Received 6 July 2016
Revised 23 March 2017
Accepted 22 July 2017

Occupational segregation and wage differences: the case of Poland

Pawel Strawinski
University of Warsaw, Warsaw, Poland
Aleksandra Majchrowska
Department of Economics, University of Lodz, Lodz, Poland, and
Paulina Broniatowska
Szkola Glowna Handlowa w Warszawie, Warszawa, Poland

Abstract

Purpose – The purpose of this paper is to analyse the relation between occupational segregation and the gender wage differences using data on three-digit occupational level of classification. The authors examine whether a statistically significant relation between the share of men in employment and the size of the unexplained part of the gender wage gap exists.
Design/methodology/approach – Traditional Oaxaca (1973) – Blinder (1973) decomposition is performed to examine the differences in the gender wage gaps among minor occupational groups. Two types of reweighted decomposition – based on the parametric estimate of the propensity score and non-parametric proposition presented by Barsky *et al.* (2002) – are used as the robustness check. The analysis is based on individual data available from Poland.
Findings – The results indicate no strong relation between occupational segregation and the size of unexplained differences in wages. The unexplained wage differences are the smallest in strongly female-dominated and mixed occupations; the highest are observed in male-dominated occupations. However, they are probably to a large extent the result of other, difficult to include in the econometric model, factors rather than the effects of wage discrimination: differences in the psychophysical conditions of men and women, cultural background, tradition or habits. The failure to take them into account may result in over-interpreting the unexplained parts as gender discrimination.
Research limitations/implications – The highest accuracy of the estimated gender wage gap is obtained for the occupational groups with a similar proportion of men and women in employment. In other male- or female-dominated groups, the size of the estimated gender wage gaps depends on the estimation method used.
Practical implications – The results suggest that decreasing the degree of segregation of men and women in different occupations could reduce the wage differences between them, as the wage discrimination in gender balanced occupations is the smallest.
Originality/value – To the best of the authors' knowledge, this study is one of the few conducted at such a disaggregated level of occupations, and one of few studies focused on Central and Eastern European countries and the first one for Poland.

Keywords Gender wage gap, Poland, Occupational segregation, Wage discrimination
Paper type Research paper

1. Introduction

Results of different studies suggest that gender differences in wages are a well-established feature of contemporary labour markets in most developed economies (Blau and Kahn, 2003). However, only a fraction of the wage gap between men and women can be explained by the

The paper was financed by National Science Center, Poland (Grant Nos. 2013/09/B/HS4/01304 and 2015/19/B/HS4/03231).

differences in their productivity potential. Researchers have therefore sought other explanations.

As one of the potential sources of gender wage differences, an occupational segregation is proposed. The research for various countries indicates that there is a relation between occupational segregation and level of wages but the shape of this relation is not obvious. Some research underlines that the relation is linear: predominately female occupations pay less than male occupations (England *et al.*, 1994; de Ruijter *et al.*, 2003; de Ruijter and Huffman, 2003). Other research reveals a non-linear relation between sex composition of occupation and wages, where the highest wages for both men and women are earned in gender balanced occupations (Magnusson, 2013). Brynin and Perales (2015) found evidence that wage effects of occupational segregation evolves into more complex processes, including differing impact for graduate and non-graduate workers.

Several studies have analysed the link between occupational segregation and the size of the gender wage gap across countries, although in many cases the conclusions are inconclusive. Blau and Kahn (2003) and Dolado *et al.* (2002) found some weak evidence that the gender wage gap and occupational segregation are positively correlated. Other research indicated a non-linear relation: gender wage gap is relatively low in mixed occupations (Hakim, 1998). There is also considerable disagreement whether this gap actually reflects gender discrimination in the labour market or rather productivity differences among workers. Most of studies use the aggregate data to measure the occupational segregation of workers as one of the explanatory variables in the gender wage gap regression.

This paper aims to add to the empirical literature by analysing the relation between occupational segregation of men and women and the size of differences in wages in detailed occupational groups. The research question is:

RQ1. Whether in some groups of occupations (male or female-dominated) the unexplained differences in wages of men and women are significantly higher than in others in a systematic way?

Based on theoretical and empirical literature the authors expect that the differences in wages between men and women should be higher in occupations with a permanent need to improve the qualifications, as a significant number of women have to interrupt the job carrier to give birth and take care of young children. Moreover, for the same reason gender wage gaps are expected to be higher in time-consuming occupations.

The study deepens the existing literature by focusing on the analysis of detailed occupations. It uses data about wages of workers and their characteristics at a three-digit level of International Standard Classification of Occupations (ISCO) for Poland. The authors estimate the gender wage gap for each of the 98 minor occupational groups and evaluates to what extent the differences are justified by workers' characteristics. Having done this, the paper examines if a statistically significant relation between the share of men and women in occupation and the size of the unexplained part of the wage gap exists. To the best of the authors' knowledge, this study is one of the few conducted at such a disaggregated level of occupations, one of few studies focused on Central and Eastern European countries and the first one for Poland.

The unexplained part of the wage gap is usually treated in the literature as an estimate of wage discrimination. However, as is well known, the discrimination effects can be overstated if some important explanatory variables are missing. On the contrary, the unexplained part will also understate discrimination if some of the explanatory variables have themselves been influenced by discrimination (Blau and Kahn, 2016). One of the strategies used to overcome these problems is to analyse more homogenous samples. This is the approach adopted in this paper. As the major occupational groups are less segregated by sex than detailed occupations, it can lead to underestimation of the occupational gender wage gap (de Ruijter *et al.*, 2003). The authors use, therefore, occupational measures that are as detailed as possible. Blau and Kahn (2016) confirm that omitted-variable biases are less severe in such homogeneous samples.

The standard methodological approach advised in the literature is followed. The adjusted wage gap for each of the occupational group is obtained based on the extended Mincer equation estimates. To answer the question to what extent the gender wage gap can be explained by different characteristics of males and females, the Oaxaca (1973) and Blinder (1973) decomposition is used. The authors then try to answer the question whether statistically significant relation between the unexplained part of the gender wage gap and share of men in the occupational group exists. To check the robustness of the results two types of reweighted decomposition are used. The first is based on the parametrically estimated odds of employees being women, and the second on non-parametrically estimated odds.

The structure of the paper is as follows. Section 2 reviews the literature related to the occupational segregation and the differences in wages between men and women. In Section 3, some methodological aspects of the estimations of the gender wage gap are discussed. Section 4 presents the data used and outlines the empirical strategy. Section 5 presents the empirical results, and Section 6 concludes the study.

2. Literature review

Numerous empirical studies confirm the negative relation between occupational segregation and level of wages: workers in female-dominated occupations earn lower wages than workers in male-dominated occupations. Most of the research explains these differences by adopting at least one of three main mechanisms (de Ruijter *et al.*, 2003).

The first mechanism is based on the Human Capital Theory, which explains the differences in wages in male- and female-dominated occupations, according to individual differences in their human capital (Mincer and Polachek, 1974). Women have a comparative advantage in the domestic sphere (as they are the birth-givers and usually the main caretakers of the children), and therefore invest less than men in human capital. For the same reason, employers generally invest less in the human capital of female employees. Therefore, to some extent, the existence of the gender wage gap may reflect women's tendency to choose occupations with more flexible working hours or smaller wage penalties for time-out-of-work, due to the necessity to balance work and family life (motherhood). Consequently, not only occupational segregation, but also the traditional perception of roles in the family, may disadvantage women in the labour market (Blau, 1996).

The second mechanism is based on the Crowding Theory, which explains the gender differences in wages according to women's restricted access to some occupations (Bergmann, 1986). Wages in each of the segments of the labour market are set according to the principles of supply and demand. Due to some restrictions, the amount of supply in female-dominated occupations is artificially high, which lowers the average level of their wages. In this theory, men and women compound "non-competing groups", which means that there are boundaries between segments that prevent a reallocation of men and women across the different segments.

The third mechanism is based on gender bias, according to which male and female labour is differently evaluated. Research shows that both men and women attribute less value to work performed by women. Therefore, it appears that occupations that require typically female skills yield lower wages (England *et al.*, 1994).

The analyses concerned with the gender wage gap in different countries indicate that wage differences vary substantially across countries (Blau and Kahn, 2003) among both experience levels and/or income quantiles. In many countries the gender wage gap is wider at the top or at the bottom of the wage distribution, suggesting the existence of "glass ceilings" or "sticky floors". There are also substantial gender differences across firms, occupations and industries.

As far as the relation between occupational segregation and gender wage gap is concerned, the empirical literature is not conclusive. Several researches confirm that the gender wage gap is larger in occupations requiring highly educated workers (e.g. Evertsson *et al.*, 2009). Dolado *et al.* (2002) and Blau and Kahn (2003) regressed the gender wage gap in several countries on (among others) a measure of the degree of occupational segregation by gender and found some weak positive relation. De Ruijter and Huffman (2003) using Dutch data found compelling evidence of men's net wage advantage across all occupations; however, this advantage decreased with a the percentage female in an occupation. Xiu and Gunderson (2015) analysed the gender earnings gap in China and found considerable heterogeneity in the effect of occupational discrimination within the sub-occupations in the different broad occupational groups.

The impact of the feminisation of occupational groups on the gender wage gap was analysed by Grönlund and Magnusson (2013). They aimed to test to what extent the gender wage gap in different occupational groups can be attributed to one of the above mentioned three theories that explain differences in the earnings of men and women. According to their research, the negative correlation between the percentage of female employees and hourly wages is inadequately explained by any of the three mechanisms. These results confirm the earlier findings of de Ruijter *et al.* (2003).

Barón and Cobb-Clark (2010) analysed the link between occupational segregation and the gender wage gap in the public and private sector in Australia. They find that gender wage gap is flat across the wage distribution in public sector and increasing with the wage level increases in private sector. Moreover, the gender wage gap among low-paid workers is totally explained by differences in wage-related characteristics. However, this is largely unexplained among high-paid workers in both sectors. Similar results were found by Murphy and Oesch (2015) for Britain, Germany and Switzerland: the wage penalties associated with working in a female occupation were larger in private than public sector.

Nonetheless, there is little research regarding the Central and Eastern European countries. Among the few, one can mention Jurajda (2003) focused on the Czech Republic and Slovakia, and Banerjee (2014) on Macedonia. Jurajda (2003) stated that employment segregation is related to over one-third of the overall gender wage differences in both countries analysed. Banerjee (2014) confirmed that women are over-represented in female-dominated occupations and that feminisation in occupations has a negative impact on earnings.

As far as Poland is concerned, the majority of the previous studies on the Polish labour market have concentrated on estimating the gender wage gap for the whole economy with occupational groups or specific industries as explanatory variables. The estimates vary significantly due to differences in databases and the method of estimation (see Majchrowska and Strawiński, 2016; Goraus *et al.*, 2017). Many studies also confirmed that gender wage gap in Poland is significantly differentiated across regions, age groups or wage quantiles. The previous research on the gender wage gap in Poland shows that part of the wage differences might be explained by differences in working hours. Moreover, the majority of the research indicates that only small part of the wage differences can be explained by differences in personal and employment characteristics. The unexplained component of the gender wage gap is much higher than the explained one and suggests the existence of some discrimination of women in the labour market in Poland.

3. Different views and interpretations of the gender wage gap

The raw wage gap can be a misleading indicator of gender inequality in the labour market as it does not take into account the existing differences between male and female employees in productivity potential. It may either overestimate the extent of discrimination if women are systematically less qualified than are men, or underestimate it. Therefore, for scientific and policy-making purposes, the numerous studies since the early 1970s focused on gender differentials have been decomposed into a part that can be explained by differences in human capital endowments, and an unexplained part (or a part explained by the difference in the value attached by the labour market to equal male and female endowments). The latter part is treated as an estimate of gender discrimination in the labour market. The decomposition was pioneered by Blinder (1973) and Oaxaca (1973), and is usually based on the Mincer wage equation (Mincer, 1974; Mincer and Polachek, 1974). In this approach logarithmic wages are regressed against individual characteristics that are relevant from the perspective of the labour market, such as years of education, work experience or time-out-of-work. In line with the underlying human capital model (Becker, 1964) the coefficients in the wage regression are interpreted as returns on investment (or loss from disinvestment). It is also customary to include among the explanatory variables some of the job characteristics such as occupation or industry.

Despite its simplicity, this approach poses several econometric problems. First of all, the estimates of discrimination effect are conditional upon the control variables included in the wage equation. If the gender differences in potential productivity are not fully accounted for by the control variables, the unexplained residual is likely to be biased upwards. On the other hand, if the explanatory variables in the wage equation are themselves the result of discrimination, the unexplained residual will be underestimated. This may be the case with job characteristics if occupational gender segregation leads to the overrepresentation of women in lower-paid professions. Another potential source of estimation bias is the unobserved heterogeneity problem resulting in the endogeneity of regressors in the wage equation. Some unobserved individual characteristics, e.g. mental abilities, affect both wages and some of the explanatory variables, e.g. educational attainment, which may lead to the inconsistency of the ordinary least squares (OLS) estimates. The estimates may also be inconsistent due to sample selection bias or measurement errors. For instance, the decision to supply labour – especially in the case of women – may be conditional on several factors, such as expected wage, partner's income and number of children.

These identification problems are widely recognised in the literature, but there is no consensus on how to handle them (Kunze, 2008). The choice of controls is often restricted by the data set available to the researcher. The endogeneity of regressors may be corrected using an instrumental variables estimator, but finding valid instruments – especially given data constraints – often poses major problems. If panel data are available, the unobserved individual effect may be captured by means of a fixed-effects estimator, and some transformations of endogenous variables may be used as instruments by applying the estimators of Hausman and Taylor (1981), Arellano and Bond (1991) or Arellano and Bover (1995). Despite the potential inconsistency of the estimates, most studies based on cross-sectional data sets nevertheless apply the OLS estimator (Kunze, 2008).

Performing the Oaxaca (1973) and Blinder (1973) decomposition requires the estimation of a hypothetical wage structure, which is called "counterfactual wage distribution" in the treatment effects literature. There exists an explicit link between these two strands. Recent applications of treatment-effect related methods have used reweighting (Barsky *et al.*, 2002; Kline, 2011; Słoczyński, 2015), matching on covariates (Black *et al.*, 2006, 2008; Ñopo, 2008) and propensity score matching (Frölich, 2007), to study intergroup differences in various outcomes.

The conventional Oaxaca-Blinder approach based on the OLS regression requires a parametric assumption about the form of the conditional expectation function. Barsky *et al.* (2002) showed the importance of the linearity assumption of the conditional expectation function for the validity of the Oaxaca-Blinder decomposition. If the assumption does not hold, then the Oaxaca-Blinder decomposition does not provide consistent estimates of the wage structure and composition effect. The second problem is the lack of a common support. As a solution, Barsky *et al.* (2002) proposed reweighting using a non-parametric estimate of the propensity score. The reweighting procedure involves the stratification and equalisation of distributions. The reweighting approach itself is a modification of the standard Oaxaca-Blinder decomposition. Instead of a direct comparison of two groups, the group of interest is compared to the weighted average of the comparison group.

Kline (2011) pointed out that the classic regression-based estimator of counterfactual means studied by Oaxaca and Blinder constitutes a propensity-score reweighting estimator, based upon a linear model for the conditional odds of being treated as a linear function of the covariates. The weighted average of the control outcomes, with weights proportional to the conditional odds of treatment, identifies the counterfactual mean of the treated population. One must note, however, that the reweighting procedure may have some undesirable statistical properties in small samples, especially when there is a problem of common support (for more details, see Fortin *et al.*, 2011).

The available alternative is to use matching or propensity score matching. These techniques solve the common support problem and relax the linearity assumption. The former method, also called non-parametric matching, has been used in the context of gender wage gaps, for instance by Black *et al.* (2006, 2008) and Ñopo (2008). The latter, for example, was used by Frölich (2007). One can think of matching as equalising the distribution of covariates (Black *et al.*, 2008). In the case of non-parametric matching, the counterfactual distribution is constructed as the weighted average wage gap for groups formed upon values of demographic characteristics. In the case of propensity score matching, the counterfactual distribution is constructed from the wages of matched (i.e. similar in demographic characteristics) individuals. Recently, Słoczyński (2015) provided an alternative solution to the comparison group choice problem by developing a new version of the Oaxaca-Blinder decomposition, whose unexplained component recovers the average wage gap.

Based on the results of the discussed research, the authors decided to use the Oaxaca-Blinder decomposition based on the extended Mincer-type wage equation as the baseline model. To strengthen our findings, two reweighting schemes were applied as robustness checks. In the first, the weights are based on a parametric estimate of the propensity score, while in the second the authors follow the non-parametric proposition of Barsky *et al.* (2002).

4. Data and empirical strategy

The data on wages and individual characteristics of employees used in this paper were obtained from the Structure of Wages and Salaries by Occupations (SWS) database in 2012 provided by the Central Statistical Office in Poland. The survey is carried out with biennial frequency. It covers entities of the national economy whose employees exceed nine persons from both public and private sector. The database includes both full- and part-time employees who worked for the whole month in October 2012.

The advantage of the SWS survey is the high reliability of its data on wages, which are reported by the accounting departments along with the number of hours worked, and the size of the database. As of 2012, the SWS survey covered 12.8 per cent of the total number of enterprises in Poland whose employees exceeded nine persons. The total number of observations in the sample is approximately 725,000[1].

The database contains information on individuals' wages and several personal characteristics, such as gender, age, level of education, work tenure and occupational group. It also includes some employers' characteristics, such as ownership sector, size of the enterprise and its location, as well as the NACE section. All the variables were used as controls in the Mincer equation.

The analysis was conducted at minor (three-digit) occupational group level in order to study homogenous group of workers who perform similar tasks at work. This level of analysis is suggested in the literature (de Ruijter *et al.*, 2003). For instance, all health professionals are allocated in the same sub-major (two-digit) occupational group. However, evidently medical doctors (minor 221 group) perform different tasks than nursing professionals (minor 222 group).

The empirical strategy was divided into several steps. In the first step, data cleaning was performed. Those observations with a declared hourly base wage below the national minimum of PLN8.15 were removed from the sample. The occupational groups are defined according to the ISCO-08 classification. By definition, the concept of occupation is defined as a "set of jobs whose main tasks and duties are characterised by a high degree of similarity[2]". The occupations are divided into groups according to the nature of performed task (i.e. specialist, craft and manual) and formal level of education.

Additionally, at least 100 observations for each gender for each occupational group in the sample were required. This condition is necessary for computational reason. This step cuts down the total number of 124 separate three-digit occupational groups in the sample to 98 (less than 10 per cent of the total observations were discarded). The main reason for removal was a huge male domination or the scarcity of the occupation in the population.

The second step was the estimation of the Mincer-type wage equation for each minor (three-digit) occupation group separately. The functional form of the equation is as follows:

$$\ln(w_i) = \beta_0 + \beta_1 \mathrm{EDU}_i + \beta_2 \mathrm{EXP}_i + \beta_3 \mathrm{EXP}_i^2 + \sum_{k=1}^{K} \gamma_k X_{k,i} + \mathrm{INTERACTIONS}_i + \varepsilon_i$$

where w_i is the base wage (PLN); EDU is the number of years of formal education (in years); EXP is the number of years of work experience (in years); and X is the job characteristics:

- size of the firm (small, medium, large, very large);
- ownership sector (1 for public, 0 for private); and
- NACE section.

INTERACTIONS include interactions of years of education with sector- and firm-size dummies and experience with sector dummies.

In the third step, the Oaxaca-Blinder decomposition for each of 98 minor occupational groups was performed. To check the robustness of the results two reweighting schemes were applied. In the first, the weights are based on a parametric estimate of the propensity score, while in the second the authors followed the non-parametric proposition of Barsky *et al.* (2002). The results of the decomposition allowed us to answer a question if statistically significant relations between occupational segregation and an unexplained part of a wage gap exist.

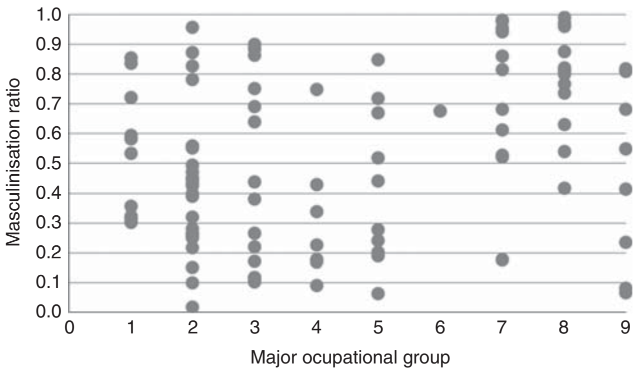
5. Occupational segregation, gender wage gaps and discrimination: empirical results

Looking at the sample of 98 minor, three-digit level occupational groups in Poland, there are 39 male-dominated groups where more than 60 per cent of employees are males, 37 female-dominated groups where men constitute less than 40 per cent of workers and 22 mixed groups with a masculinisation ratio[3] between 40 and 60 per cent[4].

The seventh and eighth major occupational groups are the most masculinised ones: they contain almost half of all minor male-dominated groups (Figure 1). The highest shares of men in employment (98-99 per cent) are noted among heavy truck and bus drivers (833), electrical equipment installers and repairers (741) and sheet and structural metal workers, moulders and welders and related workers (721).

The most feminised major occupational group is the second one (professional workers). The lowest shares of men in employment are noted among nursing and midwifery professionals (222; masculinisation ratio 0.02) and among cashiers and ticket clerks (532) and domestic, hotel and office cleaners and helpers (911; 0.07).

Among occupations with almost balanced men-to-women ratio in employment (47-52 per cent), there are: life science professionals (213), physical and earth science



Source: Authors' own calculations

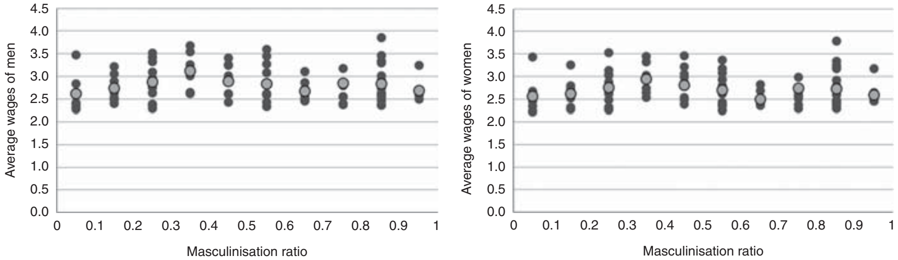
Figure 1. The distribution of masculinisation ratio among minor occupational groups in Poland in 2012

professionals (211), building and housekeeping supervisors (515) and food processing and related trades workers (751).

The differences in the degree of masculinisation within the major occupational groups are large. In almost each of the major occupational groups, there are female- and male-dominated as well as mixed occupations. This finding supports our empirical approach to conduct analysis at minor occupational group level. The highest dispersion of the masculinisation ratio is noted in second major occupational group (professionals), where the indicator varies from 0.02 (nursing and midwifery professionals, 222) to 0.96 (electrotechnology engineers, 215; Figure 1).

The level of average base salaries per hour of men and women among occupational groups shows a very clear pattern. The highest wages of both of men and women are observed in the groups where the highest level of education and qualifications is required: there are the first (managers) and second (professionals) major occupational groups (Table AI). The lowest wages appear in groups where no specific skills are required: the ninth (elementary occupations) and fifth (service and sales workers) major occupational groups.

Comparing the average base salaries per hour of men and women within the masculinisation ratio intervals (Figure 2), one can observe that both men and women have the highest average wages in occupations with a masculinisation ratio between 0.31 and 0.40. The lowest average wages of both men and women are recorded among strongly female-dominated occupations (masculinisation ratio at 0.10 or lower), and among groups with a masculinisation ratio of 0.61-0.70, as well as among the most masculinised occupational groups (0.91 and more). However, within each interval strong differences in wages can also be observed (Figure 2).



Note: Bigger black-edged circles denote the average wage level in a given interval
Source: Authors' own calculations

Figure 2. Average base wages per hour of men (left) and women (right panel) in the three-digit level occupational groups (log of PLN), within the masculinisation ratio intervals in Poland in 2012

In the next step, the gender wage gaps for each of the 98 occupational groups at the three-digit level of ISCO classification in 2012 were estimated using the Oaxaca-Blinder decomposition. The differences in the size of the gender wage gaps among occupational groups are significant. They vary from 45 per cent in the group of life science technicians and related associate professionals (314), where the highest disproportions in average base salary per hour between men and women are noted, to −21 per cent in the group of authors, journalists and linguists (264), where the average salaries of women are significantly higher than those of men.

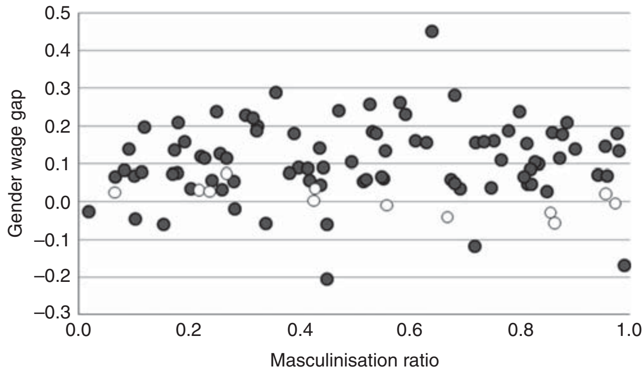
In most of the occupational groups analysed, the estimated gender wage gaps are positive and statistically different from zero. In only 12 (out of the 98) occupational groups, the estimated gender wage gaps are statistically insignificant. There are 13 groups where the differences in base salaries per hour between men and women exceed 20 per cent. Over half of the groups with the highest wage differences between men and women are placed among managers (first major group) and professionals (second major group).

On the other hand, in ten minor occupational groups the average base salaries per hour for women are statistically higher than the average wages of men. The highest positive wage differences for women were observed in the groups of authors, journalists and linguists (264; gender wage gap −0.21), heavy truck and bus drivers (833; gender wage gap −0.17) and travel attendants, conductors and guides (511; gender wage gap −0.12).

Looking at the size of the estimated gender wage gaps and the masculinisation ratio in the occupational groups analysed, some non-linear, but weak relation between them can be found. The smallest average wage differences between men and women can be observed among strongly male- and female-dominated occupations (masculinisation ratio between 0.00-0.10 and 0.91-1.00), and among the groups with an almost balanced men-to-women ratio in employment (masculinisation ratio between 0.41 and 0.50; Figure 3).

The Oaxaca-Blinder decomposition of the gender wage gaps among the 98 occupational groups analysed shows that, in most of the cases, the differences in endowments of men and women cannot explain the differences in wages. Only in 25 occupational groups the explained part of the wage gap is positive and statistically different from zero; however, in most cases it does not exceed 5 per cent (Table AI).

In the last step, the authors tried to find out if a statistically significant relation between the occupational segregation and size of the unexplained part of the gender wage gap exists. In other words – if it is assumed that unexplained part of the gender wage gap can be interpreted as a measure of discrimination – the authors wanted to check



Note: Non-significant observations are marked with white circles
Source: Authors' own calculations

Figure 3.
The relation between estimated gender wage gaps and masculinisation ratio in occupational groups at the three-digit level of classification in Poland in 2012

if there is any regularity: are women more discriminated in male-dominated or female-dominated occupations?

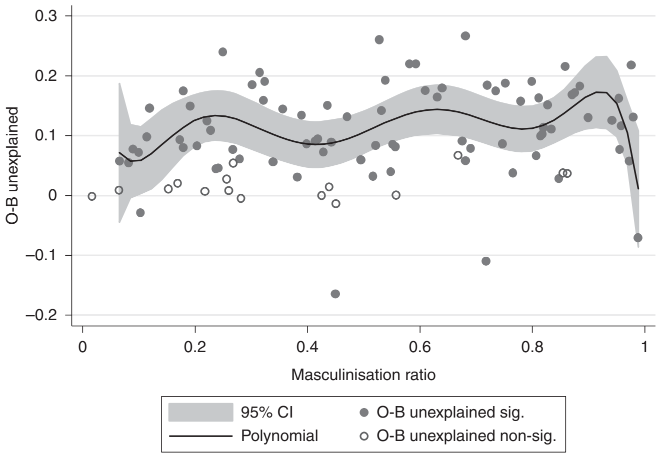
The results of the decomposition show that in the majority of the occupational groups, the unexplained part of the gender wage gap is positive and significant, indicating possible discrimination effects. The size of the unexplained part differs significantly among occupational groups: from −0.16 to 0.27 (see Figure 4).

To identify in which occupational groups the discrimination effect is the strongest, a local polynomial regression of the unexplained part of the gender wage gap on the masculinisation ratio in minor occupational group was fitted. Only statistically significant values were included in the model.

The relation between the unexplained part of the gender wage gap and the average education could be easily ruled out as the present-day labour market faces problem of over-education. This phenomenon strongly affects wages in service related occupations and does not affect the wages of specialist. Moreover, the data show that no relation between these two variables exists (see Figure A1).

The relation turned out to be evidently non-linear. The smallest unexplained part of the gender wage gap is recorded in the most feminised occupations (masculinisation ratio below or at 10 per cent), and in occupational groups with an almost balanced men-to-women ratio (masculinisation ratio between 0.41 and 0.5; see Figure 4). In these groups, the unexplained part of the gender wage gap is around 0.03-0.05, on average. The highest unexplained part of the wage gaps is noted in occupational groups with a masculinisation ratio around 0.35 and 0.75.

Looking more deeply at the groups where the unexplained differences in wages were the highest one can find some patterns. Most of the occupations with the highest unexplained part are the male-dominated ones: among 28 minor occupational groups with unexplained differences higher than 15 per cent, 17 of them were the male-dominated ones. They are mostly the occupational groups located in the seventh and eighth major occupational groups. On the contrary, among the occupational groups with relatively low unexplained differences in wages (5-10 per cent), 13 out of 30 groups were the female-dominated ones. These results seem to confirm that the relation between masculinisation ratio and unexplained differences in wages is non-linear and (slightly) positive.



Kernel = Epanechnikov, Degree = 7, Bandwidth = 0.4, pWidth = 0.6

Source: Authors' own calculations

Figure 4. The relation between the unexplained parts of the gender wage gap estimated with the Oaxaca-Blinder decomposition and masculinisation ratio

Robustness check

To check the robustness of the results obtained with the Oaxaca-Blinder decomposition two reweighting schemes were applied (the results are presented in Table AI). In the first, all the variables included in the decomposition equation were used to estimate the propensity of employees being women. Hence, based on this propensity score, the estimated odds of being a woman were used as a value to reweight the sub-sample of men. In the second, the authors followed the proposition of Barsky *et al.* (2002). For each half-decile of the distribution of women's wages, a ratio of women to men was calculated; then this ratio was used to reweight the distribution of men's wages. After applying both schemes, the Oaxaca-Blinder decomposition of wages was performed.

When the size of the gender wage gaps estimated with all three methods is compared, one can notice that the differences are relatively low in the case of estimates with the propensity score method (Figure A2), but relatively high in the case of the non-parametric proposition of Barsky *et al.* (2002). For most of the occupational groups, the estimated gaps using the latter are lower than those obtained with the traditional Oaxaca-Blinder decomposition. Looking at the differences in the gender wage gaps estimated using the three methods and the masculinisation ratio, one can notice that the lowest differences between estimates are for the strongly feminised occupational groups (masculinisation ratio at or below 0.1), and for the occupational groups with an almost balanced men-to-women ratio (masculinisation ratio between 0.41 and 0.50).

A similar pattern can be observed when one analyses the differences in the estimates of the unexplained component of the gender wage gaps (Figure A3). For most of the occupational groups, the estimates using the propensity score method are similar to those obtained with the traditional Oaxaca-Blinder decomposition. For most of the occupational groups, the unexplained part of the gender wage gaps estimated with the non-parametric method of Barsky *et al.* (2002) are significantly lower than the ones obtained with the Oaxaca-Blinder decomposition. Looking at the differences in the unexplained part of the gender wage gaps estimated with the three methods and the masculinisation ratio, one can notice that the lowest differences between estimates are again noted for the strongly feminised occupational groups (masculinisation ratio at or below 0.1), and for the occupational groups with an almost balanced men-to-women ratio.

6. Conclusions

The aim of the paper was to analyse the relation between occupational segregation and the gender wage differences using detailed data on occupations. The authors estimated the gender wage gaps separately for each of the 98 occupational groups at three-digit level of classification in Poland in 2012 and checked to what extent the differences in wages are justified by workers' characteristics. The relation between the size of the unexplained part of the gap and share of men in employment was then analysed. Choosing to analyse the differences in wages in such a big country as Poland has many comparative advantages: even at the very disaggregated level, almost all minor occupational groups are represented in the sample. In the case of small countries, significant parts of the occupations would be missing. To the best of the authors' knowledge, this study is one of the few conducted at such disaggregated level of occupations, one of few studies for the Central and Eastern European countries and the first one for Poland.

The results indicate that wages of both men and women are the highest in occupations where the highest level of education and in-the-job investments are required (managers and professionals). However, the relation between the level of average wages and the occupational segregation seems to be rather weak.

The differences in wages of men and women working in the same occupational groups are substantial. Relatively high values of the gender wage gap are observed among managers and professionals. The lowest wage differences between men and women are observed in occupations where rather basic skills are required. The range of the estimated gender wage gap across three-digit occupational level (from −21 to 45 per cent) is much higher than the estimates obtained on the more aggregated level (13-20 per cent according to Majchrowska and Strawiński, 2016 or 15-23 per cent according to Goraus *et al.*, 2017). The results obtained in this paper confirm the earlier findings that the analyses conducted on the aggregated level can lead to underestimation of the occupational gender wage gap, as they do not properly control segregation effects.

The results of the Oaxaca-Blinder decomposition show that in the majority of the groups, a very small part of the wage differences between men and women can be explained by differences in personal and firms' characteristics. In most of the groups, the majority of the gender wage gap remains unexplained.

Comparing the size of the unexplained part of the gender wage gap identified in the literature with the wage discrimination effect and the masculinisation ratio, the authors found very weak non-linear relations between these two indicators. The unexplained part of the gender wage gap is the smallest in strongly female-dominated occupations and in occupations with an almost balanced male-to-female ratio. The discrimination effects seem to be relatively high in male- and female-dominated occupations with some indications that the analysed relation is slightly positive.

Regarding the robustness of the results, two reweighting schemes were applied, based on the parametric estimate of the propensity score and non-parametric proposition of Barsky *et al.* (2002). Both methods confirm that the smallest gender wage gaps are observed for the occupational groups with an almost balanced men-to-women ratio (masculinisation ratio between 0.41 and 0.50). In these groups, the unexplained component of the Oaxaca-Blinder decomposition is the lowest, independently of the estimation method used.

The results indicate that the highest precision of the estimated gender wage gap is obtained for the occupational groups with a similar proportion of men and women in employment. In other male- or female-dominated groups, the size of the estimated gender wage gaps depends to a large extent on the estimation method used. Moreover, the results suggest that decreasing the degree of segregation of men and women into different occupations could reduce the wage differences between these two groups, as the wage discrimination in occupational groups with balanced men-to-women ratio is the smallest. These results are in line with the recommendations of the European Commission (Burchell *et al.*, 2014).

The results obtained in the paper indicate that the unexplained differences in wages between men and women in Poland are, however, not systematic, putting into question the wage discrimination phenomenon on the (Polish) labour market. The differences in wages seem to be, to a large extent, the result of other, difficult to include in the model, factors like differences in psychophysical conditions of men and women, cultural background, tradition or habits. They could explain, for example, relatively high differences in wages of men and women found among occupational groups belonging to the seventh and eighth major occupational groups (skilled workers and machine operators), where extra physical strength is required. Relatively high unexplained differences in wages of men and women in some managerial and professional occupations could also be justified by time-consuming nature of these jobs and the fact that more women than men have to reconcile work and motherhood. The failure to take into account these factors may result in over-interpreting the unexplained part as a measure of gender discrimination.

These findings are confirmed by the recent studies. Blau and Kahn (2016) analysed the importance of norms, psychological attributes or non-cognitive skills possessed by men and women in explaining the wage differences that exists between them. Women have been found to be less willing than are men to negotiate and compete and to be more risk averse. Gender differences in such characteristics have been proposed as an explanation for women's lower wages and lower representation in high-level jobs. Blau and Kahn (2016) cite several recent papers where these factors were examined: the results shows that the role of psychological traits is lower than occupation and industry effects, but they still account even for more than 25 per cent of total gender wage gap. The studies analysing the gender wage gap in the Central and Eastern European countries underline also the role of social norms and the legacy of the stereotypes from the centrally planned period concerning the occupational segregation of men and women into particular jobs, which have shaped social attitudes (Banerjee, 2014). The interpretation of the unexplained gap as gender discrimination, therefore, has to be treated with caution.

Notes

1. For more information about the sample selection scheme see: Structure of Wages and Salaries by Occupations in October 2012, CSO, www.stat.gov.pl
2. www.ilo.org/public/english/bureau/stat/isco/isco88/publ2.htm
3. Calculated as the share of men in employment in a given occupational group.
4. Defining male-dominated, female-dominated and mixed occupational groups, the authors follow the European Commission approach (Burchell *et al.*, 2014).

References

Arellano, M. and Bond, S. (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, Vol. 58 No. 2, pp. 277-297.

Arellano, M. and Bover, O. (1995), "Another look at the instrumental variable estimation of error components models", *Journal of econometrics*, Vol. 68 No. 1, pp. 29-51.

Banerjee, B. (2014), "Occupational segregation and gender differentials in earnings in Macedonia", *IZA Journal of European Labor Studies*, Vol. 3 No. 4, doi: 10.1186/2193-9012-3-4.

Barón, J.D. and Cobb-Clark, D.A. (2010), "Occupational segregation and the gender wage gap in private- and public-sector employment: a distributional analysis", *The Economic Record*, Vol. 86 No. 273, pp. 227-246.

Barsky, R., Bound, J. and Kofi, Ch. K. (2002), "Accounting for the black-white wealth gap: a nonparametric approach", *Journal of the American Statistical Association*, Vol. 97 No. 459, pp. 663-673.

Becker, G.S. (1964), *Human Capital – A Theoretical and Empirical Analysis with Special Reference to Education*, Columbia University Press, New York, NY.

Bergmann, B.R. (1986), *The Economic Emergence of Women*, Basic Books Inc., New York, NY.

Black, D., Haviland, A., Sanders, S. and Taylor, L. (2006), "Why do minority men earn less? A study of wage differentials among the highly educated", *The Review of Economics and Statistics*, Vol. 88 No. 2, pp. 300-313.

Black, D., Haviland, A., Sanders, S. and Taylor, L. (2008), "Gender wage disparities among the highly educated", *The Journal of Human Resources*, Vol. 43 No. 3, pp. 630-659.

Blau, F.D. (1996), "Where are we in the economics of gender?: the gender pay gap", NBER Working Papers No. 5664, Cambridge.

Blau, F.D. and Kahn, L.M. (2003), "Understanding international differences in the gender pay gap", *Journal of Labor Economics*, Vol. 21 No. 1, pp. 106-144.

Blau, F.D. and Kahn, L.M. (2016), "The gender wage gap: extent, trends, and explanations", IZA Discussion papers No. 9656, Bonn.

Blinder, A.S. (1973), "Wage discrimination: reduced form and structural estimates", *Journal of Human Resources*, Vol. 8 No. 4, pp. 436-455.

Brynin, M. and Perales, F. (2015), "Gender wage inequality: the de-gendering of the occupational structure", *European Sociological Review*, Vol. 32 No. 1, pp. 162-174.

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, Luxembourg.

de Ruijter, J.M.P. and Huffman, M.L. (2003), "Gender composition effects in the Netherlands: a multilevel analysis of occupational wage inequality", *Social Science Research*, Vol. 32 No. 2, pp. 312-334.

de Ruijter, J.M.P., van Doorne-Huiskes, A. and Schippers, J.J. (2003), "Size and causes of the occupational gender wage-gap in the Netherlands", *European Sociological Review*, Vol. 19 No. 4, pp. 345-360.

Dolado, J.J., Felgueroso, F. and Jimeno, J.F. (2002), "Recent trends in occupational segregation by gender: a look across the Atlantic", IZA Discussion Papers No. 524, Bonn.

England, P., Herbert, M.S., Kilbourne, B.S., Reid, L.L. and McCreary Megdal, L. (1994), "The gendered valuation of occupations and skills: earnings in 1980 census occupations", *Social Forces*, Vol. 73 No. 1, pp. 65-100.

Evertsson, M., England, P., Mooi-Reci, I., Hermsen, J., de Bruijn, J. and Cotter, D. (2009), "Is gender inequality greater at lower or higher educational levels? Common patterns in the Netherlands, Sweden, and the United States", *Social Politics*, Vol. 16 No. 2, pp. 210-241.

Fortin, N., Lemieux, T. and Firpo, S. (2011), "Decomposition methods in economics", in Ashenfelter, O. and Card, D. (Eds), *Handbook of Labour Economics*, Vol. 4a, Elsevier, Amsterdam, pp. 1-102.

Frölich, M. (2007), "Propensity score matching without conditional independence assumption – with an application to the gender wage gap in the United Kingdom", *Econometrics Journal*, Vol. 10 No. 2, pp. 359-407.

Goraus, K., Tyrowicz, J. and van der Velde, L. (2017), "Which gender wage gap estimates to trust? A comparative analysis", *Review of Income and Wealth*, Vol. 63 No. 1, pp. 118-146.

Grönlund, A. and Magnusson, C. (2013), "Devaluation, crowding or skill specificity? Exploring the mechanisms behind lower wages in female professions", *Social Science Research*, Vol. 42 No. 4, pp. 1006-1017.

Hakim, C. (1998), *Social Change and Innovation in the Labour Market: Evidence from the Census SARs on Occupational Segregation and Labour Mobility, Part-time Working and Student Jobs*, Oxford University Press, Oxford.

Hausman, J.A. and Taylor, W.E. (1981), "Panel data and unobservable individual effects", *Econometrica*, Vol. 49 No. 6, pp. 1377-1398.

Jurajda, S. (2003), "Gender wage gap and segregation in enterprises and the public sector in late transition countries", *Journal of Comparative Economics*, Vol. 31 No. 2, pp. 199-222.

Kline, P. (2011), "Oaxaca-blinder as a reweighting estimator", *The American Economic Review*, Vol. 101 No. 3, pp. 532-537.

Kunze, A. (2008), "Gender wage gap studies: consistency and decomposition", *Empirical Economics*, Vol. 35 No. 1, pp. 63-76.

Magnusson, C. (2013), "More women, lower pay? Occupational sex composition, wages and wage growth", *Acta Sociologica*, Vol. 56 No. 3, pp. 227-245.

Majchrowska, A. and Strawiński, P. (2016), "Regional differences in gender wage gaps in Poland: new estimates based on harmonized data for wages", *Central European Journal of Economic Modelling and Econometrics*, Vol. 8 No. 2, pp. 115-141.

Mincer, J. (1974), *Schooling, Experience and Earnings*, National Bureau of Economic Research, New York, NY.

Mincer, J. and Polachek, S. (1974), "Family investments in human capital: earnings of women", *Journal of Political Economy*, Vol. 82 No. 2, pp. 76-108.

Murphy, E. and Oesch, D. (2015), "The feminization of occupations and change in wages: a panel analysis of Britain, Germany and Switzerland", SOEP Papers on Multidisciplinary Panel Data Research No. 731, Berlin.

Ñopo, H. (2008), "Matching as a tool to decompose wage gaps", *Review of Economics and Statistics*, Vol. 90 No. 2, pp. 290-299.

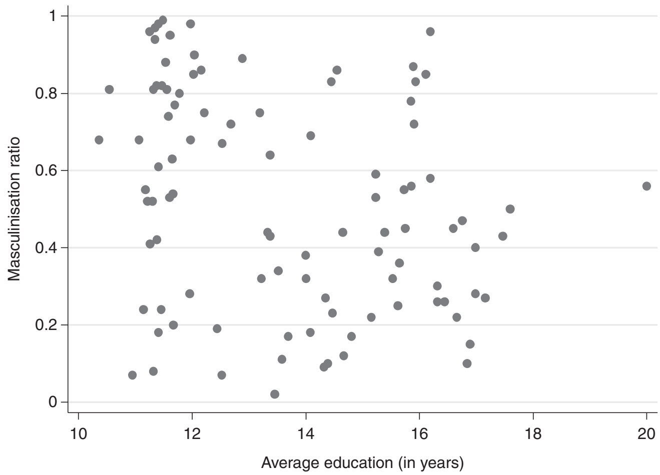
Oaxaca, R.L. (1973), "Male-female wage differentials in urban labor markets", *International Economic Review*, Vol. 14 No. 3, pp. 693-709.

Słoczyński, T. (2015), "Average wage gaps and Oaxaca-Blinder decompositions", IZA Discussion Papers No. 9036, Bonn.

Xiu, L. and Gunderson, M. (2015), "Occupational segregation and the gender earnings gap in China: devils in the details", *International Journal of Manpower*, Vol. 36 No. 5, pp. 711-732.

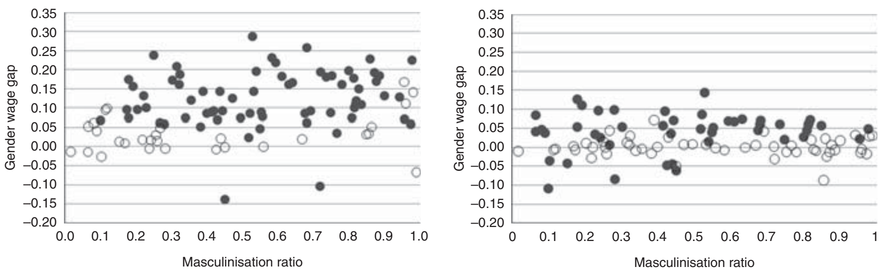
Corresponding author
Aleksandra Majchrowska can be contacted at: aleksandra.majchrowska@uni.lodz.pl

Appendix



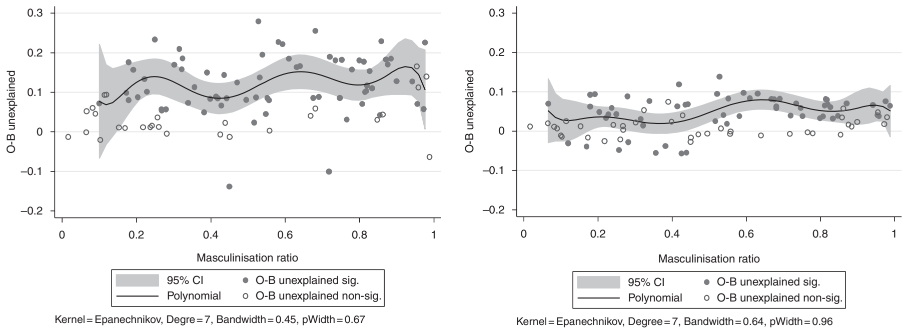
Source: Authors' own calculations

Figure A1. Masculinisation ratio and average education in minor occupational groups in Poland in 2012



Note: Weights based on the propensity score (left) and with proposition of Barsky *et al.* (2002) (right)
Source: Authors' own calculations

Figure A2. Estimated gender wage gaps and masculinisation ratio in occupational groups at the three-digit level of classification in Poland in 2012



Note: Weights based on the propensity score (left) and with proposition of Barsky *et al.* (2002) (right)

Source: Authors' own calculations

Figure A3. Estimated unexplained part of the gender wage gap

Table AI.
Results of the estimations

Code of group	Masculinisation ratio	Number of observations Men	Number of observations Women	Oaxaca-Blinder method GWG	Oaxaca-Blinder method EX	Oaxaca-Blinder method UNEX	Propensity score method GWG	Propensity score method EX	Propensity score method UNEX	Barsky *et al.* (2002) method GWG	Barsky *et al.* (2002) method EX	Barsky *et al.* (2002) method UNEX
111	0.582	849	608	0.262 (0.019)	0.042 (0.010)	0.220 (0.018)	0.231 (0.021)	0.004 (0.009)	0.228 (0.019)	−0.008 (0.021)	−0.047 (0.012)	0.038 (0.019)
112	0.720	8,171	3,172	0.155 (0.018)	−0.029 (0.009)	0.184 (0.017)	0.194 (0.026)	0.004 (0.008)	0.190 (0.025)	−0.031 (0.017)	−0.071 (0.009)	0.039 (0.017)
121	0.356	5,363	9,704	0.288 (0.012)	0.143 (0.007)	0.145 (0.012)	0.120 (0.017)	0.007 (0.006)	0.113 (0.015)	−0.004 (0.011)	0.051 (0.006)	−0.056 (0.010)
122	0.593	2,864	1,948	0.231 (0.024)	0.011 (0.011)	0.220 (0.022)	0.219 (0.026)	−0.003 (0.011)	0.222 (0.023)	0.068 (0.023)	−0.029 (0.010)	0.097 (0.022)
132	0.834	7,113	1,415	0.100 (0.018)	−0.011 (0.009)	0.111 (0.016)	0.108 (0.044)	−0.002 (0.009)	0.110 (0.043)	−0.009 (0.018)	−0.042 (0.010)	0.032 (0.016)
133	0.855	1,249	212	−0.028 (0.051)	−0.066 (0.029)	0.038 (0.046)	0.035 (0.027)	−0.008 (0.026)	0.042 (0.005)	−0.087 (0.051)	−0.078 (0.027)	−0.009 (0.045)
134	0.302	2,341	5,413	0.229 (0.020)	0.044 (0.009)	0.185 (0.019)	0.172 (0.027)	0.002 (0.007)	0.170 (0.026)	0.053 (0.018)	0.031 (0.008)	0.022 (0.019)
141	0.324	212	433	0.199 (0.045)	0.008 (0.018)	0.191 (0.044)	0.187 (0.058)	0.001 (0.015)	0.186 (0.056)	0.030 (0.040)	−0.024 (0.016)	0.054 (0.040)
142	0.316	883	1,827	0.222 (0.019)	0.016 (0.009)	0.206 (0.018)	0.209 (0.025)	−0.001 (0.007)	0.210 (0.024)	0.013 (0.016)	−0.019 (0.008)	0.031 (0.015)
143	0.532	1,477	1,289	0.183 (0.021)	0.042 (0.012)	0.142 (0.020)	0.143 (0.022)	0.005 (0.010)	0.138 (0.020)	0.007 (0.020)	0.000 (0.011)	0.007 (0.019)
211	0.495	538	551	0.104 (0.027)	0.045 (0.017)	0.059 (0.024)	0.074 (0.029)	−0.007 (0.015)	0.081 (0.024)	0.008 (0.025)	0.014 (0.016)	−0.007 (0.023)
213	0.471	1,714	1,902	0.239 (0.014)	0.108 (0.009)	0.131 (0.014)	0.126 (0.018)	0.001 (0.007)	0.125 (0.016)	0.007 (0.013)	0.016 (0.008)	−0.008 (0.012)
214	0.780	10,759	3,019	0.186 (0.011)	0.028 (0.004)	0.158 (0.010)	0.161 (0.020)	0.003 (0.004)	0.158 (0.019)	−0.013 (0.011)	−0.007 (0.004)	−0.006 (0.010)
215	0.956	2,409	110	0.021 (0.040)	−0.056 (0.026)	0.077 (0.038)	0.071 (0.026)	0.000 (0.023)	0.071 (0.011)	−0.014 (0.040)	−0.062 (0.024)	0.048 (0.037)
216	0.552	880	718	0.061 (0.023)	−0.024 (0.009)	0.086 (0.023)	0.089 (0.024)	0.003 (0.008)	0.086 (0.023)	0.052 (0.020)	−0.030 (0.009)	0.083 (0.020)
221	0.426	3,247	4,380	0.003 (0.013)	0.003 (0.005)	0.000 (0.015)	−0.007 (0.022)	0.000 (0.004)	−0.007 (0.021)	−0.047 (0.022)	0.010 (0.005)	−0.057 (0.024)
222	0.017	377	21,601	−0.027 (0.011)	−0.026 (0.005)	−0.001 (0.010)	−0.013 (0.511)	0.000 (0.005)	−0.013 (0.510)	−0.011 (0.011)	−0.022 (0.005)	0.011 (0.010)
226	0.268	106	292	0.074 (0.054)	0.020 (0.030)	0.054 (0.050)	0.048 (0.027)	−0.006 (0.025)	0.054 (0.008)	0.044 (0.056)	0.040 (0.031)	0.003 (0.049)
228	0.260	1,139	3,240	0.031 (0.017)	0.023 (0.011)	0.008 (0.016)	0.015 (0.032)	0.003 (0.008)	0.012 (0.031)	−0.018 (0.016)	0.004 (0.010)	−0.021 (0.015)
231	0.557	5,666	4,455	0.133 (0.007)	0.052 (0.004)	0.082 (0.007)	0.078 (0.008)	−0.002 (0.004)	0.080 (0.007)	−0.002 (0.007)	0.005 (0.004)	−0.007 (0.006)
232	0.450	1,467	1,793	−0.059 (0.011)	−0.046 (0.008)	−0.014 (0.008)	−0.001 (0.011)	0.012 (0.007)	−0.013 (0.009)	−0.062 (0.011)	−0.046 (0.008)	−0.016 (0.009)
233	0.282	5,866	14,917	−0.018 (0.004)	−0.013 (0.002)	−0.005 (0.003)	−0.005 (0.005)	0.000 (0.002)	−0.005 (0.005)	−0.085 (0.004)	−0.057 (0.002)	−0.028 (0.004)
234	0.100	2,225	19,914	0.069 (0.006)	−0.003 (0.004)	0.072 (0.005)	0.069 (0.026)	−0.003 (0.004)	0.072 (0.025)	−0.109 (0.010)	−0.098 (0.006)	−0.011 (0.009)
235	0.153	1,757	9,721	−0.060 (0.011)	−0.071 (0.006)	0.011 (0.009)	0.013 (0.032)	0.002 (0.005)	0.011 (0.031)	−0.042 (0.009)	−0.054 (0.005)	0.012 (0.009)
241	0.250	3,950	11,963	0.238 (0.012)	−0.002 (0.004)	0.240 (0.012)	0.238 (0.022)	0.005 (0.004)	0.233 (0.022)	0.015 (0.010)	−0.028 (0.004)	0.043 (0.010)
242	0.322	7,703	16,151	0.187 (0.007)	0.027 (0.003)	0.159 (0.007)	0.162 (0.011)	0.003 (0.003)	0.158 (0.011)	0.006 (0.007)	−0.008 (0.003)	0.014 (0.006)
243	0.436	5,886	7,562	0.141 (0.012)	−0.009 (0.005)	0.150 (0.011)	0.143 (0.013)	−0.001 (0.004)	0.144 (0.012)	0.036 (0.010)	−0.031 (0.004)	0.067 (0.010)
244	0.390	204	322	0.180 (0.051)	0.046 (0.029)	0.134 (0.046)	0.143 (0.061)	−0.007 (0.022)	0.150 (0.057)	0.072 (0.043)	−0.003 (0.025)	0.075 (0.039)
251	0.827	5,178	1,072	0.104 (0.019)	−0.047 (0.009)	0.151 (0.017)	0.150 (0.045)	−0.003 (0.008)	0.153 (0.044)	−0.006 (0.019)	−0.067 (0.008)	0.061 (0.017)
252	0.871	2,948	432	0.115 (0.028)	−0.053 (0.016)	0.168 (0.026)	0.193 (0.087)	0.010 (0.014)	0.183 (0.085)	−0.013 (0.028)	−0.083 (0.015)	0.070 (0.025)
261	0.399	1,671	2,521	0.091 (0.018)	0.005 (0.010)	0.086 (0.016)	0.086 (0.020)	0.002 (0.009)	0.084 (0.017)	0.001 (0.019)	−0.023 (0.010)	0.024 (0.016)

(continued)

Code of group	Masculinisation ratio	Number of observations Men	Women	Oaxaca-Blinder method GWG	EX	UNEX	Propensity score method GWG	EX	UNEX	Barsky *et al.* (2002) method GWG	EX	UNEX
262	0.256	315	889	0.125 (0.022)	0.098 (0.016)	0.027 (0.021)	0.030 (0.036)	−0.006 (0.010)	0.036 (0.035)	0.000 (0.019)	0.048 (0.014)	−0.048 (0.019)
263	0.218	630	2,282	0.031 (0.021)	0.024 (0.012)	0.007 (0.018)	0.018 (0.040)	0.006 (0.012)	0.012 (0.038)	−0.029 (0.018)	−0.004 (0.011)	−0.025 (0.016)
264	0.450	467	609	−0.205 (0.059)	−0.041 (0.021)	−0.164 (0.065)	−0.138 (0.053)	0.001 (0.012)	−0.139 (0.052)	−0.050 (0.059)	−0.023 (0.018)	−0.028 (0.064)
265	0.558	430	352	−0.009 (0.024)	−0.010 (0.012)	0.000 (0.022)	0.000 (0.024)	−0.003 (0.009)	0.003 (0.023)	−0.002 (0.024)	−0.001 (0.011)	−0.001 (0.022)
311	0.752	11,226	3,571	0.161 (0.008)	−0.027 (0.004)	0.188 (0.007)	0.184 (0.013)	0.002 (0.003)	0.182 (0.013)	−0.012 (0.007)	−0.050 (0.003)	0.038 (0.007)
312	0.885	3,094	388	0.208 (0.022)	0.025 (0.009)	0.183 (0.021)	0.184 (0.009)	−0.001 (0.008)	0.185 (0.003)	−0.007 (0.022)	−0.019 (0.009)	0.012 (0.020)
313	0.899	3,955	431	0.140 (0.014)	0.010 (0.009)	0.130 (0.015)	0.132 (0.052)	0.003 (0.006)	0.128 (0.052)	0.017 (0.014)	−0.006 (0.008)	0.023 (0.014)
314	0.640	1,071	547	0.450 (0.020)	0.271 (0.021)	0.180 (0.025)	0.166 (0.011)	0.000 (0.010)	0.167 (0.004)	−0.001 (0.022)	0.010 (0.017)	−0.011 (0.023)
321	0.114	426	3,265	0.077 (0.018)	−0.021 (0.009)	0.098 (0.017)	0.095 (0.093)	0.001 (0.007)	0.093 (0.093)	−0.007 (0.016)	−0.033 (0.008)	0.025 (0.016)
325	0.382	2,179	3,368	0.076 (0.010)	0.046 (0.006)	0.031 (0.009)	0.052 (0.011)	0.003 (0.003)	0.049 (0.010)	−0.017 (0.009)	0.021 (0.005)	−0.038 (0.009)
331	0.119	1,539	11,396	0.196 (0.014)	0.050 (0.006)	0.146 (0.014)	0.099 (0.073)	0.005 (0.005)	0.094 (0.073)	−0.004 (0.012)	0.027 (0.005)	−0.031 (0.012)
332	0.691	5,957	2,663	0.035 (0.012)	−0.044 (0.006)	0.079 (0.012)	0.093 (0.018)	0.005 (0.004)	0.088 (0.017)	0.042 (0.012)	−0.041 (0.006)	0.082 (0.012)
333	0.439	908	1,206	0.044 (0.022)	0.029 (0.010)	0.015 (0.021)	0.021 (0.022)	−0.002 (0.009)	0.022 (0.020)	−0.044 (0.021)	0.011 (0.010)	−0.055 (0.021)
334	0.174	1,642	7,700	0.136 (0.012)	0.043 (0.006)	0.093 (0.012)	0.096 (0.032)	−0.002 (0.005)	0.098 (0.032)	−0.007 (0.010)	0.010 (0.005)	−0.017(0.010)
335	0.222	2,137	7,473	0.119 (0.007)	−0.005 (0.004)	0.124 (0.007)	0.132 (0.014)	−0.002 (0.003)	0.134 (0.013)	0.000 (0.006)	−0.033 (0.004)	0.033 (0.006)
341	0.103	419	3,713	−0.046 (0.013)	−0.017 (0.011)	−0.029 (0.013)	−0.027 (0.053)	−0.006 (0.009)	−0.021 (0.052)	−0.036 (0.014)	−0.021 (0.016)	−0.015 (0.018)
343	0.267	672	1,828	0.115 (0.022)	0.039 (0.012)	0.076 (0.017)	0.062 (0.025)	0.005 (0.007)	0.057 (0.024)	0.007 (0.015)	−0.007 (0.008)	0.013 (0.013)
351	0.863	1,891	288	−0.056 (0.035)	−0.093 (0.019)	0.037 (0.032)	0.051 (0.106)	0.008 (0.019)	0.044 (0.104)	−0.025 (0.037)	−0.083 (0.020)	0.057 (0.034)
411	0.228	3,530	11,848	0.114 (0.009)	0.005 (0.003)	0.109 (0.008)	0.100 (0.017)	−0.001 (0.003)	0.102 (0.017)	0.034 (0.007)	−0.009 (0.003)	0.043 (0.006)
413	0.429	370	467	0.034 (0.025)	−0.038 (0.017)	0.072 (0.021)	0.070 (0.030)	0.003 (0.016)	0.067 (0.026)	0.026 (0.028)	−0.014 (0.017)	0.041 (0.023)
421	0.179	1,190	5,478	0.076 (0.013)	−0.004 (0.008)	0.081 (0.014)	0.074 (0.035)	−0.006 (0.006)	0.080 (0.035)	0.053 (0.011)	−0.009 (0.007)	0.062 (0.012)
422	0.169	794	3,815	0.073 (0.023)	0.052 (0.007)	0.020 (0.023)	0.009 (0.048)	−0.001 (0.006)	0.010 (0.048)	0.002 (0.017)	0.042 (0.007)	−0.039 (0.018)
431	0.090	482	4,765	0.138 (0.021)	0.061 (0.012)	0.077 (0.018)	0.041 (0.097)	−0.005 (0.009)	0.046 (0.096)	0.037 (0.017)	0.031 (0.011)	0.007 (0.015)
432	0.747	11,970	3,849	0.036 (0.006)	−0.051 (0.003)	0.087 (0.006)	0.087 (0.011)	0.001 (0.002)	0.086 (0.011)	0.019 (0.006)	−0.039 (0.002)	0.059 (0.006)
441	0.339	3,355	6,464	−0.059 (0.008)	−0.115 (0.005)	0.057 (0.009)	0.074 (0.016)	0.001 (0.004)	0.073 (0.015)	−0.009 (0.008)	−0.086 (0.004)	0.077 (0.008)
511	0.718	678	268	−0.118 (0.028)	−0.008 (0.018)	−0.110 (0.026)	−0.104 (0.035)	−0.003 (0.016)	−0.101 (0.031)	0.001 (0.037)	0.008 (0.021)	−0.007 (0.037)
512	0.204	793	2,356	0.035 (0.013)	−0.048 (0.008)	0.083 (0.013)	0.096 (0.043)	0.009 (0.008)	0.087 (0.039)	0.011 (0.009)	−0.038 (0.007)	0.049 (0.009)
513	0.280	486	1,181	0.053 (0.012)	−0.008 (0.007)	0.061 (0.011)	0.057 (0.025)	0.001 (0.006)	0.057 (0.023)	0.098 (0.010)	0.009 (0.007)	0.089 (0.009)
515	0.517	2,555	2,149	0.054 0.007	0.022 0.004	0.032 0.007	0.024 (0.007)	0.001 (0.002)	0.023 (0.006)	0.048 (0.006)	0.022 (0.003)	0.026 (0.006)
516	0.668	227	112	−0.040 (0.036)	−0.107 (0.035)	0.067 (0.040)	0.020 (0.041)	−0.020 (0.023)	0.040 (0.036)	0.006 (0.034)	−0.076 (0.028)	0.082 (0.035)
522	0.192	5,044	19,986	0.157 (0.005)	0.008 (0.002)	0.149 (0.005)	0.157 (0.012)	0.000 (0.002)	0.157 (0.012)	0.111 (0.003)	0.016 (0.002)	0.094 (0.003)

(*continued*)

Table AI.

Table AI.

Code of group	Masculinisation ratio	Number of observations Men	Women	Oaxaca-Blinder method GWG	EX	UNEX	Propensity score method GWG	EX	UNEX	Barsky *et al.* (2002) method GWG	EX	UNEX
523	0.065	227	3,332	0.025 (0.019)	0.016 (0.010)	0.009 (0.016)	−0.014 (0.110)	−0.013 (0.007)	−0.001 (0.109)	0.040 (0.016)	0.020 (0.011)	0.020 (0.014)
524	0.443	2,170	2,604	0.089 (0.014)	0.000 (0.007)	0.089 (0.012)	0.092 (0.012)	0.007 (0.005)	0.085 (0.011)	0.070 (0.011)	0.002 (0.005)	0.068 (0.010)
532	0.242	364	1,220	0.055 (0.011)	0.010 (0.006)	0.046 (0.011)	0.017 (0.025)	0.001 (0.005)	0.016 (0.025)	0.025 (0.010)	0.009 (0.005)	0.016 (0.010)
541	0.848	6,718	1,003	0.028 (0.008)	0.000 (0.005)	0.028 (0.008)	0.032 (0.025)	0.002 (0.004)	0.030 (0.024)	0.057 (0.009)	0.018 (0.004)	0.039 (0.009)
611	0.676	350	148	0.059 (0.025)	−0.032 (0.018)	0.091 (0.022)	0.086 (0.030)	0.000 (0.016)	0.085 (0.026)	0.044 (0.023)	−0.037 (0.016)	0.081 (0.020)
713	0.942	1,953	121	0.070 (0.026)	−0.055 (0.013)	0.125 (0.028)	0.128 (0.011)	0.001 (0.008)	0.128 (0.008)	0.032 (0.031)	−0.035 (0.010)	0.067 (0.033)
721	0.976	8,172	181	0.181 (0.018)	−0.038 (0.009)	0.218 (0.016)	0.225 (0.010)	−0.001 (0.010)	0.226 (0.003)	0.047 (0.019)	−0.029 (0.008)	0.076 (0.018)
722	0.954	17,712	791	0.147 (0.009)	−0.015 (0.003)	0.162 (0.009)	0.168 (0.095)	0.002 (0.003)	0.166 (0.095)	0.022 (0.010)	−0.019 (0.003)	0.041 (0.009)
731	0.610	1,125	645	0.160 (0.014)	−0.015 (0.008)	0.175 (0.015)	0.182 (0.017)	−0.003 (0.006)	0.185 (0.016)	0.067 (0.013)	−0.017 (0.006)	0.084 (0.013)
732	0.681	1,687	744	0.281 (0.015)	0.015 (0.008)	0.267 (0.015)	0.258 (0.018)	0.002 (0.006)	0.256 (0.018)	0.060 (0.014)	0.002 (0.006)	0.058 (0.014)
741	0.980	11,406	214	0.134 (0.020)	0.003 (0.010)	0.131 (0.019)	0.141 (0.404)	0.001 (0.009)	0.140 (0.404)	0.029 (0.020)	−0.007 (0.009)	0.036 (0.019)
742	0.859	2,071	339	0.182 (0.017)	−0.034 (0.014)	0.216 (0.021)	0.229 (0.038)	−0.001 (0.007)	0.230 (0.038)	0.022 (0.018)	−0.012 (0.012)	0.035 (0.021)
751	0.522	3,210	2,886	0.058 (0.007)	−0.026 (0.003)	0.083 (0.007)	0.087 (0.007)	0.000 (0.002)	0.087 (0.007)	0.086 (0.006)	−0.009 (0.002)	0.095 (0.006)
752	0.815	4,250	947	0.046 (0.008)	−0.053 (0.004)	0.099 (0.008)	0.101 (0.017)	0.000 (0.003)	0.101 (0.017)	0.063 (0.008)	−0.018 (0.003)	0.081 (0.008)
753	0.179	1,295	5,739	0.208 (0.011)	0.034 (0.004)	0.175 (0.011)	0.174 (0.039)	−0.002 (0.003)	0.176 (0.039)	0.126 (0.007)	0.034 (0.004)	0.092 (0.007)
754	0.528	231	197	0.257 (0.025)	−0.004 (0.024)	0.260 (0.032)	0.287 (0.044)	0.007 (0.016)	0.280 (0.041)	0.143 (0.025)	0.005 (0.023)	0.138 (0.029)
811	0.973	7,075	105	−0.005 (0.025)	−0.062 (0.008)	0.057 (0.025)	0.058 (0.008)	0.000 (0.007)	0.057 (0.005)	−0.018 (0.025)	−0.035 (0.006)	0.018 (0.025)
812	0.876	4,034	535	0.177 (0.011)	0.005 (0.005)	0.172 (0.010)	0.170 (0.035)	-0.001 (0.004)	0.171 (0.035)	0.006 (0.010)	−0.011 (0.004)	0.016 (0.010)
813	0.799	3,578	842	0.237 (0.013)	0.047 (0.008)	0.190 (0.012)	0.198 (0.025)	0.017 (0.007)	0.181 (0.024)	0.027 (0.013)	−0.006 (0.007)	0.033 (0.012)
814	0.735	4,081	1,386	0.158 (0.008)	−0.017 (0.004)	0.175 (0.008)	0.181 (0.011)	0.000 (0.003)	0.181 (0.011)	0.061 (0.007)	−0.010 (0.002)	0.071 (0.007)
815	0.419	670	847	0.057 (0.014)	−0.038 (0.008)	0.095 (0.013)	0.092 (0.017)	0.007 (0.007)	0.085 (0.016)	0.095 (0.020)	−0.025 (0.008)	0.120 (0.017)
816	0.631	1,980	1,072	0.154 (0.012)	−0.010 (0.006)	0.164 (0.011)	0.161 (0.014)	−0.003 (0.005)	0.163 (0.013)	0.075 (0.012)	−0.021 (0.005)	0.095 (0.012)
817	0.820	1,696	368	0.046 (0.012)	−0.068 (0.009)	0.114 (0.010)	0.119 (0.008)	0.002 (0.008)	0.117 (0.002)	0.019 (0.012)	−0.049 (0.007)	0.068 (0.010)
818	0.812	4,886	1,156	0.154 (0.010)	−0.009 (0.005)	0.163 (0.010)	0.178 (0.018)	0.001 (0.003)	0.177 (0.018)	0.059 (0.009)	−0.006 (0.004)	0.065 (0.009)
821	0.539	6,215	4,930	0.180 (0.006)	−0.013 (0.003)	0.193 (0.006)	0.195 (0.006)	0.000 (0.002)	0.195 (0.005)	0.016 (0.005)	−0.025 (0.002)	0.041 (0.005)
831	0.766	4,200	1,322	0.109 (0.006)	0.072 (0.005)	0.037 (0.005)	0.035 (0.003)	0.004 (0.004)	0.031 (0.001)	0.004 (0.005)	0.007 (0.003)	−0.003 (0.004)
833	0.989	12,257	152	−0.170 (0.024)	−0.099 (0.011)	−0.071 (0.021)	−0.067 (0.691)	−0.003 (0.014)	−0.064 (0.690)	0.031 (0.026)	−0.034 (0.009)	0.065 (0.024)
834	0.959	6,486	268	0.067 (0.013)	−0.049 (0.007)	0.116 (0.012)	0.112 (0.146)	−0.001 (0.006)	0.113 (0.146)	−0.006 (0.012)	−0.044 (0.005)	0.038 (0.012)
911	0.066	1,123	12,105	0.065 (0.008)	0.007 (0.004)	0.057 (0.007)	0.052 (0.049)	0.001 (0.003)	0.052 (0.049)	0.084 (0.006)	0.014 (0.003)	0.070 (0.005)

(*continued*)

Table AI.

Code of group	Masculinisation ratio	Number of observations		Oaxaca-Blinder method			Propensity score method			Barsky *et al.* (2002) method		
		Men	Women	GWG	EX	UNEX	GWG	EX	UNEX	GWG	EX	UNEX
912	0.237	192	518	0.027 (0.021)	−0.017 (0.019)	0.044 (0.020)	−0.005 (0.030)	−0.017 (0.011)	0.012 (0.030)	0.096 (0.017)	0.038 (0.018)	0.059 (0.019)
921	0.681	570	253	0.050 (0.023)	−0.008 (0.011)	0.058 (0.024)	0.061 (0.031)	0.003 (0.008)	0.059 (0.030)	0.071 (0.026)	0.008 (0.010)	0.063 (0.027)
932	0.414	3,566	4,748	0.089 (0.007)	−0.003 (0.002)	0.091 (0.007)	0.089 (0.008)	0.000 (0.002)	0.089 (0.008)	0.057 (0.006)	−0.005 (0.002)	0.063 (0.006)
933	0.818	2,484	541	0.085 (0.012)	−0.018 (0.006)	0.103 (0.012)	0.111 (0.027)	0.008 (0.006)	0.103 (0.026)	0.072 (0.013)	−0.009 (0.005)	0.080 (0.013)
941	0.082	245	1,832	0.082 (0.015)	0.027 (0.007)	0.055 (0.016)	0.061 (0.067)	0.000 (0.005)	0.061 (0.066)	0.047 (0.009)	0.035 (0.008)	0.011 (0.012)
961	0.808	1,202	252	0.066 (0.014)	0.000 (0.009)	0.067 (0.014)	0.075 (0.031)	0.004 (0.007)	0.071 (0.030)	0.044 (0.015)	0.006 (0.007)	0.038 (0.015)
962	0.549	3,739	2,115	0.066 (0.006)	0.026 (0.003)	0.040 (0.007)	0.047 (0.007)	0.002 (0.002)	0.045 (0.007)	0.039 (0.006)	0.021 (0.002)	0.018 (0.006)

Note: The insignificant results are marked in grey
Source: Authors' own estimates