Central European Journal of Economic Modelling and Econometrics

Regional Differences in Gender Wage Gaps in Poland: New Estimates Based on Harmonized Data for Wages

Aleksandra Majchrowska*, Paweł Strawiński†

Submitted: 21.09.2015, Accepted: 31.05.2016

Abstract

The aim of this paper was to estimate the gender wage gap in Poland and in the 16 NUTS2 Polish regions in 2010, and to verify the predictions of the spatial monopsony model for Poland with a newly created, harmonized database for wages of individuals in Poland. According to the model, the unexplained part of the gender wage gap, identified with wage discrimination, tend to be lower in regions with more competition between employers.

The results of the analyses performed in this paper show that in more urbanized regions the average wages are higher than in the rural ones. In each of the 16 NUTS2 Polish regions, women earn less than men. Raw differences in wages between men and women are largest in the most urbanized regions but a significant part of the differences in those regions can be explained by differences in workers' characteristics, especially by different sectoral structure of employment. The part of the gender wage gap which remains unexplained, and in the literature is commonly attached to discrimination, is the highest in rural regions of Eastern Poland in line with the predictions of the spatial monopsony model.

Keywords: gender wage gap, Poland, regional labour markets

JEL Classification: J31, R23, J16

*University of Łódź and Narodowy Bank Polski; e-mail: arogut@uni.lodz.pl
†University of Warsaw; e-mail: pstrawinski@wne.uw.edu.pl

1 Introduction

The aim of the paper was to estimate the gender wage gap in Poland and in the 16 NUTS2 Polish regions in 2010, and to verify the predictions of the spatial monopsony model for Poland. According to the model part of the gender wage gap identified with wage discrimination tend to be lower in regions with more competition between employers.

Looking at the statistical data, significant differences in the average wage level among Polish regions can be observed. The average wage level in the capital region of Poland (Mazowieckie) in 2010 was, depending on the dataset, around 40–50% higher than the wages in the least developed regions of Eastern Poland.

Nevertheless, Polish economic literature lacks empirical analyses of the regional diversity of the gender wage gap and its causes. Most of the previous empirical literature on the gender wage gap has focused on the variation of the gender pay gap between countries and its evolution over time. An aspect that has attracted far less attention is the regional variations of the gap within the same country. To the best of the authors' knowledge there is only one paper (Słoczyński, 2012) that has analysed the problem of regional differences in the gender wage gap in Poland.

This is at least partly due to the lack of statistical data for individuals' wages in Poland that is fully reliable and representative of the whole economy. The previous estimates of the average gender wage gap in Poland were based on one of the two datasets available: the Polish Labour Force Survey (PLFS; see for example Goraus and Tyrowicz, 2014, and Goraus *et al.*, 2015), or the Structure of Wages and Salaries by occupation (SWS; see for example Matysiak *et al.*, 2010, and Magda and Szydłowski, 2008). Both datasets have, however, some disadvantages (these are explained in more detail in Section 3).

In this paper, the authors attempted to estimate gender wage gaps across the 16 Polish NUTS2 regions and to explain the differences among them with the newly created, harmonized database for wages of individuals in Poland. We have linked the statistical information from the PLFS and the SWS databases and received a harmonized dataset for individual wages in Poland representative of the whole Polish economy (a more detailed description of the dataset is presented in Section 3).

We decomposed the differences in wages between men and women in order to verify to what extent they can be explained by the differences in characteristics of workers (education, work experience, place of work etc.). The differences in wages that cannot be explained by differences in endowments are identified in the literature with the effect of discrimination. According to the spatial monopsony model (Hirsch, 2009) this unexplained part of the wage differences between men and women should be lower in more urbanized and more competitive regions. To the best of the authors' knowledge this is the first paper that has tried to verify this hypothesis for Polish regions.

We started our empirical approach by estimating the gender wage gap in Poland and in each of the 16 NUTS2 regions with the three datasets available: the PLFS, the

SWS and the harmonized data, and we then compared the results. We performed the standard Oaxaca (1973) – Blinder (1973) decomposition. In the second step, we limited the sample to occupational groups with similar share of men and women in employment (masculinisation ratio between 0.4 and 0.6). Discarding the highly feminised and masculinised occupational groups we retained those groups of workers with very high probability of finding a statistical match in a second gender group. In the third step, we checked the robustness of our results by performing the Ñopo (2008) decomposition.

The structure of the paper is as follows. In the second part, we present the theoretical justification of regional differences in gender wage gaps and some previous empirical evidence. The third part describes the statistical data used in the paper and the empirical strategy. In the fourth part, the empirical results are presented. The fifth part concludes.

2 Regional differences in gender wage gaps: theoretical justification and empirical evidence

According to the Mincer model, the differences between the wages of men and women can be the result of differences in personal characteristics, especially the amount of human capital embodied in individuals (Mincer, 1997). Differences in the gender wage gap across regions may at least be partly due to differences in individuals' level of education, as well as the differences in their in-work investments, which are usually approximated by work experience.

Regional differences in gender wage gaps may also be the result of different employment structures and characteristics of employers – size of the firm, ownership sector, branch etc. Empirical analyses confirm that bigger firms pay higher wages, other factors being constant (Gibson and Stillman, 2009; Lallemand *et al.*, 2005). Empirical analyses confirm that the gender wage gap increases along the wage distribution. For instance, Arulampalam *et al.* (2007) found that for eleven European countries, gender pay gaps are typically bigger at the top of the wage distribution, a finding that is consistent with the existence of 'glass ceilings' for women.

More recently, regional differences in wages between men and women have been explained more deeply on the basis of the monopsonistic theory of discrimination. Hirsch (2009) presents a spatial monopsony model of the labour market, in which he explains the causes of the differences in gender wage gaps among regional labour markets. The model assumes that workers are located at different places, while employers do not exist at each potential location. Some workers therefore have to commute and bear some travel costs, both direct and indirect. Direct costs are concerned with travelling, whereas indirect costs follow from the fact that travelling requires time and thus imposes some opportunity costs (Hirsch *et al.*, 2010).

Since employers and the jobs they offer are not perfect substitutes to workers,

competition among employers is imperfect and firms possess some monopsony power. Moreover, the model assumes that, due to their domestic responsibilities, some women have higher average opportunity costs than men. This translates to lower spatial mobility of women and less choice of employers. Hirsch (2009) therefore arrives at the conclusion that firms have higher monopsony power over the female workers than men; therefore, they offer them lower wages, giving rise to a gender wage gap. This part of the theory explains why women earn on average less than men.

Regarding the regional dimension, Hirsch (2009) argues that the differences in wages of men and women among regions are due to differences in the monopsonistic position of employers. More competition between employers in more urbanized areas should increase wages of both men and women. Moreover, higher competition between employers in more urbanized areas should also constrain employers' ability to discriminate against women, which should lead to lower gender wage gaps in more urbanized regions. This part of the theory explains the differences in the wages of men and women among regions.

Hirsch *et al.* (2010) confirmed the findings of the spatial monopsony model for the German labour market. They analysed the regional differences in the gender wage gap in Germany, and found that the unexplained part of gender wage gap for young workers was substantially lower in large metropolitan than in rural areas. Additionally, they confirmed that the differences between regions persist over time. Guyot *et al.* (2009) also showed that the gender pay gap in Germany grow wider from core regions to periphery.

Lopez-Bazo and Motellon (2009) analysed the effect of human capital on regional wage differentials in Spanish regions. Their results not only confirmed that the regions differ in the endowment of human capital, but also that the return that individuals obtain from it varies sharply across regions. Regional heterogeneity in returns is especially intense in the case of education. The differences in return on human capital account for a significant proportion of the differences in regional wage gaps in Spain.

More recently, Castaño and Paredes (2015) confirmed a negative relationship between spatial density of firms and the gender wage gap in Chile. Their results show that the effect of a 10% increase in the number of firms in a municipality is associated with a decrease in the gender wage gap of between 0.6 pp. and 0.9 pp. Taniguchi and Tuwo (2014) analysed differences in wages in Indonesia. They found that the gender wage gap is wider in urban areas, even after controlling for socioeconomic variables. However, the unexplained part of the wage gap is larger in rural areas than in urban areas which – according to the interpretation of the authors – implies that the labour market is more efficient in urban areas than in rural areas.

As far as Poland is concerned, statistical data confirm that on average women earn less than men. According to the Eurostat data, the unadjusted gender wage gap (this is defined by Eurostat as the raw difference between the average gross hourly earnings of men and women expressed as a percentage of the average gross hourly earnings of men: `http://ec.europa.eu/eurostat/statistics-explained/`

index.php/Gender_pay_gap_statistics). in Poland is, however, relatively low in comparison with that in other European countries. In 2010, the gender wage gap in Poland (on average, in the industry, construction and services sectors) reported by Eurostat was around 5%, which was the second-lowest among the EU countries after Slovenia, whose unadjusted gender wage gap equalled 0.7%. The largest raw differences between the wages of men and women in 2010 were noted in Estonia, and amounted to 26%.

The estimates of the gender wage gap in Poland adjusted for both personal and firms' characteristics vary significantly, mainly due to the different datasets used in the analyses. Among the most recent papers where the gender wage gap in Poland has been estimated, we should mention the following: Goraus *et al.* (2015); Goraus and Tyrowicz (2014); Śliwicki and Ryczkowski (2014); Mysiková (2012); Rokicka and Ruzik (2010); Matysiak *et al.* (2010); Magda and Szydłowski (2008); and Majchrowska *et al.* (2015). Among the earlier papers, we should also mention Grajek (2003) and Adamchik and Bedi (2003). However, as they used data from the early transition period, their results cannot be directly compared with the later ones. For a more comprehensive review of the main results of these papers, see for instance Majchrowska *et al.*, 2015, or Goraus *et al.*, 2015.

Although the regional dimension has appeared in many of the previous analyses as a control variable (see for example Goraus and Tyrowicz (2014) or Matysiak *et al.* (2010)), only one paper (Słoczyński, 2012) has analysed the regional differences in gender wage gaps in Poland. Słoczyński (2012) examined the gender wage gaps in the 16 NUTS2 regions in Poland in 2008, based on the Structure of Wages and Salaries (SWS) data in October 2008. He performed the Oaxaca-Blinder and the Juhn, Murphy and Pierce decompositions. His results show very strong diversification of gender wage gaps among Polish regions – from 25% in the very urbanized Slaskie region, with a relatively high share of workers in the mining industry, to 6% in Podkarpackie, one of the less developed regions in Eastern Poland. As SWS data were used in the paper, the results are representative of employees working in firms with at least ten employees.

In this paper, we re-estimated gender wage gaps across Polish regions using a harmonized dataset for wages that is representative of the whole economy, and verified the predictions of the spatial monopsony model. To the best of the authors' knowledge this is the first paper that has tried to verify these predictions for Polish regions. In the paper we aimed to verify the following hypotheses arising from this model:

1. More competition between employers in more urbanized regions of Poland translates into higher wages of both men and women than in the more rural regions.

2. Due to higher average indirect (opportunity) costs, women in each of the 16 Polish regions earn on average less than men.

3. Raw differences in wages between men and women are higher in more urbanized regions, due to more pronounced differences in workers' characteristics.

4. The part of the gender wage gap which remains unexplained is lower in more urbanized regions, because the more competitive labour market constrains employers' ability to engage in female discrimination.

In the following empirical analysis, we will investigate to what extent these four hypotheses generated by the theoretical model are confirmed by the data.

3 Data and empirical approach

The previous estimates of the gender wage gap in Poland were based on one of the two datasets available: the Polish Labour Force Survey, and the Structure of Wages and Salaries by occupation survey. Both datasets have some limitations. The PLFS data contains full information about the employment structure in the economy, but performs much worse in the case of data on wages. Due to the high percentage of non-responses to questions about wages, especially among persons with a relatively high income, the wage distribution is shifted down. The SWS data are fully reliable as they come from the employers' accounting departments, but they are representative only of enterprises with at least ten employees. Taking into account that wages in micro-enterprises are much lower than in medium and big companies, the wage distribution in the SWS dataset is shifted upwards.

Therefore, the authors decided to create a harmonized dataset that could be representative of the whole economy. Firstly, the definitions of socio-demographics and firms' characteristics in both datasets were standardised. In order to combine data from different sets, information on province, gender, education and age of the respondents were used. The sample was divided into sub-samples formed by a combination of region, gender, educational-level group and age group. Secondly, using the multiple imputation technique, the wages from the SWS data were imputed into the employment structure given by the PLFS. Specifically, for persons working in the micro sector, the wage information given by the PLFS was retained. For individuals working in bigger enterprises, the wage data was imputed from the SWS. Wages were assigned to the dataset by a multiple imputation technique based on the extended Mincer-type wage equation, which takes into account not only the standard variables, but also the NACE section and the great occupational group in which the respondent works. To achieve the desired reliability, we used an average of forty imputations from the SWS data. As a result, the total sample of the newly created harmonized dataset comprises about 100,000 observations. Its advantage is that is uses wage information from two sources, while preserving the demographic and residence structure of the Polish population.

The multiple imputation technique is deeply rooted in statistical theory. It was first proposed by Rubin (1978) as a possible solution to the problem of a survey's

non-responses. Zhang (2003) described three common approaches used in multiple imputations: the propensity score method, the predictive model method and the Markov Chain Monte Carlo. In this research, the predictive model method is utilised. Multiple imputations usually are used as a within-data method. It is worth noting that Laaksonen (2006) stressed the need for good auxiliary data when dealing with imputation, as these increase the quality of data. A similar exercise of replacing the PLFS wage information with the SWS wage information was previously used by Myck, Morawski and Mycielski (2007). However, for their imputation purpose they used kernel estimation based on the propensity score of employment.

As far as the empirical approach is concerned, we start with estimating the gender wage gap in Poland and in the 16 NUTS2 Polish regions by means of a standard two-component Oaxaca-Blinder decomposition. This method allows us to decompose the difference in average wages of men and women into two parts. The first part is due to differences in characteristics of individuals; both personal characteristics (level of education, age, work experience etc.) and characteristics of the firms that employ them (size, ownership sector, branch etc.). The remaining (unexplained) part of the differences in wages is interpreted as the discrimination effect. In our paper (based on previous research) we assume that the eventual discrimination in Poland concerns only women. Therefore, we assume that the non-discriminatory coefficient vector is the vector of men's wages. Our equation takes therefore the following form:

$$D = \ln \overline{w}_M - \ln \overline{w}_F = \hat{B}_M \left(\overline{X}_M - \overline{X}_F\right) + \left(\hat{B}_M - \hat{B}_F\right) \overline{X}_F \tag{1}$$

where:

$\ln \overline{w}_M (\ln \overline{w}_F)$ – average wage level in the group of men (women);

$\hat{B}_M(\overline{X}_M - \overline{X}_F)$ – is the so-called 'explained' part of the wage gap – that part of the gap which is due to differences in characteristics of individuals;

$(\hat{B}_M - \hat{B}_F)\overline{X}_F$ – is the remaining (unexplained) part of the wage gap.

The above decomposition is based on an extended Mincer-type wage equation:

$$\ln (w_i) = \beta_0 + \sum \gamma X_i + \sum \theta X_j + \sum \delta_{ij} X_i X_j + \varepsilon_i \tag{2}$$

where:

X_i – vector of i personal characteristics of individuals;

X_j – vector of j firms' characteristics.

The parameters of wage equation (2) were estimated separately for the group of men and women as suggested in the literature (Oaxaca, 1973).

As far as the personal characteristics of individuals are concerned, we included

education, work experience and occupational group at 1-digit-level of classification. Education is measured as the number of years necessary to achieve a certain level of education. Work experience is measured as the implied years of work experience (age minus preschool years and years in the educational system). Implied values are used to achieve coherency in analysis for the three sets of data used in the paper. Following the standard Mincer-type model, experience appears in Equation (2) in levels and as a square, in order to take into account the fact that in-job investments tend to diminish with age due to their lower profitability (Mincer, 1997).

In the second vector, several firms' characteristics are included, such as a dummy for the public sector with the private sector as the reference level and the economic section for the main activity of the firm (according to the NACE classification). Additionally, the set of dummies describing the size of the firm is included. The population of firms has been divided into small firms (11–50 employees), medium-sized firms (51–250 employees) and big firms (251 and more employees), with establishments with no more than ten employees as a baseline category. Moreover, we allow for all possible interactions between individual's and firm characteristics.

As the aim of the paper is to estimate differences in the gender wage gap across regions, different sectoral employment structures between regions have been taken into account, in particular the fact that the share of workers employed in the agricultural sector amounts in some Eastern regions of Poland to 30% of all employees. Because most of the workers in agriculture do not receive regular salaries but the income from agricultural production, we decided to exclude the agricultural workers from our sample.

In the first step, we estimated the gender wage gap in Poland and in the 16 NUTS2 Polish regions with the Oaxaca-Blinder decomposition, using our harmonized data for wages. To check how well our data performed and to compare our estimates with the previous ones, we estimated the gender wage gaps using the other datasets available: the Polish Labour Force Survey and the Structure of Wages and Salaries by occupation.

One of the disadvantages of the traditional Blinder-Oaxaca decomposition is the fact that it estimates earnings equations for all working females and all working males, without restricting the comparison only to those individuals with comparable characteristics by recognising gender differences in the supports (Goraus and Tyrowicz, 2014). It is therefore necessary to make an 'out-of-the-support' assumption that the fitted regression surface can be extended for individual characteristics that have not been found empirically in the dataset, using the same estimators computed with the observed data.

Therefore, in the second step we adjusted the sample of all individuals by retaining only those individuals for whom the probability of finding a statistical match in a second group is relatively high. Following the approach of European Commission (see Burchell *et al.*, 2014), we decided to eliminate male-dominated and female-dominated occupational groups (at the 3-digit level of classification). By doing this we

excluded the occupational groups where the ratio of men to women working is highly asymmetric (e.g., miners or nurses). An occupation is defined as male-dominated if more than 60% of the employees in that occupation are male, female-dominated if more than 60% of employees are female, and mixed if the proportions of men and women are between 40% and 60%. The European Commission (see Burchell *et al.*, 2014) underlines that the 0.4–0.6 interval for mixed occupations is wide enough to take into account only those groups of workers with a similar share of men and women. If the interval was wider (0.3–0.7, for example), the occupation could have more than twice as many men than women and still be treated as mixed.

By making this adjustment of the sample, we expect to have eliminated the differences in wages of men and women that are not the effect of discrimination (according to the definition discrimination exists when female or minority workers — who have the same abilities, education, training, and experience as white male workers — are accorded inferior treatment with respect to hiring, occupational access, promotion, wage rate or working conditions (McConnell and Brue, 1986)) but which are due to the segregation of men and women into different occupations. Occupational segregation can be said to exist when the distribution of occupations within one demographic group is very different from the distribution in another. With respect to gender, occupational segregation is reflected in there being female-dominated occupations and male-dominated ones (Ehrenberg and Smith, 1994, p. 399). We expect that the gender wage gap in occupations with balanced men-to-women ratio in Poland will be lower than the total gender wage gap.

The difference between our approach and the standard Oaxaca-Binder decomposition could be viewed as the difference between the local average treatment effect and the average treatment effect. The former is the average difference between locally comparable units, while the latter is the sample or population difference.

As a robustness check, we performed the Ñopo decomposition, which accounts for gender differences in the supports. This decomposition consists of four additive elements (Ñopo, 2008):

$$\triangle = \triangle_M + \triangle_X + \triangle_F + \triangle_O \tag{3}$$

The first component is the part of the gap that can be explained by differences between two groups of males: those who have characteristics that can be matched to female characteristics and those who do not. This part of the gap would disappear in the event that there are no males with combinations of characteristics that remain entirely unmatched by females. Alternatively, this component would also disappear if those males with individual characteristics that are not matched by females were paid, on average, the same as the average matched males. It is computed as the difference between the expected male wages outside the common support and the expected male wages in the common support, weighted by the probability measure (under the distribution of characteristics of males) of the set of characteristics that are not shared by females (Ñopo, 2008).

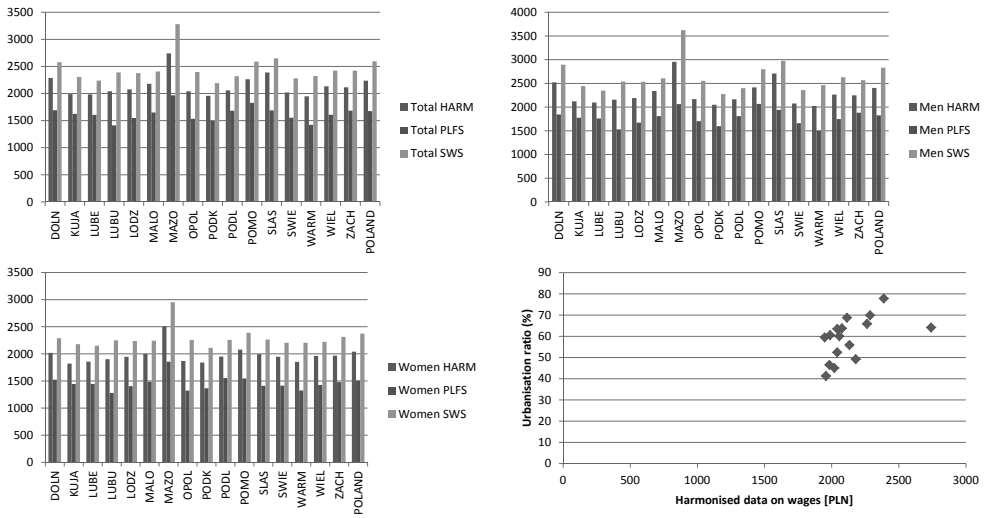
The second component (Δx) is the part of the wage gap that can be explained

by differences in the distribution of characteristics of males and females over the common support, and corresponds to the explained component of the Oaxaca-Blinder decomposition $\hat{B}_M(\overline{X}_M - \overline{X}_F)$ (Ñopo, 2008).
The third component is the part of the gap that can be explained by the differences in characteristics between two groups of females: those who have characteristics that can be matched to male characteristics, and those who do not. It accounts for that part of the gap that would disappear should it ever be the case that all females had characteristics that could be matched to the population of males. It would also disappear if unmatched females were paid, on average, as much as matched females. It is computed as the difference between the expected female wages, in and out of the common support, weighted by the probability measure (under the distribution of characteristics of females) of the set of characteristics that are not shared by males (Ñopo, 2008).
The fourth component corresponds to the unexplained part of the Oaxaca-Blinder decomposition $(\hat{B}_M - \hat{B}_F)\overline{X}_F$. It is the share of the wage gap that cannot be attributed to differences in characteristics of the individuals, and is typically attributed to a combination of both the existence of unobservable characteristics that explain earnings and the existence of discrimination (Ñopo, 2008).
Finally, we compared the unexplained part of the gender wage gap obtained with the Oaxaca-Blinder and the Ñopo decomposition with the regional urbanization ratio. If our fourth hypothesis is true, the unexplained part of the wage gap should be significantly lower in urbanized than in rural regions.

4 Gender wage gap in Poland and in the 16 NUTS2 regions: Empirical results

In the first step, we compared the average level of wages of men and women across regions using the three datasets available. Looking at the mean wage level of men and women in the three data sources (the SWS gross wages have been recalculated to net wages to achieve comparability), we can confirm that in the case of both men and women, on average the harmonized data on wages lie between the PLFS and the SWS data in Poland and in each of the 16 regions in line with authors predictions (see Figure 1).
Comparing the average wage level and the urbanization ratio, we can confirm the first hypothesis. More competition between employers in the more urbanized regions of Poland translates into higher wages for both men and women than in the more rural regions. We can also confirm the second hypothesis. Women in each of the 16 Polish regions earn on average less than men. In the next step, we will analyse the gender wage gap in Poland and verify the third and fourth hypotheses.

Figure 1: Average net wages of men and women in Poland and in the 16 NUTS2 regions according to the harmonized dataset, the PLFS and the SWS datasets in 2010, and the relationship between average wages (harmonized data) and the urbanization ratio



Source: PLFS, SWS and http://eregion.wzp.pl/wspolczynnik-urbanizacji-dla-poszczegolnych-wojewodztw-w-latach-2010-2013; own estimates.

4.1 Estimates of overall gender wage gap in Poland

In the first step, we used the harmonized dataset on wages to estimate the gender wage gap in Poland in 2010. We compared the obtained results with those estimated with the PLFS and SWS datasets.

The results show that the gender wage gap in Poland in 2010, estimated for the harmonized dataset, equals 15.7%, which is significantly lower that the number obtained with PLFS data (19.8%), and significantly higher that the gender wage gap obtained with SWS data (12.6%; see Table 1). All gender wage gap estimates are significant at the 1% significance level.

The results above are generally consistent with the authors' intuition. The SWS data provides higher estimates of average wages of both men and women than the hypothetical country average. This is a direct consequence of the limited scope of the SWS survey. However, due to the fact that more women than men work in small firms and the financial sector (not covered by the SWS data), we can expect that the upward bias is higher in the case of women. If this is true, the gender wage gap estimated with the SWS data will be lower than that estimated for the harmonized dataset. Following our intuition, one can expect that at the regional level, the average wage of women based on the SWS data should be more upwardly biased in regions

Table 1: The Oaxaca-Blinder decomposition of gender wage gap in Poland in 2010 estimated for the harmonized dataset, PLFS data and SWS data

	HARM	PLFS	SWS
$\ln(\overline{w}_M)$	7.689^{***} (0.002)	7.446^{***} (0.004)	7.933^{***} (0.001)
$\ln(\overline{w}_F)$	7.532^{***} (0.002)	7.248^{***} (0.004)	7.807^{***} (0.001)
$\ln(\overline{w}_M) - \ln(\overline{w}_F)$	0.157^{***} (0.003)	0.198^{***} (0.005)	0.126^{***} (0.002)
$\hat{B}_M\left(\overline{X}_M - \overline{X}_F\right)$	-0.013^{***} (0.002)	-0.009^{***} (0.004)	-0.015^{***} (0.001)
$\left(\hat{B}_M - \hat{B}_F\right)\overline{X}_F$	0.170^{***} (0.002)	0.207^{***} (0.005)	0.142^{***} (0.002)
Number of observations	101,947	43,957	681,747

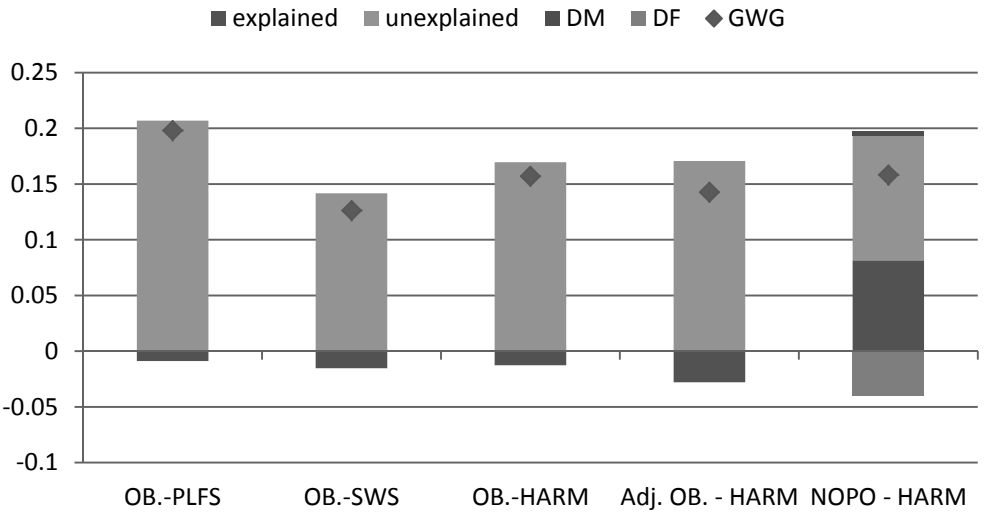
*** significant at 1% significance level.

with a relatively high share of small enterprises (less developed, more rural regions). It is more difficult to make a prediction in the case of the PLFS data. Wages reported in the PLFS data are lower on average than those in the SWS data, for at least two reasons. The first is that the data covers all of the economy, not only enterprises employing ten or more workers. The second is the well-known phenomenon that individuals with relatively high incomes frequently refuse to answer the question about wages in the PLFS survey (see for instance Myck, Morawski and Mycielski, 2007, or Strawiński, 2015). Those data, however, are not sufficient to derive conclusions regarding the distribution of the expected differences across regions and gender.

Looking at the decomposition of the gender wage gap in Poland (see Figure 1), we see that the explained part of the gap is negative (irrespective of the dataset used: PLFS, SWS or HARM), indicating that women should be better paid than men if both groups receive the same treatment. The detailed decomposition shows that women are more educated than men and more often work in the public sector. The unexplained part of the wage gap is higher than the total gap, showing that some other factors have not been taken into account, or that women are wage-discriminated in the labour market.

Strawiński *et al.* (2016) analysed segregation in the Polish labour market in 2000–2013, and found that women much more often work in jobs that require a greater degree of soft skills and are less risky. Significantly more men perform more physical jobs in industry and construction. Bearing in mind that the differences in wages between men and women may be at least partly explained by the different tasks they perform, in the next step we eliminated from the sample all employees working in the male-dominated or female-dominated occupational groups (at the 3-digit level of classification), as explained in the previous section. We were left with only those occupational groups where the probabilities of finding a statistical match for men or women were relatively high. In this section, we performed the analyses only on the harmonized dataset. Then, the results were compared with the ones obtained from the whole sample. Our intuition is that part of the differences in the gender

Figure 2: Gender wage gap (GWG) in Poland in 2010 and its decomposition by different methods and datasets



OB-PLFS – the Oaxaca-Blinder decomposition of the gender wage gap estimated on the PLFS data; OB-SWS – the Oaxaca-Blinder decomposition of the gender wage gap estimated on the SWS data; OB-HARM – the Oaxaca-Blinder decomposition of the gender wage gap estimated on the harmonised data; Adj.OB-HARM – the Oaxaca-Blinder decomposition of the gender wage gap estimated on the adjusted sample (occupational groups with masculinisation ratio between 0.4 and 0.6); NOPO-HARM – the Ñopo decomposition of the gender wage gap estimated on the harmonised data.

wage gap estimated for the whole sample of individuals is not due to discrimination, but is the result of the different employment structures of men and women. In more homogenous sub-sample the adjusted gender wage gap should then be lower than the total one.

After eliminating the male-dominated and female-dominated occupations we were left with 26 occupational groups at the 3-digit level and around 14,000 observations (out of around 100,000 in the initial sample). The estimates show that the gender wage gap in the limited sample is in fact lower than that estimated for the whole sample. The total gender wage gap amounts to 15.7%; the adjusted one equals 14.3% (Figure 2). The results confirm therefore that at least part of the total gender wage gap can be assigned to the differences in employment structure.

To check the robustness of the above results, we performed the Ñopo decomposition on the whole sample of harmonized data on wages. The results indicate that men outside the common support earn on average only 0.4 pp. more than men in the common support, but women outside the common support earn 4 pp. less than women in the common support (Figure 2). The findings are consistent with our intuition. The component that corresponds to differences in characteristics of men and women in the common support is positive (8 pp.), which indicates that in the common support,

women have different labour market characteristics than men. The differences in endowments, however, explain less than one-third of the wage differences between men and women in the common support. The overwhelming majority of the wage differences remain unexplained.

4.2 Regional differences in gender wage gaps in Poland

In this section, we aimed to estimate the gender wage gap in each of the 16 NUTS2 Polish regions, and to answer the question which factors are responsible for the differences between regions. Firstly, we analysed the gender wage gap in the whole sample of the harmonized dataset, and compared the results with those from the PLFS and SWS datasets. Then, we limited our sample to the occupational groups with the masculinisation ratio 0.4–0.6, and estimated the gender wage gaps in each region. Finally, to check the robustness of the results, we performed the Ñopo decomposition for each of the regions separately.

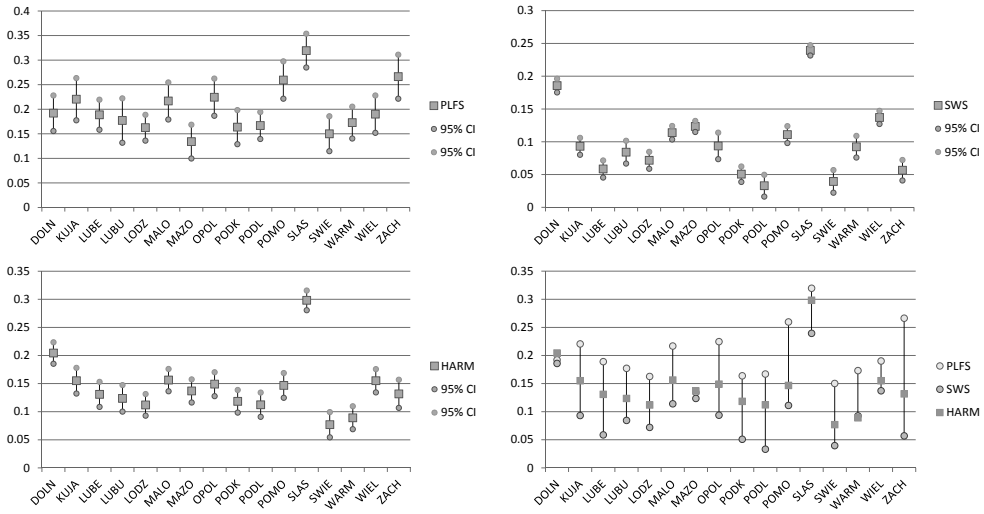
As shown in Figure 3, all the estimated regional gender wage gaps are significant at the 1% significance level. Although the estimated gender wage gaps for the three datasets show some similarities (the highest gender wage gap is in the Slaskie region for all three datasets), the differences are notable (see Figure 3). The tables with estimates of gender wage gaps and their decomposition among Polish regions are presented in Appendix.

Comparing the results obtained for the three datasets, we can see that, in most regions, the gender wage gaps estimated for the harmonized dataset lie between those estimated with the PLFS and SWS data. The regional gender wage gaps estimated with the PLFS data are the highest. In almost all regions, they are higher than the gender wage gaps estimated with the harmonized dataset. The two exceptions are Dolnoslaskie and Mazowieckie regions. The regional gender wage gaps estimated with the SWS data are the lowest. In almost all regions (with exception of Warminsko-Mazurskie), they are lower than the gender wage gaps estimated with the harmonized dataset.

The differences between the wages of men and women across Polish regions are significant. Looking at the results from the harmonized dataset, we see that gender wage gaps varied in 2010 between 7.7% in Swietokrzyskie and 29.8% in Slaskie region (the average gender wage gap in Poland was 15.7%). When looking more carefully at the results, we see that there are only two regions where differences in wages between men and women are relatively high – the Slaskie and Dolnoslaskie regions – which both have a significant share of men working in the very well-paid mining sector. In 2010, 12% of men working in Slaskie and 4% in Dolnoslaskie were employed in the coal-mining industry. In the case of women, these shares were respectively 1.3 and 0.6%.

In general, we can confirm our third hypothesis. The differences in wages between men and women are higher in more urbanized regions, due to more pronounced differences in workers' characteristics.

Figure 3: Gender wage gaps in the 16 NUTS2 regions in Poland based on different datasets as % of men's wages

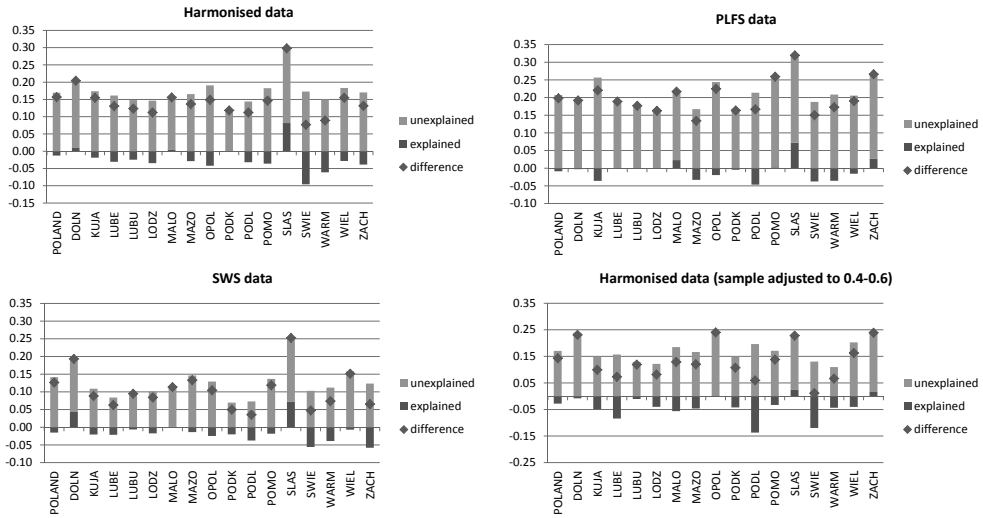


Now we look at the decomposition of gender wage gaps by regions, calculated with the Oaxaca-Blinder decomposition. We use the whole sample of harmonized data for 2010. The first conclusion is that the differences in endowments are not the factors that can explain the differences in wages between men and women among regions (Figure 4). The explained part in most of the regions is negative, indicating that women have better labour market characteristics and should be paid more than men. Only in Slaskie region is the explained part of the wage gap positive and significant (in Dolnoslaskie and Malopolskie it is positive, but not statistically significant), indicating that almost one-third (27%) of the wage differences between men and women in that region can be explained by their characteristics. More detailed decomposition indicates that the wages of men are higher, mostly because of the differences in employment structure (men working in the mining sector).

The picture is largely unchanged when we look at the results of the Oaxaca-Blinder decomposition performed on the PLFS and SWS data (see Figure 4). With the PLFS data, the explained part in most of the regions is very small and not significant. In the Slaskie region only, 23% of the gender wage difference can be attributed to different characteristics of individuals. According to the SWS data, 22% and 28% of gender wage differences in Slaskie and Dolnoslaskie respectively can be explained by differences in characteristics. In most of the regions, the explained part is negative.

Bearing in mind that part of the differences in wages between men and women can be the result of different employment structures, in the next step we excluded from the sample all the individuals in male-dominated and female-dominated occupations. In

Figure 4: The Oaxaca-Blinder decomposition of gender wage gaps in the 16 NUTS2 Polish regions in 2010 according to harmonized data, PLFS data and SWS data



Light red colour indicates the statistically insignificant explained part of the wage gap.

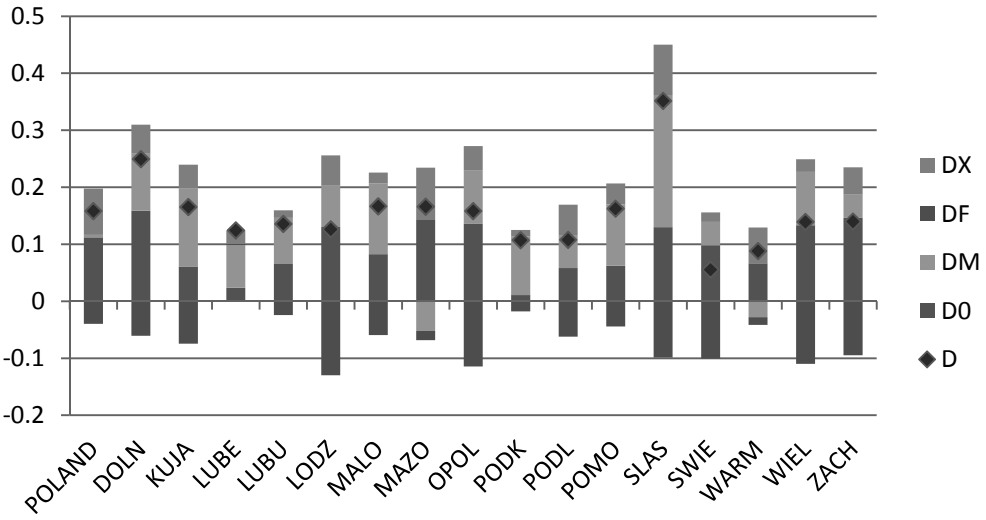
each of the regions, we retained only individuals in 3-digit-level occupational groups with a masculinisation ratio of 0.4–0.6. The results indicate that the differences in wages between men and women in most of the regions were smaller than in the whole sample (see Figure 4). They varied in 2010 from 1.1% (however not significant) in Świętokrzyskie to 24% in Zachodniopomorskie (14.3% on average in Poland). This confirms our intuition that part of the wage differences between men and women in Poland is due to different employment structure of men and women.

The results of the Oaxaca-Blinder decomposition show that differences in endowments cannot explain the differences in wages between men and women in the regional labour market. In most of the regions where the explained part is significant, it is negative indicating that women should be better paid than men if both groups receive the same treatment.

To check the robustness of the results, we performed the Ñopo decomposition for each of the 16 NUTS2 regions in Poland separately. Looking at the contribution of the first factor (DM, see Figure 5), we see that in most of the regions, men out of the common support earn more than men in the common support. The largest difference (unsurprisingly) is observed in Śląskie (the difference is 23 pp.). In other regions, the differences in men's wages vary from 4 to 12 pp. Only in two regions (Mazowieckie and Warminsko-Mazurskie), men out of the common support earned respectively 5 and 3 pp. less than men in the common support in 2010. The difference between the wages of women outside and in the common support (DF in Figure 5) is negative in almost every region. The largest differences are observed in Łódzkie (11 pp.) and

Wielkopolskie (8 pp.). Only in Lubelskie are wages of women in and out of the common support the same.
The component that shows the differences in the distribution of characteristics of males and females over the common support is highly positive in most regions, from 1 pp. in Lubuskie to 9 pp. in the Mazowieckie and Śląskie regions. However, the results of the Ñopo decomposition show that a significant part of the differences in wages between men and women cannot be explained by differences in the characteristics of individuals, which confirms the previous results of the Oaxaca-Blinder decomposition.

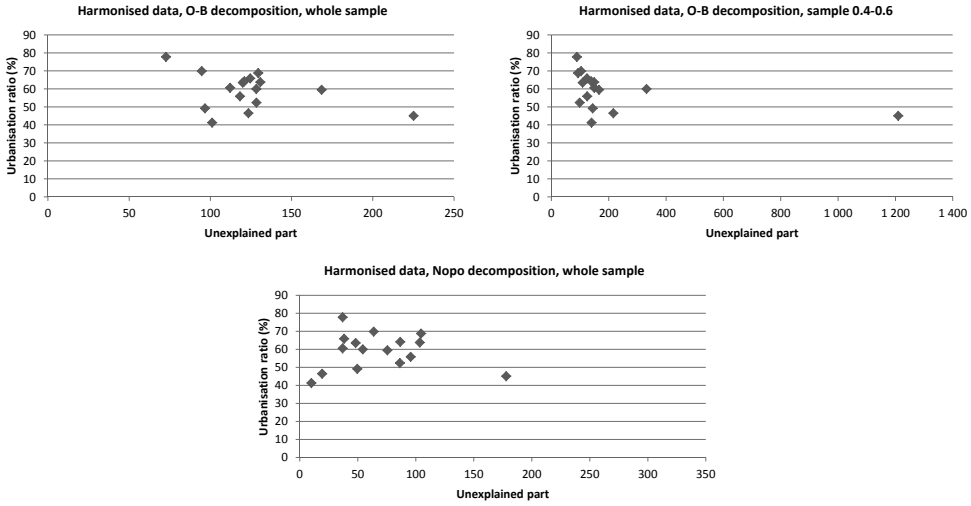
Figure 5: The results of the Ñopo decomposition for the 16 NUTS2 Polish regions in 2010



In the last step we checked the relationship between the part of the gender wage gap which remains unexplained and the urbanization ratio among Polish regions. To take into account significant differences in the size of the estimated gender wage gaps on regional labour markets we decided to change slightly the approach proposed by Hirsch *et al.* (2010). Instead of comparing the size of unexplained part of the gender wage gap measured in log pp. we look at the part of the gender wage gap which remains unexplained, measured as percentage of total wage gap in the region.
Figure 6 presents the relationship between the unexplained gender wage gap (in percentage of total difference in wages between men and women) and the urbanization ratio. According to the spatial monopsony model, in more urbanized regions the discrimination effects should be lower than in rural regions, due to greater competition between employers. We see that, irrespectively of the decomposition

method used (Oaxaca-Blinder or Ñopo), the relationship is negative, confirming that the unexplained gender wage gap is lower in more urbanized regions. According to the spatial monopsony model it is because the more competitive labour market constrains employers' ability to engage in female discrimination. We can therefore confirm our last hypothesis.

Figure 6: The relationship between unexplained gender wage gap (% of total difference in wages between men and women) and urbanization ratio (%) across the 16 NUTS2 Polish regions in 2010



5 Conclusions

This paper was the first attempt to use a harmonized dataset, newly created by the authors, to analyse the gender wage gap in Poland. It was also the first attempt to verify the spatial monopsony model for the Polish economy.

The results of the analyses performed in the paper show that the harmonized dataset describes individual data on wages in Poland very effectively. In particular, the harmonized data confirm that the distribution of Polish Labour Force Survey data on wages is shifted down in relation to average wages in the economy. Conversely, the new dataset indicates that the distribution of Structure of Wages and Salaries data is shifted upwards.

We succeeded in confirming all the initial hypotheses based on the spatial monopsony model. In more urbanized regions the average wages are higher than in the rural ones. In each of the 16 NUTS2 Polish regions women earn less that men. The differences

in wages between men and women are the highest in the most urbanized regions, but a significant part of the differences in those regions can be explained by differences in workers' characteristics, especially by different sectoral structures of employment. The part of the gender wage differences which remains unexplained and in the literature is commonly attached to discrimination, is the highest in rural regions of Eastern Poland in line with the predictions of the spatial monopsony model.

Acknowledgements

The paper was financed by National Science Centre grant DEC-2013/09/B/HS4/01304. The authors are thankful to anonymous referee for valuable comments and suggestions.

References

[1] Adamchik V. A. and Bedi A. S. (2003), Gender Pay Differentials during the Transition in Poland, *The Economics of Transition*, Vol. 11, Issue 4, pp. 697–726.

[2] Arulampalam W., Booth A. L. and Bryan M. L. (2007), Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution, *Industrial and Labor Relations Review*, Vol. 60, No. 2 (Jan., 2007), pp. 163–186.

[3] Blinder A. S. (1973), Wage discrimination: reduced form and structural estimates, *Journal of Human Resources*, 8(4), pp. 436–455.

[4] Burchell B., Hardy V., Rubery J. and Smith M. (2014), A New Method to Understand Occupational Gender Segregation in European Labour Markets, European Commission – Directorate-General for Justice.

[5] Castaño, A. and Paredes D. (2015), The effect of the spatial density of firms on gender wage gap', *Documentos de Trabajo en Economia y Ciencia Regional*, No. 55.

[6] Ehrenberg R.G. and Smith R.S. (1994), Modern Labor Economics. Theory and Public Policy, New York: Harper Collins.

[7] Gibson J. and Stillman S. (2009), Why Do Big Firms Pay Higher Wages? Evidence from an International Database, *Review of Economics and Statistics*, Vol. 91, No. 1, pp. 213–218.

[8] Goraus K., Tyrowicz J. and van der Velde L. (2015), Which Gender Wage Gap Estimates to Trust? A Comparative Analysis, *Review of Income and Wealth*, version online: `http://onlinelibrary.wiley.com/doi/10.1111/roiw.12209/abstract`

[9] Goraus K. and Tyrowicz J. (2014), Gender Wage Gap in Poland – Can It Be Explained by Differences in Observable Characteristics?, *Ekonomia*, No. 36.

[10] Grajek M. (2003), Gender Pay Gap in Poland, Economic Change and Restructuring, *Economics of Planning*, No. 36, pp. 23-44.

[11] Guyot A. Berwing S. and Lauxen-Ulbrich M. (2009), Income Differentials on Regional Labour Markets in Southwest Germany, *Panoeconomicus*, Vol. 56, No. 3 (September), pp. 379–396.

[12] Hirsch B. (2009), The Gender Pay Gap Under Duopsony: Joan Robinson Meets Harold Hotelling, *Scottish Journal of Political Economy*, Vol. 56, Issue 5, pp. 543–558, November.

[13] Hirsch B., König M. and Möller J. (2010), Is There a Gap in the Gap? Regional Differences in the Gender Pay Gap, IZA Discussion Paper, 4231 (published in 2013 in *Scottish Journal of Political Economy*, Volume 60, Issue 4, pp. 412–439).

[14] Laaksonen S. (2006), Need for High Quality Auxiliary Data Service for Improving the Quality of Editing and Imputation, [in:] UNStatistical Comission (2006) *Statistical Data Editing*, Vol 3. Impact on Data Quality, New York and Geneva.

[15] Lallemand T., Plasman R. and Rycx F. (2005), Why do large firms pay higher wages? Evidence from matched worker-firm data, *International Journal of Manpower*, Vol. 26 No. 7/8, pp. 705–723.

[16] Lopez-Bazo E. and Motellon E. (2009), Human capital and regional wage gaps, *IREA Working Papers*, 2009/24, Research Institute of Applied Economics, University of Barcelona.

[17] McConnell C. R. and Brue S.L. (1986), *Contemporary Labor Economics*, New York: McGraw Hill Book Co.

[18] Magda I. and Szydłowski A. (2008), Płace w makro i mikroperspektywie [in:] M. Bukowski (ed.) *Zatrudnienie w Polsce 2007 – Bezpieczeństwo na elastycznym rynku pracy*, Ministry of Labour and Social Policy, Warsaw.

[19] Majchrowska A., Strawiński P., Konopczak K. and Skierska A. (2015), Gender wage gap by occupational groups in Poland, *Lodz Economics Working Papers*, 3/2015, Faculty of Economic and Sociology, University of Lodz.

[20] Matysiak A., Baranowska A. and Słoczyński T. (2010), Kobiety i mężczyźni na rynku pracy [in:] M. Bukowski (ed.) *Zatrudnienie w Polsce 2008 – Praca w cyklu życia*, Human Resources Development Center, Warsaw.

[21] Mincer J. (1997), The Production of Human Capital and the Life Cycle of Earnings: Variations on a Theme, *Journal of Labor Economics*, Vol. 15, no. 1, pp. 26–47.

[22] Myck M., Morawski L. and Mycielski J. (2007), Employment fluctuations and dynamics of the aggregate average wage in Poland, 1996-2003, *The Economics of Transition*, The European Bank for Reconstruction and Development, Vol. 15, pp. 759-779, October.

[23] Mysiková M. (2012), Gender Wage Gap in the Czech Republic and Central European Countries, *Prague Economic Papers*, Issue 3, pp. 328–346.

[24] Ñopo H. (2008), Matching as a Tool to Decompose Wage Gaps, *Review of Economics and Statistics*, 90(2), pp. 290–299.

[25] Oaxaca R. L. (1973), Male–female wage differentials in urban labor markets, *International Economic Review*, 14(3), pp. 693–709.

[26] Rokicka M. and Ruzik A. (2010), The Gender Pay Gap in Informal Employment in Poland, *CASE Network Studies and Analyses*, No. 406/2010.

[27] Rubin D. B. (1978), Multiple imputations in sample surveys. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, pp. 20–34.

[28] Słoczyński T. (2012), Próba wyjaśnienia regionalnego zróżnicowania międzypłciowej luki płacowej w Polsce, *Studia Regionalne i Lokalne*, No. 3(49).

[29] Strawiński P. (2015), Krzyżowe porównanie danych o wynagrodzeniach z polskich badań przekrojowych, *Bank i Kredyt*, Vol. 46(5), pp. 433–462.

[30] Strawiński P., Broniatowska P. and Majchrowska A. (2016), Zmiany zawodowej i sektorowej struktury zatrudnienia kobiet w Polsce w latach 2000–2013, *Gospodarka Narodowa*, 1(281), Rok LXXXVI/XXVII, styczeń–luty, pp. 73–97.

[31] Śliwicki D. and Ryczkowski M. (2014), Gender pay gap in the micro level – case of Poland, *Quantitative Methods in Economics*, Vol. XV, No. 1, pp. 159–173.

[32] Taniguchi K. and Tuwo A. (2014), New Evidence on Gender Wage Gap in Indonesia, *Asian Development Bank Working Paper*, 404.

[33] Zhang P. (2003), Multiple Imputation: theory and method, *International Statistical Review*, 71(3), pp. 581–592.

Appendix

Table 2: Average level of wages in Poland and in the 16 NUTS2 regions according to the harmonized dataset, the PLFS and the SWS datasets in 2010 (PLN) and the urbanisation ratio (%)

	Harmonized dataset			PLFS			SWS			Urbanization
	Men	Women	Total	Men	Women	Total	Men	Women	Total	ratio
DOLN	2521	2018	2287	1845	1523	1691	2894	2289	2577	69,9
KUJA	2123	1822	1989	1776	1444	1625	2444	2178	2306	60,6
LUBE	2098	1857	1984	1760	1446	1606	2348	2151	2239	46,5
LUBU	2157	1902	2042	1531	1279	1412	2540	2250	2389	63,5
LODZ	2189	1945	2077	1672	1405	1549	2534	2238	2376	63,8
MALO	2338	1999	2180	1809	1486	1649	2607	2244	2406	49,2
MAZO	2955	2508	2741	2064	1856	1967	3623	2954	3281	64,2
OPOL	2169	1869	2042	1703	1323	1533	2554	2256	2398	52,4
PODK	2050	1840	1958	1599	1367	1493	2275	2112	2192	41,3
PODL	2164	1950	2058	1810	1554	1682	2400	2257	2320	60,0
POMO	2415	2079	2264	2068	1546	1829	2798	2389	2589	65,9
SLAS	2708	1996	2388	1941	1412	1688	2976	2264	2649	77,8
SWIE	2076	1947	2020	1661	1415	1554	2360	2205	2281	45,1
WARM	2025	1852	1948	1502	1324	1421	2464	2205	2322	59,5
WIEL	2264	1962	2133	1748	1427	1607	2630	2222	2422	55,9
ZACH	2249	1968	2114	1883	1479	1684	2568	2313	2422	68,8
POLAND	2404	2041	2238	1825	1507	1676	2831	2373	2594	59,4

Source: PLFS, SWS, `http://eregion.wzp.pl/wspolczynnik-urbanizacji-dla-poszczegolnych-wojewodztw-w-latach-2010-2013` and own estimates.

Table 3: The Oaxaca-Blinder decomposition of gender wage gap in Poland and in 16 NUTS2 Polish regions in 2010 estimated on harmonised dataset

	Wages of men (ln)	Wages of women (ln)	Difference in wages	Explained part	Unexplained part	No. of observations
POLAND	7.689*** (0.002)	7.532*** (0.002)	0.157*** (0.003)	−0.013*** (0.002)	0.170*** (0.002)	101947
DOLN	7.748*** (0.007)	7.544*** (0.007)	0.204*** (0.01)	0.011 (0.007)	0.193*** (0.007)	6759
KUJA	7.577*** (0.008)	7.422*** (0.009)	0.155*** (0.012)	−0.019* (0.01)	0.174*** (0.009)	5211
LUBE	7.568*** (0.008)	7.437*** (0.008)	0.131*** (0.011)	−0.031*** (0.01)	0.161*** (0.008)	5661
LUBU	7.599*** (0.008)	7.475*** (0.009)	0.123*** (0.012)	−0.025** (0.01)	0.148*** (0.009)	4799
LODZ	7.602*** (0.007)	7.490*** (0.007)	0.112*** (0.01)	−0.034*** (0.009)	0.146*** (0.007)	7125
MALO	7.677*** (0.007)	7.522*** (0.007)	0.156*** (0.01)	0.005 (0.009)	0.151*** (0.007)	6341
MAZO	7.871*** (0.008)	7.734*** (0.007)	0.137*** (0.01)	−0.029*** (0.009)	0.165*** (0.007)	8547
OPOL	7.593*** (0.007)	7.444*** (0.008)	0.149*** (0.011)	−0.042*** (0.01)	0.191*** (0.01)	7265
PODK	7.557*** (0.007)	7.438*** (0.008)	0.118*** (0.01)	−0.001 (0.009)	0.120*** (0.008)	5691
PODL	7.602*** (0.008)	7.490*** (0.008)	0.112*** (0.011)	−0.032*** (0.01)	0.144*** (0.009)	5595
POMO	7.702*** (0.008)	7.555*** (0.008)	0.147*** (0.011)	−0.036*** (0.01)	0.183*** (0.009)	6289
SLAS	7.816*** (0.006)	7.518*** (0.006)	0.298*** (0.009)	0.082*** (0.008)	0.216*** (0.006)	8880
SWIE	7.550*** (0.007)	7.474*** (0.009)	0.077*** (0.011)	−0.096*** (0.011)	0.173*** (0.01)	5669
WARM	7.550*** (0.007)	7.461*** (0.008)	0.089*** (0.01)	−0.061*** (0.009)	0.150*** (0.008)	6117
WIEL	7.642*** (0.007)	7.487*** (0.008)	0.155*** (0.011)	−0.028*** (0.009)	0.183*** (0.008)	7143
ZACH	7.624*** (0.009)	7.493*** (0.009)	0.132*** (0.013)	−0.039*** (0.011)	0.170*** (0.01)	4855

*** indicates 1% significance level; ** indicates 5% significance level; * indicates 10% significance level.

Table 4: The Oaxaca-Blinder decomposition of gender wage gap in Poland and in 16 NUTS2 Polish regions in 2010 estimated on Polish Labour Force Survey data

	Wages of men (ln)	Wages of women (ln)	Difference in wages	Explained part	Unexplained part	No. of observations
POLAND	7.446*** (0.003)	7.248*** (0.004)	0.198*** (0.005)	−0.009*** (0.004)	0.207*** (0.005)	43957
DOLN	7.452*** (0.013)	7.260*** (0.013)	0.192*** (0.018)	−0.003 (0.015)	0.195*** (0.017)	2481
KUJA	7.374*** (0.015)	7.153*** (0.016)	0.221*** (0.022)	−0.036*** (0.019)	0.256*** (0.021)	2158
LUBE	7.382*** (0.011)	7.193*** (0.011)	0.189*** (0.016)	0 (0.014)	0.188*** (0.016)	3368
LUBU	7.340*** (0.018)	7.163*** (0.015)	0.177*** (0.023)	−0.001 (0.018)	0.179*** (0.021)	1824
LODZ	7.373*** (0.01)	7.211*** (0.009)	0.162*** (0.013)	−0.002 (0.011)	0.165*** (0.013)	3474
MALO	7.458*** (0.013)	7.241*** (0.014)	0.217*** (0.019)	0.023 (0.017)	0.194*** (0.019)	1823
MAZO	7.589*** (0.012)	7.454*** (0.013)	0.134*** (0.018)	−0.033*** (0.014)	0.167*** (0.016)	3820
OPOL	7.381*** (0.014)	7.156*** (0.013)	0.224*** (0.019)	−0.019 (0.017)	0.244*** (0.018)	3095
PODK	7.322*** (0.012)	7.159*** (0.013)	0.164*** (0.018)	−0.004 (0.015)	0.168*** (0.017)	2085
PODL	7.423*** (0.01)	7.256*** (0.01)	0.167*** (0.014)	−0.047*** (0.012)	0.213 (0.013)	3743
POMO	7.57*** (0.014)	7.316*** (0.013)	0.259*** (0.019)	0.002 (0.016)	0.257 (0.019)	2923
SLAS	7.518*** (0.012)	7.199*** (0.013)	0.319*** (0.018)	0.072*** (0.015)	0.247*** (0.016)	2977
SWIE	7.325*** (0.013)	7.174*** (0.013)	0.150*** (0.018)	−0.038*** (0.017)	0.188*** (0.018)	2242
WARM	7.407*** (0.012)	7.234*** (0.012)	0.173*** (0.017)	−0.036*** (0.013)	0.208*** (0.014)	3050
WIEL	7.353*** (0.014)	7.163*** (0.014)	0.190*** (0.019)	−0.016 (0.016)	0.206*** (0.017)	2887
ZACH	7.435*** (0.016)	7.168*** (0.016)	0.266*** (0.023)	0.027 (0.02)	0.239*** (0.023)	2007

*** indicates 1% significance level; ** indicates 5% significance level; * indicates 10% significance level.

Table 5: The Oaxaca-Blinder decomposition of gender wage gap in Poland and in 16 NUTS2 Polish regions in 2010 estimated on Structure of Wages and Salaries data

	Wages of men (ln)	Wages of women (ln)	Difference in wages	Explained part	Unexplained part	No. of observations
POLAND	7.933*** (0.001)	7.807*** (0.001)	0.126*** (0.002)	−0.015*** (0.001)	0.142*** (0.002)	681747
DOLN	8.005*** (0.004)	7.813*** (0.003)	0.193*** (0.005)	0.043*** (0.004)	0.150*** (0.005)	55481
KUJA	7.862*** (0.004)	7.774*** (0.004)	0.088*** (0.006)	−0.021*** (0.005)	0.109*** (0.006)	33942
LUBE	7.829*** (0.005)	7.766*** (0.004)	0.062*** (0.006)	−0.021*** (0.005)	0.084*** (0.006)	30496
LUBU	7.909*** (0.006)	7.814*** (0.006)	0.095*** (0.008)	−0.006 (0.007)	0.101*** (0.008)	16013
LODZ	7.875*** (0.005)	7.791*** (0.004)	0.084*** (0.006)	−0.017*** (0.005)	0.101*** (0.006)	39305
MALO	7.912*** (0.004)	7.798*** (0.003)	0.114*** (0.005)	0.002 (0.004)	0.111*** (0.005)	56962
MAZO	8.155*** (0.003)	8.023*** (0.003)	0.133*** (0.004)	−0.014*** (0.003)	0.146*** (0.004)	126280
OPOL	7.900*** (0.007)	7.796*** (0.007)	0.104*** (0.01)	−0.025*** (0.008)	0.129*** (0.01)	14009
PODK	7.797*** (0.004)	7.748*** (0.004)	0.049*** (0.006)	−0.020*** (0.005)	0.070*** (0.006)	34916
PODL	7.858*** (0.006)	7.823*** (0.005)	0.035*** (0.008)	−0.038*** (0.006)	0.073**** (0.009)	18430
POMO	7.976*** (0.005)	7.857*** (0.004)	0.118*** (0.006)	−0.018*** (0.005)	0.137*** (0.006)	35687
SLAS	8.060*** (0.003)	7.808**** (0.003)	0.252*** (0.004)	0.072*** (0.003)	0.180*** (0.004)	92969
SWIE	7.839*** (0.006)	7.792*** (0.006)	0.047*** (0.008)	−0.056*** (0.007)	0.103*** (0.008)	17463
WARM	7.866*** (0.006)	7.793*** (0.005)	0.073*** (0.008)	−0.039*** (0.006)	0.112*** (0.008)	21461
WIEL	7.920*** (0.003)	7.769*** (0.004)	0.151*** (0.005)	−0.007* (0.004)	0.158*** (0.005)	66264
ZACH	7.904*** (0.006)	7.839*** (0.005)	0.065*** (0.008)	−0.058*** (0.006)	0.123*** (0.007)	22069

*** indicates 1% significance level; ** indicates 5% significance level; * indicates 10% significance level.

Table 6: The Oaxaca-Blinder decomposition of gender wage gap in occupational groups with masculinisation ratio between 0.4 and 0.6 in Poland and in 16 NUTS2 Polish regions in 2010 estimated on harmonised dataset

	Wages of men (ln)	Wages of women (ln)	Difference in wages	Explained part	Unexplained part	No. of observations
POLAND	7.744*** (0.006)	7.602888 (0.006)	0.143*** (0.008)	−0.028*** (0.007)	0.171*** (0.006)	13780
DOLN	7.790*** (0.023)	7.559*** (0.015)	0.231*** (0.028)	−0.008 (0.024)	0.239*** (0.016)	962
KUJA	7.585*** (0.021)	7.486*** (0.023)	0.099*** (0.032)	−0.049* (0.027)	0.148*** (0.022)	711
LUBE	7.590*** (0.023)	7.518*** (0.022)	0.073** (0.032)	−0.084*** (0.027)	0.157*** (0.02)	750
LUBU	7.602*** (0.021)	7.484*** (0.02)	0.119*** (0.029)	−0.01 (0.025)	0.129*** (0.02)	750
LODZ	7.627*** (0.021)	7.546*** (0.018)	0.082*** (0.028)	−0.040* (0.024)	0.122*** (0.016)	1119
MALO	7.704*** (0.02)	7.575*** (0.023)	0.129*** (0.03)	−0.056** (0.027)	0.185*** (0.017)	788
MAZO	7.976*** (0.022)	7.856*** (0.02)	0.120*** (0.029)	−0.046* (0.026)	0.166*** (0.015)	1344
OPOL	7.706*** (0.025)	7.466*** (0.019)	0.240*** (0.031)	0.003 (0.026)	0.237*** (0.021)	994
PODK	7.572*** (0.024)	7.465*** (0.024)	0.107*** (0.035)	−0.042 (0.031)	0.149*** (0.018)	617
PODL	7.610*** (0.025)	7.551*** (0.021)	0.059* (0.033)	−0.137*** (0.028)	0.196*** (0.02)	695
POMO	7.753*** (0.025)	7.615*** (0.022)	0.138*** (0.034)	−0.033 (0.028)	0.171*** (0.023)	876
SLAS	7.836*** (0.02)	7.608*** (0.018)	0.228*** (0.027)	0.025 (0.024)	0.203*** (0.013)	1099
SWIE	7.562*** (0.024)	7.551*** (0.028)	0.011 (0.036)	−0.119*** (0.033)	0.130*** (0.023)	651
WARM	7.606*** (0.023)	7.540*** (0.022)	0.066** (0.032)	−0.043 (0.027)	0.110*** (0.02)	734
WIEL	7.724*** (0.019)	7.562*** (0.02)	0.162*** (0.027)	−0.040* (0.025)	0.203*** (0.014)	1114
ZACH	7.685*** (0.027)	7.447*** (0.025)	0.239*** (0.037)	0.018 (0.032)	0.221*** (0.022)	576

*** indicates 1% significance level; ** indicates 5% significance level; * indicates 10% significance level.

Table 7: Ñopo decomposition of gender wage gap in Poland and 16 NUTS2 Polish regions in 2010 estimated on harmonised dataset

	D	D0	DM	DF	DX
POLAND	0.158	0.112	0.004	-0.040	0.081
DOLN	0.249	0.159	0.101	-0.060	0.050
KUJA	0.165	0.061	0.137	-0.074	0.042
LUBE	0.124	0.024	0.075	0.002	0.023
LUBU	0.135	0.065	0.081	-0.024	0.013
LODZ	0.126	0.130	0.072	-0.130	0.053
MALO	0.166	0.082	0.125	-0.059	0.019
MAZO	0.166	0.144	-0.052	-0.016	0.091
OPOL	0.158	0.136	0.093	-0.114	0.042
PODK	0.107	0.011	0.103	-0.018	0.012
PODL	0.107	0.058	0.058	-0.062	0.053
POMO	0.162	0.062	0.107	-0.044	0.037
SLAS	0.351	0.130	0.232	-0.099	0.088
SWIE	0.055	0.098	0.041	-0.101	0.017
WARM	0.088	0.067	-0.028	-0.013	0.063
WIEL	0.139	0.133	0.094	-0.110	0.022
ZACH	0.140	0.146	0.041	-0.095	0.048