ORIGINAL ARTICLE

Citation: Landmesser, J. (2019). Differences in income distributions for men and women in the European Union countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 14(1), 81–98. doi: 10.24136/eq.2019.004

Contact: joanna_landmesser@sggw.pl; Warsaw University of Life Sciences - SGGW, ul. Nowoursynowska 166, 02-787 Warszawa, Poland

Received: 1.12.2018; Revised: 11.01.2019; Accepted: 27.01.2019; Published online: 20.02.2019

Joanna Landmesser

Warsaw University of Life Sciences - SGGW, Poland

orcid.org/0000-0001-7286-8536

Differences in income distributions for men and women in the European Union countries

JEL Classification: J31; D31

Keywords: *gender pay gap; differences in distributions; decomposition methods*

Abstract

Research background: Recently there has been an increase in interest in the studies of income inequalities. The findings of numerous empirical studies show that males earn higher wages than females. A variety of techniques of income inequalities decomposition are becoming popular. New procedures go far beyond the Oaxaca-Blinder decomposition. They allow to study differences of income distributions for various groups of people and to decompose them at various quantile points.

Purpose of the article: The aim of the paper is to compare personal income distributions in selected countries of the European Union, taking into account gender differences.

Methods: First, we examined the income inequalities between men and women in each country using the Oaxaca-Blinder decomposition procedure. The unexplained part of the gender pay gap gave us information about the wage discrimination. Second, we extended the decomposition procedure to different quantile points along the whole income distribution. To describe differences between the incomes of men and women, we constructed the so-called counterfactual distribution, which is a mixture of a conditional distribution of the dependent variable (income) and a distribution of the explanatory variables (individual people's characteristics). Then, we utilized the residual imputation approach (JMP-approach).

Findings & Value added: In the article data from EU-SILC (Statistics on Income and Living Conditions) were used. We found that there exists an important diversity in the size of the gender pay gap across members of the European Union. The results obtained for these countries allowed us to group them into clusters. In general, there are two types of countries in Europe: the countries, where the bulk of the observed income differences cannot be explained by observed characteristics, and the countries, where the explained and the unexplained effects are both positive, with even a bigger explained effect for the lower income ranges.

Introduction

Reducing the gender pay gap is one of the key priorities of gender policies. At EU level, the European Commission prioritised "reducing the gender pay, earnings and pension gaps and thus fighting poverty among women" as one of the key areas in the framework of the Strategic engagement for gender equality 2016–2019.

The gender pay gap has been a subject of interest for a long time. The findings of numerous empirical studies show that males earn higher wages than females. There are considerable differences in earnings between EU countries, with the gender pay gap ranging from just over 5% in Romania, Italy and Luxembourg, to more than 25% in Estonia, followed by the Czech Republic and Germany (both almost 22%) (see Figure 1). In 2016 women's gross hourly earnings were on average 16.2% below those of men in the European Union (EU–28) and 16.3% in the euro area (Eurostat, 2018).

The purpose of this study is to compare personal income distributions in selected countries of the European Union taking into account gender differences. We will examine the differences in the entire range of income values by the use of the residual imputation approach (JMP-approach) (see Juhn *et al.*, 1993, pp. 410–442). In the article, data from EU-SILC will be used.

The paper is organized as follows. In the next section we provide a literature overview and show what are the results of the prior studies. In the research methodology section, we outline the econometric methods used in the analysis. In the results section we describe the empirical data set and present the obtained results. In the subsequent section, we discuss the findings and in the last section we conclude.

Literature review

The existence of a significant gap in earnings between males and females is a well-documented stylized fact of modern labour markets. There is a huge amount of literature on the topic of gender gap. Numerous studies concentrate mainly on the decomposition of the average values for incomes. For example, Jurajda (2003, pp. 199–222) analyses in this way the gender pay gap in the Czech Republic and Slovakia, Pena-Boquete *et al.* (2010, pp. 109–137) in Italy and Spain, Chatterji *et al.* (2011, pp. 3819–3833) examine the gap in Britain, Śliwicki and Ryczkowski (2014, pp. 159–173) in Poland.

A number of papers adopt a cross-country perspective. Using European Union Statistics on Income and Living Conditions (EU-SILC) data for 24

European Union members, Hedija (2017, pp. 1804–1819) showed that the gender pay gap varies among the countries. Also Oczki (2016, pp. 106–113) presented similar results for 28 EU countries. In the study conducted by Boll and Lagemann (2018) the gap is analyzed based on the Structure of Earnings Survey (EU-SES).

Sometimes, the analysis goes beyond the simple comparison of average values (e.g., for Poland Landmesser *et al.*, 2015, pp. 43–52; Landmesser, 2016, pp. 331–348). The results of numerous studies show evidence of the glass ceiling and the sticky floor effects. The glass ceiling effect means a greater wage gap at the top end of the income distribution range and the sticky floor effect means a wider wage gap at the bottom. The glass ceiling is a metaphor referring to an artificial barrier that prevents women from being promoted to managerial-level positions within an organization. The term sticky floor is used to point to a discriminatory employment pattern that keeps workers in the lower ranks of the job scale. It refers to low-paying, low-prestige, and low-mobility jobs typically held by women (e.g., clerical or service jobs). Arulampalam *et al.* (2006, pp. 163–186) examined the gender wage gap in 11 European countries using the European Community Household Panel Survey (ECHPS). The gap widened towards the top of the wage distribution in most of countries and, in a few cases, it also widened at the bottom of the distribution. Nicodemo (2009) analyzed the gap in five Mediterranean EU countries (France, Greece, Italy, Portugal and Spain), using the EU-SILC and the ECHPS datasets. She found a positive wage gap in all countries, the greater part of which cannot be explained by observed characteristics. The gender gap was larger at the bottom and smaller at the top of the distribution in most countries. Also Christofides *et al.* (2013, pp. 86–102) used EU-SILC data and estimate the unexplained part of the gender pay gap for 26 European countries.

Despite many differences among the individual studies, they all conclude that the gender pay gap exhibits a remarkable heterogeneity across European countries. The gender pay gap is the result of many factors, including occupational segregation, bias against working mothers, and direct pay discrimination. A certain part of the wage differences between men and women remains unexplained. The results of these studies depend on the used data set, the number of explanatory variables and the applied method of decomposition. An important diversity exists both in the size of the gender pay gap and in its underlying causes across transitional new members of the European Union.

In this paper, we do not compare the average income values, but examine the differences at various quantile points along the distribution using data for 19 EU countries from the EU-SILC project in 2014.

Research methodology

Our hypothesis is that there exists an important diversity in size and shape of the gender pay gap across members of the European Union. The observed differences should be analyzed in an adequate manner. In recent times, a variety of techniques of income inequalities decomposition are becoming more popular. The standard method for pay gap decomposition is the procedure proposed by Oaxaca (1973, pp. 693–709) and Blinder (1973, pp. 436–455), which is widely used to study mean outcome differences between groups. This statistical method allows to take into account the individual characteristics of male and female employees. It decomposes the gap in mean outcomes across two groups into a part that is due to group differences in the levels of explanatory variables and a part that is due to differential magnitudes of regression coefficients.

Let Y_g denote the outcome variable in group g (e.g. the personal income in men's group, $g=M$, or in women's group, $g=W$) and X_g the vector of individual characteristics of the person in group g (such as education, age, work experience). The expected value of y conditionally on X is a linear function $y_g = X_g \beta_g + v_g$, $g=M, W$, where β_g are the returns to the characteristics. The Oaxaca-Blinder decomposition for the average income inequality between two groups at the aggregate level can be expressed as

$$\bar{Y}_M - \bar{Y}_W = \bar{X}_M \hat{\beta}_M - \bar{X}_W \hat{\beta}_W = \underbrace{\bar{X}_M (\hat{\beta}_M - \hat{\beta}_W)}_{\text{unexplained effect}} + \underbrace{(\bar{X}_M - \bar{X}_W) \hat{\beta}_W}_{\text{explained effect}}. \quad (1)$$

The first component, on the right side of the equation, called the unexplained effect (or the wage structure effect), is the result of differences in the returns to observables (the estimated coefficients), and so in the “prices” of individual characteristics of group representatives. The unexplained portion of the gap is usually attributed to discrimination, but may also results from the influence of unobserved variables. The second term, called the explained effect, gives the effect of characteristics and is explained by group differences in the predictors (the so-called explained effect). A drawback of the approach is that it focuses only on average effects, which may lead to a misleading assessment if the effects of covariates vary across the income distribution.

Several authors have developed extensions of the Oaxaca-Blinder method (e.g. DiNardo *et al.*, 1996, pp. 1001–1044; Donald *et al.*, 2000, pp. 609–633; Machado & Mata, 2005, pp. 445–465; Fortin *et al.*, 2010). The proposed procedures go beyond the comparison of average values.

Let $F_{Y_g}(y)$ be the distribution function for the dependent variable Y_g , which can be expressed as the conditional distribution $F_{Y_g|X_g}(y|X)$ of Y_g and the joint distribution $F_{X_g}(X)$ of all elements of X_g (the explanatory variables):

$$F_{Y_g}(y) = \int F_{Y_g|X_g}(y|X) \cdot F_{X_g}(X) dx, \quad g = M, W. \quad (2)$$

Now, the mean decomposition analysis can be extended to the case of differences between the two distributions using the so-called counterfactual distribution:

$$F_{Y_W^c}(y) = \int F_{Y_W|X_W}(y|X) \cdot dF_{X_M}(X) \quad (3)$$

Such a mixture distribution is constructed by integrating the conditional income distribution for men with respect to the distribution of characteristics for women (the distribution of incomes that would prevail for people in group W if they had the distribution of characteristics of group M). Then the difference in the observed income distributions between men and women can be decomposed in the spirit of Oaxaca and Blinder as follows:

$$F_{Y_M}(y) - F_{Y_W}(y) = \underbrace{[F_{Y_M}(y) - F_{Y_W^c}(y)]}_{\text{unexplained effect}} + \underbrace{[F_{Y_W^c}(y) - F_{Y_W}(y)]}_{\text{explained effect}}. \quad (4)$$

The counterfactual distribution can be constructed in various ways, e.g. using the residual imputation approach (see Juhn *et al.*, 1993, pp. 410–442). In this method, one has to estimate the two equations: $y_{Wi} = X_{Wi}\beta_W + v_{Wi}$ and $y_{Mi} = X_{Mi}\beta_M + v_{Mi}$, $i = 1, \dots, n$. Then, the income y_M from the group M is replaced by a counterfactual income y_W^c , where both the returns to observables and residuals are set to be as in group W .

The implementation of the procedure is two-step. In the first step, the residuals are replaced by counterfactual residuals under the assumption of the rank preservation: $y_{Wi}^{C,1} = X_{Mi}\beta_M + v_{Wi}^{C,1}$, $i = 1, \dots, n$, where $v_{Wi}^{C,1} = F_{v_W|X}^{-1}(\tau_{Mi}(X_{Mi}), X_{Mi})$ and $\tau_{Mi}(X_{Mi})$ is the conditional rank of v_{Mi} in the distribution of residuals for M . In the second step, the counterfactual returns to observables are imputed as $y_{Wi}^{C,2} = X_{Mi}\beta_W + v_{Wi}^{C,1}$, $i = 1, \dots, n$.

In this study, after an assessment of the gender pay gap for selected EU countries, an attempt was made to group them using hierarchical clustering method. Hierarchical clustering is a widely used tool in data mining for grouping data into clusters, which exposes similarities or dissimilarities in the data. One set of approaches to hierarchical clustering is known as agglomerative, whereby in each step of the clustering process an observation or cluster is merged into another cluster. In this work, the 1-nearest neighbors algorithm for agglomerative clustering with Euclidean distance was applied.

Results

The empirical data used have been collected within the European Union Statistics on Income and Living Conditions project in 2014 (research proposal 234/2016-EU-SILC). For the analysis, we chose 19 EU countries. The main criterion for their selection was the desire to obtain the possibly diversified set of countries in terms of their size, level of economic development and the degree of discrimination on the labor market. The selected countries can be characterized with the following fundamental statistics: the highest number of inhabitants (Germany, France, United Kingdom, Italy, Spain, Poland), the highest GDP per capita in 2015 (Luxembourg, Ireland, the Netherlands, Austria, Germany), the lowest GDP per capita in 2015 (Bulgaria, Romania, Croatia, Hungary, Poland, Greece), the highest gender pay gap according to Eurostat in 2016 (Estonia, the Czech Republic, Germany, United Kingdom, Austria, Slovakia), and the lowest gender pay gap according to Eurostat in 2016 (Romania, Italy, Luxembourg, Belgium, Poland). The selected sample consisted of 122,756 observations: 62,856 men and 59,900 women.

The annual gross employee (cash or near cash) incomes (in thousands of Euro) of men were compared with those obtained by women. The gross employee income corresponds mainly to wages and salaries paid for the time worked, remuneration for the time not worked, enhanced rates of pay for overtime, payments for fostering children, supplementary payments (e.g. thirteenth month payment). It includes any social contributions and income taxes payable to social insurance schemes or tax authorities. In our empirical decomposition analysis, a logarithm of the annual income ($\log_income$) constitutes the outcome variable. The sample size and average income in selected countries are presented in Table 1.

Each person was described by the following characteristics: *age* (in years), *educlevel* (education level, 1 – primary, . . . , 5 – tertiary), *married*

(marital status, 1 – married, 0 – unmarried), *yearswork* (number of years spent in paid work), *permanent* (type of contract, 1 – permanent job/work contract of unlimited duration, 0 – temporary contract of limited duration), *parttime* (1 – person working part-time, 0 – person working full-time), *manager* (managerial position, 1 – supervisory, 0 – non-supervisory), *big* (number of persons working at the local unit, 1 – more than 10 persons, 0 – less than 11 persons).

Table 2 presents the results of the aggregate Oaxaca-Blinder decomposition of inequalities between men's and women's log incomes for 19 EU countries.

We have found that there is a positive difference between the mean values of log incomes for men and women for all 19 countries. The mean log income differential is the largest in Germany (0.625) and the smallest in Romania (0.141). The country heterogeneity is not limited to the size of the gap, but also concerns its composition. The difference between the mean log income values was decomposed into two components: the first one explaining the contribution of the different values of models coefficients (the unexplained part), and the second one explaining the contribution of the attributes differences (the explained part). The unexplained effect is huge (and positive) for the states with the low raw differential and is small for the states with the high raw differential. It ranges from 21% in Luxembourg to 125% in Bulgaria. This part of the gender pay gap gives us information about the discrimination. The explained gap is negative in six countries (among others in Bulgaria, Poland, Hungary), which have the lowest income discrepancies. The negative value of this component means that the difference of the average log incomes between men and women is reduced by the women's characteristics. In 13 countries, the explained part is positive, that is, it increases the overall gap, with a maximum explained gap in Luxembourg (79%). Only in 8 countries the explained part exceeds the unexplained part of the overall gap. However, the unexplained part is nowhere identified to be negative.

The detailed decomposition, which was also carried out, made it possible to isolate the factors explaining the inequality observed to a different extent. Because of lack of space in this paper, we present the results of the detailed decomposition only for 4 countries — Germany, the Czech Republic, Poland and Romania (see Table 3).

The strong effect of different education levels of men and women can be noticed, especially for Poland. The negative values of explained components mean that the differences of the average log incomes between men and women are reduced by the women's higher education levels. On the other hand, the values of *yearswork*, *parttime*, *manager* and *big* attributes

possessed by men and women increase the income inequality in all countries (see the positive explained component values). In all states women are discriminated against men because of their marital status (the positive unexplained components values for variable *married*) and managerial position (if possessed) but not because of the education levels or years of work.

Since the Oaxaca-Blinder technique focuses only on average effects, we carried out the decomposition of inequalities along the distribution of log incomes for men and women using the residual imputation approach (JMP-approach). The total differences between the values of log incomes are computed and the results are shown in Table 4. They are expressed in terms of percentiles (the symbols $p5$, ..., $p95$ stand for 5th, ..., 95th percentile; e.g. the 25th percentile is the log income value below which 25% of the observations may be found).

For each country, there are positive differences between the values of log incomes for men and women along the whole log income distribution. Then the calculated differences were decomposed into the sum of the unexplained and explained components (the results are presented in Figures 3, 4, 5). After assessing the gender pay gap (the raw, the explained and the unexplained gap) for all 19 countries, an attempt was made to group them using hierarchical clustering method. The use of the 1-nearest neighbors algorithm for agglomerative clustering with Euclidean distance allowed the grouping of countries into four clusters, as shown in Figure 2.

Four groups were identified:

- Group 1: the Czech Republic, Romania, Slovakia, Bulgaria, Hungary, Poland and Croatia,
- Group 2: France, the United Kingdom, the Netherlands, Austria, Luxembourg, Spain, Italy, Ireland, Belgium, Greece,
- Group 3: Estonia,
- Group 4: Germany.

The shapes of pay gap are examined in Figures 3–5, where solid lines represent the total income gap, the dashed lines denote the unexplained component and the dotted lines indicate the explained effect.

Group 1 consists mainly of the former socialist states of eastern Europe. It is characterized by the U-shaped total gender pay gap (except Romania and Bulgaria) and bigger unexplained effect than the explained one (see Figure 3). For most countries in this group, the total effect is low, but it widens at the bottom and at the top of the income distribution, suggesting sticky floor and glass ceiling effects. The share of the unexplained part is very high. The effect of coefficients is positive and constant in the whole range of the income distribution. This is the result of differences in the

“market prices” of individual characteristics of men and women, interpreted as the labor market discrimination.

For 5 countries (Romania, Bulgaria, Hungary, Poland and Croatia), the explained differential (the effect of characteristics) is negative in the middle of the distribution, which means that the properties possessed by both people’s groups decrease the inequalities. Except for Romania, the effect of characteristics is often positive at the bottom and at the top of the income distribution.

Group 2, the largest group, consists mainly of the large, highly developed countries of western Europe. In most from these countries, the total gender gap is larger at the bottom of the distribution and smaller at the top of the distribution (see Figure 4). The gender differences in characteristics are positive, which means that the different values of characteristics of men and women increase the income inequalities. The explained effect is bigger than the unexplained effect at the bottom of the log income distribution. For the higher income ranges, the unexplained effect often prevails. Both effects, the explained and the unexplained, are always positive, increasing the income discrepancies.

Group 3 consists only one country — Estonia, which is characterized by an increase of the income inequalities as we move toward the top of the income distribution (Figure 5 (a)). We notice that the glass ceiling effect is present. Moreover, the share of the unexplained effect is very high. We observe the negative explained effect at the bottom of the distribution. It means that the income inequalities decrease for the poorer.

The last group, the group 4, is made up of Germany. In this case, the large total gap and the large explained effect have a decreasing shape and are rapidly falling as we move toward the top of the income distribution (Figure 5 (b)). The unexplained part is positive and at a moderate level, presenting the existing effect of discrimination on the labor market.

Discussion

We started our analysis using the Oaxaca-Blinder method for the decomposition of the average values for log incomes. We found that there is a positive difference between the mean income values for men and women. Similar result has been documented in the previous research (e.g., compare Boll & Lagemann, 2018). The mean log income differential was the largest in Germany and the smallest in Romania. The unexplained effect, which gives us information about the discrimination, was huge for the states with the low raw differential and was small for the states with the high raw differen-

tial. The explained gap was negative in the countries with the lowest income discrepancies (e.g., in Bulgaria, Poland, Hungary). The negative value of this component means that the difference of the average log incomes between men and women is reduced by the women's characteristics.

Then, we examined the differences in the entire range of income values by the use of the JMP-residual imputation approach. We extended the decomposition procedure to different quantile points along the whole income distribution. After assessing the raw, the explained and the unexplained gaps for all 19 countries, we grouped them into four clusters using the 1–nearest neighbors algorithm for agglomerative clustering.

Group 1 consisted mainly of the former socialist states of eastern Europe (the Czech Republic, Romania, Slovakia, Bulgaria, Hungary, Poland and Croatia). The results obtained for that group indicated the low and U-shaped total effect (compare Arulampalam *et al.*, 2006, 163–186). The conducted decomposition showed that the unexplained component quantitatively dominated in the whole range of the income distribution. Our analysis confirmed that the gender pay gap was poorly explained by gender differences in observable characteristics of people. But for several countries (Romania, Bulgaria, Hungary, Poland and Croatia), the explained differential was negative in the middle of the distribution, which means that female characteristics are superior to the male ones. Similar results in this field have been obtained in Christofides *et al.* (2013, pp. 86–102).

In group 2, consisting of highly developed countries of western Europe (France, the United Kingdom, the Netherlands, Austria, Luxembourg, Spain, Italy, Ireland, Belgium, Greece), the total gender gap was larger at the bottom of the log income distribution and smaller at the top. Those findings are similar to the results obtained by Nicodemo (2009). For the lower income ranges, the explained effect was bigger than the unexplained. Both effects, the explained and the unexplained, are always positive, increasing the income discrepancies.

Group 3, made up of Estonia, was characterized by an increase of the income inequalities at the top of the income distribution (like in Christofides *et al.* (2013, pp. 86–102)). There was a very high share of the unexplained effect and the negative explained effect for the poorer.

For group 4 (Germany), we noted the large total gap and the large explained effect rapidly was falling as we moved toward the top of the income distribution (this result, however, differs from the finds in Christofides *et al.* (2013, pp. 86–102)). The unexplained part is still positive but at a moderate stable level.

Conclusions

The goal of this paper was to compare personal income distributions in selected countries of the European Union taking into account gender differences. Using data from the EU-SILC project in 2014, the gender pay gap was examined for the set of 19 European countries.

Summarizing, there exists an important diversity in the size of the gender pay gap across members of the European Union. Excluding the extreme cases of Estonia and Germany, there are two types of countries in Europe:

- the countries, where the bulk of the observed income differences cannot be explained by observed characteristics (the states with higher level of gender discrimination on the labor market),
- the countries, where the explained and the unexplained effects are both positive, with even a bigger explained effect for the lower income ranges (the states with lower level of gender discrimination on the labor market).

One should be aware that the results obtained depend on the selection of explanatory variables to the estimated models. Future work should cover all EU countries and might consider how gender differences have changed over time. The gender discrimination may lead to loss in productivity and wealth. Inequalities induced in this way pose a serious challenge for the society. Therefore, it would be very interesting to analyze whether changes in policies can affect gender differences.

References

- Arulampalam, W., Booth, A. L., & Bryan, M. L. (2006). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review*, 62(2). doi: 10.1177/001979390706000201.
- Blinder, A. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, 8. doi: 10.2307/144855.
- Boll, Ch., & Lagemann, A. (2018). Gender pay gap in EU countries based on SES (2014). European Commission. doi: 10.2838/978935.
- Chatterji, M., Mumford, K., & Smith, P. N. (2011). The public-private sector gender wage differential in Britain: evidence from matched employee-workplace data. *Applied Economics*, 43(26). doi: 10.1080/00036841003724452.
- Christofides, L., Polycarpou, A., & Vrachimis, K. (2013). Gender wage gaps, 'sticky floors' and 'glass ceilings' in Europe. *Labour Economics*, 21. doi: 10.1016/j.labeco.2013.01.003.
- DiNardo, J., Fortin, N. M., & Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973–1992: a semiparametric approach. *Econometrica*, 64. doi: 10.2307/2171954.

- Donald, S. G., Green, D. A., & Paarsch, H. J. (2000). Differences in wage distributions between Canada and the United States: an application of a flexible estimator of distribution functions in the presence of covariates. *Review of Economic Studies*, 67. doi: 10.1111/1467-937X.00147.
- Eurostat (2018). The unadjusted gender pay gap, 2016. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php/Gender_pay_gap_statistics (10.01.2019).
- Fortin, N., Lemieux, T., & Firpo, S. (2010). Decomposition methods in economics. *NBER Working Paper*, 16045. doi: 10.3386/w16045.
- Hedija, V. (2017). Sector-specific gender pay gap: evidence from the European Union countries. *Economic Research-Ekonomska Istraživanja*, 30(1). doi: 10.1080/1331677X.2017.1392886.
- Juhn, Ch., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101. doi: 10.1086/261881.
- Jurajda, Š. (2003). Gender wage gap and segregation in enterprises and the public sector in late transition countries. *Journal of Comparative Economics*, 31(2). doi: 10.1016/S0147-5967(03)00040-4.
- Landmesser, J. M., Karpio, K., & Łukasiewicz, P. (2015). Decomposition of differences between personal incomes distributions in Poland. *Quantitative Methods in Economics*, 16(2).
- Landmesser, J. (2016). Decomposition of differences in income distributions using quantile regression. *Statistics in Transition - new series*, 17(2). doi: 10.21307/stattrans-2016-023.
- Machado, J. F., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20. doi: 10.1002/jae.788.
- Nicodemo, C. (2009). Gender pay gap and quantile regression in European families. *IZA Discussion Paper*, 3978.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14. doi: 10.2307/2525981.
- Oczki, J. (2016). Gender pay gap in Poland. *Ekonomia międzynarodowa*, 14. doi: 10.18778/2082-4440.14.03.
- Pena-Boquete, Y., De Stefanis, S., & Fernandez-Grela, M. (2010). The distribution of gender wage discrimination in Italy and Spain: a comparison using the ECHP. *International Journal of Manpower*, 31(2). doi: 10.1108/01437721011042232.
- Śliwicki, D., & Ryczkowski, M. (2014). Gender pay gap in the micro level – case of Poland. *Quantitative Methods in Economics*, 15(1).

Annex

Table 1. Sample size and average annual income (in Euro) in selected countries

Country	<i>n</i>	<i>n</i> men	<i>n</i> women	Average Y_M	Average Y_W
LU	3932	2110	1822	56353.29	43849.64
NL	4912	2433	2479	49364.79	30024.32
IE	3759	1773	1986	44876.80	31165.20
AT	4798	2547	2251	44253.14	28841.96
BE	4677	2334	2343	43554.01	32171.61
DE	10128	4999	5129	42368.01	24741.80
UK	8179	3965	4214	39785.26	25168.27
FR	9251	4589	4662	33494.49	24012.36
IT	12715	6741	5974	30007.71	22655.93
ES	8493	4378	4115	25225.28	18640.89
EL	3687	2059	1628	17658.02	13934.03
EE	5506	2663	2843	13064.08	8899.18
CZ	6501	3443	3058	12413.62	8865.36
HR	3601	1983	1618	10350.23	8871.80
SK	5755	2847	2908	10292.61	7980.63
PL	9908	5180	4728	9619.24	7947.84
HU	8054	4061	3993	7565.97	6208.86
BG	4058	2018	2040	4646.59	3747.83
RO	4842	2733	2109	4306.87	3768.30
Total	122756	62856	59900		

Table 2. The results of the aggregate Oaxaca-Blinder decomposition of gender pay gap in selected countries

Country	Raw differential	Unexplained effect	Explained effect	% unexplained	% explained
DE	0.625	0.186	0.439	29.8%	70.2%
NL	0.520	0.167	0.354	32.1%	67.9%
AT	0.485	0.109	0.376	22.5%	77.5%
UK	0.479	0.250	0.230	52.1%	47.9%
FR	0.376	0.198	0.179	52.5%	47.5%
ES	0.361	0.149	0.212	41.3%	58.7%
IE	0.353	0.144	0.209	40.8%	59.2%
CZ	0.343	0.248	0.095	72.3%	27.7%
BE	0.329	0.135	0.195	40.9%	59.1%
EE	0.328	0.372	-0.044	113.4%	-13.4%
LU	0.296	0.062	0.235	20.9%	79.1%
IT	0.287	0.133	0.155	46.2%	53.8%
SK	0.265	0.247	0.019	92.9%	7.1%
EL	0.206	0.106	0.099	51.7%	48.3%
PL	0.183	0.209	-0.026	114.2%	-14.2%
BG	0.178	0.221	-0.044	124.9%	-24.9%
HU	0.176	0.200	-0.024	113.6%	-13.6%
HR	0.167	0.175	-0.008	104.8%	-4.8%
RO	0.141	0.155	-0.013	109.2%	-9.2%

Source: own elaboration using the Stata command 'decompose'.

Table 3. The results of the detailed Oaxaca-Blinder decomposition

Variable	Detailed decomposition							
	Unexplained effect				Explained effect			
	DE	CZ	PL	RO	DE	CZ	PL	RO
<i>age</i>	0.667	0.192	-0.175	0.081	0.006	0.003	0.000	0.000
<i>educlevel</i>	-0.098	-0.107	-0.091	-0.088	0.028	-0.003	-0.087	-0.038
<i>married</i>	0.174	0.101	0.091	0.061	-0.004	0.001	0.001	0.000
<i>yearswork</i>	-0.546	-0.261	-0.050	-0.089	0.045	0.017	0.021	0.012
<i>permanent</i>	0.141	-0.072	-0.048	-0.082	0.012	0.014	-0.004	-0.001
<i>parttime</i>	0.003	0.000	-0.005	0.000	0.270	0.024	0.032	0.001
<i>manager</i>	0.013	0.003	0.015	0.001	0.030	0.030	0.007	0.004
<i>big</i>	-0.099	-0.008	-0.005	0.013	0.053	0.009	0.004	0.008
<i>cons</i>	-0.069	0.399	0.478	0.254	0.000	0.000	0.000	0.000
Total	0.186	0.248	0.209	0.155	0.439	0.095	-0.026	-0.013

Source: own elaboration using the Stata command 'decompose' .

Table 4. The total differences between the log incomes along the distribution

	<i>p5</i>	<i>p10</i>	<i>p25</i>	<i>p50</i>	<i>p75</i>	<i>p90</i>	<i>p95</i>
AT	0.601	0.609	0.617	0.463	0.357	0.366	0.397
BE	0.534	0.558	0.363	0.255	0.250	0.296	0.337
BG	0.105	0.153	0.141	0.173	0.189	0.234	0.266
CZ	0.616	0.401	0.370	0.264	0.279	0.285	0.349
DE	0.752	1.073	0.791	0.555	0.412	0.422	0.505
EE	0.120	0.262	0.255	0.351	0.370	0.443	0.513
EL	0.195	0.258	0.235	0.139	0.146	0.219	0.301
ES	0.514	0.528	0.524	0.321	0.241	0.205	0.306
FR	0.775	0.729	0.358	0.246	0.264	0.324	0.406
HR	0.434	0.195	0.156	0.152	0.108	0.088	0.204
HU	0.325	0.239	0.122	0.151	0.157	0.226	0.286
IE	0.351	0.379	0.395	0.325	0.314	0.344	0.349
IT	0.395	0.383	0.417	0.234	0.219	0.254	0.302
LU	0.644	0.558	0.298	0.271	0.206	0.198	0.219
NL	0.593	0.690	0.606	0.477	0.419	0.452	0.478
PL	0.324	0.304	0.125	0.143	0.162	0.205	0.254
RO	0.032	0.143	0.119	0.133	0.168	0.122	0.086
SK	0.393	0.200	0.304	0.231	0.266	0.274	0.307
UK	0.537	0.627	0.550	0.446	0.395	0.391	0.455

Figure 1. The unadjusted gender pay gap, 2016 (difference between average gross hourly earnings of male and female employees as % of male gross earnings)

Source: Based on https://ec.europa.eu/eurostat/statistics-explained/index.php/Gender_pay_gap_statistics.

Figure 2. The result of grouping the countries using the 1-nearest neighbors algorithm for agglomerative clustering

Source: own elaboration using Stata.

Figure 3. The differences between the log income distributions for men and women vs. quantile rank in the group 1

Source: own elaboration using Stata.

Figure 4. The differences between the log income distributions for men and women vs. quantile rank in the group 2

Source: own elaboration using Stata.

Figure 5. The differences between the log income distributions for men and women vs. quantile rank in the group 3 (a) and the group 4 (b)

Source: own elaboration using Stata.