore, I researched whether the inclusion of other variables could create a principal component that compresses a significant portion of the information. For the preliminary screening of potential background variables, the set of TOP7 strongest and most correlated background variables were compiled.

The main components produced using the 2018 data was subjected to a K-means cluster analysis in order to form groups of Member States along these dimensions. I examined the relationship between the established clusters and the 2018 LMP expenditures of each Member State.

DATA AND METHODS

During the preparation of the study, the emphasis was basically on the analysis of statistical data collected from secondary sources. The territorial basis of the analysis is the 28 Member States of the European Union. The data used were provided by the EUROSTAT database. Due to the temporal and spatial availability of the vast majority of the databases used (e.g. LMP statistics); the data at the Member State level in 2008 and 2018 were processed in the study. Data for 2008 illustrate the pre-crisis labour market situation; while 2018 already represents a return to 2008 (unemployment reached pre-crisis levels in this year).

As a first step of the research, the changes in the level of labour market expenditures as a % of GDP and the unemployment rate in the 20–64 age group were compared at the Member State level. I then analysed the structure of LMP expenditures at 2008 and 2018. Data on LMP expenditure by type of measure are not available for the United Kingdom for both years and for Croatia in 2008. The data are given in Tables 8 and 9 in the Appendix.

In the next phase of the research, I examined the changes in the employment and qualification indicators determining the inclusive growth of the Europe 2020 strategy between 2008 and 2018. Data for the selected indicators for 2008 and 2018 are presented in Tab. 10 in the Appendix. The direction and magnitude of the correlations between the indicators were revealed by Pearson's correlation analysis in both years studied. The value of the coefficient varies between +1 and -1; the stronger the relationship is, the closer the absolute value of the

coefficient is to 1. The coefficient is strong in absolute terms in the range of 0.7–1, medium at 0.3–0.7 intervals, and indicates a weak correlation at 0–0.3 intervals (Nemes Nagy, 2005). Depending on the results obtained in the study, the chosen significance levels were 1% and 5% (i.e., $p = 0.01$ and $p = 0.05$).

In the next research unit – based on the 2018 data – I researched the possibilities of condensing the Europe 2020 indicators and background variables into a composite indicator by factor analysis (Principal components – PCA). A basic database of 25 variables was compiled for the analysis. I collected the variables of the basic database around the topic of the main indicators defined in the Europe 2020 strategy, supplemented by two accessibility indicators (Tab. 1).

Of the 25 variables included in the study, 5 lacked data for 2018 data (*LMP, job vacancy rate, people at risk of poverty or social exclusion, road/rail and navigable inland waterways networks, individuals regularly using the internet*), which totals represented 1.6% of the values. The variables examined showed a general pattern of lack of data. In the case of a general lack of data, the pattern has no specialty (Oravecz, 2008).

The multiple imputation (MI) proposed by Rubin (1987) was first, selected as a method to address data gaps. Multiple imputation is one of the most widely used missing data management techniques (Chung & Cai, 2018). The method can be applied to virtually any data structure and model type (Allison, 2003). Ginkel, Kroonenberg & Kiers (2014) demonstrated in their study that multiple imputation can be safely applied in the context of PCA. MI is already effective for a small number of imputations, depending on the percentage of data missing (Allison, 1999). In general, we can use imputation for variables where a maximum of 30–40% of the data per variable is missing, but the lack of data in the entire database does not exceed 10–15%. To address the lack of data, a linear regression model was developed for non-deficient observations as predictors.

Pearson's correlation analysis was also performed to examine the variables in the base and imputed databases. The strength of the relationships showed a minimal improvement of a few hundredth % as a result of imputation for one or two variables. In the further stage of the research, the basic data of the analyses were provided by the imputed database. The basic descriptive statistics (minimum, maximum, mean, standard deviation) of the imputed basic database are given in Tab 1.

Table 1 Descriptive statistics of examined indexes (2018)

Examined indicators (2018)	Min.	Max.	Mean	Std. dev.
Intramural R&D expenditure % of GDP	0,50	3,32	1,64	0,85
GDP/capita in PPS (% of EU27)	51,00	261,00	101,79	41,85
Income quintile share ratio (S80/S20)	3,03	7,66	4,89	1,21
Median income (€)	6849	27698	16724	5758,27
Labour market policy (LMP), % of GDP	0,09	2,80	1,19	0,74
Labour cost index (LCI)	5,40	43,90	22,35	12,12
Life expectancy (years)	75,00	83,50	80,25	2,73
Impact of social transfers (other than pensions) on poverty reduction %	16,07	53,67	34,24	10,46
Employment rate (20–64 aged)	59,50	82,40	73,82	5,29
Activity rate (20–64 aged)	70,40	87,30	78,80	4,02
Employed ICT specialists, % of total employed	2,20	6,80	4,12	1,24
Job vacancy rate	0,60	5,50	2,16	1,12
Tertiary educational attainment (30–34 aged)	24,60	57,60	42,71	8,79
Lifelong learning – Adult participation in % of unemployed	0,90	31,40	11,55	7,61
Early leavers from education and training	3,30	17,90	9,24	3,94
Less than primary, primary and lower secondary education (levels 0–2, 20–64 aged)	3,80	27,30	14,45	6,07
People at risk of poverty or social exclusion	12,20	32,80	21,76	5,39
Unemployment rate (20–64 aged)	2,20	19,30	6,37	3,58
Long-term unemployment (12 months and more) % of unemployed	13,70	70,10	34,97	13,47
Less than primary, primary and lower secondary education (levels 0–2) % of unemployed	11,80	52,60	27,08	11,36
At risk of poverty rate (cut-off point: 60% of mean equivalised income)	10,80	28,70	20,87	4,52
Severely materially deprived people	1,30	20,90	6,60	4,84
People living in households with very low work intensity	4,50	14,60	8,50	2,54
Accessibility indicators				
Road/rail and navigable inland waterways networks (km/1000km ²)	0,00	81,00	22,00	20,11
Individuals regularly using the internet	62,50	95,20	81,52	8,26

Source: Own construction and calculation based on Eurostat data

In order to group the designated indicators of the Europe 2020 strategy in 2018 and to reduce the number of variables, I performed a factor analysis after data-standardizing. The method of extraction was Principal components analysis. Principal component analysis (PCA) is widely used in data processing and downsizing (Zou, Hastie & Tibshirani, 2006). In the analysis, I used the results of the correlation test as a starting point; for the analysis is based on the Pearson correlation matrix. The essence of the method is the existence of linear correlations between the individual variables, on the basis of which we form new indicators by linear regression. From the set of linearly correlated variants in pairs, uncorrelated principal components were generated by orthogonal transformation (Shlens, 2014; Kovács, 2014; Liu, Singleton & Arribas-Bel, 2019). These new uncorrelated variables maximize variance (Jolliffe, 2005). To achieve well-interpretable results, the condition that the number of observation units be at least twice the number of variables must be met (Bottlik, 2008). The properties of PCA have some

undesirable features when the variables have different units of measurement. To overcome this undesirable feature, it is common practice to begin the analysis by standardizing the variables (Jolliffe & Cadima, 2016). It is worth omitting a variable from the analysis if the explained proportion would be too low. If the communality (multiple coefficient of determination) is less than 0.25, then the variable does not correlate moderately with any main component (Kovács, 2014). Fabrigar et al. (1999), the value of communality above 0.7 can be considered high, below 0.4 it can be considered low. One of the most important issues in factor analysis is the question of the appropriate number of factors (principal components) to be extracted (Hakstain & Muller, 1973; Cattell & Vogelman, 1977). This is a critical decision because the number of factors has a direct influence on the subsequent parameter estimates and the interpretation of the solution (Lambert, Wildt & Durand, 1990).

To ensure the validity of the factorization, both the Kaiser-Meyer-Olkin measurement and the Bartlett spherical test were performed. The Kaiser-Meyer-Olkin (KMO) criterion is used to judge the suitability of variables for factor analysis, with a value between 0 and 1. Kaiser (1981) originally recommended that the baseline criterion for factorability should be 0.50. Based on Kaiser's recommendation, I interpreted the KMO index as follows: $KMO > 0.9$ marvellous, $0.8-0.9$ meritorious, $0.7-0.8$ middling, $0.6-0.7$ mediocre, $0.5-0.6$ miserable, $KMO < 0.5$ unacceptable. The basic hypothesis of Bartlett's chi-square test is that the original variables are independent (Arsham & Lovric, 2011), the variables are suitable for factor analysis if the homogeneity test hypothesis can be rejected.

Principal component analysis was performed with several parameters and components. To pre-screen the potential background variables of the analysis, a set of TOP7 strongest and most correlated background variables was compiled. I influenced the number of significant principal components on the one hand directly and on the other hand by giving my own values (Eigenvalues) based on the Kaiser criterion. For a factor to have positive Kuder-Richardson reliability (Cronbach's alpha), it is necessary and sufficient for the associated eigenvalue to be greater than 1 (Kaiser, 1960). Finally, to compile the principal components of the Europe 2020 indicators extended with other components, I used the group of factors for which the Cronbach's alpha value is greater than 0.70.

I performed a K-means cluster analysis with the principal components created during the empirical research. Cluster analysis is suitable for arranging (clustering) data arrays into homogeneous groups, so it essentially functions as a dimension-reducing method (Rao, 1971). The essence of clustering is that the data within each cluster are similar in some dimension, and

in this respect they differ from the elements of other clusters (Bardhoshi, Um & Erford, 2021). K-means clustering (MacQueen, 1967) is a commonly used method for automatically partitioning a data set into k groups. In the K-mean cluster analysis, the grouping of the data set is based on the selection of the central point and the calculation of the Euclidean distance (Bansal, Sharma & Goel, 2017). In a research with a small sample, K-means clustering can be used successfully (Székelyi-Barna, 2005). The number of clusters needs to be given before the analysis (Birkner, Peter & Fehérvölgyi, 2012). In the present research, the determination of the number of clusters was based on previous analysis (hierarchical cluster analysis). There is no clear specification for the sample elements and variables involved in the clustering study, the researcher should be careful to consider only those criteria in the grouping procedure that can be considered relevant based on theoretical considerations (Simon, 2006).

Our new CLU variable obtained by clustering is nominal, which allows the use of the ETA coefficient. ETA^2 in this case is the quotient of the sum of squares between the groups and the total deviation. The discriminant power of the variables was checked in ANOVA (analysis of variance) standard deviation resolution table. F test performed in the framework of ANOVA is a procedure for testing the equality of standard deviations, in which the null hypothesis is that the variance of two normally distributed samples is the same. Several studies have confirmed that with a given experimental error rate, F test is the strongest statistical test for examining variance (Ramsey, 1978; Harper, 1984; Ramsey & Ramsey, 2013). As a control study, the rank-based Kruskal-Wallis test was run as an effective alternative to one-way analysis of variance (Breslow, 1970; Vargha & Delaney, 1998; Ostertagová, Ostertag & Kováč, 2014).

RESULTS

Change in LMP expenditures 2008–2018

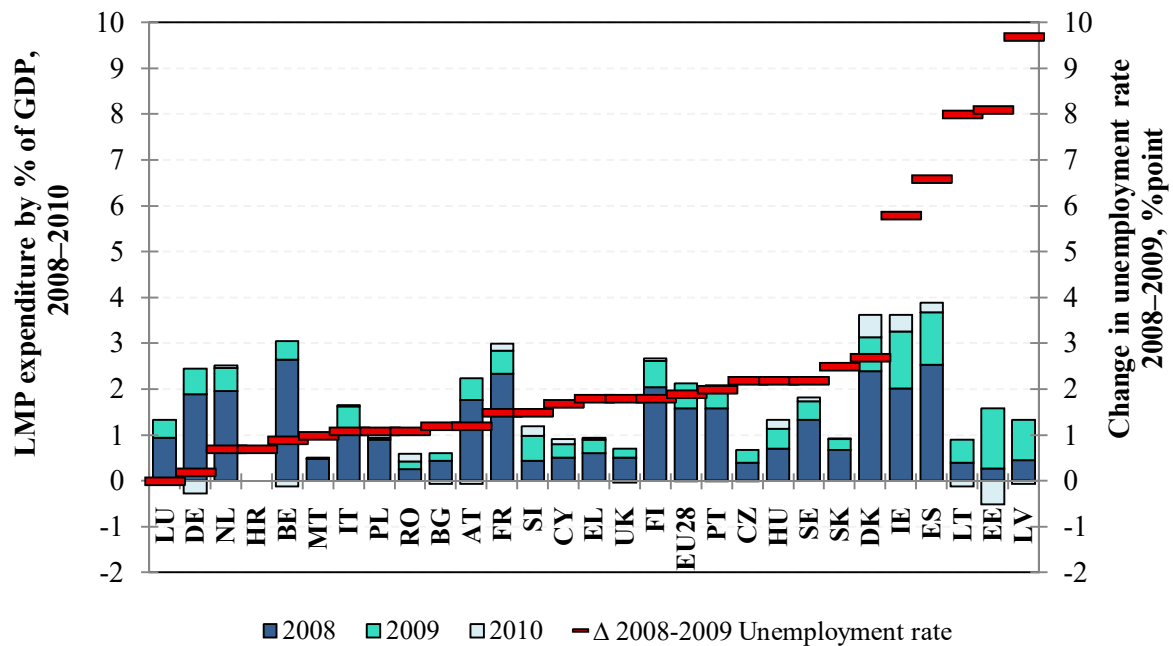
I first examined the extent to which unemployment increased in the first year of the crisis and, in this context, how the level of LMP spending changed between 2008 and 2010 at Member State level. LMP expenditure as a percentage of GDP related to crisis management measures ranged from 0.427% (Romania) to 3.675% (Spain) in the first year of the crisis. Expenditure on interventions has risen in all Member States after the crisis.

Fig. 3 illustrates the cumulative development of expenditures as a % of GDP between 2008 and 2010 and the change in the number of unemployed between 2008 and 2009, expressed in percentage points. In those Member States where unemployment has risen sharply (Lithuania,

Estonia, Latvia, Spain, Ireland), spending has also risen sharply. Member States' expenditure as a share of GDP is presented in ascending order based on changes in the unemployment rate of the 20–64 age groups.

Compared to 2008, labour market expenditure as a % of GDP in 2009 increased almost 6 times (by 1.31 %points) in Estonia, more than 1.5 times (by 1.25 %points) in Ireland and by almost 1.5 (1.15 %points) in Spain. In most Member States, expenditure has already increased to a lesser extent in 2010. On average in the European Union, LMP expenditure peaked at 2.124% this year. There was a significant decrease compared to the previous year in Estonia, Denmark, and Belgium; furthermore, to a lesser extent in Lithuania, Latvia, Luxembourg, Austria and Bulgaria.

Figure 3 LMP expenditure by percentage of GDP and unemployment rate (2008–2010)



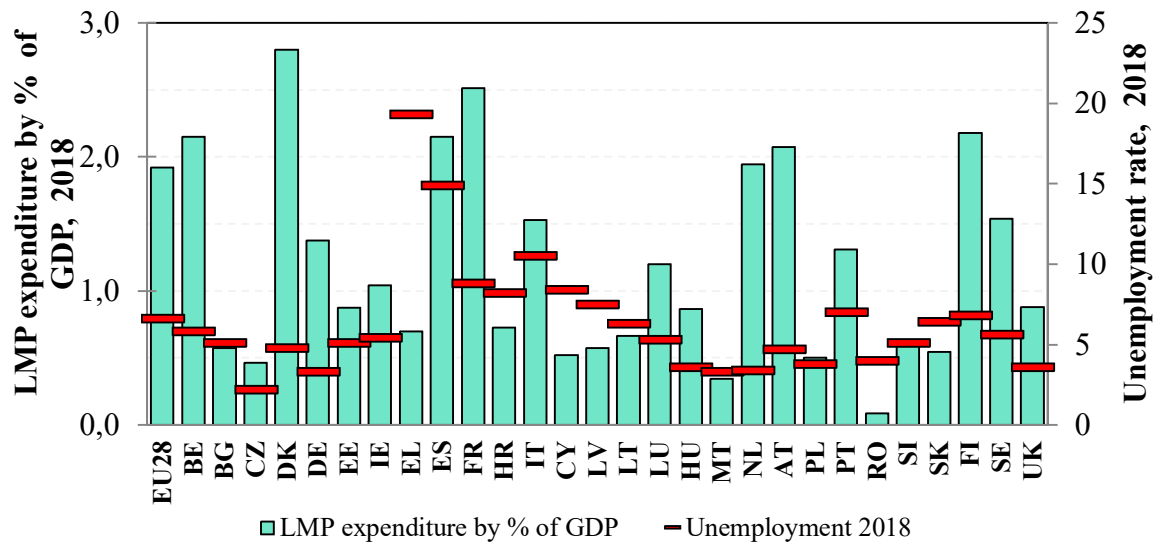
Source: Own editing by Eurostat data

After 2010, labour market expenditures as a share of GDP decreased with differences by Member State, but gradually at the EU level, and according to EUROSTAT data, by 2018, the EU as a whole fell below the pre-crisis level (Fig. 4). The significant variance in the expenditure ratio is due to the values of the following Member States:

- Over 2%: France (2.84%), Denmark (2.80%), Finland (2.18%), Spain (2.15%), Belgium (2.15%) and Austria (2.07%)
- Less than 0.5%: Romania (0.09%), Malta (0.35%) and the Czech Republic (0.46%).

The unemployment rate for the 20–64 age group is the highest in Greece (19.3%), while it ranks only 17th in terms of LMP expenditures. In addition to Greece, unemployment is also above 10% in Spain and Italy.

Figure 4 LMP expenditure as a percentage of GDP in EU Member States (2018)

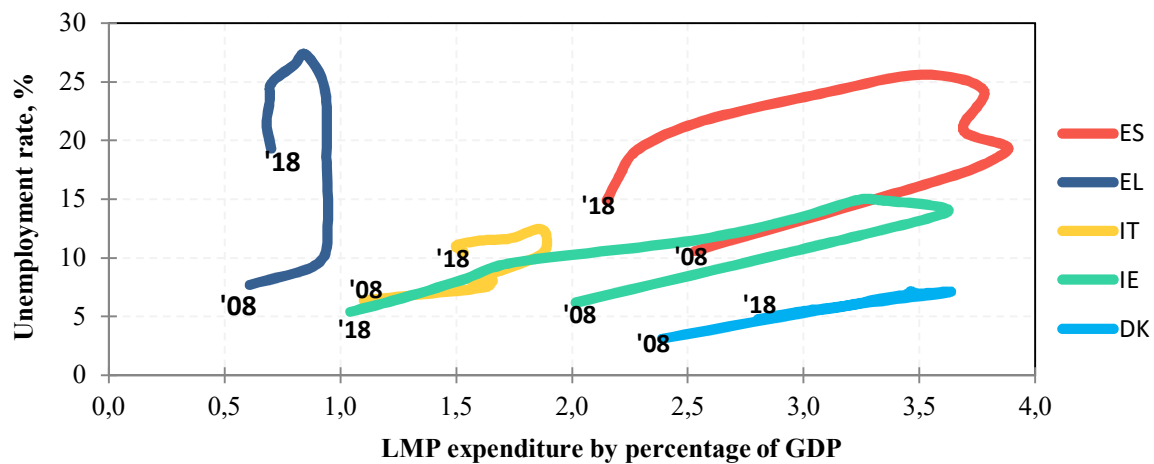


Source: Own editing by Eurostat data

Five Member States were selected in the study. The selection was based on LMP expenditure as % of GDP (<3.5%) and the unemployment rate (<10%). For the Member States selected on the basis of the criteria (Denmark, Ireland, Spain, Greece and Italy), unemployment was plotted as a function of LMP expenditure (Fig. 5).

A similar trend emerged in the curves illustrating the variables for Spain, Italy, Ireland and Denmark. Along with rising unemployment, the volume of LMP expenditures is increasing, and the curves are moving outwards. After 2013, there was a change, with the curves shifting to the origin, so in addition to improving unemployment data, Member States spent less on labour market interventions.

In Greece, the labour market developed differently from in the Member States examined above. In this Member State more severely affected by the crisis, intervention spending has not risen with soaring unemployment. The unemployment rate has also been on a declining trend in Greece since 2013, but there is certainly a lack of intervention. The unemployment rate for the 20–64 age group has barely fallen below 20%; here was the worst situation at EU level in 2018, next to Spain and Italy.

Figure 5 LMP expenditure by percentage of GDP and unemployment (2008–2018)

Source: Own editing by Eurostat data

Within the toolkit to address labour market anomalies, the weight of each measure varies significantly. In the framework of the research, I examined the structure of LMP expenditures in 2008 and 2018, which showed a very small change at the EU level, while at the Member State level there were more significant shifts in the share of each type of measure.

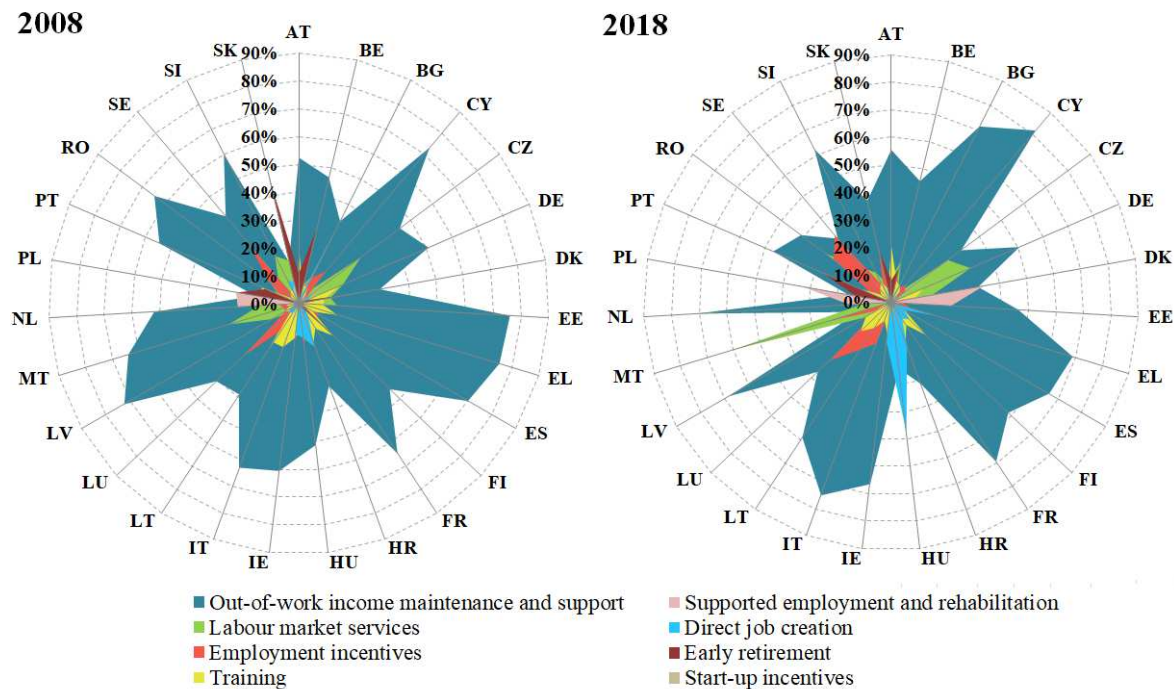
On average, *out-of-work income maintenance and support* expenditures are the most significant in the EU, with a slight change in this category between 2008 and 2018. The share of *early retirement* expenditures within expenditures decreased by 3.3 percentage points, while *employment incentives* increased by 2.6 percentage points, the change in further measures was not significant, hovering around +/- 1 percentage points.

The LMP expenditure (as a % of GDP) of most Member States is the highest in both years in the *out-of-work income maintenance and support* category (Fig. 6). The composition of LMP expenditure in Bulgaria, Poland and Slovakia in 2008 was different from the average. In Bulgaria, when examining the composition of the labour market toolkit, *direct job creation* stands out (37%), In Slovakia, *early retirement* (47%), while in Poland, *early retirement* and *supported employment and rehabilitation* interventions were prioritized by 23–23%. In Belgium, too, more than average, 27% of resources were spent on *early retirement*. There are also discrepancies in the Czech Republic and Malta, where more than a quarter of spending goes to the *labour market services* category.

In the decade following the crisis, the proportions of types of measures changed in several Member States. In those countries where the structure of expenditures was different from the average in 2008, there was equalization towards average values. In Bulgaria, *direct job creation*, which previously stood at 37%, fell to 10% in 2018, accompanied by an increase in *out-of-work income maintenance and support* (72%). In Poland, less was spent on *early*

retirement and increased *supported employment and rehabilitation* and *employment incentives* spending by 10–10%. In Slovakia, *early retirement*, which was previously high, also declined and *out-of-work income maintenance and support* and *employment incentives* expenditures increased. In Denmark, the largest share of spending was on *supported employment and rehabilitation* in 2018, while 17% less was spent on *early retirement*. In addition to Denmark, Poland also spends one third of spending on *supported employment and rehabilitation*. *Labour market services* spending remains significant in Malta and the Czech Republic. Within the system of tools for dealing with labour market anomalies, the expansion of the possibility of public employment in Hungary has been given priority. *Direct job creation* increased from 12% to 48%, while *out-of-work income maintenance and support* spending was halved.

Figure 6 Composition of LMP expenditure by EU Member State (2008, 2018)



Source: Own editing by Eurostat data

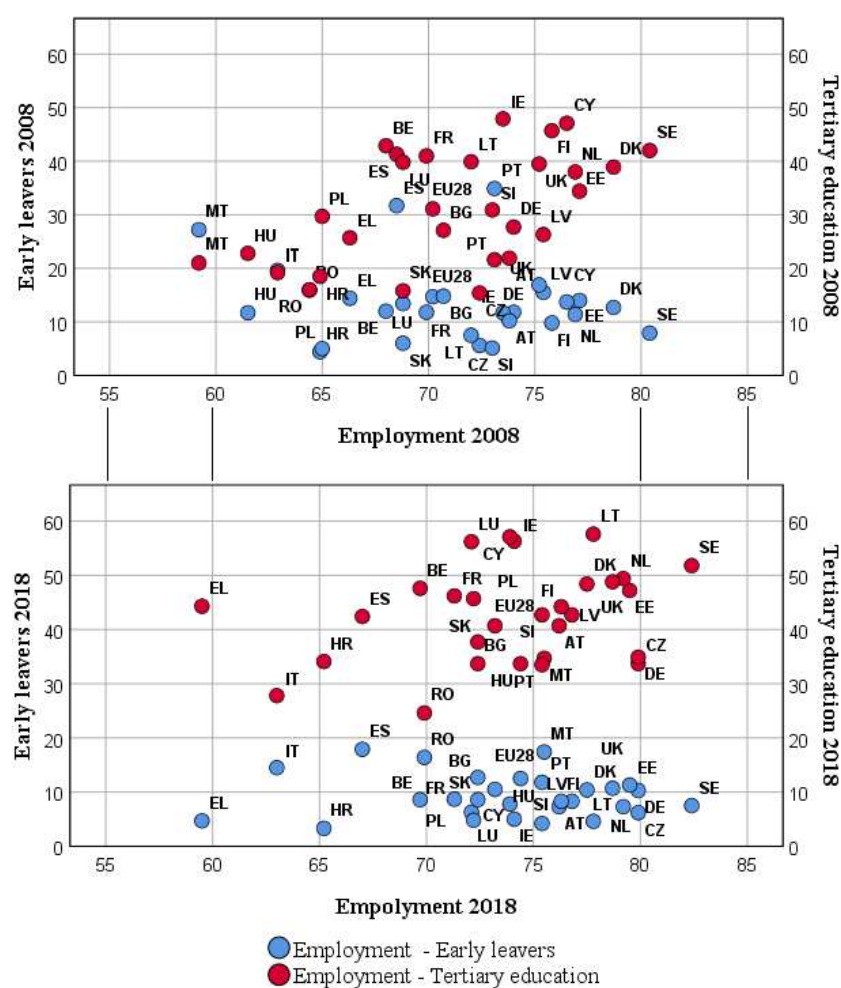
Following, I examined the changes in the indicators of inclusive growth in the Europe 2020 strategy between 2008 and 2018, as well as the correlations of the indicators.

Change and correlation of Europe 2020 indicators (2008-2018)

Indicators of inclusive growth are of paramount importance for the labour market. The change of the indicators between 2008 and 2018 is illustrated in Fig.7. The figure shows the qualification variables at the national level (the proportion of those who drop out of school and the proportion of those with a higher education degree) as a function of the employment rate.

The first scatter plot illustrates the initial state in 2008 and the second the state in 2018. The employment rate in most Member States ranged from 65–80%, while the share of tertiary graduates was between 20–45% and the rate of early school leavers was typically below 20% in 2008. The indicators show a significant improvement over the period under review.

Figure 7 Employment and qualification indicators by EU Member State (2008, 2018)



Source: Own editing by Eurostat data

In 2018, the standard deviation of the Member States will take place at a much smaller interval in all three dimensions, which justify the reduction of territorial differences. The number of Member States affected by employment rates below 70% has halved, with only 6 countries having lower employment rates in 2018. The number of early school leavers was remarkably high, close to 30% or more in Portugal, Spain and Malta. By 2018, the value of the indicator had been reduced to below 20% in all Member States. The preferred value of the indicator at EU level is 10%, which has already been reached by most Member States. The

proportion of people with tertiary education has shifted upwards, ranging from 30–60% in 2018; proportions below 30% were reported only from Italy and Romania.

In the examined years, the correlations between the Europe 2020 indicators were revealed by Pearson's correlation. The coefficients of the indicators of R&D, employment, skills and impoverishment are presented in Table 2. Only the coefficients that were significant at the $p=0.01$ and $p=0.05$ levels are included in the correlation matrix.

It was found that there is no explicitly strong correlation between the indicators for either the 2008 or 2018 data. In both years, a moderately strong positive correlation ($r_{2008}=0.545$, $r_{2018}=0.395$) was found between R&D expenditure as a % of GDP and employment. While there is a negative, medium-strong correlation ($r_{2008}=-0.590$, $r_{2018}=-0.597$) between poverty and R&D expenditure.

Table 2 Correlation of Europe 2020 indicators (2008, 2018)

Indicators		Intramural R&D expenditure	Employment rate	Tertiary educational attainment	Early leavers from education and training	People at risk of poverty or social exclusion
2008	Intramural R&D expenditure	1	,545**	,425*		-,590**
	Employment rate	,545**	1	,531**		
	Tertiary educational attainment	,425*	,531**	1		
	Early leavers from education and training				1	
	People at risk of poverty or social exclusion	-,590**				1
2018	Intramural R&D expenditure	1	,395*			-,597**
	Employment rate	,395*	1			-,512**
	Tertiary educational attainment			1	-,518**	
	Early leavers from education and training			-,518**	1	
	People at risk of poverty or social exclusion	-,597**	-,512**			1

** Correlation is significant at the 0.01 level.

* Correlation is significant at the 0.05 level.

Source: Own construction and calculation by Eurostat data

Several differences were detected between the coefficients of the two examined time points. While the data for 2008 show a moderate correlation between the share of tertiary graduates and R&D expenditure and employment, in 2018 these correlations no longer exist. For the 2018 data, there is a significant correlation between the proportion of people with tertiary education and early school leavers and between employment and poverty.

In the next phase of the research, composite indexes were created from the selected indicators of the strategy, and then by expanding them. Factor analysis was performed to compress the content of difficult-to-compare data with minimal information loss.

Principal component analysis of Europe 2020 indicators

Following the Z standardization of the 2018 data of the indicators, a factor analysis was performed, in which the factors were subtracted using the principal component method (Tab. 3). The number of significant principal components was determined based on the eigenvalue of the components. The minimum value of the Kaiser criterion is 1.5, i.e., I considered only those components with a variance greater than 1.5 to be significant.

Table 3 Principal component analysis of Europe 2020 indicators (2018)

Parameters	Composite index
Extraction Method:	Principal Component
Kaiser criterion	1,5
Bartlett's test (significance)	0,000
KMO	0,606
Sums of Squared Loadings	Component 1
Total	2,312
% of Variance	46,234
Cumulative %	46,234
Components (2018)	
Zscore: Intramural R&D expenditure	0,720
Zscore: Employment rate	0,717
Zscore: Tertiary educational attainment	0,593
Zscore: Early leavers from education and training	-0,521
Zscore: People at risk of poverty or social exclusion	-0,810

Source: Own construction and calculation by Eurostat data

Based on the Bartlett chi-square test, I discarded at all standard significance levels that the original variables were independent. That is, the test confirmed that the variables are suitable for factor analysis; the significance level is less than 0.05. However, the value of the Kaiser-Meyer-Olkin (KMO) criterion is only 0.606, which means only mediocre. The variance explained from the variance of the variables is only 46.234%, i.e. our obtained factor retains only this much information. The factor with the lowest weight has the proportion of *early leavers from education and training*. The *people at risk of poverty or social exclusion* indicator retain 65.7% of the original information, while the rate of *early leavers from education* is only 27.1%.

In order to improve the KMO values and to increase the information content of the principal components and the explained variance, other variables were included in the principal component analysis in the next phase of the research.

Principal component analysis supplemented with background variables

As potential background variables, 20 national-level variables were identified, which were subjected to a preliminary correlation analysis (Tab. 4). The largest numbers of selected potential background variables show a significant correlation with the indicator *people at risk of poverty or social exclusion*. This indicator is also outstanding in terms of the strength of the relationships, the value of the correlation coefficients for 4 potential background variables (*S80/S20*, *impact of social transfers on poverty reduction*, *at risk of poverty rate*, *severely materially deprived people*) is greater than 0.7. The indicator of *people at risk of poverty or social exclusion* explains the *income quintile share ratio* measure in 80.10%. *Intramural R&D* is also related to a large number of background variables, the indicator is in close direct proportion to LMP expenditures and *labour costs*. *Lifelong learning* shows the strongest correlation with this indicator.

The coefficients indicated an almost equal number of significant correlations between the employment rate and *tertiary educational attainment* indicators. The employment rate has several strong correlations. Of course, there is a strong, direct relationship between employment and the activity rate, and the indicator is inversely proportional to unemployment and long-term unemployment. The coefficients related to the *tertiary educational attainment* indicator confirmed the existence of only a moderate relationship in all cases.

The *early leavers from education and training* indicator showed the least number of correlations with background variables. The indicator has a close, significant correlation only with indicator of the *less than primary, primary and lower secondary education* and the *less than primary, primary and lower secondary education % of unemployed*. For the other variables, the results obtained did not prove to be significant.

Table 4 Correlation between EU2020 indicators and other selected indexes (2018)

Indicators (2018)	Intramural R&D expenditure	Employment rate	Tertiary educational attainment	Early leavers from education and training	People at risk of poverty or social exclusion
GDP/capita in PPS (% of EU27)			,520**		
Income quintile share ratio (S80/S20)	-,475**				,895**
Median income (€)	,652**		,493**		-,552**
Labour market policy (LMP), % of GDP	,733**				-,370*
Labour cost index (LCI)	,751**		,417*		-,470*
Life expectancy (years)	,437*				-,379*
Impact of social transfers on poverty reduction %	,574**	,423*	,411*		-,759**
Activity rate (20-64 aged)	,401*	,871**	,435**		
Employed ICT specialists, % of total employed	,598**	,483**	,408*		-,537**
Job vacancy rate	,596**	,533**			-,643**
Lifelong learning	,667**	,502**	,465*		-,504**
Less than primary education (levels 0-2, 20-64 aged)				,741**	
Unemployment rate (20-64 aged)		-,740**			,482**
Long-term unemployment % of unemployed		-,678**	-,440*		,472**
Less than primary education % of unemployed			-,420*	,763**	
At risk of poverty rate	-,416*	-,400*			,875**
Severely materially deprived people	-,577**	-,502**			,829**
People living in households, very low work intensity		-,540**			,394*
Road/rail networks (km/1000km2)					
Individuals regularly using the internet	,646**	,586**	,542**		-,659**

** Correlation is significant at the 0.01 level.

* Correlation is significant at the 0.05 level.

Source: Own construction and calculation by Eurostat data

In addition to the Europe 2020 indicators, the relationships between the 20 other indices selected for the factor analysis were also evaluated. The TOP7 matrix contains the coefficients of the 7 indicators with the strongest correlations (Tab. 5). Only the coefficients within the 99% and 95% confidence intervals were displayed in the matrix. The TOP7 indicators, with one exception, show a strong or at least moderate correlation with each other. Of the indicators examined, I found no significant correlation between *lifelong learning* and the *impact of social transfers on poverty reduction* alone. I calculated the strongest and most moderate correlations for LCI. The coefficients of the *labour cost* index indicated a strong ($r > 0.7$) positive correlation with 6 indicators, and a medium strong correlation for another 3. The proportion of *individuals regularly use the internet* is also closely related to 6 other indicators and moderately strong to 2 variables. *Median income*, *severely materially deprived people*, *employed ICT specialists*, *lifelong learning* and *impact of social transfers on poverty reduction* also showed a similarly high number of strong and medium correlations.

The value of the correlation coefficient is the largest between the *median income* and the *labour cost* ($r = 0.914$). *Labour cost* determines the variance of median earnings by 83.54%. *Median income* showed a strong significant correlation with almost all indicators.

There was a significant negative correlation ($r = -0.832$) between *severely materially deprived people* and *individuals regularly using the internet*. That is, the higher the proportion of *individuals regularly use the internet*, the lower the proportion of *severely materially deprived people*. Internet use is also strongly correlated with other indicators, for example *median earnings*, *labour costs*, the proportion of *employed ICT specialists* and the *lifelong learning*. The proportion of *employed ICT specialists* explains *median income* in 56.40% and internet use in 55.95%.

Percentage of the *impact of social transfers on poverty reduction* in the TOP7 matrix is most closely related to the proportion of *employed ICT specialists* and moderately inversely related with the proportion of *people living in severely materially deprived*.

Table 5 TOP7 – Variables with many strong correlations (2018)

Indicators (2018)	Median income	LCI	Lifelong learning	Impact of social transfers on poverty reduction	Employed ICT specialists	Severely materially deprived people	Individuals regularly using the internet
Median income (€)	1	,914**	,733**	,582**	,751**	-,754**	,799**
LCI	,914**	1	,711**	,564**	,686**	-,619**	,715**
Lifelong learning	,733**	,711**	1		,480**	-,565**	,408*
Impact of social transfers on poverty	,582**	,564**		1	,632**	-,585**	,582**
Employed ICT specialists	,751**	,686**	,480**	,632**	1	-,700**	,748**
Severely materially deprived people	-,754**	-,619**	-,565**	-,585**	-,700**	1	-,832**
Individuals regular using the internet	,799**	,715**	,408*	,582**	,748**	-,832**	1

** Correlation is significant at the 0.01 level.

* Correlation is significant at the 0.05 level.

Source: Own construction and calculation based on Eurostat data

I expanded the principal component analysis with the 5 indicators of the Europe 2020 strategy with another 5 indicators, which were selected from the TOP7 variables, and the number of factors to be subtracted was determined in several ways. In the empirical research conducted, the criteria related to the models are as follows: the value of the KMO index exceeds 0.7, and that the variance explained by the factors should be close to 70%. The parameters of the final model that meets the criteria are shown in Tab. 6. The null hypothesis of the Bartlett test can be rejected for the model because the significance level is less than 0.05, i.e. the variables are suitable for factor analysis.

During the factor analysis, based on these parameters, 3 principal components were generated; the variance explained by the first component was 40%; the second factor compresses 24% and the third compacts 15.94% of the information. The variance explained by the principal components has improved significantly, collectively compressing more than 80% of all information.

In compiling the first principal component, Labour costs, Median income, Employed ICT specialists, R&D expenditure, and regular Internet use are significant variables based on factor weights. The second factor includes the employment rate and indicators related to poverty, while the principal component three summarizes the indicators related to qualifications. The principal components based on the information they compress are: 1. Innovation environment, 2. Employment, and 3. Education.

Table 6 Principal component analysis of EU2020 indicators and other selected indicators

Parameters	Final model		
Extraction Method:	Principal Component		
Kaiser criterion	1		
Bartlett's test (significance)	0,000		
KMO	0,760		
Sums of Squared Loadings	Component 2/1	Component 2/2	Component 2/3
% of Variance	40,211	24,068	15,943
Cumulative %	80,222		
Components (2018)	Rotated Component Matrix_5		
	1	2	3
Zscore: LCI	0,967	0,039	0,132
Zscore: Median income (€)	0,913	0,218	0,186
Zscore: Employed ICT specialists	0,733	0,421	0,093
Zscore: Intramural R&D expenditure	0,724	0,366	-0,020
Zscore: Individuals regularly using the internet	0,700	0,544	0,284
Zscore: Employment rate	0,092	0,875	0,113
Zscore: People at risk of poverty or social exclusion	-0,422	-0,743	-0,107
Zscore: Severely materially deprived people	-0,629	-0,638	-0,138
Zscore: Early leavers from education and training	0,038	-0,100	-0,866
Zscore: Tertiary educational attainment	0,341	0,124	0,812

Source: Own construction and calculation based on Eurostat data

As a final step in the research I grouped the Member States along the three principal components of the final model with cluster analysis, so I explored the position of the Member States within the dimensions by creating homogeneous groups.

Cluster analysis of principal components

In order to facilitate the interpretability of the obtained results, I grouped the Member States into groups in three dimensions of the final model using K-means cluster analysis. Previous

hierarchical cluster analysis and structure exploration analysis identified 4 clusters. Detailed test results and descriptive statistics for clusters are shown in Table 7.

Based on the F-test and p significance level in the ANOVA table, the distinctive power of the principal components was found to be significant over a 99% confidence interval. The Kruskal-Wallis Test performed as an alternative to the test also gave the same result. The ETA values calculated for the nominal index of the cluster confirmed the close positive correlation between the individual principal components and the formed clusters, the quotient of the sum of squares of the total difference between groups value was higher than 0.7 in all dimensions. Based on the within-group variance, the formed clusters can be considered homogeneous in the proportion of 83.33%. The standard deviation in 2 cases slightly exceeded the total standard deviation.

Table 7 Summary statistical table of cluster analysis

Analysis of variance		Innovation environment	Employment	Education
F-test		18,671	7,821	22,791
Significance		0,000	0,001	0,000
Kruskal-Wallis Test		18,405	14,362	21,299
Asymptotic Significance		0,000	0,002	0,000
ETA		0,832	0,701	0,856
Descriptive statistics of clusters		Innovation environment	Employment	Education
Mean (Final Cluster Centres)	1	1,032	0,238	0,106
	2	-0,713	-0,400	1,081
	3	-0,323	-1,262	-1,438
	4	-0,726	0,938	-0,678
Std. Deviation	1	0,400	0,633	0,566
	2	0,561	1,094	0,483
	3	1,147	0,384	0,441
	4	0,422	0,567	0,645

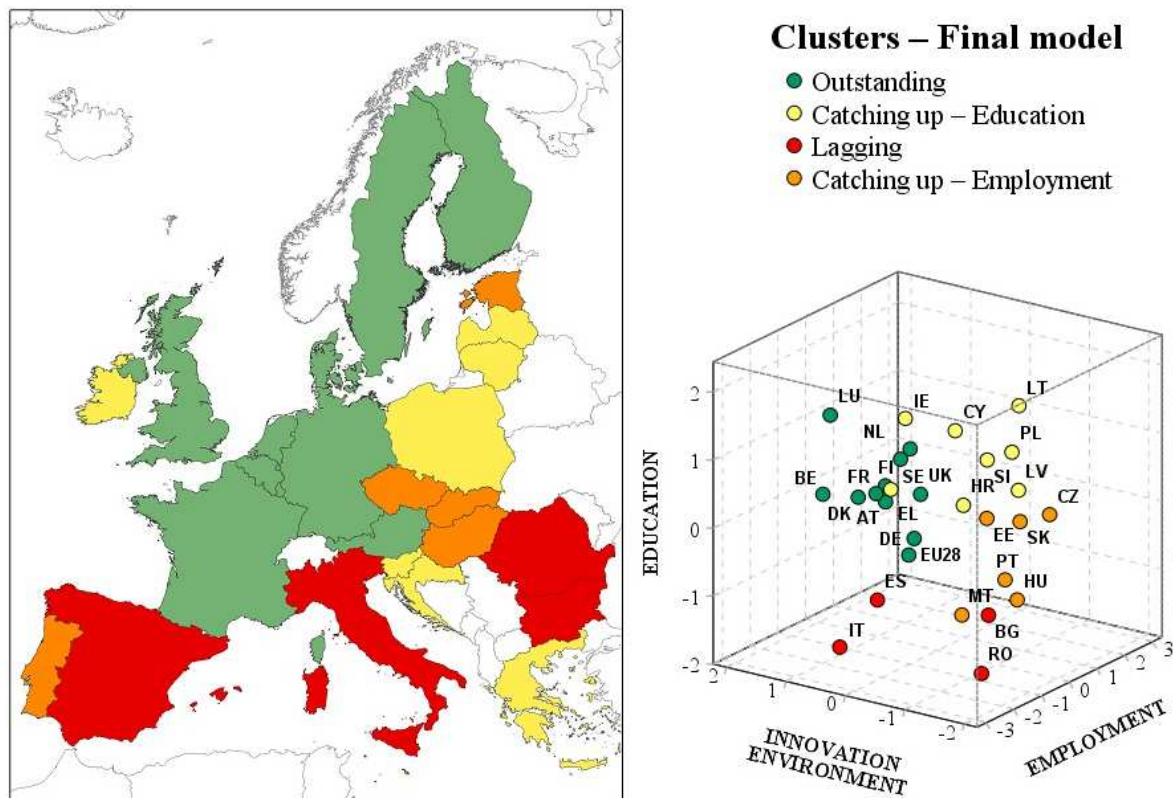
Source: Own editing by own calculation

Clusters were interpreted based on a comparison of dimension means. The first cluster included 10 Member States and the EU28, the second 8 country, the third 4 country and the fourth 6 countries (Fig. 8). I named the clusters based on their relationship to each dimension. The Member States in the first cluster also perform above average in terms of the innovation environment, employment and education, which is why I named the cluster Outstanding. The second cluster consists of countries with above-average levels of education (Catching up – Education) but with below-average levels in the other two dimensions. The third group includes countries that remained below average for all three composite indicators, so the group was named Lagging. The fourth cluster includes countries where employment is above average (Catching up – Employment) but the innovation environment and qualifications are lower. It is

important to note that progress in one dimension is not necessarily to the detriment of another dimension of sustainable competitiveness.

I plotted the clustering on a 3D point cloud diagram. The Outstanding group includes more developed Member States such as France, Belgium, Sweden, Finland, Austria, Germany, Luxembourg, etc. The Lagging cluster consists of 4 Member States: Spain, Italy, Romania, and Bulgaria. The Visegrad Group (V4) did not fit into a common cluster, but all four Member States are characterized by catching up, as they performed above average in some dimensions. In addition to Hungary, Czech Republic and Slovakia are also in the group with better-than-average employment rates, while Poland is above average in the dimension of education. The Catching up and Lagging groups are typically made up of member states that joined after 2004.

Figure 8 Final model (2018) – Innovation environment, employment, education



Source: Own editing by own calculation

Analysis of variance revealed the relationship between the CLU variable of the created clusters and the 2018 LMP expenditures of each member state. The research confirmed a statistically significant ($p=0.01$), strong correlation ($\text{ETA}=0.749$) between LMP expenditures and clusters. Further correlations were also found with the time of accession to the European Union (before/after 2004). While there is a strong correlation ($\text{ETA}=0.782$) between the time

of accession and LMP expenditures, a moderate ($\text{ETA} = 0.590$) correlation was shown in the context of clusters.

DISCUSSION

The aim of the study was to explore the state of the structure of labour market interventions in the context of unemployment after the 2008 crisis before the crisis (2008) and in a priority year for crisis management (2018). In her study, Kálmán (2015) explains that social spending increased after the crisis and that individual countries developed a differentiated system of tools to deal with the general crisis in the labour market.

The research confirmed that expenditure on interventions increased significantly in 2009 and has been gradually declining at EU level after 2010. In general, spending has increased significantly in those Member States where unemployment has risen sharply (Lithuania, Estonia, Latvia, Spain and Ireland). With the exception of Greece, which has been hit harder by the crisis, where intervention spending has not increased, despite a sharp rise in unemployment. By 2018, the shares of measure types changed in several Member States compared to the pre-crisis composition. In those countries where the structure of expenditures was different from the average in 2008, there was equalization towards average values. However, in several Member States a different type of measure was prioritized than in the previous period. In Hungary, the expansion of public employment opportunities was given priority, so *direct job creation* increased from the previous 12% to 48%, while *out-of-work income maintenance and support* expenditures were halved.

The objectives of the Europe 2020 strategy were set on the basis of the factor influencing economic growth. According to Jankó (2010), the most important element of the long-term solution of economic problems is to increase the education of the population. In their study, Mankiw, Romer & Weil (1992) highlight that several research have already demonstrated a positive correlation between GDP per capita and human capital. The relationship between economic development and human capital was also confirmed by the analysis, although only at a moderate strength. In connection with the education of the population were included the indicators *tertiary educational attainment* and *early leavers from education and training* in the research. The proportion of people with tertiary education showed a moderate, positive correlation with employment and R&D expenditure. The correlation analysis of other background variables also confirmed a moderate, positive correlation ($r_{2008} = 0.481$, $r_{2018} = 0.520$) between the ratio of tertiary graduates and GDP/capita in PPS (% of EU27) in both years.

According to the Employment Report of the Commission and the Council (European Commission, 2018/a), the reform efforts introduced in the EU have contributed to job-creating economic growth.

The analysis carried out supports the above statement, although it should be noted that the achievement of the targets set in the Europe 2020 strategy is very diverse at the national level. The key indicators improved significantly between 2008 and 2018. A striking positive change can be observed in the field of employment, the employment rate for the 20–64 age group shifted predominantly to 70–85% by 2018. The proportion of people with tertiary education has also risen and is concentrated around 30–60%, creating a more coherent picture. Overall, the rate of early school leavers also improved compared to 2008 and was around 10% in most Member States.

CONCLUSION

The variance explained by the single composite index resulting from the 2018 data of the Europe 2020 indicators is only 46.234%. Thus, the composite index cannot be considered suitable for a significant compression of the information content of the indicators. The 3 main components of the final model developed with the inclusion of other background variables compress a significant part of the information, 80.222%. In the clusters formed along the dimensions of the final model (Education, Employment, Innovation environment); the member states were well separated based on their differences in development (Outstanding, Catching up and Lagging). Overall, the countries of the Outstanding group are characterized by development, the countries of the Lagging cluster are characterized by lagging behind, and the Catching up groups are characterized by catching up. The three dimensions (Education, Employment, Innovation environment) are interacting. However, synergy between dimensions does not mean that progress in one dimension can only be at the expense of the other two dimensions.

In the Catching up – Employment group (e.g. Czech Republic, Slovakia and Hungary); economic growth was mainly extensive, based on a significant increase in employment. In order to catch up more quickly with more developed countries the quality of resources (human capital) also needs to be improved. The Catching up – Education group (e.g. Poland, Latvia, Lithuania) moved towards the development of human capital, where the qualification indicators were better than average in 2018.

A correlation can be found in several Member States between the 2018 LMP expenditures and the established clusters. LMP expenditures in 2018 are typically the lowest in East-Central

Europe member states (Romania, Czech Republic, Poland, Slovakia, and Bulgaria). In general, it was below 1% in the Member States that joined the European Union after 2004.

In the context of clusters, it can be declared that those Member States where LMP expenditure as a proportion of GDP below 1% (countries joining after 2004) typically belong to the Catching up and Lagging clusters. Lagging group's 2018 LMP expenditures show significant discrepancies. While Romania and Bulgaria are among the 10 least spending Member States, Italy and Spain have labour market interventions in proportion to GDP are more significant. The members of the Outstanding group which joined before 2004 typically have higher LMP expenditures due to their more developed economic situation. The exceptions are Italy (1958), Greece (1981) and Spain (1986), which joined the Union very early, yet were among the lagging economies in 2018. Despite the fact that LMP expenditures in Spain is 2.15%, which much higher than average and also in Italy it was well above 1%. In Greece, LMP spending is low despite very high unemployment. The unemployment rate for the 20–64 age group fell to just fewer than 20% by 2018, with next to Spain and Italy being the worst off at EU level in this respect.

Strengthening territorial cohesion is an important element of the Europe 2020 strategy; in this context, they seek to extend the benefits of economic growth to peripheral areas. In the period under review, a small degree of equalization between the indicators can be observed, which, with the reduction of territorial differences, has the effect of strengthening cohesion.

Overall, the EU has the capacity to act in times of crisis and to adapt its economies and societies to change. The experience gained during crisis management can be used to plan for dealing with recessions in the near future. Europeans today must once again prepare for transformation in order to cope with the effects of crises, overcome the EU's structural weaknesses and growing global challenges. It is indisputable that the coronavirus epidemic that will appear in 2019 will have an impact on the economy of the European Union. As a result of the economic recession caused by the epidemic, meeting more of the Europe 2020 targets will be an even greater challenge for Member States.

REFERENCES

- Allison, P. (1999). Multiple Imputation for Missing Data: A Cautionary Tale. *Sociological Methods & Research*, 28(3), 301–309. DOI: 10.1177/0049124100028003003.
- Allison, P. (2003). 'Missing data techniques for structural equation models'. *Journal of Abnormal Psychology*, 112(4), 545–557. DOI: 10.1037/0021-843X.112.4.545112
- Arsham, H. & Lovric, M. (2011). Bartlett's Test. In: Lovric M. (eds.), *International Encyclopedia of Statistics Science*, Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-04898-2_132

- Bansal, Arpit & Sharma, Mayur & Goel, Shalini. (2017). Improved K-mean Clustering Algorithm for Prediction Analysis using Classification Technique in Data Mining. *International Journal of Computer Applications*, 157(6), 35-40. DOI: 10.5120/ijca2017912719.
- Bardhoshi, G., Um, B. & Erford, B. T. (2021). Conducting a Cluster Analysis in Counseling Research: Four Easy Steps. *Counseling Outcome Research and Evaluation*, 12(1), 54-62. DOI: 10.1080/21501378.2020.1768522
- Birkner, Z., Peter, E. & Fehérvölgyi, B. (2012). A possible approach of the concept of 'GLOCAL' through innovation. *Deturope*, 4(3), 4-21.
- Bottlik, Zs. (2008). A regionális és etnikai különbségek összefüggései Bulgáriában az ezredfordulón. *Területi Statisztika*, 48(3), 334-343. DOI: https://doi.org/10.1007/978-3-642-04898-2_132
- Breslow, N. (1970). A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship. *Biometrika*, 57(3), 579-594. DOI: <https://doi.org/10.1093/biomet/57.3.579>
- Capello, R. (2007). A forecasting territorial model of regional growth: the MASST model. *The Annals of Regional Science*, 41(4), 753-787. Retrieved February 20, 2021, from <http://doi.org/c5wxvk> DOI: 10.1007/s00168-007-0146-2.
- Cattell, R. B. & Vogelman, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research*, 12(3), 289-325. DOI: https://doi.org/10.1207/s15327906mbr1203_2
- Chung, S. & Cai, L. (2018). Alternative Multiple Imputation Inference for Categorical Structural Equation Modeling. *Multivariate Behavioral Research*, 54(3), 323-337. DOI: 10.1080/00273171.2018.1523000
- Diez Roux, A. V., Borrell, L. N., Haan, M., Jackson, S. A. & Schultz, R. (2004). Neighbourhood environments and mortality in an elderly cohort: Results from the cardiovascular health study. *Journal of Epidemiology & Community Health*, 58, 917-923. DOI: <https://doi.org/10.1136/jech.2003.019596>
- Egri, Z., Töröcsik, V. & Tánczos, T. (2009). Regional HDI as a territorial and social differentiation index in Central Europe. In Káposzta J. (eds.), *New elements and research in spatial economy*, 167-180.
- Egri, Z. (2017). Regional health inequalities in the European macroregion from the East Central European perspective. *Regional Statistics*, 7(1), 197-224.
- European Commission (2010). Labour market and wage developments in 2009. In: *European Economy*, Brussels. DOI:10.2765/36582
- European Commission (2013). Sustainable development in the European Union — 2013 monitoring report of the EU sustainable development strategy. Eurostat Statistical Books. Luxembourg: Publications Office of the European Union. ISBN 978-92-79-31155-0 DOI: 10.2785/11549
- European Commission (2018/a). Draft joint employment report from the Commission and the Council. Retrieved September 27, 2021, from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0761&from=EN>
- European Commission (2018/b). Labour market policy statistics – Methodology 2018. Luxembourg: Publications Office of the European Union. ISBN 978-92-79-91299-3 DOI: 10.2767/291515
- European Commission (2020). Proposal for a joint employment report from the Commission and the Council. Retrieved September 27, 2021, from https://eur-lex.europa.eu/resource.html?uri=cellar:d5c1e490-29b7-11eb-9d7e-01aa75ed71a1.0001.02/DOC_1&format=PDF

- Ékes, I. (2011). Foglalkoztatás és munkahelybővítés - nemzetközi kitekintéssel. *Statisztikai Szemle*, 86(10-11), 970-994.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- Ginkel, J. R. V., Kroonenberg, P. M. & Kiers H. A. L. (2014). Missing data in principal component analysis of questionnaire data: a comparison of methods, *Journal of Statistical Computation and Simulation*, 84(11), 2298-2315, DOI: 10.1080/00949655.2013.788654
- Hajdú, D. (2020). A munkanélküliség és felnőttképzés területi eloszlása az Észak-magyarországi régióban. *Észak-magyarországi Stratégiai Füzetek*, 17(2), 62-69. DOI: <https://doi.org/10.32976/stratfuz.2020.14>
- Hakstian, A. R., & Muller, V. J. (1973). Some notes on the number of factors problem. *Multivariate Behavioral Research*, 8(4), 461-475. DOI: https://doi.org/10.1207/s15327906mbr0804_4
- Harper, J. F. (1984). Peritz' F test: Basic program of a robust multiple comparison test for statistical analysis of all differences among group means. *Computers in Biology and Medicine*, 14(4), 437-445. DOI: [https://doi.org/10.1016/0010-4825\(84\)90044-1](https://doi.org/10.1016/0010-4825(84)90044-1).
- Jankó, B. (2010). A tudás szerepe a gazdasági növekedésben — az alapmodellek bemutatása. *Statisztikai Szemle*, 88(1), 47-72.
- Jolliffe, I. T. (2005). Principal component analysis. 2nd ed. John Wiley & Sons. New York. DOI: 10.1002/0470013192.bsa501
- Jolliffe, I. T. & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society*, 374(2065), 1-16. DOI: <http://dx.doi.org/10.1098/rsta.2015.0202>
- Kaiser, H. F. (1960). The application of electronic computers to factoranalysis. *Educational and Psychological Measurement*, 20, 141–151. DOI: <http://dx.doi.org/10.1177/001316446002000116>
- Kaiser, H. F. (1981). A revised measure of sampling adequacy for factor-analytic data matrices. *Educational and Psychological Measurement*, 41(2), 379-381.
- Kálmán, J. (2015). A közfoglalkoztatási programok háttére és nemzetközi tapasztalatai. In: *Munkaerőpiaci tükrök, 2014* (pp.42-58). MTA Közgazdaság- és Regionális Tudományi Kutatóközpont Közgazdaság-tudományi Intézet, Budapest.
- Kovács, E. (2014). Többváltozós adatelemzés. Budapesti Corvinus Egyetem, Operációkutatás és Aktuáriustudományok tanszék. Typotex, Budapest. Retrieved February 20, 2021, from http://etananyag.ttk.elte.hu/FiLeS/downloads/14_KOVACS_E_Tobbvalt_adatelemzes.pdf.
- Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1990). Assessing sampling variation relative to number-of-factors criteria. *Educational and Psychological Measurement*, 50(1), 33–48. DOI: <https://doi.org/10.1177/0013164490501004>
- Lengyel, I. (2010). Regionális gazdaságfejlesztés. Versenyképesség, klaszterek és alulról szerveződő stratégiák. Budapest, Akadémiai Kiadó. ISBN 978-963-05-8837-9
- Liu, Y., Singleton, A. & Arribas-Bel, D. (2019). A Principal Component Analysis (PCA)-based framework for automated variable selection in geodemographic classification. *Geo-spatial Information Science*, 22(4), 251-264. DOI: 10.1080/10095020.2019.1621549
- Mankiw, N. G., Romer, D. & Weil, D. N. (1992). A contribution to the empirics of economy growth. *The Quarterly Journal of Economics*. Retrieved February 15, 2021, from https://eml.berkeley.edu/~dromer/papers/MRW_QJE1992.pdf 407-437. p.
- MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Symposium on Math, Statistics, and Probability (pp. 281–297). Berkeley, CA: University of California Press.

- Nemes Nagy, J. ed. (2005). Regionális elemzési módszerek. In: ELTE Regionális Földrajzi Tanszék, Budapest. (*Regionális Tudományi Tanulmányok, 11.*)
- Oravecz, B. (2008). Hiányzó adatok és kezelésük a statisztikai elemzésekben. *Statisztikai szemle*, 86(4), 365-384.
- Ostertagová, E., Ostertag, O., & Kováč, J. (2014). Methodology and Application of the Kruskal-Wallis Test. *Applied Mechanics and Materials*, 611, 115–120. DOI: <https://doi.org/10.4028/www.scientific.net/amm.611.115>
- Rao, M. R. (1971). Cluster Analysis and Mathematical Programming. *Journal of the American Statistics Association*, 66(335), 622-626. DOI: 10.1080 / 01621459.1971.10482319
- Ramsey, P. H. (1978). Power Differences between Pairwise Multiple Comparisons. *Journal of the American Statistical Association*, 73(363), 479-485. DOI: 10.1080/01621459.1978.10480038
- Ramsey P. H. & Ramsey P. P. (2013). Pairwise testing of group mean vectors in MANOVA with small samples, *Journal of Statistical Computation and Simulation*, 83(6), 1037-1057. DOI: 10.1080/00949655.2011.648936
- Rechnitzer, J., Barsi, B., Szabó, P. & Németh N. (2003). A gazdasági térszerkezet vizsgálatát elősegítő új dimenziók illetve az ezzel kapcsolatos módszerek kutatása. In: VÁTI Magyar Regionális Fejlesztési és Urbanisztikai Kht. Területfejlesztési igazgatóság, Elemző és értékelő iroda. Budapest.
- Robinette, J. W., Charles, S. T. & Gruenewald, T. (2017). Neighborhood socioeconomic status and health: A longitudinal analysis. *Journal of Community Health*, 42(5), 865–871. DOI: <https://doi.org/10.1007/s10900-017-0327-6>
- Romer, P. M. (1994). The origins of endogenous growth. *The Journal of Economic Perspectives*, 8(1), 3–22. Retrieved January 12, 2021, from <http://doi.org/bbs8mn> DOI: 10.1257/jep.8.1.3.
- Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. John Wiley & Sons. New York.
- Shlens, J. (2014). A Tutorial on Principal Component Analysis. Retrived September 27, 2021, from arXiv: 1404.1100
- Simon, J. (2006). A klaszterelemzés alkalmazási lehetőségei a marketingkutatásban. *Statisztikai Szemle*, 84(7), 627-651.
- Székelyi, M., Barna, I. (2005). Túlélőkészlet az SPSS-hez, Typotex Kiadó. Budapest.
- Vargha, A., & Delaney, H. D. (1998). The Kruskal-Wallis Test and Stochastic Homogeneity. *Journal of Educational and Behavioral Statistics*, 23(2), 170–192. DOI: <https://doi.org/10.3102/10769986023002170>
- Zou, H., Hastie, T. & Tibshirani, R. (2006). Sparse Principal Component Analysis. *Journal of Computational and Graphical Statistics*, 15(2), 265-286. DOI: <https://doi.org/10.1198/106186006X113430>