Comparative Economic Research. Central and Eastern Europe
Volume 26, Number 4, 2023
https://doi.org/10.18778/1508-2008.26.30

COPE
Member since 2018
JM13708

The Gender Gap in the Visegrád Group Countries Based on the Luxembourg Income Study

Alina Jędrzejczak https://orcid.org/0000-0002-5478-9284
Ph.D., Associate Professor, University of Lodz, Department of Statistics and Demography, Lodz, Poland
e-mail: alina.jedrzejczak@uni.lodz.pl

Kamila Trzcińska https://orcid.org/0000-0002-4714-4074
Ph.D., University of Lodz, Department of Statistics and Demography, Lodz, Poland
e-mail: kamila.trzcinska@uni.lodz.pl

Abstract

Gender equality is a fundamental human right and one of the core values of the European Union (EU). Great efforts have been made to defend this right and to promote gender equality within the member states and around the world. However, there are still significant differences between men and women, especially in terms of income. The main objective of the paper is to compare income distributions for gender groups across four Central European countries, Poland, Slovakia, Czechia and Hungary, i.e., the members of the Visegrád Group (V4). These countries share similar histories and similar economic development, but there are substantial differences between their approaches to economic reforms, including labour market policy. This, in turn, is reflected in different income distributions and income inequality patterns. There is a debated research issue regarding the methodology of measuring the gender gap – the traditional methods based on comparing means and medians seem unsatisfactory as they do not consider the shape of income distributions. The paper's novelty lies in the application of the relative distribution concept, which goes beyond the typical focus on average income differences toward a full comparison of the entire distribution of women's earnings relative to men's. In the paper, we implement a parametric approach for estimating the relative distribution, which allows us to compare and visualise the "gap" between the gender groups at each distribution quantile. The basis for the calculations was the microdata from the Luxembourg Income Study (LIS). The statistical methods applied in the study were appropriate to describe the gender gap over the entire income range. The results of the empirical analysis helped to reveal similarities and substantial differences between the countries.

creative commons

© by the author, licensee University of Lodz – Lodz University Press, Poland.
This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Received: 8.11.2022. Verified: 16.05.2023. Accepted: 15.09.2023

Keywords: income inequality, gender gap, Dagum distribution, relative distribution method, Visegrád Group

JEL: C4, D63, J16

Introduction

Although average incomes across countries have been converging for over two decades (World Bank 2016), COVID–19 directly offset the reduction in the gap between countries. As a result, income inequality and substantial regional disparities are still a great challenge for policymakers in many European countries. One of the critical elements of this phenomenon is the inequality between the income distributions of men and women. The gender pay gap can be a problem from a public policy perspective because it reduces economic output and means that women are more likely to be dependent upon welfare payments, especially in old age.

Reducing inequalities for European Union (EU) citizens and promoting upward convergence in living conditions is high on the policy agenda for the European Commission. Although the EU allocated 347 billion euros (over one-third of its budget) in the period 2007–2013 to transfers for regional policy to reduce economic and social disparities within and among member states, income disparity still grew in both EU and OECD countries, and in 2022 reached its highest level for the past half-century (OECD 2015). Several studies were conducted on the issue for the purpose of social and economic policies, including *Divided We Stand. Why Inequality Keeps Rising* (OECD 2011) and *In It Together: Why Less Inequality Benefits All* (OECD 2015). The trend of rising inequality has become a priority for policymakers, and there have been calls for the analysis of various aspects of income inequality, including its measurement and decomposition by regional area, income source, and, recently, gender (see Jędrzejczak 2015; Zenga and Jędrzejczak 2020).

Gender equality is one of the fundamental values of the EU. The European Commission's work on gender equality policy is based on the *Strategic engagement for gender equality 2016–2019*, which focuses on five priority areas, including increasing female labour-market participation, reducing the gender pay, earnings and pension gaps, combating gender-based violence, improving gender balance in decision-making, and promoting gender equality within the Member States and across the world. Although it is generally illegal for employers in the EU to pay men and women different amounts for doing the same job, there are many reasons why, on average, there are substantial income differences between men and women. The income differences capture differences across many possible dimensions, including education levels, working hours, experience, and occupation, among others. Gender equality, often understood only in terms of income, should be viewed as multidimensional. Gender equality also means equal economic independence for women and men; it refers to equality in decision-making,

and, in the broader setting, it requires equal dignity, integrity, and the ending of gender-based violence.

On the one hand, occupational segregation is perhaps the main reason: men are prevalently in higher-paid industries, while women are mostly in lower-paid industries. There is vertical segregation, too. Few women work in senior, and hence better-paying positions. Finally, some barriers to entry into the labour market are related to the education level and single parenting rate. Blau and Kahn (2000; 2003) developed an in-depth analysis that showed that differences in pay are caused by many concurring factors. Meanwhile, Leythienne and Ronkowski (2018) studied gender gaps in many countries based on the Structure of Earnings Survey (SES) data. Greselin and Jędrzejczak (2020) studied the gender gap in Poland and Italy, comparing data provided by Eurostat for Poland and Italy through the relative distribution approach. Greselin, Jędrzejczak, and Trzcińska (2023) proposed a new parametric approach to gender gap analysis based on different theoretical distributions, and it was applied to EU-SILC data for Poland and Italy. Various dimensions of the discrepancy between men and women have recently been considered in the literature for many countries (e.g., Doorley and Claire 2020; Avram and Popova 2022; Boneva et al. 2022; Cordova, Grabka, and Sierminska 2022; Glaubitz, Harnack-Eber, and Wetter 2022; and Cozarenco and Szafarz 2023).

The research issue of measuring the gender gap is a subject of debate regarding the methodology used. The Eurostat database contains an indicator called "unadjusted gender pay gap", defined as the relative difference between average gross hourly earnings, from the four-yearly Structure of Earnings Survey (Eurotat n.d.). The gender pay gap in the EU in 2019 was 14.1% (i.e., women earn 14.1% less per hour than men on average), and it had only changed minimally over the last decade. Another summary measure used by Eurostat, called "the gender overall earnings gap", stood at 36.7%. It measures the combined impact of average hourly earnings, the monthly average number of hours paid (before any adjustment for part-time work) and the employment rate. Similar indicators can easily be obtained based on EU-SILC (Survey of Income and Living Conditions) data by comparing mean or median incomes for gender groups. However, this approach is unsatisfactory, as the gender gap is related to the whole distribution of incomes in a population, so it is difficult to capture the full range of experiences using a single metric. To reveal the factors that contribute to the gender discrepancy, one should adopt a variety of tools, consider concomitant variables, and go beyond the typical focus on average or median earnings differences toward a full comparison of the entire distribution of women's earnings relative to men's.

The present paper focuses on income distributions across four Central European countries: Poland, Slovakia, Czechia and Hungary, i.e., the members of the Visegrád Group (V4). These countries share not only similar histories but also similar economic development (measured by Gross Domestic Product (GDP)). The idea was to compare the neigh-

bouring countries with different paths of economic transformation from a centrally planned to a market-based economy. Even though state influence was radically weakened in favour of market liberalisations, the effective transformation of these economies was based on country-specific institutional reforms. Differences between the national approaches to economic reforms, including labour market policy, are particularly reflected in different income distributions and income inequality patterns.

According to a World Bank report (World Bank 2000), the first study to include Central European transition countries, income disparities between the rich and poor increased in virtually all transition economies during the 1990s. However, the extent of this increase varied considerably across countries. In explaining the main causes of the observed changes in income inequality, the study highlighted the role of increased inequality of labour earnings, which could be traced to a rapid rise in returns to education. Government tax and transfer policies were also found to have had a huge impact on income distribution, dampening the rises in income inequality due to increased dispersion of earnings. It was visible in Central European countries much more than in many other countries.

On the other hand, public policy in these countries played an important role in reducing income inequalities, both through national taxation as well as benefits systems. This impact was so large in post-socialist countries because of greater inequality aversion, which resulted in social policies against the rise in inequality (World Bank 2000). A Tárki Social Research Institute study on intolerance to income inequality across countries confirmed a markedly lower level of acceptance of inequality in the post-socialist bloc than in the other European countries, with high national differentiation also within the bloc (TARKI 2009). Zaidi's (2009) study of the main drivers of income inequality in Central European and Baltic Countries revealed that Slovakia and the Czech Republic had relatively low income inequality because of the strong redistributive role of taxes and benefits. By contrast, the role of direct taxes and public transfers in redistributing incomes was much smaller in Poland and, to some extent, Hungary.

During the transformation, Slovakia had relatively low average tax shares (around 17–18% of equivalised disposable incomes). In Hungary and the Czech Republic, the share was about 21%, while Poland had the highest share (33%). Meanwhile, in Poland and Hungary, the shares of social benefits in disposable incomes (27% and 33%, respectively) were higher than in the Czech Republic and Slovakia (25% and 23%). Ultimately, however, income inequalities are clearly higher in Poland and Hungary than in the Czech Republic and Slovakia, which can be explained by a more effective tax and benefit policy (see Zaidi 2009). In another study on V4 countries, using aggregated data from Fraser Institute, Eurostat, and the OECD database, Szczepaniak and Szulc-Obłoza (2020) identified the impact of many labour market institutions when trying to limit income inequalities. Using a taxonomic analysis based on aggregated sta-

tistical data coming from Fraser Institute, Eurostat and OECD, they confirmed that the Czech Republic and Slovakia are classified in one group of EU countries, while Poland and Hungary are classified in another one, taking into consideration both labour market institutions and income inequalities. However, the introduction of social programs in Poland in 2016 (e.g., the Family 500+ child benefit program) significantly reduced income inequalities, especially in families with children (Jędrzejczak and Pekasiewicz 2020a). Further changes are expected when the minimum wage increases.

A phenomenon strictly related to income inequalities is the discrepancy between male and female incomes, known as the gender gap. The main objective of this paper is to analyse this discrepancy in the V4 countries using theoretical income distributions and the relative distribution approach. A nonparametric version of the method has already been applied by Greselin and Jędrzejczak (2020), who compared income distributions in Poland and Italy based on EU-SILC data. The current analysis was conducted for 2015–2020, as it was the most recent data available. Due to the relative stability of income distributions, huge changes between particular countries are not observed over time. The statistical analysis was performed using our own numerical procedures implemented in the R-project environment. The calculations were based on microdata from the LIS (Luxembourg Income Study).

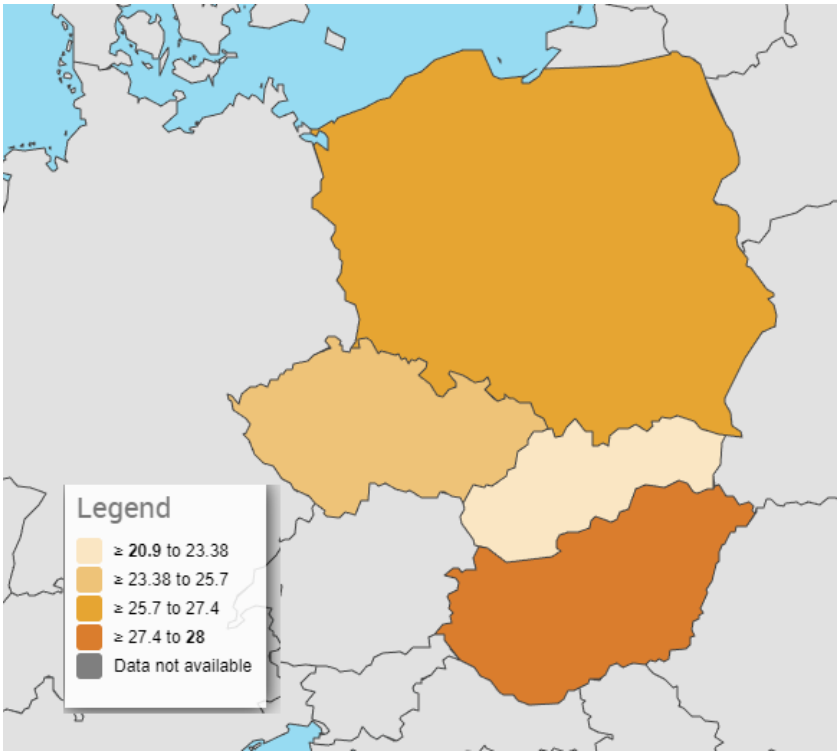

Figure 1. Gini coefficient of equivalised disposable income in 2020 based on EU-SILC survey

Source: own elaboration based on Eurostat database n.d.

This paper aims to show the similarities and differences in the income distributions of women and men in the Visegrad group countries, which differ significantly in terms of their overall income inequality measured by the Gini index (see Figure 1). A new parametric approach for estimating the relative distribution was used, which makes it possible to compare and visualise the "gap" between the gender groups at each distribution quantile. This approach goes beyond the typical focus on average or median earnings differences toward a full comparison of the entire distribution of women's earnings relative to men's.

Theoretical income distributions

Since Pareto proposed his first income distribution model in 1896, many economists and mathematicians have tried to describe empirical distributions using simple mathematical formulas with few parameters. These formulas can be useful for many reasons. Firstly, applying a theoretical model simplifies the analysis because different distribution characteristics can be performed using the same parameters. Secondly, a theoretical model that is well-fitted to the data can be used to predict wage and income distributions in different divisions. Additionally, approximating empirical wage and income distributions using the theoretical curves can smooth irregularities from the data-collecting method, which is often the case for income data. Two different economic size distributions widely employed in the literature for fitting income data are the two-parameter Lognormal model and the three-parameter Dagum model. We provide a definition and basic information for making inferences from survey data.

The Lognormal distribution is a two-parameter model frequently applied for fitting income distributions in many countries, mainly due to its simplicity and the straightforward interpretation of its parameters (see Aitchison and Brown, 1957). It fits lower income levels better than the Pareto distribution, but its fit towards the upper tail is far from satisfactory. Nevertheless, it can be applied to approximate selected empirical income distribution, especially in post-socialist countries. A Lognormal random variable Y has the following density function:

$$f(y)=\frac{1}{y\sigma\sqrt{2\pi}}exp\left(-\frac{(\ln y-\mu)^2}{2\sigma^2}\right) \text{ for } y\in R, \tag{1}$$

where: $\mu\in R-$ the expected value of the logarithms of income Y,

$\sigma>0-$ the standard deviation of the logarithms of income Y.

The Lognormal random variable cannot be explicitly expressed by the formula; it can only be written as:

$$F(y)=\frac{1}{y\sigma\sqrt{2\pi}}erf\left(-\frac{\ln y-\mu}{\sigma\sqrt{2}}\right), \tag{2}$$

where: $erf(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$.

The methods most frequently applied to estimate the parameters μ and σ of the Lognormal model $LN(\mu,\sigma)$ based on empirical data are the following are the maximum likelihood method (ML), the method of moments (MM), and the method of quantiles. The ML estimators, which present the best statistical properties for large samples $Y_1,Y_2,\ldots,Y_n$, are given by the following formulas:

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}\ln Y_i. \tag{3}$$

$$\widehat{\sigma^2}=\frac{1}{n-1}\sum_{i=1}^{n}\left(\ln Y_i-\hat{\mu}\right)^2. \tag{4}$$

Both estimators are unbiased and most efficient; what is more, their large sample variances are given by simple formulas: $D^2(\hat{\mu})=\frac{\sigma^2}{n}$ and $D^2\left(\widehat{\sigma^2}\right)=\frac{2\sigma^4}{n}$.

The Dagum model takes its name from Camilo Dagum, who introduced it in the 1970s when looking for a statistical distribution that closely fit empirical income and wealth distributions. To mimic the characteristic properties observed in such datasets, Dagum (1977) searched for a model that simultaneously permitted an interior mode (like the Lognormal) and could handle heavy tails (like the Pareto). He based his proposal on the empirical observation that the income elasticity $\eta(F, y)$ of the cumulative distribution function (*cdf*) of income is a decreasing and bounded function of *F*, and, therefore, of *y*. After decades of applications to real data in different divisions, the Dagum model has proven to be an appropriate candidate to model male and female income distributions.

We can say that F belongs to the Dagum family if its probability density function (*pdf*) is given by the following formula:

$$f(y)=\begin{cases} \dfrac{ap}{b}\left(\dfrac{y}{b}\right)^{ap-1}\left(1+\left(\dfrac{y}{b}\right)^{p}\right)^{-a-1} & for \quad y>0, \\ 0 & for \quad y\le 0. \end{cases} \tag{5}$$

The *cdf* of the Dagum distribution takes the form:

$$F(y)=\begin{cases} \left(1+\left(\dfrac{y}{b}\right)^{-p}\right)^{-a} & for \quad y>0, \\ 0 & for \quad y\le 0, \end{cases} \tag{6}$$

where: $a>0, b>0$ i $p>0$.

The quantile function of the Dagum distribution has the form:

$$F_D^{-1}(u;a,b,p)=b\left[u^{-1/p}-1\right]^{-1/a}, \quad \mathrm{u}\in(0,1). \tag{7}$$

The Dagum model $D(a, b, p)$ can be seen as a special case of the generalised beta distribution of the second kind (GB2); it is also a member of the Burr family, equivalent to the Burr type III distribution (for more detail on this distribution in the framework of economic size distributions, see Kleiber and Kotz 2003; Kleiber 2008, Jędrzejczak and Pekasiewicz 2020b).

This model allows for various degrees of positive skewness and leptokurtosis; moreover, it has built-in flexibility to be unimodal (to approximate income distributions) or zero modal (to describe wealth distributions). The ML estimators of the Dagum model parameters, namely $\hat{a}_0, \hat{b}_0, \widehat{p}_0$, can only be obtained using numerical procedures, but they present good inferential properties for large samples (for detail, see Jędrzejczak, Pekasiewicz, and Zieliński 2021).

Quantifying the gap between income distributions

Let Y denote a random variable that represents the analysed income distribution with density $f(y)$ and distribution function $F(y)$, i.e., the comparison distribution. Let Y_0 denote a random variable that represents a reference distribution with density $f_0(y)$ and distribution function $F_0(y)$. Our objective is to study the differences between distributions Y and Y_0, using Y_0 as the reference distribution.

The relative distribution of random variable Y with respect to the variable Y_0 is defined as the distribution of the transformed random variable R (see Handcock and Morris 2006):

$$R = F_0(Y). \tag{8}$$

Therefore, variable R determines the ranks of random variable Y with respect to the distribution of random variable Y_0. In other words, the value of the income of an individual from one population is assigned to the rank that this income would have in the distribution of the other population.

The distribution function of random variable R (i.e., *relative distribution function*) is expressed as:

$$G(r) = F\left(F_0^{-1}(r)\right) \text{ for } r \in [0,1]. \tag{9}$$

The respective density of random variable R (i.e., *relative density*) can be obtained as the integral of $G(r)$, and takes the form:

$$g(r) = \frac{f\left(F_0^{-1}(r)\right)}{f_0\left(F_0^{-1}(r)\right)} \text{ for } r \in [0,1]. \tag{10}$$

For the Dagum distribution $D(a, b, p)$, assumed as an underlying income distribution model, the relative distribution function can be expressed as:

$$G_{D(a,b,p;a_0,b_0,p_0)}(r) = \left[1 + \left(\frac{b}{b_0}\right)^{p} \left(r^{-1/p_0} - 1\right)^{p/a_0}\right]^{-a}. \tag{11}$$

Given two Dagum distributions, well fitted to the empirical data for both gender groups, the function $G(r)$ can be expressed using the estimates of their parameters, namely $\hat{a}, \hat{b}, \hat{p}$ and $\hat{a}_0, \hat{b}_0, \hat{p}_0$:

$$G_{D(\hat{a},\hat{b},\hat{p};\hat{a}_0,\hat{b}_0,\hat{p}_0)}(r) = \left[1 + \left(\frac{\hat{b}}{\hat{b}_0}\right)^{\hat{p}} \left(r^{-1/\hat{p}_0} - 1\right)^{\hat{p}/\hat{a}_0}\right]^{-\hat{a}}. \tag{12}$$

Formula (12) will further be applied to estimate the gap between the income distributions of men and women for the Visegrád Group countries.

Applications to LIS data

The Luxembourg Income Study Database (LIS) is a cross-national data centre which serves a global community of researchers, educators, and policymakers. LIS is the largest available income database of harmonised microdata collected from about 50 countries in Europe, North America, Latin America, Africa, Asia, and Australia, spanning five decades.

Harmonised into a common framework, LIS datasets contain household- and person-level data. The focus of the survey is to measure the standard of living and to gather information about household income. The survey acquires datasets with income, wealth, employment, and demographic data from many high- and middle-income countries, harmonises them to enable cross-national comparisons, and makes them publicly available in two databases, the Luxembourg Income Study Database (LIS) and the Luxembourg Wealth Study Database (LWS).

We used individual monthly disposable household income in all calculations, but households with negative or zero incomes were excluded from the statistical analysis for methodological and interpretation reasons. The analysis of the Visegrad Group was based on the latest data available for those countries, i.e. from 2020 for Poland, from 2018 for Slovakia, from 2016 for the Czech Republic and from 2015 for Hungary (Luxembourg Income Study (LIS) Database n.d.).The results are given in Tables 1–3 and Figures 2–9.

Tables 1 and 2 reveal remarkable differences between the estimates of basic statistical characteristics for men and women, which are visible for all the countries considered. The biggest discrepancies between means and medians were observed for the Czech Republic, where men's incomes were, on average, almost 50% higher. Poland and Slovakia had the next highest results (men out-earned women by approximately 30%, on average), while for Hungary, this difference was relatively small (20%). When comparing income inequality within the gender groups for each country, there are relatively big differences between the Gini ratios for the Czech Republic and Hungary (they are about 10% higher for men), while for Poland and Slovakia the Gini indices for men and women are more similar. It is worth noting that the Gini indices for personal income are visibly different than those observed for equivalised household incomes in the Eurostat database (see Figure 1).

Table 1. Lognormal model parameters and estimates of basic statistical characteristics for income distributions in V4 countries by gender

Country		Sample size	Lognormal model parameters		Mean [in euros or national currency]	Median [in euros or national currency]	Gini coefficient
			μ	σ^2			
Czech Republic	Male	7,124	9.9146	0.4567	25,412.12	20,224.46	0.3672
	Female	7,780	9.5605	0.3974	17,312.63	14,192.51	0.3442
Hungary	Male	2,275	11.6899	0.3618	143,030.05	119,363.65	0.3294
	Female	2,766	11.5222	0.3243	118,701.87	100,932.78	0.3128
Poland	Male	26,811	7.8562	0.3681	3,103.37	2,581.70	0.3321
	Female	29,070	7.5922	0.2975	2,300.68	1,982.69	0.3002
Slovakia	Male	5,228	6.5478	0.4235	862.24	697.71	0.3546
	Female	6,075	6.2931	0.3861	656.01	540.85	0.3396

Source: own elaboration based on LIS individual monthly disposable household income data: LIS Cross-National Data Center in Luxembourg n.d.

Table 2. Dagum model parameters and estimates of basic statistical characteristics for income distributions in V4 countries by gender

Country		Sample size	Dagum model parameters			Mean	Median	Gini coefficient
			a	b	p			
Czech Republic	Male	7,124	3.152	24,999.1	0.696	25,233.52	21,104.35	0.3478
	Female	7,780	3.499	16,999.8	0.705	16,905.15	14,675.13	0.3138
Hungary	Male	2,275	3.537	141,000	0.700	138,352.24	122,587.08	0.2940
	Female	2,766	4.279	120,000	0.658	113,470.07	103,719.92	0.2647
Poland	Male	26,811	3.581	3,002.4	0.732	3,022.00	2,643.55	0.3034
	Female	29,070	3.755	2,203.1	0.796	2,274.29	2,017.13	0.2831
Slovakia	Male	5,228	4.448	1,036.8	0.433	828.92	761.44	0.2992
	Female	6,075	4.190	731.1	0.526	633.31	575.38	0.2900

Source: own elaboration based on LIS individual monthly disposable household income data: LIS Cross-National Data Center in Luxembourg n.d.

Table 3. Accuracy of the parametric estimates based on the lognormal and Dagum models

Income distribution		Dagum model		Lognormal model	
		Relative difference [%]		Relative difference [%]	
		Mean	Median	Mean	Median
Czech Republic	Male	1.09	3.62	1.81	1.22
	Female	0.72	11.68	1.68	8.01
Hungary	Male	1.77	1.93	1.55	4.51
	Female	2.76	0.27	1.72	2.95
Poland	Male	1.98	1.67	0.66	0.70
	Female	0.70	0.86	0.36	0.87
Slovakia	Male	0.38	1.17	4.42	9.44
	Female	0.08	3.53	3.67	2.68

Source: own elaboration based on LIS individual monthly disposable household income data: LIS Cross-National Data Center in Luxembourg n.d.

The results obtained using the Dagum model (Tab. 2) are more reliable, as this model exhibits a better fit to the empirical data (see Tab. 3 and Figures 2–5). The differences between the observed and estimated means and medians mostly do not exceed 3% of the parameters, indicating a close-to-perfect fit. For the Lognormal model, the differences between the empirical and theoretical distributions are higher. In some cases, they are not acceptable for further analysis, which is connected with the number of parameters and the shape of the density curve. The Lognormal distribution has light tails and only two parameters, making it generally less flexible and not well suited to high-income groups, in particular. In Figures 2–5, the heavy-tailed three-parameter Dagum model behaves better for higher income groups, and the difference between genders is more visible.

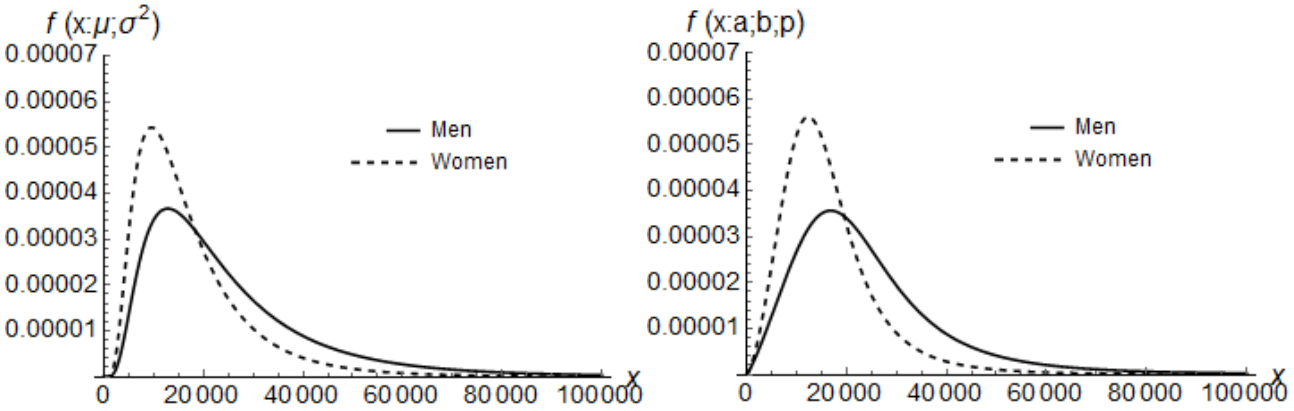

Figure 2. Estimated models: Lognormal (left) and Dagum (right) for men and women in the Czech Republic

Source: own elaboration based on LIS individual monthly disposable household income data: LIS Cross-National Data Center in Luxembourg n.d.

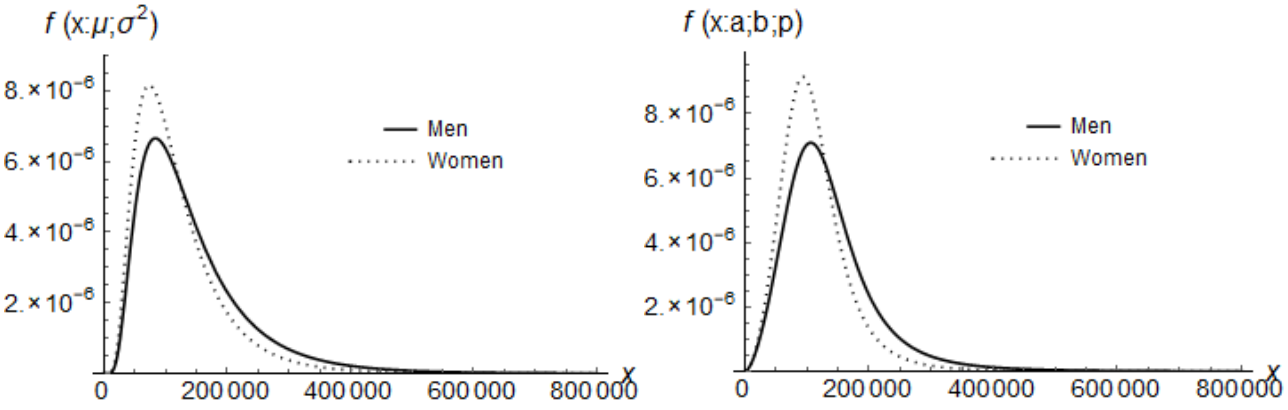

Figure 3. Estimated models: Lognormal (left) and Dagum (right) for men and women in Hungary

Source: own elaboration based on LIS individual monthly disposable household income data: LIS Cross-National Data Center in Luxembourg n.d.

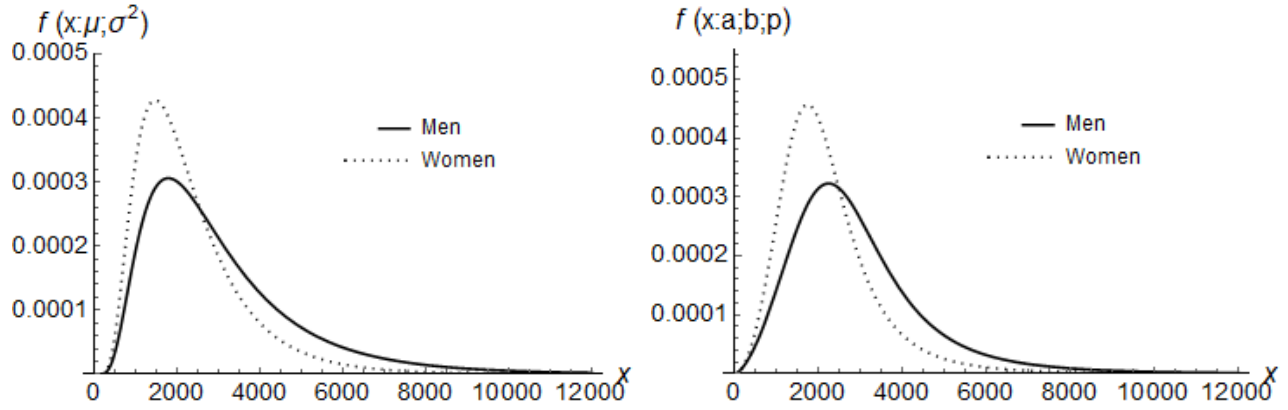

Figure 4. Estimated models: Lognormal (left) and Dagum (right) for men and women in Poland

Source: own elaboration based on LIS individual monthly disposable household income data: LIS Cross-National Data Center in Luxembourg n.d.

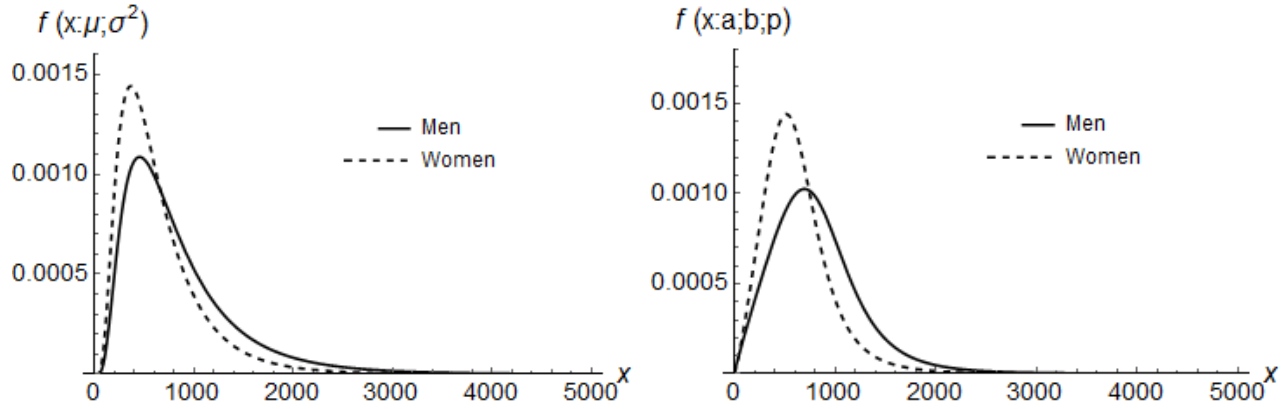

Figure 5. Estimated models: Lognormal (left) and Dagum (right) for men and women in Slovakia

Source: own elaboration based on LIS individual monthly disposable household income data: LIS Cross-National Data Center in Luxembourg n.d.

Figures 2–5, which show the theoretical income distributions for each country, show the same regularity: the income distributions of men are all shifted to the right and show a higher dispersion than for women. This suggests that, in all these countries, there is a large income gap between the genders. It is difficult to fully recognise this phenomenon based on a comparison of density curves (Figs. 2–5) as they are not scale-invariant. To that end, to measure and visualise the gap across the entire income range, we applied the relative distribution method, as described in section 3. The results are shown in Figures 6–9.

Figures 6–9 show the gender gaps using the relative distribution functions (9), estimated for the Dagum distribution as the underlying income distribution model (11). The gap is understood as the discrepancy between the curve and the diagonal line, which represents perfect equality between the groups. The values of the third, sixth and ninth income deciles, which are visible on the upper and right axes, are given for each country in its national currency or in euros (for Slovakia). The distances between the euro values on the right-hand scale are measured in units of persons rather than in euro (or national currency). Therefore, the distance between, e.g., 0 and 1000 euros is larger than the distance between 1000 and 2000 euros because a larger fraction of people have an income falling in the former range of incomes than in the latter. The curve of the relative income distribution provides rich and detailed information about the two income distributions being compared. Each point on the curve has a precise interpretation – in particular, the intersecting inner lines show the gap that corresponds to the third decile of the male distribution, which means the value below which 30% of the lowest earners are.

There are remarkable differences between the V4 countries concerning the gap between the income distributions of men and women. The gender gap is the most pronounced for the Czech Republic (Figure 6), where we can observe that at the third decile of the male distribution, i.e., $p = 0.3$, it holds $G(0.3) = 0.54$. This means that approximately 54% of women earn less than the third decile male's income. The result is even more striking for the median, i.e., $p = 0.5$, where this share equals 76%. The second biggest result, which is obviously unfavourable for women, was observed for Poland (Figure 8), where 50% of women earn less than the third decile male's income. This result is very close to Slovakia, where it was 49%. Hungary (Figure 7) had the smallest gender gap, as only 41% of women earned less than the "third decile man".

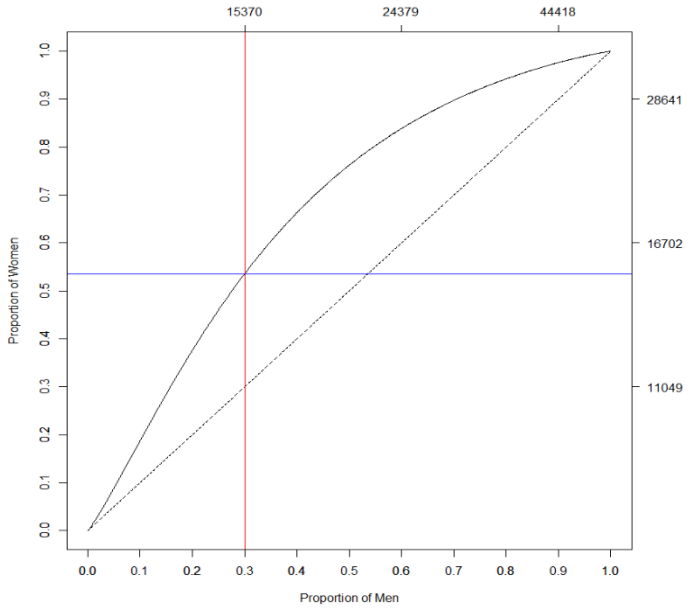

Figure 6. Gender gap between women (comparison) and men (reference) distributions in the Czech Republic

Source: own elaboration based on LIS data.

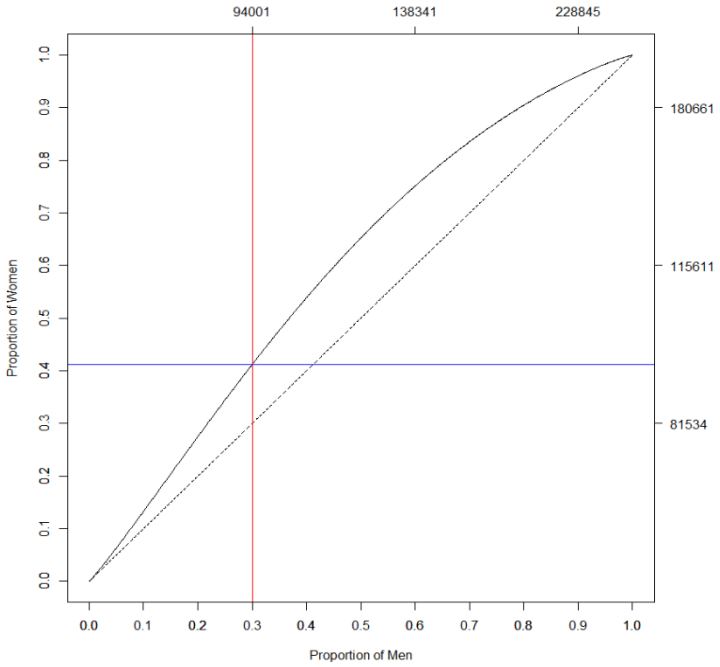

Figure 7. Gender gap between women (comparison) and men (reference) distributions in Hungary

Source: own elaboration based on LIS data.

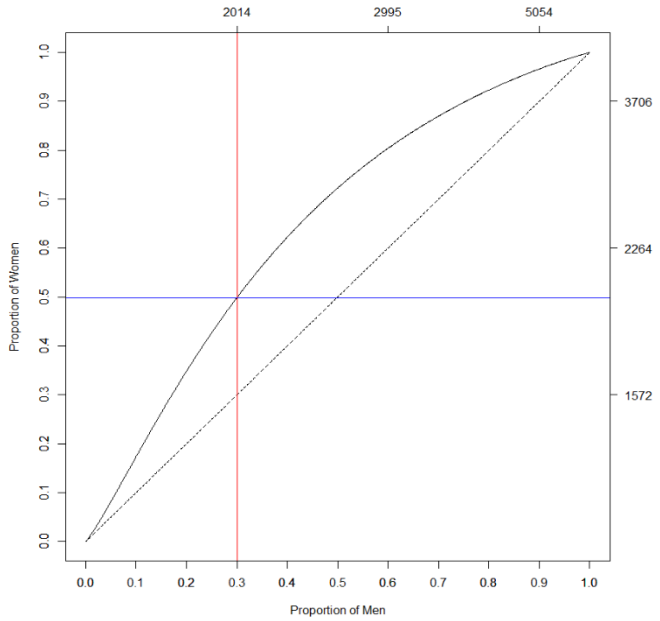

Figure 8. Gender gap between women (comparison) and men (reference) distributions in Poland

Source: own elaboration based on LIS data.

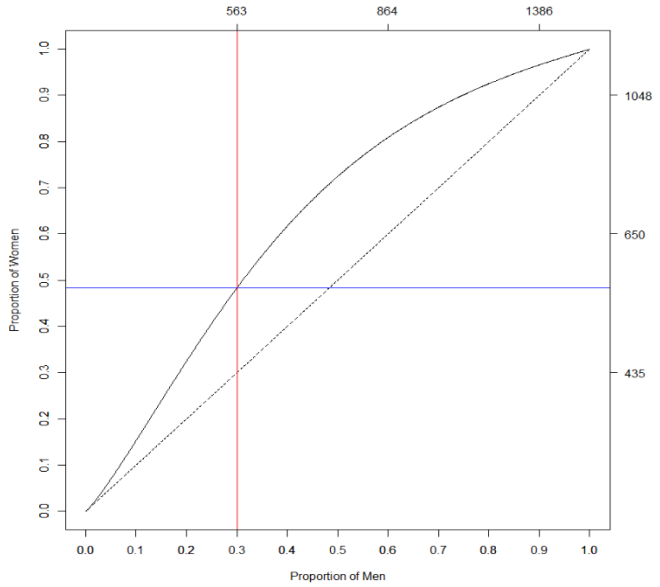

Figure 9. Gender gap between women (comparison) and men (reference) distributions in Slovakia

Source: own elaboration based on LIS data.

In all the countries except Poland, the gaps are asymmetric, increasing for the higher income group, while in Poland, the curve is symmetrical. It can be interpreted as the result of social programmes which smoothed out income inequality and made income distribution more symmetrical (see also Figs. 2–5). Moreover, for the Czech Republic and Hungary, income inequality measured by the Gini index was higher for men (see Tab. 2), suggesting that most of the discrepancies between the gender groups in these

countries came from extremely high male incomes. In Slovakia, by contrast, the gender gap could be a result of a "shift" in the distributions, connected with different means, rather than distributional inequality.

Conclusion

In this paper, we investigated the gender gaps in the Visegrád Group countries using the relative distribution method based on a parametric approach. The idea was to fully compare income distributions of men and women. In the parametric approach, we utilised the Dagum distribution, which is well-fitted to empirical data and performs better than its competitor, the Lognormal model. This approach allowed us to smooth out irregularities due to imperfections in the sampling method and to consider the full range of incomes, including high-income groups that were key to the analysis. The relative distribution methods applied in the study were appropriate to describe the gender gap for the entire income range and helped to detect and highlight important similarities and differences between the V4 countries.

- There are large discrepancies between income distributions of gender groups in all V4 countries – the economic advantage of men is visible not only in aggregate measures such as means, medians and the Gini ratios, but also across the entire income range.
- Compared to popular gender gap indices, the parametric approach based on the Dagum model made it possible to better describe the gender gaps in the V4 countries, especially at the tails.
- The relative distribution method has proven to be a useful tool for displaying and visualising gaps between income distributions.
- The parametric approach can also be helpful to smooth out the irregularities due to sample data.
- Within-group inequalities, measured by the inequality measures, were diverse across the countries, although they have a small impact on the discrepancy between gender groups.
- The size of the gap between the personal incomes of gender groups in each country does not reflect the overall level of household income inequality in these countries.

The next natural step of the analysis would be to evaluate the impact of the main drivers of income discrepancies between men and women using a decomposition approach.

References

Aitchison, J., Brown, J.A.C. (1957), *The Lognormal Distribution*, Cambridge University Press, Cambridge.

Avram, S., Popova, D. (2022), *Do taxes and transfers reduce gender income inequality? Evidence from eight European welfare states*, "Social Science Research", 102, 102644, https://doi.org/10.1016/j.ssresearch.2021.102644

Blau, F.D., Kahn, L.M. (2000), *Gender Differences in Pay*, "Journal of Economic Perspectives", 14 (4), pp. 75–99, https://doi.org/10.1257/jep.14.4.75

Blau, F.D., Kahn, L.M. (2003), *Understanding International Differences in the Gender Pay Gap*, "Journal of Labor Economics", 21 (1), pp. 106–144, https://doi.org/10.1086/344125

Boneva, T., Buser, T., Falk, A., Kosse, F. (2022), *The Origins of Gender Differences in Competitiveness and Earnings Expectations: Causal Evidence from a Mentoring Intervention*, "CEPR Discussion Paper", no. DP17008, https://doi.org/10.2139/ssrn.4009409

Cordova, K., Grabka, M.M., Sierminska, E. (2022), *Pension Wealth and the Gender Wealth Gap*, "European Journal of Population", 38 (4), pp. 755–810, https://doi.org/10.1007/s10680-022-09631-6

Cozarenco, A., Szafarz, A. (2023), *Financial inclusion in high-income countries: gender gap or poverty trap?*, [in:] V. Hartarska, R.J. Cull (eds.), *Handbook of Microfinance, Financial Inclusion and Development*, Edward Elgar Publishing, Cheltenham–Northampton, pp. 272–296, https://doi.org/10.4337/9781789903874.00024

Dagum, C. (1977), *A new model of personal income distribution: specification and estimation*, "Economie Appliquée", 30 (3), pp. 413–437.

Doorley, K., Claire, K. (2020), *Tax-Benefit Systems and the Gender Gap in Income*, "IZA Discussion Paper", No. 13786, https://doi.org/10.2139/ssrn.3713627

Eurostat (n.d.), *Structure of earnings survey*, https://ec.europa.eu/eurostat/web/microdata/structure-of-earnings-survey (accessed: 25.09.2023).

Glaubitz, R., Harnack-Eber, A., Wetter, M. (2022), *The Gender Gap in Lifetime Earnings: The Role of Parenthood*, "DIW Berlin Discussion Paper", No. 2001, http://doi.org/10.2139/ssrn.4071416 (accessed: 23.05.2023).

Greselin, F., Jędrzejczak, A. (2020), *Analyzing the Gender Gap in Poland and Italy, and by Regions*, "International Advances in Economic Research", 26, pp. 433–447, https://doi.org/10.1007/s11294-020-09810-3

Greselin, F., Jędrzejczak, A., Trzcińska, K. (2023), *A new parametric approach to Gender Gap with application to EUSILC data in Poland and Italy*, "Statistical Analysis and Data Mining", 16 (3), pp. 1–17, https://doi.org/10.1002/sam.11623

Handcock, M.S., Morris, M. (2006), *Relative distribution methods in the social sciences*, Springer Science & Business Media, New York.

Jędrzejczak, A. (2015), *Regional Income Inequalities in Poland and Italy*, "Comparative Economic Research. Central and Eastern Europe", 18 (4), pp. 27–45, https://doi.org/10.1515/cer-2015-0027

Jędrzejczak, A., Pekasiewicz, D. (2020a), *Changes in Income Distribution for Different Family Types in Poland*, "International Advances in Economic Research", 26 (2), pp. 135–146, https://doi.org/10.1007/s11294-020-09785-1

Jędrzejczak, A., Pekasiewicz, D. (2020b), *Teoretyczne rozkłady dochodów gospodarstw domowych i ich estymacja*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Jędrzejczak, A., Pekasiewicz, D., Zieliński, W. (2021), *Confidence interval for quantile ratio of the Dagum distribution*, "REVSTAT Statistical Journal", 19 (1), pp. 87–97, https://www.ine.pt/revstat/tables.html (accessed: 25.09.2023).

Kleiber, C. (2008), *A Guide to the Dagum Distributions*, [in:] D. Chotikapanich (ed.), *Modeling Income Distributions and Lorenz Curves*, Springer, Berlin, pp. 97–117, https://doi.org/10.1007/978-0-387-72796-7_6

Kleiber, C., Kotz, S. (2003), *Statistical Size Distributions in Economics and Actuarial Sciences*, John Wiley & Sons, Hoboken, https://doi.org/10.1002/0471457175

Leythienne, D., Ronkowski, P. (2018), *A decomposition of the unadjusted gender pay gap using structure of earnings survey data*, Publications Office of the European Union, Luxembourg.

LIS Cross-National Data Center in Luxembourg (n.d.), *LIS Database*, https://www.lisdatacenter.org/our-data/lis-database/ (accessed: 25.10.2022).

Luxembourg Income Study (LIS) Database (n.d.), http://www.lisdatacenter.org (accessed: 25.10.2022).

OECD (2011), *Divided We Stand. Why Inequality Keeps Rising*, OECD Publishing, Paris, https://read.oecd-ilibrary.org/social-issues-migration-health/the-causes-of-growing-inequalities-in-oecd-countries_9789264119536-en#page1 (accessed: 25.07.2023).

OECD (2015), *In It Together: Why Less Inequality Benefits All*, OECD Publishing, Paris, https://read.oecd-ilibrary.org/employment/in-it-together-why-less-inequality-benefits-all_9789264235120-en#page1 (accessed: 25.07.2023).

Szczepaniak, M., Szulc-Obłoza, A. (2020), *Labour Market Institutions and Income Inequalities in the Visegrad Group Countries*, "Comparative Economic Research. Central and Eastern Europe", 23 (3), pp. 75–90, https://doi.org/10.18778/1508-2008.23.21

TARKI (2009), *European Social Report 2009*, TÁRKI Inc., Budapest.

World Bank (2016), *Poverty and Shared Inequality 2016. Taking on Inequality*, https://doi.org/10.1596/978-1-4648-0958-3

World Bank (2020), *Making Transition Work for Everyone: Poverty and Inequality in Europe and Central Asia*, World Bank, Washington.

Zaidi, S. (2009). *Main drivers of income inequality in central European and Baltic countries-some insights from recent household survey data*, World Bank Policy Research Working Paper, (4815), https://doi.org/10.1596/1813-9450-4815

Zenga, M., Jędrzejczak, A. (2020), *Decomposition of the Zenga inequality index I(Y) into the contributions of macro-regions and income components – an application to data from Poland and Italy*, "Argumenta Oeconomica", 1 (44), pp. 101–125, https://doi.org/10.15611/aoe.2020.1.05

Luka dochodowa w krajach Grupy Wyszehradzkiej na podstawie danych pochodzących z Luxembourg Income Study

Równość płci jest jedną z podstawowych wartości Unii Europejskiej (UE). Wiele wysiłków włożono w obronę tego prawa i promowanie równości płci w państwach członkowskich i na całym świecie, jednak nadal obserwuje się znaczne różnice między mężczyznami i kobietami, które dotyczą między innymi dochodów. Wciąż dyskutowana jest także kwestia badawcza dotycząca metodologii pomiaru zjawiska tzw. luki dochodowej – tradycyjne metody oparte na porównywaniu średnich i median wydają się niezadowalające, ponieważ nie uwzględniają całego rozkładu dochodów. Celem artykułu jest analiza rozkładów dochodów obserwowanych w czterech krajach Europy Środkowej: Polsce, Słowacji, Czechach i na Węgrzech, należących do Grupy Wyszehradzkiej (V4). Kraje te mają podobną historię, podobny rozwój gospodarczy, ale istniały i wciąż istnieją znaczne różnice między ich podejściami do reform gospodarczych, w tym do polityki rynku pracy, co znajduje odzwierciedlenie w odmiennych rozkładach dochodów i ich nierównomierności. W artykule proponujemy parametryczne podejście oparte na rozkładzie relatywnym, które umożliwia porównanie i wizualizację „luki" między grupami płci dla każdego kwantyla rozkładu. To nowe podejście wykracza więc poza typowe analizy oparte na średnich lub medianach, w kierunku porównania całego rozkładu dochodów kobiet z rozkładem dochodów mężczyzn. Podstawą obliczeń były mikrodane pochodzące z LIS (*Luxembourg Income Study*). Zastosowane w badaniu metody statystyczne okazały się odpowiednie do opisu luki między mężczyznami a kobietami w całym przedziale dochodów. Wyniki analizy empirycznej pozwoliły ujawnić zarówno podobieństwa, jak i istotne różnice między krajami.

Słowa kluczowe: nierówności dochodowe, luka płci, rozkład Daguma, metoda rozkładu względnego, Grupa Wyszehradzka