University of Warsaw
Faculty of Economic Sciences

Working Papers

No. 28/2013 (113)

LUCAS VAN DER VELDE
JOANNA TYROWICZ
KAROLINA GORAUS

What is the true gender wage gap? A comparative analysis using data from Poland

Warsaw 2013

What is the true gender wage gap? A comparative analysis using data from Poland

LUCAS VAN DER VELDE
Faculty of Economic Sciences,
University of Warsaw
e-mail: l.van-der-velde@student.uw.edu.pl

JOANNA TYROWICZ
Faculty of Economic Sciences,
University of Warsaw
National Bank of Poland
e-mail: j.tyrowicz@uw.edu.pl

KAROLINA GORAUS
Faculty of Economic Sciences,
University of Warsaw
e-mail: kgoraus@wne.uw.edu.pl

Abstract

Given the proliferation of methods to estimate gender wage gap, practical issues arise. The aim of this paper is to compare estimates of the adjusted wage gap from different methods and sets of conditioning variables. We apply available parametric and non-parametric methods to LFS data from Poland for 2012. While the raw gap amounts to nearly 10% of the female wage, after the correction for the endowments, the adjusted wage gap estimates range between 15% and as much as 25% depending on the method and the choice of conditional variables. The differences across methods and conditioning variables do not exceed 3pp. The largest differences emerged between methods estimating gap at the mean and those operating at quantiles. Within the same moment, methods which account for selection into employment yielded higher estimates of the adjusted wage gap. When expanding the conditioning set, to account for possible sorting of women into lower paid jobs, estimates of gap increase. While the actual point estimators of adjusted wage gap are slightly different, all of them are roughly twice as high as the raw gap, which corroborates the policy relevance of this methodological study.

Keywords:

gender wage gap, transition, Poland, decomposition methods

JEL:

C24, J22, J31, J71

Acknowledgments:

Authors would like to thank Tymon Słoczyński for insightful suggestions. The support of National Center for Science (grant UMO-2012/05/E/HS4/01510) is gratefully acknowledged. Usual disclaimer applies.

Introduction

For a long time, the literature on the gender gap in wages has been dominated by just a handful of techniques, namely the Oaxaca-Blinder decomposition and dummy variables in pool regressions of different types (OLS, IV, etc.). The estimates were referred to as adjusted wage gap, i.e. the size of the gender wage gap *controlling* for differences in characteristics important for productivity (such as age, education, industry, occupation, firm characteristics, etc.). However, these methods are troubled with weaknesses well recognized in the literature: the estimates cannot be easily applied if the characteristics diverge; they cannot measure the differences outside the mean; and they cannot correct for selection into employment. Either of the problems may generate a significant bias in the results.

The last two decades brought about the expansion of the available toolset with the objective to address one or more of the three problems. Starting with the Juhn, Murphy and Pierce (1993) decomposition a new wave of techniques was developed. The new methods attempt to address many weaknesses associated with traditional parametric approach. One strand of the literature goes beyond the analyses of mean wages, employing quintile and sampling methods to be able to estimate adjusted wage gaps along the distribution. Another strand focuses on assuring comparability, by implementing what is referred to as the common support constraint. Both these strands aim to address explicitly the problem of selection bias. Clearly, each of the methods provides not only econometric advancement, but refines also the way adjusted gender wage gap can be interpreted for policy purposes.

Though the proliferation of methods is welcome from a methodological perspective, it also introduces confusion from a practitioner's perspective. Are results susceptible to a method? How do the estimates of the gaps compare to each other? These questions were partially addressed by Wichselbaumer and Winter-Ebmer (2006), who conducted a meta-analysis of the gender wage gap literature. They find that estimates with Heckman correction are in principle higher, but the results from different databases using different controls and wage measures are not directly comparable to each other. In our paper we propose to fill this gap by offering a comparative analysis of the most popular methods and conditioning variables.

We perform a comparative analysis of various methods to estimate the adjusted gender wage gap. This comparative exercise is performed for Poland, a country which passed two decades of economic transition from a centrally planned to a market economy, but which is characterized by extremely low female labor force participation. While we do not discuss access to labor market, nor to the professions, we analyze comprehensively current extent of the gender wage gap unexplained by the differences in endowments.

The characterization of the gender wage gap for Poland is an important side product of our analysis, as the previous literature for the country is rather scarce. Some authors explored the topic in the aftermath of the transition, and the incorporation to the EU; however, there is a lack of recent analysis on the topic. Moreover, we also profited from the most recent Labor Force Survey (LFS), which includes information on the field of study, a variable which has not been included in previous studies. Given the low female labor force participation, analyses of the gender wage gap are of clear policy relevance. If "discrimination" is indeed a prevalent phenomenon, it can partly explain the low female employment rates.

The paper is structured as follows. First, we present the decomposition techniques used in the literature, analyzing their interpretational advantages and disadvantages. Second, we briefly discuss the earlier work on gender wage gap in Poland. We employ data from the Polish Labor Force Survey, which are described in the third section. The results are discussed synthetically in the subsequent session, but the details for each of the employed method were moved to the appendix for brevity. The final conclusion derives the policy recommendations and suggests potential avenues for further research in this field.

Section 1. Different methodological approaches to measuring gender wage gap

The simplest way to estimate the differences between genders is to use a Mincer equation for wages expanded with a gender dummy. Usually, Mincer equations is estimated as a semi-elasticity to reduce the problems due to the skewness of the wage distribution. This procedure is extremely simple, but at the same time flexible. i.e. allows for inclusion of variables which are of interest, such as age, education, profession, marital status or race. Moreover, when the results are presented in an interesting way, it can be used as an element for policy recommendation. An example of this is found in Watson (2010), who uses the "simulated change approach"[1] to decompose the differences in wages due to different factors.

However, linear regressions have several shortcomings, such as the presence of unobservable effects which might be correlated with the error term, the sample selection bias, and more importantly, the fact that man and women might receive different rewards for the same characteristics. This can be controlled for in the linear regression context by adding interaction variables with the gender dummies. This procedure is in practice equivalent to the estimation of two separate equations, and then decomposing the absolute differences in wages into the component attributable to differences in characteristics and component that cannot be explained by objective differences. The latter is conceptualized as adjusted gender wage gap, often identified with discrimination.

The seminal decomposition technique proposed by Oaxaca (1973) and Blinder (1973) consisted of taking the advantages of the OLS properties. Namely, differences across genders can be separated into differences due to the characteristics and the differences due to the rewards. The first step is to estimate two separate regressions for each gender, save the estimates and create a counterfactual mean value for one of the groups. Then the size of the pay difference can be decomposed from:

$$W_m - W_f = \beta^*\left(X_m - X_f\right) + X_m(\beta_m - \beta^*) + \ X_f(\beta^* - \beta_f)$$

The first term in RHS represents differences in characteristics between males and females. The second and third term in the RHS represent the differences in coefficients or the unexplained component, written in the most general form. This means that no assumptions are made concerning the "true" wages in the absence of discrimination effects. In this formulation, the first

[1] The procedure to compute the simulated change approach is quite simple. First, estimate a pooled regression for the wage including all possible variables and save the coefficients. Second, estimate the means of the variables of interest for men and women (notice that the subset variables of interest can be smaller than the total number of variables used) and calculate the gender differences. Third, multiply the differences in means by the coefficients and sum it up. Finally, obtain the percentage contribution of each characteristic.

equation represents the male (dis)advantage while the second represents the female (dis)advantage.

Following Oaxaca (1978) and Blinder (1978), a number of variations have been developed for this decomposition. Namely, the important question concerns the definition of β^*. Several alternatives were proposed, as summarized in Table 1. All these procedures are commonly used. However, the two most popular are the Oaxaca (1978) and the Neumark (1988). The decision to include one or the other is non-trivial, as it necessitates the index variable problem, affecting the interpretational issues[2].

Table 1. Literature approaches to determining β^*

Oaxaca (1973)	$\beta^* = \beta_m$	Male regression coefficients
Blinder (1973)	$\beta^* = \beta_f$	Female regression coefficients
Cotton (1988)	$\beta^* = 0.5\beta_m + 0,5\beta_f$	Simple average of the coefficients in both groups
Reimers (1983)	$\beta^* = \left(\frac{N.Males}{N}\right)\beta_m + \left(\frac{N.Female}{N}\right)\beta_f$	Weighted average of the coefficients in both groups
Neumark (1988)	$\beta^* = \beta_p$	The coefficients from a pool regression

Source: own elaboration

Downsides of these decomposition techniques, though, are threefold. First, this approach implies that the average wage gap may be estimated for men and women whose characteristics are starkly different. Second, this approach only looks at average difference between male and female compensations. Third, the selection bias problem is neglected. In the reminder of this section we discuss how the literature has so far developed to address these shortcomings[3].

The problem of common characteristics

Parametric Oaxaca-Blinder decomposition and its variations do not take into consideration the interplay between characteristics and how they are priced. In particular, if certain characteristics are relatively more abundant among one gender than among other, their price is likely to reflect the abundance along with the market valuation. For example, higher compensation to miners is confused with lower wages of women in general if there are few (or no) female miners.

This problem involves also another debate from the literature on which variables should be included in explanatory vectors. One example is the inclusion of occupation fixed effects or not. Methodologically, it would be inaccurate to omit these variables as they are usually significant; however, in some cases they have been excluded because they were considered endogenous, and a part of the "discrimination".

More importantly, the inclusion of occupational (and industrial) dummies does not help to address another problem, i.e. unavailability of wages for female miners which biases the average adjusted gender wage gap measure. This problem – often referred to as common support

[2] A note of caution is also needed regarding the Neumark (1988) decomposition. A dummy variable for gender shall be included in the pooled regression in order to avoid the omitted variable bias; however, in most cases the results are not reported.

[3] For the reminder of the literature review we only refer to the literature on the estimation of the (continuous) wages. We thus leave aside contributions concerning non-standard measures of wages (e.g. dichotomous or discrete indicators of wages), e.g. Fairlie (2003) or Bauer and Sinning (2006)

problem – is neglected by all of the discussed methods, whereas the severity of this issue is large irrespectively from the choice of baseline β^*. If men and women are not strictly comparable or if these comparisons make no sense, reliance on estimated β's to evaluated gender wage gap is particularly misleading. Think for instance in the case of women with low work experience due to maternity leave and men with similar work experience, which resulted from the lack of attachment to the labor market. In OLS based estimations, the lack of common support implies that the out-of-the sample prediction is biased, which undermines the use of the estimated coefficients for any reliable interpretations.

The simple solution to this issue is to estimate the gender gap only on the observations where the characteristics of men and women are comparable, i.e. within the common support. In order to determine the common support, authors relied on nonparametric techniques. Barsky et al (2002) propose to use a reweighting equation, in which the weights attached to every observation are the ratio of probabilities of finding an individual with a given level of income in each of the two groups. An alternative weighting scheme, proposed by Black *et al* (2007) assigns a weight of zero to all unmatched observations, and a weight P(x)/[1-P(X)] to the matched observations, where P(X) is the probability of finding an individual of the disadvantaged group with the same characteristics. Both weighting schemes, then, require an exact matching between members of both groups prior to the estimations. One disadvantage of this approach, is that it focuses only in outcomes, and not on the reasons behind them. Thus, the problem of different reasons for lower experience by women (having children) and some men (with low attachment to the labor market) remains unaddressed. More importantly, the conclusions cannot be generalized to the whole population. Furthermore, such estimates are likely to represent only a lower bound, if the categories for which one of the genders is missing are related to systematically higher or systematically lower wages.

An alternative was proposed by Black *et al* (2002) and Ñopo (2008), who put special emphasis on the comparability of the male and female group *before* the calculation of the adjusted gender wage gap. Both authors use a non-parametric approach employing a matching estimator. Implicitly, these techniques assume that the wage distribution function can be divided in two sections: one where the characteristics of the members of the two groups coincide (common support) and another which represents the deviations of each group from the common support. The logic for this decomposition is that only on the common support the characteristics of the males might be rewarded different than those of females, so only this part of the wage differential can be explained by *discrimination*. Males for which identical females do not exist may earn (discriminatorily) high wages if access to these professions is constrained for women. Intuitively, these males are not “used” to compute gender wage gap for the currently employed women.

Black *et al* (2002) uses exact matching (on observable covariates), but both male and female non-matched observations are discarded. The decomposition relies on:

$$\Gamma(G_j) = \sum_x P(G_{jx})\{E(y|X = x, G_j) - E(y|X = x, \sim G_j)\} - \sum_x \{P_{\sim G} - P_G\}\, E(y|X = x, \sim G_j),$$

where G_j represents the disadvantaged group; x represents a combination of covariates and $P(G_{jx})$ is the probability of finding individuals with this subset of characteristics in each of the groups. The first term represents thus the unexplained component attributable to the discrimination while the second term is the difference in wages justifiable by the differences in characteristics[4].

Dropping the unmatched individuals removes important part of the information from the data set. To avoid that Ñopo (2008) develops a decomposition which allows to identify the part of the wage gap that is attributable to the characteristics of the “unmatched” men and a part of the gap that is attributable to characteristics of the “unmatched” women. In fact, unlike Black *at al* (2002), Ñopo (2008) constructs a counterfactual population of women and the rewards from men. This is done by sampling each time one woman and matching it to all statistically identical men. This way a synthetic counterfactual female wage observation is created, which equals the average wage of the matched men. This procedure is implemented for all women, but – clearly – not all them will have an exact match. Ñopo (2008) uses information on the unmatched men (those who were not identical to any of the woman in the sample) and unmatched women (those for whom a match among men could not be found) to construct the following decomposition:

$$\Delta \equiv E[Y|M] - E[Y|F] = \Delta\mathrm{M} + \Delta\mathrm{X} + \Delta\mathrm{O} + \Delta\mathrm{F}$$

where, $\Delta\mathrm{O} = \left(E_{F,matched}\,[Y|M] - E_{F,matched}[Y|F]\right)$, is the part due to differences in unobservable characteristics, or discrimination, whereas $\Delta\mathrm{X} = \left(E_{M,matched}\,[Y|M] - E_{F,matched}[Y|M]\right)$ is the part due to differences in observable characteristics, where the second term is the synthetic position. Additionally $\Delta\mathrm{M} = \mu^{\mathrm{M}}\left(E_{M,unmatched}\,[Y|M] - E_{Mmatched}[Y|M]\right)$, is the difference due to men out of common support, and μ is the probability of men being not matched. Finally, $\Delta\mathrm{F} = \mu^{\mathrm{F}}\left(E_{F,matched}\,[Y|F] - E_{F,unmatched}[Y|F]\right)$ is the part of the gender gap which can be explained by women having different endowments than men.

Even though the Ñopo (2008) decomposition posits a reliable estimation of the gender gap, it has some shortcomings. First, the matching technique implies a trade-off between the number of characteristics to control for and the ratio between matched and unmatched observations for both men and women. In fact, the more covariates are used for matching, the less likely the exact match. Thus, more reliable estimate of adjusted gender wage gap is obtained, but for a smaller fraction of sample, which limits the external validity of the finding. This problem is referred to as the “dimensionality curse”. Second, the creation of the counterfactual distribution of salaries using the means is probably biased if the overall distribution is skewed. Also, it prohibits profiting from information on differences in wage dispersion between men and women with the same characteristics[5]. Finally, sampling is done over the entire distribution, which makes it challenging to analyze the different gender wage gaps along the wage distribution[6].

[4] Standard errors are obtained via bootstrapping in Black et al (2002) and are derived analytically in Ñopo (2008)

[5] Shorrocks (2012) proposes a non-parametric method employing Shapley value do propose a decomposition in a matching framework which allows to analyze the impact of a given variable in the results of some analysis of inequalities.

[6] This method is also useless if wages are not continuous (e.g. coded within bands). Another problem is that Ñopo decomposition does not create a continuous counterfactual distribution of wages for women. The distribution is full of “jumps”, this reflects the changes in the cell of reference for women.

The problem of uninformative mean

Average gender wage gap may be uninformative if there are large discrepancies in discrimination depending on profession and/or wage level. In extreme scenario, gender wage gap could average to zero if high income women were overpaid, whereas low income ones were heavily underpaid. Notice also that the overly simplistic solution of breaking the earning distribution into several bands not only underutilizes an important amount of information, but it is plainly inconsistent.

A first alternative to deal with the decomposition at different quartiles was proposed by Juhn, Murphy and Pierce (1993) and later developed by Blau and Kahn (1996). This method is parametric, as requires an estimation of a Mincerian wage regression. Coefficients form the advantaged (male) group regression are used to obtain a counterfactual wages distribution for the disadvantaged (female) group

$$Y^j = a^m + \beta_i^m x_i^j + \theta^j \sigma^m \quad ,j = f, m$$

where σ^m represents the standard deviation of the error term in the male regression, while θ^j is a standard error term. In the case of the male equation, $\theta \sim N(0,1)$. In the case of the female equation θ is the difference between the actual value of wages for women, and that predicted by using the male coefficients and the female characteristics (for interpretational convenience, both terms are divided by the standard deviation from male equation). Thus, adjusted gap is defined as:

$$D = Y^m - Y^f = \left(X^m - X^f\right) * \beta^m + (\theta^m - \theta^f)\sigma^m$$

The first term represents the differences in observable characteristics, while the second term represents the unobservable differences between men and women, with σ^m interpreted as the price of the unobservable characteristics. This equation can be used to estimate the differences at several quantiles[7]. The main difference between Juhn, Murphy and Pierce (1993) and Oaxaca-Blinder (1978) decomposition is that in the former X's and θ's do not represent the values at the mean, but rather the average value of the selected characteristics at any given quantile.

However, the method has several problems which undermine its usability. First, the estimations are still done at the mean, which as mention implies that the factors have the same rewards all over the earning distribution. A second problem that the method implicitly assumes conditional rank preservation, i.e. residuals (ordered) follow the same pattern in both the male and the female distribution. This assumption is hard to test, but one should also rarely expect it to hold in practice, as for example the residuals may reflect problems with the method and not just unobserved effort. Additionally, it is difficult to rank the residuals when there are more observations in one of the groups.

Some of these issues are addressed by the Machado and Mata (2005), which is also much more flexible than other decomposition techniques. However, greater flexibility of the M-M estimator comes with a cost as the calculation of the gender gap is considerably more difficult and there is no easy way to attribute the "explained" part of the adjusted wage gap to a particular explanatory variables (i.e. individual and firm level characteristics), which makes it less useful

[7] An example of this is Zhang (2008) and Cho and Cho (2011). Jimenez-Rubio and Hernandez Quevedo (2009) provide an example of the flexibility of the JMP as they incorporate some non- parametric function among the covariates.

for policy recommendations. The approach of Machado and Mata (2005) involves simulating a population of the disadvantaged (female) group with the rewards of the advantaged (male) group. With subsequent simulations, total gender gap at each given quantile (q) is given by:

$$x^m\beta^m(q) - x^f\beta^f(q) = \left(x^m - x^f\right)\beta^f(q) + x^m\left(\beta^m(q) - \beta^f(q)\right).$$

The quality of Machado and Mata (2005) decomposition relies on the number of simulations, but the model also suffers from path dependence. More importantly, this technique effectively estimates a Mincerian wage regression for each simulation, which brings about all of the problems associated with sample selection and functional form. Moreover, it assumes an identical functional form at each quantile.

The efficiency issue was addressed by Melly (2006), who also contributes by deriving the variance of the Machado and Mata estimator. Albrecht *et al* (2008) allow the selection correction at every quintile. Even though these methods represent an improvement over the original Machado and Mata (2005) decomposition, they have both a common disadvantage, which is the arbitrary (and untestable) choice of the functional form in the original Mincerian regression. In other words, the researcher's' choice of functional form affect the final result in an intractable way[8].

An alternative to simulation techniques are again non-parametric estimators, such as proposed by Di Nardo, Fortin and Lemieux (1996). They suggest using the actual, yet counterfactual density of wages (e.g. wages that would have prevailed in a given year if the characteristics were the same as those prevailing in another year, country, group, etc.). The general procedure is based on the premise that the structure of wages can be decompose in two fairly independent parts: a structure of wages, which determines the premiums given to the characteristics and a structure of attributes. Given this assumption[9], the counterfactual conditional distribution can be obtained via a reweighting procedure through which the attributes obtained in group, country or period i are converted to those in group, country or period $j \neq i$. The most important part of the estimator is thus the reweighting procedure, as it is the responsible for the conversion of one structure of attributes into the second. In their setup, following Di Nardo, Fortin and Lemieux (1996), weights should be calculated using a probit model, in which the probability that an observation with characteristics equal to x comes from the group i[10]. The results are in fact two distribution functions[11].

Since wage distribution is by definition a metric, the integral of a difference between them over the whole wage range is by construction zero. However, domains within that range can be characterized by excessively high proportion of women. If that is true for lower percentiles of the distribution (conditional on individual characteristics), the "local" gap between the distributions of wages for women and distribution of wages for men if they had characteristics

---

[8] See for example Weichselbaumer and Winter-Ebmer (2006) meta-analysis of the impact of the methodological choices in the computation of the gender gap.

[9] The assumption implies that the law of demand and supply does not apply to the labor market. It also rules out the possibility that agents are maximizers over time and that they might adjust their attributes to the demands of the market.

[10] Garsch and Paarsch (2000) offer an alternative weighting procedure to obtain the counterfactual distribution, which builds on hazard function models.

[11] The choice of the shape of the distribution is fairly irrelevant, but the size of the bin for density function approximation has important consequences. In addition, in principle the distributions should be defined on exactly the same domain, which effectively requires common support condition to hold.

like women can be interpreted as the adjusted wage gap. In order to obtain a synthetic measure of discrimination other statistics should be used (e.g. a comparison of the mean of the three distributions –male, female and counterfactual female- will be equivalent to a Oaxaca-Blinder decomposition; alternatively quantiles can be used to obtain an equivalent of Machado-Mata decomposition).

One downside of the DiNardo, Fortin and Lemieux (1996) is that its application of this method to the case of detailed decomposition is very limited. Firpo, Fortín and Lemieux (2007) proposed a technique that enables to separate the impact of particular covariate on the explained and unexplained part of the gap for any quantile of the unconditional distribution of dependent variable. This method has important advantages in comparison to Machado and Mata (2005) as it provides detailed decomposition not only of wage structure effect, but also of composition effect. In addition, the values of components corresponding to particular covariates are path independent. Firpo, Fortín and Lemieux (2007) thus provided a tool to obtain results of detailed decomposition for any distributional statistic that has desirable features of Oaxaca-Blinder decomposition results for the mean[12]. This method – referred to as RIF-regression decomposition method, recentered influence functions – is most frequently used to obtain results of unconditional quantile regressions.

The problem of selection bias

Common to any Mincerian estimation of wage equation is the issue of sample selection. Namely, if data on wages is only available for some non-random subsample of individuals, bias is likely to emerge for the characteristics which drive both the likelihood of working (i.e. availability of wage data) and productivity (i.e. particular value of wage). If we rely on the estimated parameters to compute the adjusted wage, sample selection bias undermines reliability of this approach.

The most common method to solve this problem is to employ the Heckman (1979) procedure[13], which relies on the idea, that self-selection bias can be treated as an omitted variable problem and solved by recovering that variation from the available data. Typically, one uses determinants of labor force participation, which should not drive directly the wages. Such candidates are marital status, household structure or availability of non-earned income within the household. These variables are used as instruments in the first stage probit regression of employment, which delivers a correction term for the second stage wage regression[14]. Given this methodological limitation, Heckman (1979) correction is only as strong as the instruments are powerful in predicting the employment status.

In the context of empirical application of the Oaxaca-Blinder decomposition, the sample selection correction is traditionally done only for women, as they are considered to be selected out due to the gender status and the household role division. Given the large and growing size of

[12] However, it lacks a clear control for the self-selection into employment, thus the results do not have features of OB decomposition with Heckman correction.

[13] Most common but not the only. Alternative approaches include the imputation of wages to the non-labor market participants based on their characteristics, or limiting the estimation of the wages equations to a subset of the population were the prevalence of unemployment is relatively low, and hence the self-selection problem can be ignored.

[14] This correction term is in fact a measure of probability that individual is part of the employed subpopulation, conditional on individual characteristics.

the so-called NEETs (Not in Employment, Education or Training) across both genders in industrialized countries, this assumption is not likely to match well the data.

Recently, authors tried to go beyond the Heckman procedure for selection. An example is Machado (2012), who proposes an extension allowing for differences in the selection process at different levels of the income distribution, for men and women. This alternative consists of dividing the population into several groups depending on whether and how they have changed their "decision" to be employed when some circumstances (the instruments) changed. These instruments are subsequently incorporated in the wage regression. The disadvantage of this method is that it requires a panel data for the estimation, which is not usually the case.

Recognizing employment is a matter of (constrained or unconstrained) choice, one needs to acknowledge that also the occupation, industry and the form or employment are (at least partially) endogenous. Omitting them from a gender wage gap analysis makes the estimate of adjusted wage gap flawed, whereas their inclusion makes the parametric estimates unreliable. The solution proposed by Brown *et al* (1980) and Appleton *et al* (1999) is a modification of the Heckman procedure, only that instead of using a probit model in the first stage, a multinomial logit is preferred[15]. Similarly to the Neumark decomposition (1988), Appleton *et al* assume the existence of a non-discriminatory sectoral structure and evaluates the impact of having different selection probabilities for each gender. The equation of the decomposition is given by:

$$W_m - W_f = \sum_{j-1}^{j} p_j^*(\overline{W}_{mj} - \overline{W}_{fj}) + \sum_{j-1}^{j} W_{mj}\left(\overline{p}_{mj} - \overline{p}_j^*\right) + \sum_{j-1}^{j} W_{fj}\left(\overline{p_j^*} - \overline{p}_{fj}\right)$$

$$+ \sum_{j-1}^{j} W_{mj}(p_{mj} - \overline{p}_{mj}^*) + \sum_{j-1}^{j} W_{fj}(\overline{p_{fj}^*} - p_{fj})$$

where j indicates the number of sectors (j-1, as one is the reference category) and p is the proportion of workers in each j sector. The first element in the RHS is just a shortcut for the Neumark decomposition mentioned above. The third and fourth terms represent the differences in earnings due to differences in distributions which can be explained by differences in the characteristics of men and women; while the last two terms (where p_{mj} represents just the sample proportion, represent the part of the differences due to occupation segregation which is not explained by differences in the characteristics[16].

Clearly, one downside of these methods is that women are underrepresented in some professions along the entire wage distribution, not just in high paid jobs. Equally important, this method cannot account for the differences in wages to each of the occupations, which leaves glass ceiling and sticky floor unaccounted for.

[15] One of the few applications include Nordeman and Roubaud (2009) for an analysis of Madagascar.

[16] Brown (1980) method is actually simpler as only one multinomial logit is estimated (for men) and then it is used to create a counterfactual probability distribution for women (applying the coefficients to the mean). In fact it employs the unconditional probabilities for men and women, but the conditional probabilities for the counterfactual distribution.

Summary of different methods

Given this brief review of the literature, it is clear that there is a multiplicity of computing gender wage gap, but also that no perfect method to compute the adjusted gender wage gap exists. In fact, the standard parametric approaches, such as Oaxaca-Blinder (1978) decomposition have a variety of shortcomings, but the subsequent econometric developments address only some of them at a time. The perfect method needs to (i) address selection issues, (ii) compare wages within the common support only and (iii) allow to account for wage differences at different percentiles of the earnings distribution.

On the one hand, methods based on matching – especially Ñopo (2008) – allow to address the problem of common support. Using this method one is able to specifically identify the role of characteristics and the role of "unexplained" components comparing adequately men and women. This method is also immune to the selection issues, but here the causal interpretation needs to be careful. The drawback of Ñopo (2008) is that distributional analysis is not effectively possible. There is also a path dependency problem, i.e. the contribution of each variable depends on the removal order, which constrains the extent of policy relevance.

On the other hand, methods relying on regression and sampling allow to account for selection issues and keep the power to deliver an analysis along the income distribution, but have difficulty in assuring that the comparison only concerns individuals with comparable characteristics. Namely, reweighting is used to balance potential under- or over-representation of one group. However, reweighting does not provide informational content and has limited reliability if weight within one of the groups is strictly (or close to) zero. Also, they make extensive use of the "error term" in interpretation.

The fact that no "perfect" method exists for the time being is why a comparative analysis of these methods is useful. The review of the advantages and disadvantages of each of the presented methods suggests that an ample view of the different possible measures of the gender wage gap is needed.

Section 2. Gender wage gap in Poland – previous studies

The literature on gender wage gap in Poland is not vast. As pointed out by Grajek (2003) Poland had a significant delay in having their academic, business, and political elites recognize this issue. In fact, gender wage gap has been analyzed in Poland mostly in the context of transition period as performed by Grajek (2003) or Adamchik and Bedi (2003). While nearly all wage related studies control for gender, few make a contribution on that particular topic. In fact, Adamchik and Bedi (2003) underline that the relative economic welfare of women is one of the measures of nation's well-being and they doubted if the economic position of females in Poland has improved along with the positive economic performance of the country. They also pointed out that among several indicators – such as income, employment, or social benefits, wages are probably the most important determinant of economic well-being and personal success, and they should be analyzed to assess relative situation of females.

Majority of earlier studies analyzing the gender wage gap focused on comparing the raw wage differentials or on estimating the linear wage regression with a gender dummy. Examples of such studies include Kotowska and Podogrodzka (1994), Kalinowska-Nawrotek (2005), Zwiech (2005)[17] or Mazur-Łuczak (2010). Poland was also included in a number of cross-country studies, such as Brainerd (2000), Pailhé (2000), Blau and Kahn (2001) as well as Newell and Reilly (2001). Without exceptions, all studies find lower wages for women in Poland along with better characteristics, such as higher educational attainment. Unfortunately, few studies dedicated to Poland employed decomposition techniques.

Using data from 1990s Kot, Podolec and Uhlman (1999) employed the parametric Oaxaca-Blinder (1979) decomposition and found that the adjusted gender wage gap was about the double of the raw gap. Adamchik and Bedi (2003) have used both the standard Blinder-Oaxaca method and its modified version presented in Neumark (1998) on data from the same period. According to their findings, the percentage of the wage gap explained by differences in observed characteristics varies across methods, but in both it is quite limited over the analyzed period 1993-1997. What is more, for each year the explained portion of the gap is considerably higher for the modified version, than for standard Blinder-Oaxaca decomposition. The contribution of Adamchik and Bedi (2003) is also important, as it discusses the characteristics that could be used in wage equations. The basic set of regressors in their paper included conventional human capital characteristics (e.g. education or experience), personal characteristics (e.g. marital status), and regional labor market conditions, like information on the region, or whether the area is urban or rural. In the second specification of the set of characteristics they have also included job characteristics, like information on type of industry, occupation, branch of economy (high-paying or low-paying), or firm size.

What is more, the authors discussed possible criticism of inclusion of job characteristics in an earning equation. For instance, a number of job-related characteristics might be endogenous on the labor market. It is not clear if differences in job characteristics for males and females reflect employment discrimination, or different tastes and preferences, or both. At the end, they have followed the convention and treated job characteristics as factors explaining the wage differential between females and males, rather than manifestation of employment discrimination. We follow the same approach in this work, as job characteristics will be also considered as explanatory variables in the empirical analysis.

Grajek (2003) applied the John, Murphy, and Pierce decomposition to analyze data on Polish employees from Household Budget Survey for the period 1987 – 1996. He also found that explained component is relatively small and rises slowly over the analyzed period. Similar method was employed by Łatuszyński and Woźniak (2008), who confirmed the findings of Grajek (2003) using data for 2004.

More recently Magda and Szydłowski (2008) as well as Matysiak, Słoczyński and Baranowska (2010) provide parametric decompositions, focusing on the life cycle aspects. In contrast, Goraus (2011) analyzed the gender wage gap using Ñopo (2008) non-parametric decomposition. Employing quarterly data from the Polish Labor Force Survey over 1995-2010, she finds fairly stable unexplained gender wage gap of approximately 20% of female wage, which is double of the raw wage gap. Słoczyński (2012) employs an innovative technique of population averaged treatment effects analyzes the regional differentiation of the gender wage gap. However, this

[17] Oddly enough, this topic is a frequent subject of MA theses, with students exploring both various methods (as discussed earlier in this paper) and various sources of data (including LFS, household budget surveys, Social Diagnosis, PGSS, etc.). These manuscripts remain unpublished, which prevents general access or citation in this study.

study employs Structure of Earnings Survey, which prevents accounting for the selection effects[18].

In addition, Rokicka and Ruzik (2010) analyze gender wage gap in informal employment in Poland, using Melly (2006) decomposition. They find greater extent of adjusted gender wage gap in informal employment than in formal employment. In addition, discrimination is larger at the bottom of informal sector earnings and in the top of the distribution for the formal sector employees. However, their analysis relies on an unrepeated and unrepresentative survey focused on informal employment, which limits the external validity of the findings.

Summarizing, empirical evidence is surprisingly scarce, yet consistent. Gender wage gap is a general phenomenon in Poland, reflected in both raw and adjusted gender wage gaps. Compared to the previous studies on Poland, our paper contributes in three major ways. First, we provide a comparison of the estimated adjusted gender wage gap for various methods. Second, we employ a rich data set, which permits controlling for a large number of observable characteristics as well as fairly accurately account for the selection bias. Finally, we provide estimates of adjusted gender wage gap on a large and representative sample of nearly 250 000 individuals for a recent period, thus filling the gap between the studies from late 1990s and early 2000s and the current times.

Section 3. Data and method

This paper uses the most recent information on Polish labor market, as the data was extracted from the four rounds of the Polish LFS of 2012. The total sample consisted of 296,427 observations, corresponding to individuals between the ages of 18 and retirement age. These ages were considered as the lower and upper limit for presence in the labor market. Additionally, the analysis is limited to the subsample of men and women that are neither self-employed nor so-called helping family members. The sample is evenly split between men and women, as the latter represent approximately 50,6% of the total observations.

All variables were constructed following the standard measures. The dependent variable is hourly wage, which was obtained by dividing monthly wages by the number of worked hours. Working hours equal the number of hours worked on average during the week times four.

The data contains standard demographic variables for all individuals, which in addition to gender include age (in years) and marital status (in relationship, single, widowed, divorced/separated). We can also identify whether the place of residence is a rural area, a large city or neither. All analyses show that both the capital region (Mazovia) and large cities (above 50 thou. inhabitants) are characterized by consistently higher wages. These variables were thus indispensable for the study.

We capture human capital by the measures of highest educational attainment. We use three levels: primary education or less, secondary and vocational education and tertiary education. In

[18] SES data contain information on employees only.

addition, the data set is rich enough to contain information on the actual field of education[19]. In addition, the data contains also declared overall work experience and tenure with the current employer. Both these measures are declared in years and months (integers) and have been recoded to years (natural numbers). In addition, ISCO coding of occupations is available at one digit level of disaggregation.

The data is also relatively rich on employers characteristics. We are able to identify the size of the firm in which an individual is employed. In addition, the data covers industry where it pursues its activities. The original variable has 11 levels, following NACE categorization. However, as a robustness check we have also grouped these categories into broader ones: agriculture, construction, manufacturing and services. In addition to industry, the data also contains information on the public or private ownership.

The labor market status is identifiable directly, based on self-reported indication of employment or unemployment. As is standard in LFS type data, individuals are asked if they have worked for at least 2 hours in a week preceding the survey. If they have not, they are asked about willingness to undertake employment and active seeking[20]. Only if the respondent is non-employed but seeking for an employment, we consider him as unemployed. Otherwise, individuals are characterized as inactive. We use only data on active individuals. Since wage data is missing for the unemployed individuals, correction for selection into employment is required.

As discussed earlier, the reliability of the Heckman correction relies on the relevance of the instrument variables in the decision to work. In this study, we use the available data on the structure of the household and its sources of income. First, we employ data on the presence of children and small children (younger than five years old). We evaluate the results with both variables and decided to include the former on the selection equation. Second, LFS provides information on the presence of other sources of income available for the household. We include three sources a retiree (with pension benefits), other earners and social benefits received by the household. Intuitively, these variables are likely to affect the alternative cost of working and thus the labor supply decision. However, these variables are not related to individual productivity. One notable exception may be presence of small children in the household, as they may require more attention from the parents. Thus, not only labor supply but also effective productivity are likely to be affected. To address this problem, children were used in both selection and wage regression as well as in wage comparisons for the non-parametric methods.

Table 2 provides the descriptive statistics for our data. In fact, for nearly all characteristics, the difference between men and women is non-zero in a statistically significant way (although the economic relevance of that difference may be low in many cases). Male workers earn on average approximately 25% more than their female counterparts(using male wage as a reference), they also work slightly more hours a week. Hourly wages are also higher for men, though the difference is much smaller, approximately 9% of male wages. Female workers are better educated, as almost a third of them has a tertiary degree, while in the case of male only 16% belong to the highly educated category. Male workers are also a bit more experienced, by around one year; but they have a slightly lower tenure with the current employer.

---

[19] There are 9 categorical variables for tertiary education: pedagogy and teaching, human sciences (including art and languages); social sciences (which includes law and economics); natural sciences, exact sciences, engineering; agricultural sciences and veterinary; medicine; services, and others.

[20] In fact, those who have not worked are subsequently asked a follow up question if lack of work is associated with holidays, sickness, strike etc. Only if not, question leading to determining the unemployment status are asked.

Table 2: Descriptive statistics

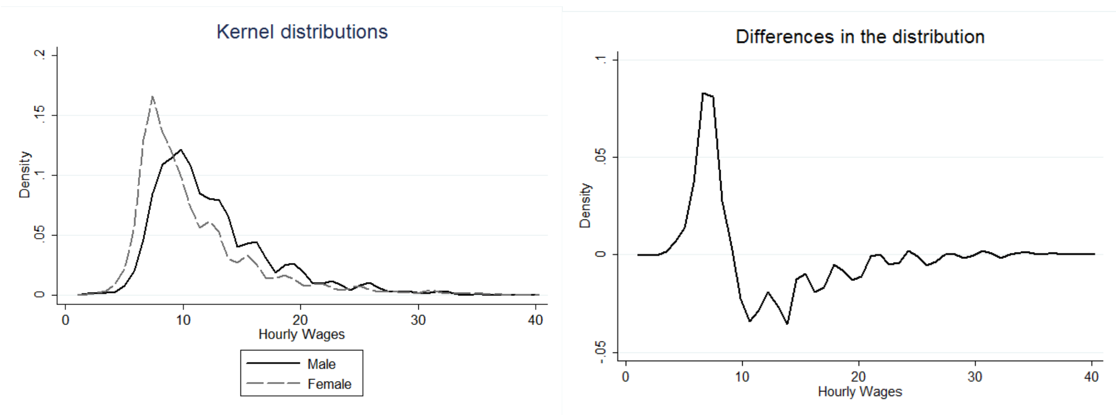
	Overall	Male	Female	t-stat	Support
% active		0.59	0.49		
Among active, % employed		0.87	0.86		
Average monthly wage (in PLN)	1 861.60	2 032.83	1 677.70	32.17***	0.24
Weekly hours	39.96	41.43	38.31	40.69***	0.31
Average hourly wage	11.48	11.91	11.00	15.19***	0.12
Age (in years)	40.56	40.27	40.87	-4.96***	0.04
Primary education (%)	1.00	1.00	1.00	.	.
Secondary educaton (%)	0.68	0.74	0.61	26.08***	0.19
Tertiary education (%)	0.25	0.17	0.33	-36.82***	0.27
Experience (in years)	18.62	19.23	17.96	10.20***	0.08
Tenure (in years)	9.51	9.18	9.86	-6.71***	0.05
Married (%)	0.68	0.69	0.67	3.31***	0.02
Divorced (%)	0.05	0.03	0.07	-18.22***	0.13
Residence in rural areas (%)	0.40	0.43	0.36	14.65***	0.11
Residence in large cities (%)	0.32	0.30	0.35	-10.18***	0.08
Residence in Mazovia (%)	0.10	0.10	0.11	-1.56*	0.01
Second earner (% of individuals)	0.91	0.91	0.92	-5.56***	0.04
Children in household (as above)	0.19	0.21	0.17	10.36***	0.08
Pension in the household (as above)	0.04	0.05	0.04	2.64***	0.02
Market services (%)	0.39	0.35	0.43	-16.20***	0.12
Non-market services (%)	0.21	0.08	0.35	-65.85***	0.49
Construction (%)	0.08	0.15	0.01	50.64***	0.39
Manufacturing (%)	0.31	0.41	0.21	42.64***	0.32
Private firm ownership (%)	0.67	0.76	0.58	38.40***	0.28
Firm size (<50 employees) (%)	0.50	0.47	0.54	-11.8***	0.09
Firm size (50 to 250 employees) (%)	0.27	0.27	0.26	1.05	0.01
Firm size (>250 employees) (%)	0.18	0.20	0.16	9.25***	0.07
High skills (ISCO 1-3, %)	0.30	0.21	0.4	-40.98***	0.3
Low skills (ISCO 5-8, %)	0.62	0.74	0.49	49.98***	0.37
Arts and humanities (%)	0.02	0.01	0.03	-15.21***	0.11
Engineering & construction(%)	0.39	0.59	0.18	86.91	0.65
Social sciences and law(%)	0.18	0.07	0.28	-54.06***	0.40
Medicine (%)	0.04	0.01	0.08	-34.65***	0.26

Source: own calculation, Polish LFS, four quarters of 2012. *Note*: numbers refer to the working population. The common support test follows Imbens and Rubin (2007), and was constructed as the absolute value of the ratio of the difference in the means to the square root of the sum of the variances. Agriculture represented only a small part of paid employment, and hence was not presented.

Finally, we can notice a few things about the work characteristics of men and women. Men tend to be more concentrated in the private sector and work in larger companies. As expected the percentage of male workers in rural areas is larger than in the case of women, which are concentrated in cities. Women, by contrast are more present in the public sector. We also observe a sort of occupational segregation into different industries –sectors- and in different fields of study. For example, women are more concentrated in social sciences and law, while in the case of male workers, the most common educational field is engineering and construction.

The last column presents a test to evaluate how much the distributions of characteristics of the two genders overlap. This test was proposed by Imbens and Rubin (2009). A rule of thumb for interpreting this index is that values over .25 should be a source of concern, as they indicate important differences in the location of the distribution of covariates (the distributions do not overlap enough). In our case, we observe that only for a few variables, we might encounter difficulties: the amount of hours worked, the high education and the private sector. The largest differences are on the industry and the field of study. These results highlight the importance to use methods that control for the common support, especially in the models including industries and field of education.

Graph 1: Estimates of kernel distributions for hourly wages across genders.



Source: Polish LFS 2012, results for hourly wages

In principle, average wage differential may reflect both a shift between the distributions and differences in the shape of the wage distributions. To identify which of the two effects is present in Polish LFS data, kernel distributions of hourly wage were obtained, Graph 1. The graph on the left shows the percentage of workers (densities) at every point of the income distribution for hourly wages[21]; while on the right the line represents the difference between female and male densities. The graph on the right indicates that the differences in wages are not constant along the distribution. At the lower levels, for hourly remuneration between 3PLN and 10PLN (approximately), there is a higher concentration of women than men. From then on, the difference is negative along the rest of the distribution, although it gets smaller for higher levels of income. The female curve is taller and thinner, which indicates that the women's salary distribution is more condensed that the corresponding function for men. Moreover, the distribution of men is farther to the right, which is consistent with men receiving higher average wages. The spikes in the graph reflect the fact that people tend to report round numbers for salaries, a problem common to survey self-reported wage declarations. The differences in the

[21] Here we only look at hourly wages lower than 40PLN

distribution visible in the right hand graph suggest that besides the mean, we should also focus on other statistics of interest.

In total, we use 7 decomposition methods and 7 different conditioning sets (observables we use to control for differences in endowments). More specifically, we compare the results from a regression with a gender dummy, Oaxaca-Blinder (1979) decompositions, Juhn, Murphy and Pierce (1993) decompositions, Machado-Mata (2005) decompositions, recentered influence function approach by Firpo, Fortin and Lemieux (2007) as well as Ñopo (2008) decomposition[22]. In each case we perform two analyses: one for the total sample and the other one for the sample constrained to the common support, as derived from Ñopo (2008)[23].

These decomposition techniques were performed for different specifications of the conditioning variables. First, the basic conditioning set, regardless of the method[24], includes age, education, experience, tenure, marital status and geographical indicators, as described in Table 2. Second, we use household level information as labor supply controls. This group of variables is used as exclusion criterion in the Heckman corrected parametric estimations. We also use information on children in non-parametric estimations. The second group adds the occupation of the individual to the basic variables. In the third and fourth group we include firm related factors, such as industry (in both) and the size and type of ownership (only in the fourth). The following group includes tenure as an additional covariate. In the next specification, we added the information on the field of education. Finally, we repeat the estimations for all the variables combined.

Section 4. Results

We present the results in three substantive parts. First, we compare the estimates of the adjusted gender wage gap for the basic specification. This specification includes age, education, geographical indicators, marital status, experience, tenure and presence of children in the small household. In the case of Heckman corrected coefficients, also sources of income in the household are used. This choice of variables is widely acknowledged in the literature, cfr. Belzil (2007).

Second, we compare the estimates across the methods, depending on the inclusion of additional control variables. Namely, we include separately industry, firm characteristics, occupation and finally the field of education. Again, when relevant, these results are presented with and without common support restriction.

In the third part of this section we discuss the differentiation in the obtained estimates across methods and conditioning sets. Namely, we diagnose the range and the reliability of estimates from the employed method,.

Adjusted gender wage gap for narrow choice of conditioning variables

---

[23] Unless otherwise stated, regression-based methods account for Heckman correction. Estimates without Heckman correction are used only in Table 5 for comparison purposes..,

[24] In the Ñopo (2008) decomposition, we recoded the variables age, experience and tenure in 10 year-groups, The size of each group was selected to maximize the number of matched observations without losing explanatory power.

In Table 3 we compare the values from different decompositions, with and without the common support exclusion. In all the specifications the gap is significant and it points in the direction of lower wages for female employees. The different estimations were also consistent, as they signal a 10% gap at the mean. In the quartile regressions, the results indicate that the raw gap is larger at higher quartiles of the income distribution. The adjusted gender wage gap is in all cases larger than the raw gap. This means that if we consider only their characteristics, women should earn more than men. Sample selection bias has an sizeable impact on the estimates of the gender wage gap. Also, estimates within the common support are typically higher than for the total sample. Thus, it seems the more specification accounts for comparing only the "comparable" the higher the estimates of the adjusted wage gap.

Estimates that operate at the mean yield surprisingly similar estimates – whether the approach is parametric or not, the estimates of the adjusted wage gap within the common support fall into the range of 15.5%-19.6% with the nonparametric approach by Ñopo (2008) exactly in the middle of this range. Results are less consistent when we compare estimates for quartiles. The adjusted gap at the median ranges between 13.9% and 23.6%. However, it should be underlined that the exceptionally high value was obtained only from the reweighting method of DiNardo, Fortin and Lemieux (1996), while other methods provided estimates similar to those for the mean. On the other hand, within each method there are important differences between quartiles of income distribution, thus concentrating on the mean would not provide the full picture of gender inequalities in Poland. Graph 2 yields further corroborates this assertion. DiNardo *et al* (1996) decomposition allows the construction of a counterfactual distribution of wages for either male or female workers. We used the same set of regressors as in Table 4. On the left side, the graph depicts the distribution of hourly wages (in logs) for women and how it would look like if they were to be treated as men[25], and the reverse on the right hand side. The dashed line represents actual wages while the straight line represent how the distribution would like if distribution of the other gender was to be used. Both graphs speak of a gender pay gap which cannot be explained solely by the characteristics of the workers[26].

The fact that the counterfactual curves resemble so much the actual curves for the opposite gender is due to the similarities in the observed covariates used in their construction. The counterfactual graphs indicate that women would have a flatter structure of wages if they were to be paid like men, with a mode to the right of the actual structures. Their wage would on average be higher and also with a larger dispersion. We interpret the distance (difference) between the actual and the counterfactual distributions as a case of missing workers: given their characteristics, there should be a larger proportion of females in the upper side of the distribution.

[25] This counterfactual is the reweighted male distribution, so its characteristics matched those of women, but keeping the male salaries.

[26] The gender gap can be calculated from either distribution pair. The differences between these two calculations are reminiscent of the existing between the male and the female component in the OB decomposition. In addition, it would be possible to use the values from a pooled probit as reweighting factors, in which case the results would resemble the Neumark decomposition,

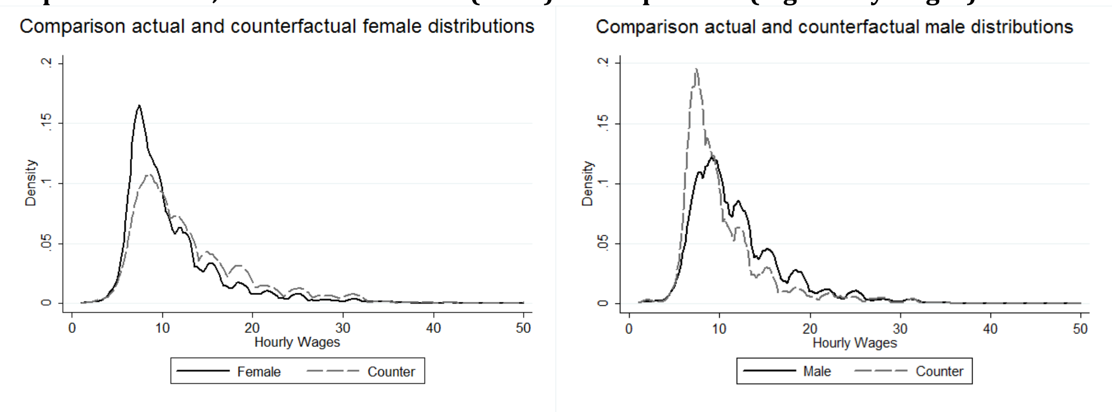
Table 3. Gender wage gap from different methods

	Total sample		Common support	
	Raw	Adjusted	Raw	Adjusted
Linear estimates				
OLS		15,6		15,9
Heckman corrected		16,1		16,4
Parametric (linear) decompositions(with Heckman correction)				
Blinder	9,1	17,1	9,5	17,3
Oaxaca	9,1	15	9,5	15,6
Reimers	9,1	16,1	9,5	16,5
Neumark	9,1	16,1	9,5	16,4
Quartile decompositions - Juhn, Murphy and Pierce (with Heckman correction)				
p25	11,8	16,9	11,8	17,1
p50	13,2	18	13,4	18,3
p75	11,8	17,6	14	19,6
Quartile decompositions - Machado Mata (2005)				
p25	8,7	13,6	10,1	14
p50	12,5	16,	12,5	16
p75	10,5	17,1	10,7	17
Conditional quantile decompositions - DiNardo, Fortin and Lemieux (1996)				
p25	8,7	18,9	10,1	18,9
p50	12,5	23,6	12,5	23,6
p75	10,5	22,3	10,8	22,3
Unconditional quantile decompositions - RIF regressions				
p25	7,3	9,8	7,4	10,2
p50	8	13,9	8,9	15
p75	12,3	22,1	12,4	21,6
Nopo (2008)				
Mean			7,6	17,5
% of matched male				96,1
% of matched female				94,4
no of observations				33 928

Source: own calculation, Polish LFS, four quarters of 2012.

Note: basic specification includes age, education, marital status, experience, tenure, children in the household, region and residence characteristics. Specifications with Heckman (1979) correction comprise household characteristics (children under 5 years of age, other earner in the household, other unearned source of income in the household). For details, see Table 2. Ñopo (2008) decomposition has 94,4% of females and 96,1% of males matched. These estimates are used for the common support. The Reimers and Neumark decomposition divide the gap into male advantage and female disadvantage. In this table we present only the sum of both. The number of observations in the common support is the average of the percentage matched in each group times the total observations.

Graph 2: DiNardo, Fortin and Lemieux (1996) decomposition (log hourly wages)



Source: Polish LFS 2012, results for hourly wages

Our main conclusion is that the different methods yield consistent results which fall into a fairly narrow range of estimates for the gender wage gap, while also the estimates of the gap tend to increase as we focus attention on more similar men and women. Even though the results are consistent across different methods and along the whole distribution, one could argue that our specifications are susceptible to bias resulting from an omitted variable problem. For one, if men perform different jobs then women (as a reflection of their preferences), one would expect to see such outcomes, regardless of educational attainment. This relates to both occupations and industries, but also to the characteristics of the employer and the fields of education. In the subsequent section we extend our specifications to include these variables and test how vulnerable were different methods were to narrow model specification.

Adjusted gender wage gap with extended set of conditioning variables

The results for the different conditioning sets are displayed in Table 4. We include the specifications for the total sample as well as for the common support. In Table 4, this is important because as the number of variables increases, the common support restriction eliminates a larger share of the observations. This undermines the generality of the conclusions. In the extreme case, when all variables are included in the model, the percentage of male and female matched is below 15%, which makes the conclusions very specific.

These results allow to identify the role of three sources of variations in the size of the estimated adjusted gender gap: the introduction of new variables, the different methods and the common support restriction. In all specifications, the gap increases in the common support which corroborates the finding that the wage differences are larger among more comparable men and women. In other words, there are unmatched women at the top of the distribution and unmatched men at the lower points. An extreme example are the results from the JMP decomposition. When all women are considered, it looks like the gap is present only at the lower quartile. After we controlled for the common support, it shows a glass ceiling effect in half of the specifications.

Table 4. Adjusted gender wage gap from different methods with extended conditioning set

	Industry		Industry +		Occupation		Tenure		Education		All	
	All	CS	All	CS	All	CS	All	CS	All	CS	All	CS
Linear regression												
OLS	16.1	17.1	15.8	16.1	15.0	15.7	16.5	17.1	18.4	18.4	15.2	15.4
Heckman corrected	16.6	17.6	16.2	16.6	15.5	16.2	16.9	17.5	18.8	18.9	15.5	15.7
Parametric (linear) decomposition (with Heckman correction)												
Blinder	15.9	19.0	15.6	17.9	18.3	18.9	17.9	18.3	22.0	22.3	18.3	17.6
Oaxaca	15.0	16.0	14.9	15.3	14.3	15.1	15.9	16.8	15.2	16.1	13.7	14.6
Reimers	16.6	17.6	16.2	16.6	15.5	16.2	16.9	17.5	18.8	18.9	15.5	15.7
Neumark	15.5	17.5	15.3	16.6	16.3	17.1	16.9	17.6	18.7	19.6	16.1	16.0
Quartile decomposition - Juhn. Murphy and Pierce (1993) (with Heckman correction)												
p25	16.1	18.4	16.4	16.7	15.5	16.7	17.6	18.6	18.9	18.2	15.0	17.3
p50	18.0	20.5	18.8	18.5	17.5	18.4	18.2	20.1	20.8	19.2	18.1	19.7
p75	16.1	19.0	17.3	16.6	15.9	20.1	18.4	21.3	20.2	20.2	17.9	14.5
Quartile decomposition - Machado Mata (2005)												
p25	11.4	14.0	10.6	14.0	9.0	14.0	14.1	14.0	12.3	14.0		13.0
p50	13.7	17.0	13.4	17.0	13.5	17.0	17.2	16.0	15.0	16.0		18.0
p75	15.0	18.0	14.8	19.0	15.1	18.0	18.3	18.0	15.8	17.0		12.0
Cond. quantile decomposition - DiNardo, Fortin and Lemieux (1996)												
p25	13,2	15,4	12,5	11	8	15	15,4	15	11,8	12,7	8	10,1
p50	18,2	18,2	18,2	18	18,2	18	18,2	18	18,2	18	18,2	18
p75	22,3	22,3	22,3	22	22,3	22	22,3	22	22,3	22	22,3	22
Unconditional quantile decompositions - RIF regressions												
p25	8.8	12	8.2	11	10.9	11.9	11.4	12	8.7	8.4	7.3	7.4
p50	12.1	18	14.7	19	14.9	17.5	17.2	20	14.1	20	8	8.9
p75	21.3	20	24.5	26	23.9	18	25.2	21	23.3	21	12.3	12
Nopo (2008)												
Mean	17.0		16.0		16.0		17.0		18.0		16.0	
% of matched male	78		51		75		87		90		8	
% of m atched female	84		57		78		72		89		9	
No of observations	33 574		33 574		33 921		33 928		33 928		33 567	

Source: own calculation, Polish LFS, four quarters of 2012.

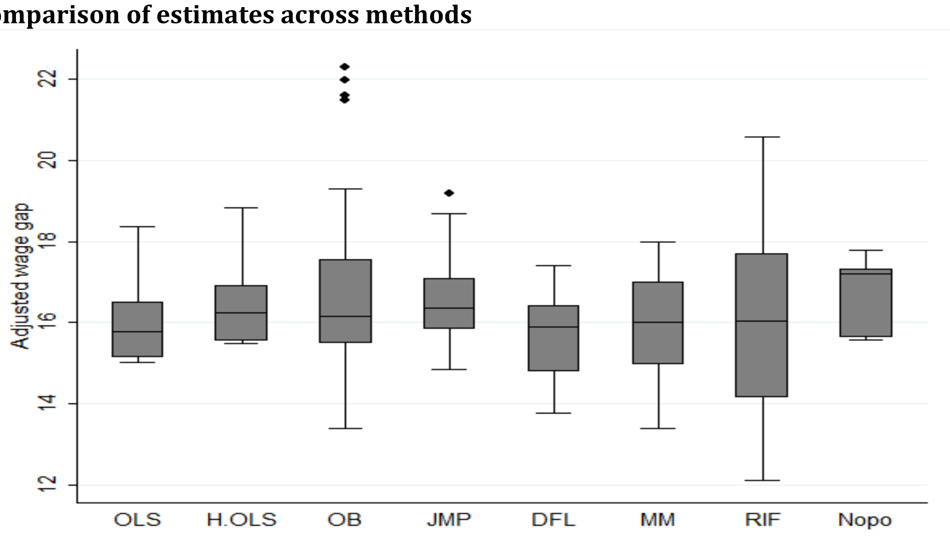
Note: please, refer to Table 3. Industry (column 1) includes 4 dummy variables (agriculture, manufacture, construction, services). Industry + (column 2) includes industries and other firm level variables (private ownership dummy and size). Occupation (column 3) comprises 9 dummy variables for each ISCO code at 1 digit. Tenure (column 4) comprises tenure in the current job (in years). Education (column 5) comprises dummies for the field of education, see footnote 17. Finally, in the last column we included all the variables together. The model failed to converge in the last specification of the Machado and Mata (2005) decomposition.

In terms of variables, including industry and firm level characteristics (the "industry +" specification) among the covariates lowers the estimates of the adjusted gender wage gap, by app. 2-4 percentage points with respect to the basic specification. The addition of occupation dummies has negligible effect on the estimates. Comprising tenure tends to increase the adjusted gap estimates, but quantitatively the effects are not large. In the case of education dummies, the effect is stronger. In 29 out of the 38 estimations the gap raises after we incorporated the field of education to our basic set. Moreover, in most cases the maximum values of the gender gap are found in this column. By contrast, in the few estimations where the value of the gap falls, the size of the drop is relatively small in absolute and in relative terms, when compared to "industry+" The inclusion of all explanatory variables in a single estimation does not reduce the adjusted wage gap significantly in most specifications. The sole exception is the Machado-Mata, where the value of gap decreased by 5 percentage points at the higher quartile. Notice, however, that this coefficient was calculated in a rather small sub-sample

Quality of the gender wage gap estimations

Table 4 also allows a comparison of the different methods. These are synthesized in Graph 3. The results were on average higher in the Juhn, Murphy and Pierce (1993) method, both in and out of the common support. This might result from the treatment given to the residual imputations, which can have a larger impact when there is a great dispersion of wages, or when the coefficients are not constant. The results from the RIF offered a clearer distinction between the low income quantiles, where the adjusted gap is smaller, and the top of the distribution, where it is significantly higher. Notice that unlike the rest of the estimations, the RIF decompositions resulted in larger gaps in the whole sample than in the common support. This highlights the importance of controlling for the common support in this method, as otherwise the results will be much larger than in any other method.

Graph 3: Comparison of estimates across methods

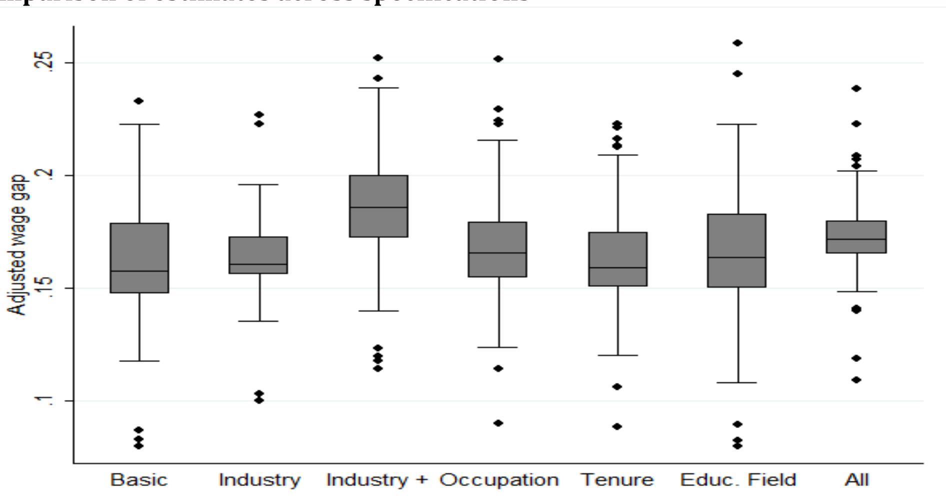


Source: estimates from Table 4, H.OLS stands for Heckman corrected OLS,; OB for Oaxaca Blinder, in all its variations; JMP, for Juhn Murphy and Pierce (at the mean); DFL, for DiNardo, Fortín and Lemieux (at the mean); MM, for machado and Mata, estimated at the median; RIF, for the recentered influence functions at the median. Nopo stands for Ñopo decomposition, of course.

We also observe that the DiNardo, Fortin and Lemieux (1996) results are similar across the specifications, in fact the median and the third quartile share the value, which we interpret as a reflection that the changes between specifications are relatively small, which is consistent with the rest of the table[27]. With respect to the methods based on means, we observe that the values are on average lower than those from the quantile analysis, revealing that the mean partially masks the differences along the distribution. The selection of the reference wage structure is also significant, as we observe that the maximum values when the female structure is taken as a reference (Blinder), and the minimum when male structure is used as a reference (Oaxaca). We do not observe a significant difference on average between parametric and non-parametric methods.

We can also analyze which methods are more sensible to the changes in the specifications. For this, we calculated the variance of the results and the range over the mean. First, we observe that the results varied less inside the common support, which is somehow surprising given the changes in the sample composition. But at the same time, it provides some additional support to the idea that the gap is constant across specifications. Among the quartile methods, Machado-Mata (2005) produced the more stable results. If we only consider the common support, it was the method with the least variation. This indicates that Machado-Mata (2005) is less sensitive to the omitted variable bias than the Juhn, Murphy and Pierce (1993) method and therefore it should be preferred when there is a fear of this bias.

Graph 4: Comparison of estimates across specifications



Source: estimates from Table 4

A similar comparison can be made between mean-based methods. The Ñopo (2008) decomposition was the more stable method across different specifications. Among parametric methods, the Oaxaca (1979) decomposition, based on the male structure, presented the smaller variance. We conclude that this methods are better fitted in the presence of an omitted variable bias. By contrast, the more volatile methods, so to say, were those based on the female wage structure. OLS regressions produced in-between results, which were very similar regardless of the control for sample selection. That being said, the Heckman corrected equations resulted on slightly higher coefficients (0.4 pp difference on average) and exhibited smaller variability.

---

[27] This is a problem of data quality, as the answers of the salary present clear spikes for some levels of wages, because people tend to provide round numbers for their salary. This resulted in points with a high frequency, which means that the estimation at those levels were more inaccurate. We could have dealt with this problem using kernel estimations, but in that case we would have to decide the level of bandwidth to maximize the differences. In practice, it would have meant a great level of arbitrariness .

Last but not least, it is often argued that lower wages of women reflect self-selection into lower paying occupations, industries and even fields of education. Our results do not confirm this contention, regardless of the method used. In fact, when controlling for industry characteristics the estimates of the gender wage gap are somewhat lower, but adjusted wage gap remains substantially higher than the raw gap. Inclusion of tenure and fields of education yields estimates somewhat larger, making the difference between raw and adjusted gap even more – not less – pronounced. Finally, occupations do not play a role in the gender wage gap: women and men in the same position are still paid differently. On the other hand, the changes in the common support are large, reflecting the larger concentration of men in low paying sectors. We thus reject the hypothesis that women self-selection into "lower paid" jobs and fields can explain the wage gap.

Table 5. Exploring the adjusted wage gap

Dependent variable: adjusted wage gap	(1) All estimations		(2) Mean based methods		(3) Distribution based methods	
	Selection bias					
Common Support	0.59***	(0.21)	0.38**	(0.15)	0.78**	(0.37)
Heckman correction	1.06***	(0.19)	0.59***	(0.14)	1.43***	(0.31)
	Method (Ñopo in columns 1-2 and JMP in column 3 are the reference levels)					
OLS	-0.07	(0.58)	-0.22	(0.323)		
Oaxaca-Blinder	0.12	(0.32)	0.07	(0.176)		
JMP	0.14	(0.41)	0.06	(0.243)		
DFL	-0.54	(0.46)	-0.63*	(0.323)	-0.58	(0.41)
MM	-2.97***	(0.53)			-2.94***	(0.46)
RIF	-1.97***	(0.53)			-2.01***	(0.46)
	Variables (all = reference level)					
Basic	0.46	(0.34)	0.39	(0.26)	0.55	(0.55)
Field of Education	2.45***	(0.34)	2.90***	(0.26)	2.13**	(0.55)
Industry	0.88**	(0.34)	0.75***	(0.26)	1.02	(0.55)
Industry +	0.37	(0.34)	0.15	(0.26)	0.56	(0.55)
Occupation	0.46	(0.34)	0.43	(0.26)	0.77	(0.55)
Tenure	2.45***	(0.34)	1.35***	(0.26)	1.46*	(0.55)
Control for quantiles	Yes		No		Yes	
Constant	14.74***	(0.004)	15.15***	(0.002)	14.55***	(0.006)
Observations	424		217		249	
R-squared	0.577		0.513		0.597	

Source: own calculation, results reported in Table 4. Column (2) comprises also the estimates from Juhn, Murphy and Pierce (1993) and the DiNardo, Fortín and Lemieux (1996) estimations at the mean. *Note:* The conventional p-levels are represented by the asterisks (** for 5% significance and *** for 1%)

In addition to the visual inspection of the results, we proceed to a more quantitative analysis by estimating a simple OLS to explore the determinants of the changes in the measures of the adjusted gender wage gap. The results are presented in Table 5. First, we compare all specifications. Second, we compare only the estimation methods which operate at the mean of distribution and separately we analyze the results from the distributional methods. To a large extent, the OLS regressions confirmed our previous analysis; moreover, most of the variables were significant predictors of the differences in the gap, which we observe in the low p-values and the high R-squared.

In each of the specifications, the estimates of the adjusted wage gap are the smallest when we include all variables, as can be inferred from the positive signs in all other set variables; though in some cases the differences are not significantly different from zero. The relations between the sets of characteristics ratify what was explored in the previous section. Inclusion of industry, tenure and education field in the specification is associated with a larger adjusted wage gap, with respect to the basic specification. This once again corroborates our earlier claim: the hypothesis that self-selection of women into low pay jobs occurs cannot explain the persistence of the adjusted wage gap.

We also ratify our intuitions with respect to the common support and the Heckman correction mechanisms, although in the latter case, the effect was much smaller. Both variables are related with higher estimations of the adjusted wage gap, which means that women experience more selection and that the unmatched women have higher wages than unmatched males. Moreover, the quantile regression methods seem to be more affected by the calculation in the common support than the methods based on means.

The comparison across methods shows that the estimations at the mean are similar to each other, as indicated by the insignificant coefficients for the different methods in the first regression. The only exception is the Machado-Mata decomposition, which produces lower estimates. When we compared inside each group (mean and distributional), we discovered that the mean based methods produce roughly the same results, except for the DiNardo, Fortin and Lemieux (1996), whose estimations for the mean were lower than. We also observe that as expected the OLS resulted in lower estimates, though it is not significantly different from zero.

In the distributional analysis, the selection of the method has bigger influence on the results. Both Machado-Mata and RIF decomposition resulted in smaller values of the gap than the JMP and the DFL decompositions, which were pretty similar. Also, this methods seemed to be less affected by the selection of the variables, with the exception of education and tenure, which produced the larger results. This confirms the advantages of including the field of education.

Conclusions

There is a multiplicity of the methods to estimate the adjusted gender wage gap. The differences between them are both methodological and interpretational. Our objective was not to determine which of the methods is a “perfect” one in determining the true gender wage gap. Instead, we performed a comparative analysis of the available alternative for computing the gender wage gap. To the best of our knowledge, such analysis has not been conducted before, and hence this is a valuable contribution of this paper, especially as it demonstrates that the selection of a method and the set of conditioning variables is far from trivial. A valuable side-product of our research was a characterization of the gender wage gap in Poland.

We used data from Polish LFS of 2012 and hourly wage as measure of compensation. The raw gap amounts to roughly 10% of male wage. In order to obtain the adjusted wage gap we applied 7 different estimation methods and employed different sets of conditioning variables. In spite of these differences, the methods and specifications provided similar results: adjusted wage gap falls in the range between 15% to 20% per cent of the male wage. All of the estimations showed that the adjusted gap was larger than the raw gap, which means that given their endowments, women should have received a larger pay than their male counterparts

A number of studies emphasizes that the adjusted wage gap tends to be overestimated if the choice of variables is too narrow. Specifically, women tend to locate in occupations, industries and even fields of study where the returns to individual characteristics are lower. Also, women are believed to have less experience. Our results suggest that the inclusion of these variables – if anything – raises the estimates of the adjusted wage gap, which would indicate that the variation inside each of these categories is larger than between. Consequently, it seems that in the case Poland, even if segregation occurs, accounting for it does not help to explain the gender wage gap Notwithstanding, these changes are not large economically.

In addition to the selection of variables, our paper analyzed the role played by the methodological decision taken by the researchers. Our results highlighted the value of the common support and distributional analysis. The adjusted wage gap is typically larger inside the common support than outside, which indicates that unmatched women are better endowed than their male counterparts. Also, the adjusted gap was not constant at different percentiles, an impression which was ratified using three different methods: the Juhn, Murphy and Pierce (1993); the Machado Mata (2005) decomposition and the DiNardo Fortín and Lemieux (1996). Women are missing at higher levels of the wage distribution, *ceteris paribus*, while the extent of the adjusted wage gap is higher in top deciles. These conclusions could be interpreted as guidelines for improving the current framework for equality opportunities implementation.

Our work provides an applied comparative analysis of the different methods on one sample of data. Thus, in a sense, we extend the findings of Wichselbaumer and Winter-Ebmer (2006) meta-analysis. First, our results are fully comparable across methods. Second, we directly control the inclusion of additional explanatory variables on the estimated size of the adjusted wage gap.

Our findings can be summarized as follows. While all the methods provide similar results, even across several subsets of variables, there are also some differences among them. First the estimations within the common support resulted in adjusted wage gaps estimates, with smaller dispersions. Thus, it seems that "comparing the comparable" makes the results less sensitive to the changes in the conditioning set. Inside the common support, Ñopo (2008) decomposition showed the least variation, and should be preferred when there is fear of an omitted variable bias. The OLS tends to produce somewhat smaller results, hence we might interpret those estimates as a lower bound of the adjusted wage gap. The results from the parametric decompositions differed by as little as +/- 1,5 pp. This implies these methods are fairly interchangeable. Methods based at distributional analysis are characterized by much higher variation across specifications. Moreover, they are more sensitive to the calculation inside the common support. Notwithstanding, the differences between estimations were more statistical than economical. In other words, a practitioner could extract useful good estimations of the gap, from relatively simpler methods.

A useful extension of our research is to check the generality of our findings. Namely, different methods emphasize particular aspects of the gender wage gap. If some of the labor market deficiencies are missing in the case of Poland, part of the general finding could stem solely from the data properties. Thus, it seems that a study covering a wider selection of countries could corroborate the findings demonstrated in this paper. This study should be welcomed given the constant evolution of the methods to measure the adjust wage gap and their increasing complexity. A research of this characteristics will allow other economist to select a method and conditioning set that provide reliable estimates well suited given the particular research needs and the data constraints.

References

Albrecht, J., van Vuuren, A. and Vroman, S., 2008, Counterfactual Distributions with sample selection adjustments: econometric theoru and an application to the Netherlands.

Appleton, S. Bigsten, A. and Manda, D. 1999, “Educational expansion and economic decline: returns to education in Kenya 1978-1995. WPS/99-6, Centre for the Study of African economies.

Bauer, T. and M. Sinning, 2006. "An Extension of the Blinder-Oaxaca Decomposition to Non-Linear Models," RWI Discussion Papers 0049, Rheinisch-Westfälisches Institut für Wirtschaftsforschung.

Belzil, C. 2007, “Testing the specification of the Mincer wage equation” Iza DP No 2650

Black, D., A., Haviland, S. Sanders and L. Taylor, 2008 “Gender wage disparities among the highly educated”. Journal of Human Resources.

Blau, Francine D., Lawrence M. Kahn, “Understanding international differences in the gender pay gap”, NBER Working Paper No. w8200, 2001.

Blinder, A. S. 1973. “Wage Discrimination: Reduced Form and Structural Estimates.” The Journal of Human Resources 8: 436–455

Cotton, J. 1988. “On the Decomposition of Wage Differentials.” The Review of Economics and Statistics 70: 236–243

Donald, S. D., Green and H. Paarsch, 2000, “Differences in Wage distributions between Canada and the United States: An application of a flexible estimator of distribution functions in the presence of covariates. Review of economic studies 67, 609 -633

DiNardo, J. , N. Fortín, and T. Lemieux, 1996 “Labor market institutions and the distribution of wages, 1973-1992: a Semi-parametric approach” Econometrica, Vol. 64, No.5, 1001 - 1044.

Fairlie, R. 2003, “An extension of the Blinder-Oaxaca decomposition technique to logit and probit models> Economic growth center, University of Yale, Discussion paper No 873

Firpo, S., Fortín, N., and Lemieux, T. 2009 Unconditional Quantile regressions, Econometrica, Vol. 77, No. 3, 953-973

Fortín, N., T. Lemieux and S. Firpo, 2010 “Decomposition methods in Economics” NBER Working paper 16045

Goraus Karolina, 2012, Gender Wage Gap in Poland – Can It Be Explained by Differences in Observable Characteristics, AIEL Conference Material, http://www.aiel.it/bacheca/Capua/papers/Goraus.pdf

Imbens and Rubin, 2009, Causal Inference, Part III, Chapter 14.Kalinowska-Nawrotek Barbara, Dyskryminacja kobiet na polskim rynku pracy. Akademia Ekonomiczna w Poznaniu, 2005.

Kot S.M., Podolec B., Ulman P., 1999, „Problem dyskryminacji płacowej ze względu na płeć”, w: S.M. Kot (red.), Analiza ekonometryczna kształtowania się płac w Polsce w okresie transformacji, Warszawa–Kraków: Wydawnictwo Naukowe PWN.

Kotowska Irena Elżbieta, Podogrodzka Maria, Dyskryminacja kobiet na rynku pracy, „Kobieta i biznes”, nr 2–3, 1994

Łatuszyński K., Woźny Ł.P., 2008, „Zróżnicowanie wynagrodzeń kobiet i mężczyzn na polskim rynku pracy w 2004 roku”, w: W. Pacho, M. Garbicz (red.), Wzrost gospodarczy a bezrobocie i nierówności w podziale dochodu, Warszawa: Szkoła Główna Handlowa.

Machado, C. 2012, „Selection, heterogeneity and the gender wage gap” IZA DP No. 7005

Magda I., Szydłowski A., 2008, „Płace w makro i mikroperspektywie”, w: M. Bukowski (red.), Zatrudnienie w Polsce 2007 – Bezpieczeństwo na elastycznym rynku pracy,Warszawa: Ministerstwo Pracy i Polityki Społecznej.

Matysiak A., Słoczyński T., Baranowska A., 2010, „Kobiety i mężczyźni na rynku pracy”, w: M. Bukowski (red.), Zatrudnienie w Polsce 2008 – Praca w cyklu życia, Warszawa: Centrum Rozwoju Zasobów Ludzkich

Mazur-Łuczak Joanna, „Kobiety na rynku pracy" Wyższa Szkoła Nauk Humanistycznych i Dziennikarstwa, 2010
Melly, B. 2006, „Estimation of counterfactual distributions using quantile regression" Working paper.
Neumark, D. 1988. "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination". The Journal of Human Resources 23: 279–295
Newell Andrew, Reilly Barry, "The gender pay gap in the transition from communism: some empirical evidence", Economic Systems, vol. 25 issue 4, 2001
Oaxaca, R. 1973, "Male-female wage differentials in Urban Labor markets" International economic review, Vol. 14, Issue 3.
Pailhé Ariane, "Gender discrimination in Central Europe during the systemic transition", Economics of Transition, vol 8 issue 2, 2000
Rokicka Magdalena, Ruzik Anna, 2010, "The Gender Pay Gap in Informal Employment in Poland", CASE Network Studies and Analyses 406
Reimers, C. W. 1983. Labor Market Discrimination Against Hispanic and Black Men. The Review of Economics and Statistics 65: 570–579.
Shorrocks, A. 2013, "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value" Journal of Economic Inequality 11: 99-126
Watson, I. (2010) "Decomposing the Gender Pay Gap in the Australian Managerial Labour Market" Australian Journal of Labour Economics, 13(1): 49-79.
Weichselbaumer Doris, Winter-Ebmer Rudolf, 2005. "A Meta-Analysis of the International Gender Wage Gap," Journal of Economic Surveys, vol. 19(3), pages 479-511
Yun, M. 2007, "Wage differentials, discrimination and inequality: a cautionary note on the Juhn, Murphy and Pierce decomposition method." IZA DP No. 2937,
Zwiech, Patrycja, Różnice w wynagrodzeniach kobiet i mężczyzn na rynku pracy w Polsce, Zeszyty Naukowe Uniwersytetu Szczecińskiego, nr 9, 2005

Faculty of Economic Sciences
University of Warsaw
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl