Croatian Operational Research Review
CRORR **14**(2023), 53-64

Multivariate analysis of post-transition OECD countries in the context of inequality measures

Tomislav Korotaj[1,*], **Nataša Kurnoga**[1] and **Nika Šimurina**[1]

[1] *Faculty of Economics and Business, University of Zagreb, Trg J. F. Kennedyja 6, 10000 Zagreb, Croatia*
E-mail: ⟨{*tkorotaj, nkurnoga, nsimurina*}*@efzg.hr*⟩

Abstract. The aim of this paper is to classify the post-transition OECD countries according to the Gini coefficient for income inequality, the S80/S20 ratio, income share of the bottom 40% of the population, educational attainment – tertiary education, and labor force participation rate using factor and cluster analyses. Factor analysis resulted in two extracted factors, and factor scores were calculated. Hierarchical and non-hierarchical cluster analysis was performed on factor scores to classify eight post-transition OECD countries and three candidate countries. The research question of this paper is to investigate whether there are similarities/differences between existing members and candidates for membership in the OECD, among the selected post-transition countries of Europe, in the context of income inequality. Based on the dendrogram obtained by the hierarchical Ward's method, a three-cluster solution was selected. The non-hierarchical k-means method for the three-cluster solution clustered Croatia with Bulgaria and Romania. These three countries are OECD candidate countries. Our findings confirm that the three candidate countries remain behind because of historical reasons and the non-implementation of structural reforms.

Keywords: cluster analysis, factor analysis, inequality measures, OECD countries

Received: November 29, 2022; accepted: April 20, 2023; available online: July 10, 2023

DOI: 10.17535/crorr.2023.0005

1. Introduction

One of the most noticeable implications of the transition from a planned to a free-market economy is the rise of income inequality. In the scientific and professional literature, there is no consensus about the completion of the transition process in the countries of Central and Eastern Europe (CEE). Some authors [13, 22, 30] suggest that the process is still active, while others [8, 15, 23] close the door and turn to the observation of post-transition effects. However, public concern over the growing income gap between the rich and other parts of the distribution makes this issue an important topic. The end of the transition is usually associated with the concepts of post-communist, post-socialist and post-transition state organization, which are interchangeable terms. In this sense and in a simplified view, if the concept of transition is reduced only to the transition from communist to the post-communist regime in a country, it is possible to single out 11 post-transition countries in Europe that are currently of special interest to researchers of income inequality.

The Organization for Economic Co-operation and Development (OECD) was founded in 1960 as a global successor to the previously formed Organization for European Economic Co-operation (OEEC). In 2022, the OECD has a total of 38 member countries. Twenty of the thirty-eight members are the founders of the OECD. Of the remaining eighteen members, eight

*Corresponding author.

are from the area of the former communist regime. The Czech Republic, Hungary and Poland joined the OECD in the mid-1990s. Slovakia became a member in 2000, while Slovenia and the Baltic countries (Estonia, Latvia and Lithuania) were admitted to membership between 2010 and 2018. As of January 2022, three more European countries have acquired the status of candidates for full membership in the OECD: Bulgaria, Croatia and Romania.

The focus of this paper is on income inequality in the aforementioned 11 post-transition countries; members and candidates for membership in the OECD. The idea of the article is to conduct a cluster analysis of the above-mentioned countries based on selected variables that reflect education, labor market and income inequality in 2020. Bulgaria, Croatia and Romania are among the last countries to join the EU. From 2022, they also jointly acquired the status of candidate for membership in the OECD. The question arises whether there are similarities between these countries and the current post-transition members of the OECD. In other words, where do these countries stand in relation to other post-transition countries in terms of measures of inequality? In the next few paragraphs, an overview of available research on this topic is given.

If inequality is not perceived in society, it may not have political repercussions; but if perceived, it can have repercussions, possibly severe. Socioeconomic development can contribute to building the image of a more equal society. For that reason, it is crucial to comprehend how society is perceived by its members.

Members of different societies have different views on inequality. Some individuals see a privileged class at the top of their society with most impecunious people at the bottom, while others see an affluent society with most people in-between. Evans and Kelley [6] investigated socioeconomic development and income inequality in 43 countries of the world based on data from the ISSP (International Social Survey Programme) questionnaire administered to more than 110,000 respondents. Applying the multilevel generalized least square (GLS) method to data for the period between 1987 and 2009, they came to the conclusion that the perception of inequality increased drastically after the collapse of communism in Europe. The majority of people in post-communist societies (Bulgaria, Croatia, Hungary, Latvia, Slovakia) believe that they live in a predominantly elitist type of society. On the other hand, certain countries are identified with a pyramidal structure of society and a slightly lower perception of inequality (Czech Republic, Estonia, Slovenia). The image of income inequality is largely conditioned by socioeconomic status, so people of higher status usually perceive society as relatively more equal, while people of lower status perceive society as relatively more unequal. Looking at the countries of post-communist Europe as a whole, the result of the research implies that on average 42% of people in more advanced countries see their homeland as egalitarian, while this result for poor countries is equal to 28%.

Josifidis et al. [15] investigated the distributional effects of foreign direct investment (FDI) in 10 CEE countries (all except Croatia). Applying the seemingly unrelated regressions (SUR) method on a panel of data of a timespan of 25 years (1990-2014), they found strong proof of a non-linear nexus between FDI and income distribution. Skilled workers were recognized as short-term gainers of FDI, whereas long-term gainers were top-level managers and investors associated with foreign companies and foreign capital, respectively. The accumulation of FDI in the observed countries did not bring changes in the distribution of income that would benefit the bottom 50% of the population. The tertiary education expansion aggravated the income status of the bottom 50% of the population and bettered the income status of the class belonging to the middle 40% of the population.

Kuzmar and Piatek [16] investigated the institutional factors that determine the movement of inequality. The authors conducted an econometric panel analysis on 21 selected post-transition countries, on data for the period 1992-2015. A negative relationship between political freedom and measures of income inequality in the observed countries has been proven. The greater difference between market income and disposable income can be attributed to greater political freedom in the post-transitional CEE countries (the 11 countries that are in the focus of this paper). On the other hand, the opposite was found to be true in countries with limited political freedom, where the aforementioned differences were smaller (member countries of the former Soviet Union).

Szczepaniak and Geise [25] investigated the relationship between income inequality and 5 aspects of well-being in 8 CEE countries. The panel autoregressive distributed lag (ARDL) model and in-depth analysis of data for the period 2004-2018 were used for that purpose. It was found that there is a significant difference between the results obtained on annual and quarterly data. By searching for short-term and long-term links between the observed variables, a long-term negative relationship was found between income inequality and the happiness index, health indicators and the natural environment. The unequal distribution of income in the long term is positively related only to indicators of education and household disposable income. In the short term, only 1 out of 5 aspects proved to be crucial in shaping the unequal distribution of income, and that is the educational aspect.

Jianu et al. [14] compare EU member states grouped into two clusters, using the median value of GDP per capita as a criterion; thus creating advanced and emerging economies. The GLS method was applied to panel data for the period 2010-2017. The emphasis is on recording the relationship between inequality and several other variables such as social transfers, tertiary education and unemployment. The dynamic panel model determined that the current level of inequality strongly depends on the historical values of inequality, especially in the cluster with emerging economies. Social transfers have a greater effect in the cluster with advanced economies, which can be attributed to better-constructed plans and larger budgets and financial capabilities of governments. Tertiary education has an opposite effect on clusters, positive on advanced economies and negative on emerging economies, while the link between the unequal distribution of income and unemployment is unequivocal and in a positive direction for all countries.

Some authors (Goedeme et al., [9]) propose the application of new indicators of inequality in research studies. More specifically, they are advocating a change in methodology in the calculation of poverty indicators through the development of extended composite measures such as the extended headcount ratio. Whether we should turn to the design of new inequality measures is just one of the open questions in the further search for a more credible picture of inequality in society.

The future picture of income inequality depends on a number of factors. According to some authors (Dolls et al., [2]), the demographic transformation of the population is one of the crucial factors and challenges for EU countries in the coming years. The aging of the European workforce has been identified as a major factor in increasing income inequality by 2030. By applying the decomposition method with the aim of separating demographic effects from wage effects, an increase in inequality was projected even in the Scandinavian countries. Moreover, further growth of inequality in Europe is inevitable. The main question in this context is what the dynamics in each particular region will be and whether it is possible to put an end to demographic pressures with different reforms than those that promote the raising of the retirement age.

Taking all of the above into account, it can be concluded that the area of income inequality is a very hot topic in the academic community, especially research on countries in Southeast Europe that were exposed to a complex process of government transition. The provided literature review leaves room for additional contributions to this research topic through a cluster analysis of countries associated with the OECD. It is precisely this context of observing countries that has been neglected so far, therefore, there is an interest in analyzing post-transition countries and finding the reasons that lead to the specific grouping of these economies.

2. Data and methodology

The data used for the analysis were taken from Eurostat, OECD and International Labour Organization (ILO) databases. As mentioned in the first section, 11 countries were selected, and the time point in the analysis is 2020. Variables used include three income inequality measures (the Gini coefficient of equivalized disposable income (*gini_coeff*), the S80/S20 income quintile ratio (*s80_s20*) and income share of the bottom 40% of the population (*bottom_40*), the measure of educational attainment (*edu_attain_C*) – tertiary education, and labor force participation rate (*lfpr*) as the measure of the labor market.

The Gini coefficient is the most commonly used measure of inequality that ranges from 0 to 100, with a lower value indicating greater equality in the distribution of income in a given society. According to Eurostat's and the OECD's classification, the calculation is based on the concept of equivalized disposable income [4], which in a nutshell represents the disposable income per household member, after taxes and related deductions, available for consumption and savings. The S80/S20 income quintile ratio represents the relationship between the two ultimate income shares, the richest 20% and the poorest 20% of members of the population [5]. It is considered an important indicator of inequality and is, along with the Gini coefficient, an indispensable part of EU statistical publications. The income share of the bottom 40% is an indicator that reflects the share of income earned by the 40% of members of the population with the lowest income. Within the framework of the United Nations' Sustainable Development Goals (Goal 10 – "Reduce inequality within and among countries", [29]), it is listed as the first indicator for monitoring progress at the bottom of income distribution. Alongside Gini and the S80/S20 ratio, it often forms the backbone for analyzing trends in inequality. The share of the population with tertiary education is based on the International Standard Classification of Education (ISCED 2011, [28]) and covers people with the highest level of education (Bachelor's, Master's or Doctoral level). The labor force participation rate [12] reflects the situation in the labor market through the ratio between the active population (labor force) and the working-age population.

According to the ISCED 2011 [28], there are three main categories of educational attainment to distinguish the population: primary, secondary and tertiary education. The educational attainment of an individual is defined as the highest ISCED level completed by an individual. For analytical purposes, educational attainment is typically measured with regard to the highest education programme successfully completed, which is normally certified by a recognized qualification. For the purpose of multivariate analysis of inequality, tertiary education was chosen due to the belief that it can best serve in explaining inequality. According to the OECD [21], the Baltic countries (Estonia, Latvia and Lithuania) stand out as economies with the highest shares of highly educated people (more than 40%), while Croatia, the Czech Republic and Romania have the lowest values (less than 26%) in 2020. Similar to the education indicator, the Baltic countries are also distinguished with the highest values (around 66%) for the labor force participation rate in 2020 [12]. Croatia is the only country with a rate lower than 55%, which implies the smallest proportion of the labor force in the working-age population. No major differences are noticeable between other post-transition countries.

The main idea of this paper is to examine whether there are similarities/differences between existing members and candidates for membership in the OECD, among the selected post-transition countries of Europe, in the context of income inequality. The empirical analysis was carried out in several parts by using multivariate methods on aforementioned variables. For that purpose, factor analysis with the principal component method and cluster analysis were chosen as appropriate methods.

Three of the five selected variables in the analysis are inequality measures. Therefore, the starting point was to check the presence of multicollinearity. Multicollinearity is one of the critical issues in cluster analysis. It endangers the clustering results because each variable is equally weighted in cluster analysis [10]. If there is a certain degree of multicollinearity among the clustering variables (VIF greater than 5.00), a problem might arise if some variables are highly correlated while others are relatively uncorrelated. In that case, highly correlated variables impact the cluster solution with a greater effect than uncorrelated variables.

The Variance Inflation Factor (VIF) was calculated for all combinations of dependent variables in the model and resulted in values greater than 5.00 for all income inequality measures (the lowest VIF = 34.88). Due to the serious multicollinearity problem, a factor analysis emerged as an adequate method for satisfying cluster analysis assumptions since there is no multicollinearity between the extracted factors. Finally, cluster analysis was applied to the factor scores serving as variables for the classification of the observed countries.

Factor analysis is an interdependence technique whose basic function is to explore the structure among a set of variables and to reduce the data to a smaller set of variables known as factors. Exploratory factor analysis was carried out on the data to group the original five variables into a smaller number of factors. The benefit of this approach is generating factors with factor scores as a composite measure representing the relative contribution of all original variables to a particular factor [10]. Another advantage, as already mentioned, is avoiding drawbacks caused by multicollinearity. On the opposite side, the disadvantage of producing factor scores is often related to the interpretability due to scores' composite character, which in this case was not an issue because factor scores were just a stepping stone to the next stage of multivariate analysis.

The model for factor analysis with the principal component method applied in this paper is defined by the following function:

$$z_i = \delta_{i1} F_1 + \delta_{i2} F_2 \ , \quad i = 1, 2 \tag{1}$$

where both observed variables are described linearly in terms of two new uncorrelated components $F_j \ , \ j = 1, 2$. It is a model in which the factor extraction is based on the total variance. The coefficients (δ_{ij}) of the components are called factor (component) loadings and represent the correlations of the variables with the components [11].

The initial (unrotated) factor matrix is rarely satisfying in terms of structure simplicity and factor interpretability. For that reason, rotational methods are often applied to reduce the complexity and improve the interpretation. Factor rotation is a process of transforming the structure of the factor matrix in such a way that the variance contained in earlier factors is reallocated to later factors in order to obtain a more meaningful factor solution. There are two basic types of methods: orthogonal and oblique [10]. Orthogonal transformation approaches are preferred as they retain orthogonality among the factors, which is not the case with the latter type of method. The most commonly used rotation method is varimax, which belongs to orthogonal methods. Varimax rotation simplifies the columns (factors) of the initial factor matrix by maximizing the variance of their squared loadings [11]. The most simplified structure is reached if there are only ones and zeroes in a column (factor).

Hierarchical and non-hierarchical methods are interdependence techniques often combined as a part of a two-step approach in cluster analysis. This way one method serves as both control and compensation for the deficiencies of the other. The previous was achieved by applying Ward‘s and k-means method. Ward‘s method is based on a hierarchical clustering algorithm in which clusters are generated by minimizing the within-cluster variation. It provides a dendrogram which displays possible clustering solutions. The k-means method is based on a non-hierarchical clustering algorithm in which “k” represents the number of pre-defined clusters. Some advantages of applying hierarchical procedures include simplicity of the clustering process, different similarity measures and a wide range of possible cluster solutions. On the other hand, too many alternative cluster solutions can be misleading, clustering results are also sensitive to outliers in the data and large samples. However, some of these shortcomings can be handled by non-hierarchical clustering, especially outliers. The k-means method assumes spherical density and approximately equal number of observations in each cluster which is considered to be its main weakness [1]. Thus, none of the methods is perfect, but their combination is often a compromise solution. All things considered, in order to determine the final cluster solution, it is recommended [10] to use Ward‘s method in the first step, and then proceed with the selected number of clusters to the second step which is the k-means method. The described two-step approach has been used as a methodological framework in this paper.

Ward‘s minimum variance method relies on squared Euclidean distances. It is an agglomerative approach in which the distance (d_{ij}^2) between two objects, a and b, with j dimensions, is calculated [24, 26] as:

$$d_{ij}^2 = \|a-b\|_2^2 = \sum_{i=1}^{j}(a_i - b_i)^2 \ , \quad j = 1,2 \tag{2}$$

In our case, the objects are post-transition OECD countries, and the dimensions are factors, generated by factor analysis. The squared Euclidean distance is advised as a suitable similarity measure for the centroid and Ward‘s method of classification [10]. It provides convenient results, avoids the computation of the square root and consequently accelerates the clustering calculation which is considered to be the advantage.

3. Results and discussion

As described earlier, multivariate analysis was performed using a combination of factor and cluster analysis methods. Factor analysis with the principal component method aims to group similar variables into dimensions. The factor model produced two factors that explained a total of 92.87% of the variance, with the first factor accounting for 60.29% (Table 1).

Value	**Eigenvalues** (Post-transition OECD) Extraction: Principal components			
	Eigenvalue	%Total	Cumulative 1	Cumulative 2
1	3.014352	60.28704	3.014352	60.28704
2	1.629192	32.58384	4.643544	92.87088

Table 1: *Eigenvalues*

Due to the complexity of the factor loadings, the initial (unrotated) factor matrix was not interpretable; therefore, it was necessary to perform orthogonal (varimax) rotation. The factor matrix obtained by varimax rotation of the factor loadings clearly separated the three measures of inequality (Factor 1) from the measures of education and the labor market (Factor 2). Factor loadings in Table 2 represent the relationship between the variables and the associated factors. Gini and the S80/S20 ratio are positively correlated with Factor 1, while the indicator of the bottom 40% of income distribution is negatively correlated with the same factor. On the other hand, tertiary education and labor force participation rate are negatively correlated with Factor 2.

Variable	**Factor Loadings** (Varimax normalized) (Post-transition OECD) Extraction: Principal components (Marked loadings are >0,700000)	
	Factor 1	**Factor 2**
edu_attain_C	0.133879	-0.903600
lfpr	-0.047680	-0.914264
gini_coeff	0.993379	-0.081462
s80_s20	0.993280	0.004137
bottom_40	-0.993755	0.058028
Expl.Var	2.981152	1.662392
Prp.Totl	0.596230	0.332478

Table 2: *Factor loadings after the factor rotation*

Note: *edu_attain_c* – tertiary education, *lfpr* – labor force participation rate, *gini_coeff* – the Gini coefficient of equivalized disposable income, *s80_s20* – the S80/S20 income quintile ratio, *bottom_40* – income share of the bottom 40% of the population

Data summarization in a simplified form was achieved by generating factor scores, which were then used as inputs for cluster analysis. Two methods were applied, Ward's (hierarchical) and k-means (non-hierarchical).

Figure 1 shows the dendrogram obtained by Ward's method. Graphically, it is possible to choose between 2 to 5 clusters. In a solution with two clusters, OECD member countries would be located on one pole, and candidate countries for OECD membership on the other. When choosing a five-cluster solution, OECD member countries are divided into three groups (the Baltic countries separately; Hungary and Poland separately; and the Czech Republic, Slovenia and Slovakia separately), while in the case of candidate countries, Bulgaria is separated as a special cluster. The dendrogram constructed below suggests choosing a three-cluster solution as the most reliable representation of post-transition countries. Bulgaria, Croatia and Romania are clearly outlined and grouped in Cluster 1. These are candidate countries for membership in the OECD. Furthermore, the Czech Republic, Slovenia, Slovakia, Hungary and Poland are classified in Cluster 2. These are member countries of the so-called Visegrád Group (all but Slovenia) which are pro-Western oriented. Finally, Cluster 3 gathered the Baltic countries. This way of clustering (Table 3) indicates that Bulgaria, Croatia and Romania are specific in terms of income inequality compared to other post-transition OECD member countries. Their common grouping can be attributed to the political and historical background and non-implementation of institutional reforms as indicated by some researchers [16, 22].

A clearer understanding of the selection of the optimal cluster solution requires a deeper look at the procedure itself. The dendrogram displays the graphical merging of clusters at each step of the hierarchical procedure until all clusters are merged into one cluster. It is based on the agglomerative procedure which produces N-1 cluster solutions, where N stands for the number of countries. There are 11 separate countries at the beginning of the analysis. The dendrogram illustrates how solutions from eleven to one cluster are created. The horizontal axis in Figure 1 shows the distance used in merging clusters. The objective is to combine countries into clusters and proceed with their further joining as long as there is a small level of heterogeneity. According to Hair [10], the measure of heterogeneity should be zero at the beginning and then increase at each stage of cluster merging to reflect the level of heterogeneity between cluster solutions. Big jumps in heterogeneity values imply that two rather distinct clusters were merged at that stage. This is considered a stopping rule in the selection of a final cluster solution. There are different ways to determine the optimal number of clusters in hierarchical cluster analysis. In our case, we opted for the Pseudo t-squared and Pseudo F measures which indicate the changes in heterogeneity when merging clusters, as well as the homogeneity of each created cluster solution, respectively [10]. For the former measure, the smaller the value, the better the distinction between clusters. For the latter measure, the largest increases in values are preferred as they reflect more distinct clusters. The Pseudo t-squared (with the lowest value of 3.13 for a three-cluster solution) and Pseudo F (with the largest increase to the value of 16.39 for a three-cluster solution) results, based on Ward's clustering method, revealed that the optimal solution is to choose three clusters (Table 4).

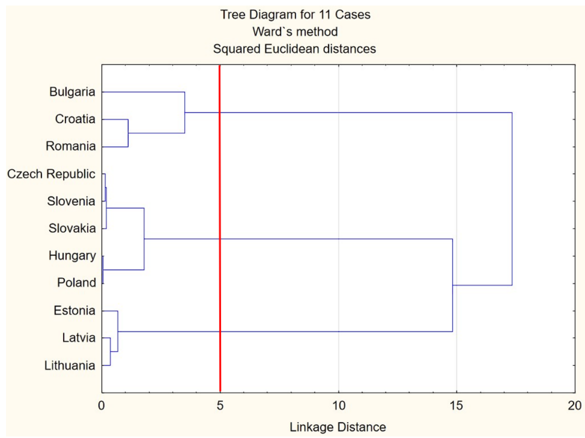

Figure 1: *Dendrogram*

Cluster 1	Cluster 2	Cluster 3
Bulgaria	Czech Republic	Estonia
Croatia	Slovenia	Latvia
Romania	Slovakia	Lithuania
	Hungary	
	Poland	

Table 3: *Results of hierarchical cluster analysis, classification of countries*

Number of clusters	Pseudo t-squared	Pseudo F
1	6.88	.
2	27.56	6.88
3	3.13	16.39
4	14.22	19.15
5	.	22.05

Table 4: *Stopping rule measures for determining the optimal number of clusters*

The number of clusters determined by hierarchical analysis can be used as the initial number of clusters in non-hierarchical analysis [10]. Guided by this approach, the results of the k-means method are consistent with those obtained by Ward's method. The composition of all clusters is identical (Table 5). Therefore, non-hierarchical analysis based on the k-means algorithm confirmed the results achieved by hierarchical analysis using Ward's minimum variance criterion. The three-cluster solution can be accepted as the final cluster solution for post-transition OECD countries in 2020.

Classification	Countries	Distances
Cluster 1 (”OECD candidates”)	Romania	0.139957
	Croatia	0.742907
	Bulgaria	0.764050
Cluster 2 (“the Visegrád Group + Slovenia”)	Czech Republic	0.129098
	Slovenia	0.294008
	Poland	0.315830
	Slovakia	0.399077
	Hungary	0.426317
Cluster 3 (“the Baltic countries”)	Lithuania	0.230119
	Latvia	0.310553
	Estonia	0.336519

Table 5: *Results of non-hierarchical cluster analysis, classification of countries and distances*

Analysis of variance (Table 6) highlights Factor 2, composed of education and labor market variables, as the dimension that contributes the most to the separation between post-transition OECD countries (higher F-value; 33.5099). Factor 1 (F-value; 9.9999), on which inequality measures are grouped, also significantly contributes to the differences between clusters, but less than Factor 2.

Concerning the conventional hypothesis testing, there are certain limitations in cluster analysis. Dubes and Jain (1979, [3]), Milligan and Mahajan (1980, [18]) as cited in Milligan and Hirtle (2012, [17]) claim that the ANOVA table after k-means analysis should be used only for descriptive purposes. Using it for statistical hypothesis testing would be a mistake since the data is not randomly assigned to the clusters. An ANOVA table nearly always provides significant results regardless of the structure in the data. Therefore, only the magnitude of F-values was interpreted to reflect how well the respective factors differentiated the clusters.

Variable	**Analysis of Variance**					
	Between SS	df	Within SS	df	F	signif. p
FACTOR 1	7.14283	2	2.85717	8	9.9999	0.00666
FACTOR 2	8.93362	2	1.06638	8	33.5099	0.00013

Table 6: *ANOVA*

Variable mean values for each of the three clusters are graphically shown in Figure 2. Cluster 1, representing the OECD candidate countries (Bulgaria, Croatia and Romania), has the highest mean values of inequality measures, tertiary education and labor force participation rate. Furthermore, Cluster 2, composed of the Visegrád Group countries (Czech Republic, Hungary, Poland, Slovakia) and Slovenia, has the lowest mean values of inequality measures. Finally, the Baltic countries, covered by Cluster 3, have the lowest mean values of tertiary education and labor force participation.

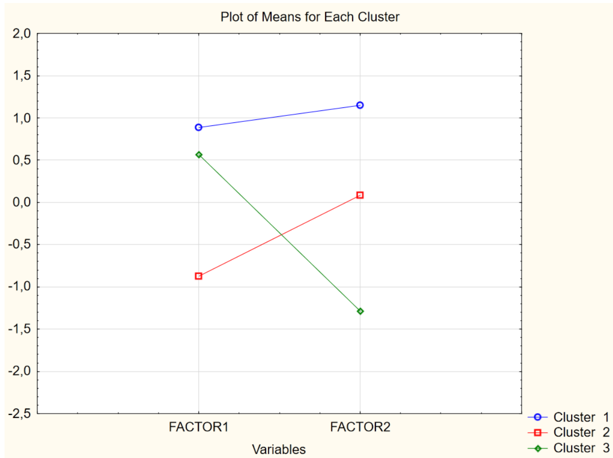

Figure 2: *Plot of Means for a three-cluster solution*

The findings obtained through multivariate analysis provide a significant contribution to existing research on post-transition OECD countries in the academic community. It seems that the geographical position of countries also shapes their current position in dealing with income inequality.

4. Conclusion

The results of the multivariate analysis led to the division of post-transition countries into three clusters. The Baltic countries have experienced substantial economic success. Apart from being geographically connected, their economic path in the transition and after the end of the process had similar outcomes. It can be said that their joint grouping is expected and in accordance with the development of events in that region.

The grouping of countries of the so-called Visegrád Group and Slovenia in the same cluster could be characterized as a pro-Western grouping. These are post-transition countries that until recently carried the label of the former Eastern Bloc, but with their economic performance, they have shown intentions to be treated as Western-oriented.

Institutional legacy and the lack of institutional reforms can be attributed to Bulgaria, Croatia and Romania as one of the main reasons why these countries are grouped in a common cluster. As of January 25, 2022, these countries started OECD membership negotiations in the hope that they themselves will experience and live as a part of a community that strives to achieve the idea contained in the Organization's motto: "Better policies for better lives" [20].

Countries must demonstrate serious intentions in order to gain candidate status for full membership in the OECD. For membership itself, it is necessary to meet strict criteria, which include, among other things, the evaluation of the candidates by the existing OECD member countries. There are currently 38 member countries in the Organization. The accession process itself goes beyond signing an accession agreement. It is extremely important to present "willingness" and "commitment" in an effort to bring the country into line with the Organization's requirements [27]. This implies (a) a democratic government with an orderly society that respects the rule of law and the protection of human rights; and (b) the image of free-market economies featuring openness and transparency.

It is a fact that income distribution is socially negotiated at the institutional level. In the same way, history and experience have shown that political elites have mostly managed institutional changes so far. This is also the reason why some post-transition countries do not have the opportunity to benefit from the system of meritocratic values today. The key drivers of change – democratic and efficient national governments, plus quality and independent institutions – are a prerequisite for the success of more equal income distribution in post-transition countries.

The OECD [7] and World Inequality Lab [19] at the Paris School of Economics mostly report the differences between Eastern and Western Europe. It is not easy to get a precise picture of inequality in such a way because some Eastern countries covered by these reports (different from the 11 in our analysis) have not completed the transition process (e.g., Western Balkan countries). The contribution of this paper is reflected in the identification of the candidate countries for the OECD as a separate group that differs from the existing post-transition OECD members, from the perspective of inequality. In 2020, Bulgaria, Croatia and Romania are outlined as countries that have a common picture of inequality. Through this empirical analysis, it became obvious that the existing European post-transition OECD members can be divided into two groups of countries. Additionally, the candidate countries were formed as a third circle of countries that do not fit into either of the two groups of existing post-transition OECD members. Finally, this confirmed greater heterogeneity between countries with a communist history.

The research discussed in this paper is an excellent starting point for understanding the current situation in CEE countries. In future research, emphasis should be placed on a more detailed analysis of post-transition countries, as well as the OECD in a wider scope. Convergence club analysis is one of the options for examining convergence and divergence within and between countries. The OECD is still under-researched in the context of convergence clubs, which is potentially a topic of importance for future analyses. It would be interesting to see how the entire OECD integration, which by its economic nature is very heterogeneous, moved through history and what the similarities, i.e., differences between and within the member countries of that spectrum are. A comparative cluster analysis covering several points in time would be an ideal complement to this scientific work. This could determine the very dynamics of income inequality in post-transition OECD countries.

References

[1] de Craen, S., Commandeur, J., J., F., Frank, L., E. and Heiser, W., J. (2006). Effects of Group Size and Lack of Sphericity on the Recovery of Clusters in K-means Cluster Analysis. Multivariate Behavioral Research, 41, 127–145. doi: 10.1207/s15327906mbr4102$_2$

[2] Dolls, M., Doorley, K., Paulus, A., Schneider, H., Sommer, E. (2019). Demographic change and the European income distribution. Journal of Economic Inequality. 17(3), 337-357. doi: 10.1007/s10888-019-09411-z

[3] Dubes, R., Jain, A., K. (1979). Validity studies in clustering methodologies. Pattern Recognition, 11, 235–254. doi: 10.1016/0031-3203(79)90034-7

[4] Eurostat – Statistics Explained. Retrieved from: ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Equivalised_disposable_income [Date of Access: 15 October 2022]

[5] Eurostat – Statistics Explained. Link: Retrieved from: ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Income_quintile_share_ratio [Date of Access: 15 October 2022]

[6] Evans, M. D. R. and Kelley, J. (2016). Communism, Capitalism, and Images of Class: Effects of Reference Groups, Reality, and Regime in 43 Nations and 110,000 Individuals, 1987-2009. Cross-Cultural Research, 51(4), 315-359. doi: 10.1177/1069397116677963

[7] Förster, M., Nozal, A., L. and Thévenot, C. (2017). Understanding the Socio-Economic Divide in Europe. OECD Centre for Opportunity and Equality. Social Policy Division. Retrieved from: oe.cd/cope-divide-europe-2017

[8] Ganic, M. (2020). Are determinants of International Financial Integration in the European Transition Countries Different from Post-transition Countries? Studies in Business and Economics, 15(1), 40-54. doi: 10.2478/sbe-2020-0005

[9] Goedeme, T., Decerf, B. and Van den Bosch, K. (2022). A new poverty indicator for Europe: The extended headcount ratio. Journal of European Social Policy, 32(3), 287-301. doi: 10.1177/09589287221080414

[10] Hair, J., F., Black, W., C., Babin, B., J. and Anderson, R., E. (2018). Multivariate Data Analysis. 8^{th} Edition. Andover, Hampshire, United Kingdom: Cengage Learning EMEA.

[11] Harman, H. H. (1976). Modern Factor Analysis, 3rd Edition. The University of Chicago Press, (pp. 15).

[12] ILO Data Catalogue. Labor force participation rate by sex and age (%). Retrieved from: ilo.org

[13] Iwasaki, I. and Kumo, K. (2019). J-Curve in Transition Economies: A Large Meta-analysis of the Determinants of Output Changes. Comparative Economic Studies, 61(1). 149-191. doi: 10.1057/s41294-018-0058-4

[14] Jianu, I., Gavril, I., A., Iacob, S., E. and Hrebenciuc, A. (2021). Income Inequalities and Their Social Determinants: An Analysis over Developed vs. Developing EU Member States. Economic Computation and Economic Cybernetics Studies and Research. 55(2), 125-142. doi: 10.24818/18423264/55.2.21.08

[15] Josifidis, K., Supic, N. and Doroskov, N. (2020). Foreign Direct Investment and Income Distribution: Evidence from Post-Communist New EU Member States. Eastern European Economics, 58(6), 497-516. doi: 10.1080/00128775.2020.1762496

[16] Kuzmar, S. and Piatek, S. (2019). Institutional determinants of inequality in chosen post-socialist countries: The role of political freedom. Ekonomia i Prawo-Economics and Law, 18(3), 295-315. doi: 10.12775/EiP.2019.021

[17] Milligan, G., W. and Hirtle, S., C. (2012). Clustering and Classification Methods. In J. A. Schinka, W. F. Velicer, and I. B. Weiner (Eds.) Handbook of psychology: Research methods in psychology (pp. 189–210). John Wiley & Sons, Inc. doi: 10.1002/9781118133880.hop202007

[18] Milligan, G., W. and Mahajan, V. (1980). A note on procedures for testing the quality of a clustering of a set of objects. Decision Sciences, 11, 669–677. doi: 10.1111/j.1540-5915.1980.tb01168.x

[19] Morgan, M. and Neef, T. (2020). What‘s New About Income Inequality in Europe (1980-2019). The World Inequality Lab. Issue Brief 2020/04. Paris School of Economics, Paris, France. Retrieved from: wid.world

[20] OECD (2020). The OECD at 60. Paris, France, pp. 8. Retrieved from: read.oecd-ilibrary.org

[21] OECD.Stat Data. Educational attainment and labor-force status. Retrieved from: stats.oecd.org

[22] Piatek, D., Pilc, M. and Szarzec, K. (2019). What determines the institutional change in transition economies? Argumenta Oeconomica, 42(1), 235-272. doi: 10.15611/aoe.2019.1.10

[23] Prascevic, A. (2020). The Applicability of Political Business Cycle Theories in Transition Economies. Zagreb International Review of Economics & Business, 23, 73-90. doi: 10.2478/zireb-2020-0024

[24] Scitovski, R., Sabo, K., Martínez-Álvarez, F. and Ungar, S. (2021). Cluster Analysis and Applications. Springer Nature Switzerland AG. Cham, Switzerland, (pp. 5). doi: 10.1007/978-3-030-74552-3

[25] Szczepaniak, M. and Geise, A. (2021). Examining the relationships between income inequalities and different dimensions of well-being in selected Central Eastern European (CEE) countries. Plos One, 16(4), 1-19. doi: 10.1371/journal.pone.0250469

[26] Tabaghi, P., Dokmanic, I. and Vetterli, M. (2019). Kinetic Euclidean Distance Matrices. IEEE Transactions on Signal Processing, 68, 452-465. doi: 10.1109/tsp.2019.2959260

[27] The Trade Union Advisory Committee to the OECD – TUAC (2018). OECD Membership and the Values of the Organisation. OECD Week 2018, Paris, France, pp. 1-12.

[28] UNESCO Institute for Statistics (2012). International Standard Classification of Education – ISCED 2011. Montreal, Quebec, Canada.

[29] United Nations. SDG Indicators. Goal 10 – Reduce inequality within and among countries. Retrieved from: unstats.un.org

[30] Vujacic, I. and Petrovic-Vujacic, J. (2020). Incomplete Transition – Is there a "Mid-Transition Trap"? Zagreb International Review of Economics & Business. 23, Special Conference Issue, 57-71. doi: 10.2478/zireb-2020-0023