GRAPE Working Paper # 20

Estimating gender wage gap in the presence of efficiency wages -- evidence from European data

Katarzyna Bech, Joanna Tyrowicz

FAME | GRAPE 2017

Foundation of Admirers and Mavens of Economics
Group for Research in Applied Economics

Estimating gender wage gap in the presence of efficiency wages -- evidence from European data

Katazyna Bech
FAME|GRAPE
Warsaw School of Economics

Joanna Tyrowicz
FAME|GRAPE, IZA
University of Warsaw

Abstract
Gender wage gap (adjusted for individual characteristics) as a phenomenon means that women are paid unjustifiably less than men, i.e. below their productivity. Meanwhile, efficiency wages as a phenomenon mean that a group of workers is paid in excess of productivity. However, productivity is typically unobservable, hence it is proxied by some observable characteristics. If efficiency wages are effective only in selected occupations and/or industries, and these happen to be dominated by men, measures of adjusted gender wage gaps will confound (possibly) below productivity compensating of women with above productivity efficiency wage prevalence. We propose to utilize endogenous switching models to estimate adjusted gender wage gaps. We find that without correction for the prevalence of efficiency wages, the estimates of the adjusted gender wage gaps tend to be substantially inflated.

Keywords:
efficiency wages, gender wage gap, endogenous switching regressions

JEL Classification
J16, J33, J71

Corresponding author
Joanna Tyrowicz, j.tyrowicz@grape.org.pl

Acknowledgements
At earlier stage of this work, Olga Pilipczuk provided wonderful research assistance. We are grateful for the valuable comments to an earlier version of this paper by Iga Magda, Filip Premik, Dafni Papoutsaki, David Neumark, and Michal Myck. We received valuable comments during SOLE 2017, EEA 2017 and WSE seminars. This research was supported by a grant from the National Science Center, UMO-2012/05/E/HS4/01510. All opinions expressed are those of the authors and have not been endorsed by NSC. The remaining errors are ours

Published by: FAME | GRAPE
ISSN: 2544-2473

Foundation of Admirers and Mavens of Economics
ull. Mazowiecka 11/14
00-052 Warszaw
Poland

W | grape.org.pl
E | grape@grape.org.pl
TT | GRAPE_ORG
FB | GRAPE.ORG
PH | +48 799 012 202

1 Introduction

Abundant anecdotal evidence suggests that men and women are not subject to the same incentives when already working. Econometric evidence typically emphasizes sorting to field of education, occupations and industries (Bertrand 2011, Dohmen and Falk 2011, Long and Conger 2013, Leibbrandt and List 2014). Thus, sorting may sometimes reflect gender-specific constraints in combining the family and professional roles (Goldin 2014, Pan 2015). Experimental evidence argues that women are not as eager to compete, nor take risk (Buser et al. 2014, Flory et al. 2015)[1]. Finally, wages are only one of the possible instruments to incentivize work commitment, while these different instruments tend to vary in effectiveness between men and women (Clark 2001, Bandiera et al. 2005, 2010). Given these empirical premises, it appears plausible that efficiency wages may be more relevant for men – whereas women may value more some other attributes of work both as incentive to avoid shirking. This would imply that a part of wage differential typically unexplained by observable characteristics could actually reflect wages of men in excess of marginal productivity (efficiency wages) and not discriminatory pricing of women's work.

While this hypothesis is by no means new, it is challenging to address empirically. Typically, efficiency wages are not identified directly (Schmitz 2005, Macpherson et al. 2014). In standard wage datasets, such as linked employer-employee data or labor force survey, prevalence of efficiency wages may be confirmed or rejected, but usually not attributed to respective workers (Murphy and Topel 1990, Blackburn and Neumark 1992). Indeed, individual productivity is rarely observed, thus making it impossible to judge if wage exceeds it. Moreover, clearly, if there is sorting, even identifying productivities is not going to help much due to endogeneity. A class of full information maximum likelihood estimators with endogenous switching provides consistent estimators of returns to individual characteristics, accounting simultaneously for selection and wage determination, but these models require that the data comprises assignment between the markets, e.g. unionized vs non-unionized workers, public vs. private sector, etc. (Lee 1978, Maddala 1983, 1986, Stelcner et al. 1989, Adamchik and Bedi 2000).

We propose an estimator of gender wage gap, which accounts for bias stemming from a separation between a privileged and standard labor markets, when this separation is endogenous and a priori unknown (unobservable). We analyze estimates of the gender wage gaps in European countries using linked employer-employee data for the European countries (EU SES). Thus, we address an important concern implicit in the previous literature that the estimates of adjusted gender wage gap are inflated by the incidence of efficiency wages (e.g. Jirjahn and Stephan 2004, Ichino and Moretti 2009).

The contribution of this paper is twofold. First, we propose a maximum likelihood estimator which jointly optimizes endogenous selection into a privileged or standard labor market and two wage regressions, allowing for a correlation between error terms from selection and wage regressions. Second, we offer new insights into the interplay between efficiency wages and gender inequality in the labor market. When we allow the model to account for endogenous selection between privileged and standard markets, we find that (a) women experience barriers accessing the privileged market (b) estimates of adjusted gender wage gaps differ substantially between the two markets with standard market being characterized by substantially lower gaps than inferred from the pooled regressions; and (c) when accounting for efficiency wages, adjusted gaps are economically different than the estimates on the same data without a correction for the possible incidence of efficiency wages. These phenomena exhibit certain extent of heterogeneity across countries.

[1] Admittedly, it is still hotly debated to what extent are the experimental results a consequence of design, to what extent the consequence of nurture and to what extent consequence of "nature" (see for example Gneezy et al. 2012, Azmat and Petrongolo 2014). In addition, occupational case studies emphasize slower adoption of innovation in less masculinized professions (Schumacher and Morahan-Martin 2001).

The paper is structured as follows. We discuss empirical evidence on efficiency wages and gender wage gaps in the subsequent section, paying particular attention to the methodology employed in the earlier studies. Section 3 describes in detail the model and estimation. In section 4 we describe the data utilized in this study and we move to discussing results in section 5. The methodological and policy implications conclude the study.

2 Empirical evidence on efficiency wages and gender wage gaps

Systematic wage gaps, that cannot be attributed to the factors determining individual productivity are typically tagged as discrimination. However, systematic departures from the unexplained wage differentiation may stem from a selective prevalence of the efficiency wages (Shapiro and Stiglitz 1984, Akerlof 1984, Yellen and Ackerlof 1984). Since efficiency wages imply (selective) pay above productivity, then it follows that a part of wage differential that cannot be attributed to observable worker characteristics need not be a payment *below productivity.* In fact, if efficiency wages are at play, some workers may receive wages exactly equal to their productivity and nonetheless lower than otherwise identical workers, because they do not obtain the anti-shirking premium. If this is the case, estimators for the extent of discrimination could not only be biased but also inconsistent, insofar as the use of efficiency wages has an imperfect overlap with the prevalence of being disadvantaged. Given that the range of both efficiency premium and discrimination penalty are of roughly similar magnitude[2], the potential bias may be severe.

The challenge lies in the fact that productivity is typically not observable. Given this constraint, a large share of the empirical literature on efficiency wages employed an indirect identification strategy (e.g. Weiss 1980, Abowd and Ashenfelter 1981, Krueger and Summers 1988, Konings and Walsh 1994, Goldsmith et al. 2000)[3] or relied on case studies (Cappelli and Chauvin 1991, Campbell 1993). This literature typically does not exploit the disadvantage margin, focusing on identifying the prevalence of efficiency wages *per se.* An alternative strand of literature has grown at the junction of labor economics and management, in the analysis of performance pay and its effects (earlier literature was reviewed by Prendergast 1999, Boeri et al. 2013, give an updated overview). This literature puts more emphasis on potentially disadvantaged groups, e.g. immigrants in Canada (Fang and Heywood 2010), blacks in the US (Heywood and O'Halloran 2005, Heywood and Parent 2012) or women (Maas and Torres-González 2011, Kangasniemi and Kauhanen 2013, Chiang and Ohtake 2014). Yet, the performance-related pay does not necessitate, nor preclude efficiency wages. In fact, it is just stating that wages are proportional to some proxy for output, but the proportionality does not have to imply no additional – anti-shirking – incentives. Hence, although this literature provides valuable insights, it cannot determine the extent to which prevalence of efficiency wages biases the estimates of wage disadvantages and *vice versa.*

Meanwhile, wage gap literature has focused largely and fruitfully on developing reliable estimation methods (see Fortin et al. 2011, Goraus et al. 2017). This amazing progress, however, has usually abstained from the labor market institutions in general and the wage setting mechanisms in particular (compare the meta-analysis Weichselbaumer and Winter-Ebmer 2005). However, Bulow (1986) already proposes efficiency wage theory as an explanation for the gender wage gaps. Offering wages in excess of productivity (to encourage creativity or deter shirking Eaton and White 1983, Shapiro and Stiglitz 1984) in the primary market may lead to adjusted gender wage gaps if women are disproportionately absent in this market or present in the secondary market, even if it pays wages in line with productivity and thus does not discriminate. According to Bulow (1986) if secondary market jobs offer some other intrinsic

[2]Typical estimates of gender or racial wage gaps suggest approximately 10 to 25% of penalty (e.g. Weichselbaumer and Winter-Ebmer 2005, Blau and Kahn 2016), whereas efficiency premia are usually estimated to about 15-30% (Krueger and Summers 1988, Konings and Walsh 1994, e.g.).

[3]A large body of literature employs the efficiency wage concept in micro-founded simulation general equilibrium models, remaining outside the scope of interest in this study.

value, such as employment stability or higher compatibility with engagement in household production, sorting may in fact be consistent with preferences even if it results in adjusted gender wage gaps, which partially may explain their prevalence, see also Goldin (1986).

Despite these strong theoretical foundations, there was little empirical inquiry into the interaction between the efficiency wage theory and the gender wage gaps. Jirjahn and Stephan (2004), Ichino and Moretti (2009) study the link between work effort as proxied by abstenteism and wages. A valuable study by Dohmen and Falk (2011) is a laboratory experiment[4], whereas Antonczyk et al. (2010) analyze centralized wage bargaining rather than efficiency wages per se. A very interesting approach is proposed by Bartolucci (2013) who develops a search and matching model calibrated closely to the case of Germany and introducing differences in productivity, disparities in friction patterns, segregation and residual wage discrimination.

Our paper is partly similar to Bartolucci (2013), as we also use matched employer-employee data. However, our intention is to study the scope of overlap between the gender wage gaps and the efficiency wages. The estimates for both these phenomena are roughly 10-20% in the empirical literature. If indeed the Bulow (1986) hypothesis holds, the (adjusted) gender wage gaps are not actually gaps relative to productivity but rather a sign that the primary market premium is lacking. This would be possible if selection to the primary market was not confined to occupational and industry sorting, but was unobservable. In the reminder of this paper we will describe how we operationalize the unobservable sorting between primary and secondary market and how accounting for this sorting affects the empirical estimates of the adjusted gender wage gap in a large selection of the European countries.

3 Model

We allow for the labor market to be divided into two parts: privileged and standard market. The privileged labor market offers a wage premium above marginal productivity. The assignment between the two markets is not observed in the data. By assumption, no workers are excluded from either of the markets. The complete model is defined by the following set of equations:

$$\begin{aligned} Y_{1,i} &= X_i\beta_1 + u_{1,i} \quad (1)\\ Y_{0,i} &= X_i\beta_0 + u_{0,i} \quad (2)\\ Y^*_{s,i} &= W_i\alpha - v_i, \quad (3) \end{aligned}$$

with $Y_{1,i}$ and $Y_{0,i}$ denoting (log) wages paid in each of two regimes. The vector of regressors X_i contains the standard Mincerian productivity controls such as education, tenure, experience, industry, region, etc. Note that the explanatory variables are the same in both regimes. Both Y_{1i} and $Y_{0,i}$ are only partially observed: $Y_{1,i}$ of individuals in the privileged market and $Y_{0,i}$ for individuals in the basic market. However, the fully observed variable is the (log) wage Y_i defined as

$$Y_i = \begin{cases} Y_{1,i} & \text{iff } Y^*_{s,i} > 0, \\ Y_{0,i} & \text{iff } Y^*_{s,i} \leq 0. \end{cases}$$

The latent variable $Y^*_{s,i}$ assigns the observations to regimes and the vector of observables W_i determines the individuals' likelihood of being in the privileged labor market. β_1, β_0 and α are the vectors of unknown parameters. We assume that the disturbance terms u_{1i}, $u_{0,i}$ and v_i are jointly normally distributed with

[4]In the reminder of this paper we abstract from substantial literature on differences in performance in laboratory experiments in ever expanding literature on risk attitudes, competitiveness, etc. These characteristics and efforts are usually not observed in the standardized data on wages used for estimating the gender wage gaps

mean vector 0 and variance-covariance matrix given by

$$\begin{pmatrix} \sigma_1^2 & 0 & \sigma_{1v} \\ 0 & \sigma_0^2 & \sigma_{0v} \\ \sigma_{1v} & \sigma_{0v} & \sigma_v^2 \end{pmatrix}.$$

Note that the covariance between $u_{1,i}$ and $u_{0,i}$ is by construction equal to 0, as $Y_{1,i}$ and $Y_{0,i}$ are never observed together[5]. An additional necessary assumption is that σ_v^2 =1, as α and σ_v cannot be identified separately. If v_i, $u_{1,i}$ and $u_{2,i}$ were pairwise uncorrelated (an exogenous switching regression) and if the sample separation was known then β_1 and β_2 could be separately estimated by OLS. The known sample separation means that there exists an observed classifying variable I_i defined as

$$I_i = \begin{cases} 1 & \text{iff } Y_i = Y_{1,i}, \\ 0 & \text{iff } Y_i = Y_{0,i}, \end{cases} \tag{4}$$

i.e. $I_i = 1$ if an individual earns in the privileged market, 0 otherwise. If equations (1), (2) and (3) were independent, the most straightforward estimation method is to apply OLS in (1) and (2), and probit regression to model (3) and (4). The estimates from two regimes could be then compared to check whether the gender wage gap differs across markets. If regimes were known, but the estimates of α were unknown, the model describes an endogenous switching with known sample separation (such as analyzed by Lee 1978, Adamchik and Bedi 2000, Lokshin and Sajaia 2004, among others). However, the assumptions that the sample split is known and that the error terms are uncorrelated are excessively demanding towards reality. Firstly, one is typically unlikely to know which workers are employed in the privileged market. As the privileged market pays wages above productivity and individual productivity is not directly measurable, identification in the data is challenging. Secondly, part of the allocation between the privileged market and the rest of the labor market is likely to depend on unobservable individual characteristics, making the error terms correlated. Hence, a method is needed for endogenous switching with an unknown sample split.

3.1 Model estimation

Given the model structure and the multivariate normal distribution of the disturbances, the objective log-likelihood function is given by:

$$\begin{aligned} \ln L = \sum_{i=1}^{n} \Bigg\{ & \left[\ln \phi\left(\frac{u_{0,i}}{\sigma_0}\right) - \ln \sigma_0 + \ln \left\{ 1 - \Phi\left(\frac{W_i\alpha - \rho_0 \frac{u_{0,i}}{\sigma_0}}{\sqrt{1-\rho_0^2}} \right) \right\} \right] \\ & + \left[\ln \phi\left(\frac{u_{1,i}}{\sigma_1}\right) - \ln \sigma_1 + \ln \Phi\left(\frac{W_i\alpha - \rho_1 \frac{u_{1,i}}{\sigma_1}}{\sqrt{1-\rho_1^2}} \right) \right] \Bigg\} \end{aligned} \tag{5}$$

Normal errors $u_{1,i}$ and $u_{0,i}$ in equation (5) are substituted for maximization purposes by $Y_{1,i} - X\beta_1$ and $Y_{0,i} - X\beta_0$ respectively, reformulating equations (1) and (2). The ϕ and Φ denote the probability density and the cumulative distribution functions of the standard normal distribution, ρ_j denotes the correlation coefficients between errors from the wage equation j and the switching equation, with $\rho_j = \frac{\sigma_{jv}}{\sigma_j}$. Note that we deal here with an "incomplete" likelihood function as the division into regimes is not known. Thus, the feasible estimation requires a transformation into a complete-data setting, in which the log-likelihood

[5] Also, it cannot be estimated, because it does not appear in the likelihood function (see Maddala 1983).

subject to maximization is given by:

$$\ln L = \sum_{i=1}^{n} \left\{ (1 - I_i) \left[\ln \phi\left(\frac{u_{0,i}}{\sigma_0}\right) - \ln \sigma_0 + \ln \left\{ 1 - \Phi\left(\frac{W_i\alpha - \rho_0 \frac{u_{0,i}}{\sigma_0}}{\sqrt{1-\rho_0^2}}\right)\right\}\right] \right.$$
$$\left. + I_i \left[\ln \phi\left(\frac{u_{1,i}}{\sigma_1}\right) - \ln \sigma_1 + \ln \Phi\left(\frac{W_i\alpha - \rho_1 \frac{u_{1,i}}{\sigma_1}}{\sqrt{1-\rho_1^2}}\right)\right] \right\} \tag{6}$$

where the actual value of (I_i) remains unknown. We estimate the endogenous switching regression model with unknown sample separation via expectation maximization algorithm (Dempster et al. 1977, Hartley 1978). Similar methods were used by Neumark and Wascher (1994a,b) to estimate minimum wage effects on employment and by Hovakimian and Titman (2006) to link investment expenditures to proceeds from asset sales in financially constrained firms. This approach has several advantages. First, it does not require the data to include any direct identification of the sample split, which makes it particularly suitable for some economic hypotheses, such as efficiency wages. Second, it produces intuition about the underlying separation between the samples. In each iteration, $(1 - I_i)$ and (I_i) are replaced by the estimated probabilities that a given observation belongs to either of the samples. Thus, the terms $(1-I_i)$ and (I_i) and their estimated determinants have an economic interpretation. Third, the EM algorithm might overcome the problem of the unboundedness of the log-likelihood function in this framework, as shown by Maddala and Nelson (1975). This algorithm is computationally intensive.

An alternative approach to deal with the unknown sample split is to apply the grid search method (Quandt 1958). Both these algorithms utilize (6) and rely on a log-likelihood function, but employ alternative algorithms to find an optimum: expectation maximization considers parameters for the whole expression in (6), whereas grid search identifies which among possible sample splits produces the highest log-likelihood. Under mild regularity conditions both techniques produce the same asymptotic results. However, the grid search approach is even more computationally demanding, as it requires performing calculations for numerous possible sample splits. Hence, this paper utilizes the EM algorithm.

Both methods require choosing the initial sample split and the initial values for the parameters of the log-likelihood function. The initial sample split is based on the residuals from the standard OLS regression of (log) wages on the set of explanatory variables for the whole sample. Individuals with positive residuals are initially assigned to the privileged market (for the starting iteration). Once the initial split is obtained, the starting values for the maximization procedure are calculated using OLS methods for each of two market equations.

This procedure naturally entails a challenge with respect to obtaining the standard errors. Since we cannot estimate the incomplete maximum likelihood, we estimate a complete maximum likelihood, iteratively updating the probabilities assigning a given person to privileged market, under the assumption that what maximizes the complete likelihood, would also maximize the incomplete likelihood function. Given this setup, it is suitable and internally coherent to utilize computational Hessian matrix of the complete likelihood to obtain the standard errors, which is the approach we follow.

3.2 Gender wage gap decomposition

The parametric Oaxaca-Blinder decomposition divides the raw gender wage gap

$$\ln \bar{W}_M - \ln \bar{W}_F = \underbrace{\beta^*(\bar{X_M} - \bar{X_F})}_{\text{characteristics}} + \underbrace{\overbrace{\bar{X_M}(\beta_M - \beta^*)}^{\text{male advantage}} + \overbrace{\bar{X_F}(\beta^* - \beta_F)}^{\text{female disadvantage}}}_{\text{coefficients}} \tag{7}$$

into a part determined by differences in characteristics and a part attributed to differences in coefficients. There are many ways to construct the counter-factual distribution of wages to obtain the wages that would have prevailed, had the coefficients been the same for men and women, i.e.

$$\beta^* = \lambda * \beta^M + (1 - \lambda) * \beta^F. \tag{8}$$

The original Oaxaca-Blinder decomposition assumed male or female coefficients as reference, subsequently other approaches were postulated, see Table A.1 in the Appendix, reflecting somewhat philosophical differences in conceptualizing the gender wage gap.

Given the structure of equation (6), our estimation determines contemporaneously the parameters of the selection equation and the parameters of the wage equation. Hence, every estimation strategy is likely to affect our results. For example, if one estimates equation (6) separately for men and women, one automatically assumes that the sorting for the privileged and standard market differs across genders, hence possibly inflating the effects of characteristics on total wage differential. To address this issue, we proceed by interacting each control variable in the selection equation and in the wage equation with the male dummy. However, this implies that obtaining β^* from pooled regressions as suggested by Neumark (1988), Fortin (2008) is inefficient, because to this end equation (6) would need to be re-estimated, possibly with alternative allocation across privileged and standard market at least for some observations. As a shortcut to avoid that internal inconsistency in the model we follow Słoczyński (2015). Hence, the obtained estimates of the control variables are equivalent to the female coefficients, whereas the estimates of the interaction terms denote the difference between the male advantage and the female disadvantage. The advantage of this approach is that it allows for a straight forward interpretation of whether – and which – coefficients differ across genders for the selection equation in a statistically significant way. Note that this inference is separate from decomposing the wage differential.

Given the rich setup of the model specified in equation (6), we may effectively decompose the raw difference in wages between men and women into six components. First, there are the two components from the selection equation: explained (attributable to differences in characteristics) and unexplained (stemming from differences in coefficients). The remaining four come from estimates from the two markets: privileged and standard yield differences in characteristics and differences in coefficients. The advantage of our approach is that each of the six components may be estimated, which is substantially richer than the standard approach which does not allow for the prevalence of efficiency wages.

We can also test explicitly if two separate regimes indeed exist (by comparing the estimates of two wage equations via Wald test) and if gender matters in the selection equation (the joint significance test on the interaction terms in the selection equation).

4 Data

This paper utilizes standardized data from the Structure of Earnings Survey of the European Union (EU SES). This choice was motivated by a number of advantages of this data over the alternative data sources. First, the wage data is detailed. Unlike labor market surveys, EU SES data is reported by the employers and thus reflects exact paid out compensations as well as exact number of hours worked (responders in surveys tend to report rounded numbers). It is relevant for us because rounding the figures could mask the split between the privileged and standard markets. Second, EU SES sample sizes are 10 to 20 times larger than labor force survey, for example. Third, EU SES data is detailed in individual characteristics as well as firm level characteristics, which permits us to compare estimates with fairly general coding to the detailed ones.

This data has also week points. First, except for wages and hours, most variables are categorical.

Age is coded by age groups and education is coded by achievement rather than years of education. Both combined imply also that measures of experience are impossible. Second, firms employing under 9 workers are usually not included.[6] Third, the sample design differs across countries. In some cases, the data is a full census (all employees from all firms), in some cases data is a quasi-full census (random sample of employees from all firms) and in some countries it is a hybrid. For example, this is a full census for firms employing between 9 and 49 workers and a random sample of employees from larger firms. These differences in sample design are not likely to be consequential for our study, because each segment of the labor market is fully represented, hence if efficiency wages prevail, the estimator would be able to identify them. However, the size of the privileged and standard market cannot be adequately measured, because sample design may sometimes leave aside a large number of workers from each of the markets.[7] We discuss the details of sample selection in each analyzed country in Appendix B.

We use data from 2006 waves for the available countries. The choice of the year was motivated by the availability: data for 2014 wave have not been released yet, whereas the data for 2010 come from a crisis year, which could introduce additional context to both efficiency wage prevalence and gender wage gap. In 2010 wave, data for Germany and Italy do not provide industry classification of employers, and thus could not be used in our study. Table C.3 reports the sample properties. The presented sample sizes are large. Typically, roughly 20%-30% of salaried workers in the enterprise sector have tertiary degree, but countries differ in whether this is equal across genders or not. Men are substantially more frequently employed in blue-collar occupations, whereas women in white-collar occupations. The largest proportion of workers in all countries is employed in the service sector for both genders. In most of the analyzed economies roughly a third of the salaried workers is in prime age, i.e. between 25 and 45 years of age.

5 Results

The endogenous switching model used in the case of the unknown sample split produces three sets of estimates: coefficients from the switching regression and for the two markets, the coefficients of the wage regression for both markets. However, the model cannot produce the actual sample split, i.e. assign observations to markets. This stems from the fact that the switching regression produces a probability of the split rather than actual split. Figure D.1 plots the cumulative distribution functions for the available samples. Clearly, the distributions differ across countries, but even within country – they differ across years. The difference concerns both the range and the slope. For example, the distribution for France has the lowest value well above 0.5, whereas for Latvia almost no observations exceed that value.

To obtain the estimates of gender gaps a sample split has to be imposed on the samples (otherwise no statistics within the two markets can be obtained and they are needed for virtually every decomposition method). However, given this heterogeneity of outcomes, there are no clear heuristics to follow. The empirical literature on the prevalence of efficiency wages suggests, that roughly 10%-20% of workers enjoy a premium to their wage (Prendergast 1999, Boeri et al. 2013). Following these findings we set the threshold at 85th percentile of the distribution in the estimated probability of sample split.

Clearly, the choice of a percentile is arbitrary and has little justification in economic theory. Moreover, by applying such split one cannot provide estimates on how big/small the privileged market is in a given economy. Finally, it may well be that for a given economy the split between the two markets is substantially different from the chosen one. To address these issues we do the following. First, we

[6] Neither are the self-employed, but this is irrelevant for our study

[7] Naturally, EU SES provides weights which allow generalizing the estimates to the total population of firms employing salaried workers in the enterprise sector. At this stage our estimator does not permit utilizing weights. However, in the case of many countries this is not likely to affect our results. For example, Czech Republic includes data for a total population of salaried workers from plants employing above 10 workers (in full-time equivalents). In many other cases, the sample covers the total population for firms employing between 10 and 49 workers (FTE), sampling workers only for larger plants.

explicitly test if indeed two separate markets prevail in a given sample. This is done by the means of a Wald test, which utilizes only coefficients and does not need to know the actual sample split. In each analyzed sample the test shows that indeed the wage equations differ in a statistically significant manner, results are reported in Table D.5 in the Appendix. Second, we apply a sensitivity check, setting the threshold value to 75th and alternatively 95th percentile. Hence, we test if the arbitrary choice of the threshold percentile affects quantitatively and qualitatively the estimated adjusted gender wage gaps. Third, we check whether the parameters on interactions in the switching equation are jointly and simultaneously equal to zero. The results are summarized in Table D.4 in the Appendix. In all cases the null hypothesis is rejected indicating the significant impact of gender on the market assignment.

Finally, we also follow a data driven approach. The endogenous unknown switching part of the regression yields an estimate of the average of the predicted value ($\bar{\hat{Y}}^*$). We use this estimate for the final way to split the data to compute the raw and adjusted gender wage gaps. As the mean of predicted probabilities gives the proportion of ones in the sample, we follow the Cramer rule to assign observations to markets, i.e. if the predicted probability exceeds this threshold it is classified as the privilidged market. We report the percentile of the distribution of calculated probabilities at which this threshold occurs in Table 1.

5.1 Gender wage gap in segmented market – a comparison to a pooled sample

To provide estimates of the gender wage gap we utilize the most commonly used method: Oaxaca-Blinder decomposition. In principle, any parametric or nonparametric decomposition method may be applied. We compare two types of estimates: from a simple pooled model for all workers and from our endogenous unknown switching model. The specifications allow both the selection equation and the wage equation to deliver gender specific coefficients. Hence, we may compare estimates of GWG with and without control for efficiency wages and show the relative contribution of the possibly gendered selection to the privileged market. Note that the estimates for the privileged and standard market are presented for the total wage, i.e. we do not provide the estimates of the efficiency wage premium.[8] The results are summarized in Table 1.

All estimates are obtained with Słoczyński (2015) decomposition, based on the following premises. The counter-factual distribution of wages is obtained from reweighing the coefficients of men and women by adequate population shares. As demonstrated by Słoczyński (2015), the implicit weights are opposite of what they should be. So long as participation of men and women in the total population are roughly equal, this makes little difference for the counter-factual distributions of wages. However, participation in the privileged market need not be balanced with respect to gender, which intensifies the risks associated with taking inappropriate weights.

Few immediate observations can be made. First, the estimates of the adjusted gap with endogenous split are relatively stable irrespective of the threshold for a sample split. Although raw gap measures adjust to sample change due to alternative thresholds, the adjusted gaps are of the same magnitude for both the privileged and standard market alike. Second, the adjusted gender wage gap in all the countries is substantially higher in the privileged market. The estimates for this market are also often higher than ones which come from the pooled OLS. Importantly, the discrepancy between the raw and the adjusted gap is much higher for the estimates from the privileged market than in the pooled OLS hinting strong selectivity patterns. The standard market by contrast, offers estimates of the adjusted gap below the adjusted gap of the OLS. Third, there is gender specificity in the access to the privileged market (we have no evidence to argue whether that specificity stems from a barrier or conforms with differentiated preferences of men and women). The specificity is relatively heterogeneous across countries and in many

[8] That would be possible with our approach, albeit only with the bootstrapped standard errors.

Table 1: Adjusted and raw gender wage gaps

Country	Split	OLS Raw (1)	OLS Adjusted (2)	Privileged market Raw (3)	Privileged market Adjusted (4)	Standard market Raw (5)	Standard market Adjusted (6)	Switching regression Raw (7)	Switching regression Adjusted (8)
Czech Rep.	75th			6.8%	26.3%	25.1%	7.5%		
	85th	24.4%	29.0%	11.9%	27.2%	25.4%	7.5%	28.9%	25.1%
	95th			12.6%	25.5%	24.8%	7.9%		
	Cramer (43th)			19.1%	26.5%	31.3%	6.7%		
Finland	75th			23.2%	25.9%	14.1%	12.5%		
	85th	22.8%	26.3%	22.5%	26.7%	20.6%	12.9%	17.6%	5.3%
	95th			35.9%	25.9%	20.2%	13.2%		
	Cramer (72th)			23.2%	25.9%	14.1%	12.5%		
France	75th			5.6%	25.7%	23.0%	15.9%		
	85th	21.3%	23.2%	18.5%	26.3%	20.1%	15.3%	4.6%	1.8%
	95th			-34.7%	24.4%	24.6%	15.4%		
	Cramer (52th)			4.8%	27.0%	36.9%	14.7%		
Greece	75th			18.3%	24.8%	8.1%	6.9%		
	85th	16.9%	18.6%	-0.5%	27.1%	10.4%	7.2%	6.4%	10.1%
	95th			.	.	9.9%	7.8%		
	Cramer (52th)			9.4%	24.5%	7.9%	6.4%		
Hungary	75th			38.8%	25.0%	-29.2%	3.7%		
	85th	10.3%	20.1%	54.0%	24.1%	-10.2%	4.7%	17.9%	20.0%
	95th			.	.	-8.2%	4.6%		
	Cramer (50th)			32.1%	25.6%	-37.6%	2.1%		
Lithuania	75th			13.3%	22.7%	25.9%	4.1%		
	85th	12.2%	22.0%	6.2%	23.2%	27.9%	3.6%	-8.6%	-0.9%
	95th			24.0%	22.9%	10.5%	3.2%		
	Cramer (59th)			6.1%	25.0%	14.9%	4.3%		
Luxembourg	75th			38.5%	13.2%	11.3%	20.6%		
	85th	9.8%	16.7%	31.1%	14.3%	4.6%	20.4%	-4.2%	12.3%
	95th			.	.	3.6%	20.5%		
	Cramer (42th)			16.5%	12.7%	18.0%	24.7%		
Norway	75th			-4.5%	17.9%	20.6%	7.6%		
	85th	15.5%	18.1%	1.1%	18.2%	16.3%	7.5%	7.1%	3.3%
	95th			1.5%	14.4%	15.6%	7.2%		
	Cramer (63th)			14.8%	17.6%	17.9%	7.9%		
Poland	75th			-46.7%	27.4%	21.6%	8.1%		
	85th	5.0%	23.6%	-51.8%	28.3%	13.8%	8.3%	3.9%	6.9%
	95th			.	.	1.8%	7.7%		
	Cramer (58th)			-23.7%	26.8%	11.6%	7.8%		
Portugal	75th			26.1%	19.4%	22.1%	17.3%		
	85th	2.3%	27.4%	.	.	31.6%	17.4%	-9.6%	6.2%
	95th			.	.	2.2%	17.8%		
	Cramer (62th)			36.6%	22.9%	24.3%	17.3%		
Romania	75th			7.4%	0.5%	7.2%	0.2%		
	85th	2.0%	9.1%	10.5%	0.2%	11.9%	0.4%	-5.0%	-0.2%
	95th			-4.0%	0.3%	11.3%	0.5%		
	Cramer(69th)			-13.8%	3.0%	4.7%	0.3%		
Slovakia	75th			6.0%	30.0%	28.9%	5.3%		
	85th	29.2%	31.0%	.	.	24.9%	5.6%	32.2%	31.2%
	95th			.	.	25.0%	6.4%		
	Cramer (55th)			35.1%	28.7%	6.8%	6.8%		
Spain	75th			15.3%	27.7%	18.8%	19.4%		
	85th	20.9%	28.4%	19.9%	27.7%	19.5%	19.5%	3.7%	5.8%
	95th			30.7%	26.5%	18.5%	19.4%		
	Cramer (49th)			11.6%	29.0%	18.1%	19.1%		
Sweden	75th			7.3%	18.8%	16.6%	5.9%		
	85th	13.7%	20.4%	17.8%	20.8%	11.9%	6.1%	13.1%	1.7%
	95th			-0.5%	21.3%	14.3%	6.1%		
	Cramer (76th)			9.7%	18.7%	14.6%	5.8%		

Notes Full estimates for every country reported in the Appendix. OLS reports a decomposition for a pooled market. Columns (3) and (4) report decomposition in the privileged market, columns (5) and (6) report estimates for the standard market and columns (7) and (8) report estimates for the selection equation. All estimates with the use of Słoczyński (2015) decomposition. Raw and adjusted gaps not estimated when there was not enough women in the privileged market.

cases the adjusted gap in the switching regression is low. The exceptions in absolute terms are Czech Republic, Hungary and Slovakia. In Poland, Spain and Greece, despite relatively low raw gaps, after adjustment for individual and job related characteristics the gap increases substantially, hinting that the unexplained component in the labor market segmentation is larger than the explained one.

Comparing to the OLS estimates, the adjusted gender wage gap in the standard market – i.e. for the majority of the workers – is substantially lower. Moreover, in many cases the standard market is characterized by lower adjusted gaps than raw gaps, which hints that gender inequality in wages can to a large extent be explained away by differences in the individual characteristics and to a smaller extent stems from unexplained component, typically associated with discrimination (Altonji and Blank 1999). Hence, an important policy implication of our study: average estimates of the gender wage gap are inflated relative to those which allow for labor market segmentation, which hints that segmentation itself is reinforcing gender inequality, masking the nature of the disparities between men and women. In fact, only in two of the analyzed countries, adjusted gaps are larger than raw gaps in the standard market: Luxembourg and Spain. In three countries, by contrast, there are virtually no women in the privileged market (Portugal, Romania and Slovakia).

Our model specification is on purpose relatively parsimonious. We only separate white collar from blue collar workers in therms of occupations and in terms of industries, we only separate manufacturing (with construction) and market services (agriculture is the base level). By such modeling choice, the estimates of the switching model are not a proxy for occupational sorting, as often analyzed in the literature (Bayard et al. 2003, Shatnawi et al. 2014, Card et al. 2016).

The adjusted gender gap in sorting between the two markets may be both a consequence of choice or lack thereof. For example, there may be inherent gender differences in the propensity to shirk, which affects the incentives for the employers to implement efficiency wages in (fe)male dominated workplaces (e.g. Mastekaasa and Melsom 2014, Johansson et al. 2014). There may also be gender differences in the effectiveness of the efficiency wages as opposed to other incentives at work (Bandiera et al. 2005, 2010). Particularly non-wage benefits appear to be relevant in job valuation for women relative to men (Clark 2001, Kalleberg and Marsden 2013). Finally, men and women may internalize differently the risk of loosing job in case of shirking (Croson and Gneezy 2009, Jung 2014). This short look at the literature suggests that strong gender imbalance in the switching regression need not signify barriers in access to privileged market for women. Notwithstanding, wages in the privileged market are higher than those in the standard market. Hence, it is the gender bias in access to the privileged market that deepens the gender wage gap – to a much lesser extent than unexplained inequality in the standard market.

Our model provides the estimates of the wage regression – and thus any decomposition of interest – accounting for an endogenous and unobserved split between two segments of a labor market. Our motivation stems from the efficiency wages hypothesis, with premises formulated by earlier empirical contributions. Similarly, Hovakimian and Titman (2006) attribute their endogenous unobserved split of firms to being financially constrained. Admittedly, our interpretation need not be the only one. In principle, segmentation could follow other unobserved and endogenous separations, provided that they are systematic in individual characteristics and gender specific. Such examples could include a health premium (e.g. Devaro and Heywood 2016), a beauty premium (e.g. Doorley and Sierminska 2015, Oreffice and Quintana-Domeque 2016), an aspirations premium (e.g. Busch-Heizmann 2014) or other unobservable characteristics in our sample (e.g. more technologically advanced firms).

A possible corroboration for the efficiency wage argument stems from the nature of the identified gaps. If beauty was the source of the premium, for example, there should be little explanatory power in the switching equation, because beauty is random across educational attainments, sectors or blue/which collar type of job. Similar argument should hold for aspirations. If these were the firm level characteristics, then one should see little effects of individual characteristics in the switching equation. Against these theoretical implications, we find that not only is switching regression significant in virtually all analyzed cases, but also that the coefficients for men and women differ substantially in this equation. Indeed, efficiency wage premium and health premium may partly overlap.

5.2 Accounting for household characteristics

The estimations above rely on matched employee-employer data. Hence, they cannot account for household level characteristics, nor selectivity of employment. We test the validity of our results using alternative data and model specification for the few selected countries, for which we obtain quality micro-level data. Of the countries included in Table 1 we repeat the estimation using data from Labor Force Survey from the same year for Poland and France. For these countries we include children in the household as an additional control to test to what extent possible lack of hourly flexibility as well as other factors influencing labor supply can explain away the conclusions from Table 1. The results are reported in Table 2, with the full specification of estimated equations reported in Table E.23 for France and Table E.24 for Poland, in the Appendices.

Table 2: Adjusted and raw gender wage gaps: LFS

		OLS		Privileged market		Standard market		Switching regression	
Country	Split	Raw	Adjusted	Raw	Adjusted	Raw	Adjusted	Raw	Adjusted
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Poland – LFS	75th			-17.8%	22.7%	13.0%	4.6%		
	85th	7.5%	19.4%	-17.8%	22.7%	13.0%	4.6%	16.0%	16.5%
	95th			.	.	11.6%	5.4%		
	Cramer (62th)			-33.7%	19.8%	28.8%	4.3%		
Poland – SES	75th			-46.7%	27.4%	21.6%	8.1%		
	85th	5.0%	23.6%	-51.8%	28.3%	13.8%	8.3%	3.9%	6.9%
	95th			.	.	1.8%	7.7%		
	Cramer (58th)			-23.7%	26.8%	11.6%	7.8%		
France – LFS	75th			27.8%	16.7%	6.1%	15.3%		
	85th	10.5%	16.7%	16.0%	17.4%	12.1%	15.2%	2.3%	-2.0%
	95th			14.4%	16.4%	12.3%	15.4%		
	Cramer (54th)			19.8%	18.3%	-1.7%	12.9%		
France – SES	75th			5.6%	25.7%	23.0%	15.9%		
	85th	21.3%	23.2%	18.5%	26.3%	20.1%	15.3%	4.6%	1.8%
	95th			-34.7%	24.4%	24.6%	15.4%		
	Cramer (52th)			4.8%	27.0%	36.9%	14.7%		

Although point estimators are somewhat different, the main conclusions prevail: adjusted gender wage gaps are lower in the standard market, the estimated Cramer percentiles for the split are similar as well. However, the contribution of the switching regression is somewhat different between the two types of data. Notably, LFS data comprises also micro-firms and public sector employment, which are missing in the SES data.

6 Discussion and conclusions

Literature on gender wage gaps so far paid a lip service to the fact that efficiency wages may potentially be gender-specific. In the presence of efficiency wages, some workers are over-compensated. Meanwhile, gender wage gap typically denotes insufficient compensation, i.e. below fair wage (or marginal productivity). If the over-compensated workers happen to be more frequently male – not because they are men, but because efficiency wages are more prevalent in the industries and occupations where men are more frequently employed – the estimates of the gender wage gap may be biased as "causal" identification of unequal pay. Indeed, both productivity and prevalence of efficiency wages are not directly observable, so separating their effects on wages remains a challenge.

In this paper we estimate gender wage gaps and prevalence of efficiency wages jointly, utilizing endogenous switching models. These models may be used to identify unknown sample splits. Neither do they require any instrument splitting the sample between two regimes. Hence, this class of models is particularly well suited for the task.

We find that indeed access to privileged market is gendered in a sense that there is unexplained gender inequality. Quantitatively more relevant, however, is the extent of unexplained gender wage inequality in the privileged market – substantially above the estimates from the pooled sample, i.e. without allowing for the prevalence of the efficiency wages. We find that this result is relatively robust to how the markets are effectively split. There appears also to be a substantial amount of country heterogeneity in how privileged and standard markets reward women's work.

One of the possible directions for future research in the field is to improve the power of these estimates. In preliminary Monte Carlo experiments, Lokshin and Sajaia (2004) and Chiburis and Lokshin (2007) show that even with known sample split, full information maximum likelihood estimation of the endogenous switching regressions could perform better. Although FIML outperforms the two-step estimator in efficiency and does not appear to be biased, its usefulness is limited in the absence of strong exclusion criterion, as is typically the case for efficiency wages or other unobservable segmentations. Another possible direction of future research concerns the definition of segmentation itself. Our segmentation is dual, in line with many labor economics theories, but the actual labor markets may be characterized by more than two segments. If that it is the case, estimates proposed in this study, relative to these multiple segments, have similar bias as the OLS relative to our estimates. Knoef and Been (2015) propose a way for a multi-level data segmentation. However, their solution is applicable to the known sample splits. Providing more flexibility in terms of the number of segments could enrich our understanding of the unobservable labor market barriers.

References

Abowd, J. M. and Ashenfelter, O. C.: 1981, Anticipated unemployment, temporary layoffs, and compensating wage differentials, *in* S. Rosen (ed.), *Studies in labor markets*, University of Chicago Press, pp. 141–170.

Adamchik, V. A. and Bedi, A. S.: 2000, Wage differentials between the public and the private sectors: Evidence from an economy in transition, *Labour economics* **7**(2), 203–224.

Akerlof, G. A.: 1984, Gift exchange and efficiency-wage theory: Four views, *American Economic Review* **74**(2), 79–83.

Altonji, J. G. and Blank, R. M.: 1999, Race and gender in the labor market, *in* D. Card and O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 3, Elsevier, pp. 3143–3259.

Antonczyk, D., Fitzenberger, B. and Sommerfeld, K.: 2010, Rising wage inequality, the decline of collective bargaining, and the gender wage gap, *Labour Economics* **17**(5), 835–847.

Azmat, G. and Petrongolo, B.: 2014, Gender and the labor market: What have we learned from field and lab experiments?, *Labour Economics* **30**, 32–40.

Bandiera, O., Barankay, I. and Rasul, I.: 2005, Social preferences and the response to incentives: Evidence from personnel data, *Quarterly Journal of Economics* **120**(3), 917–962.

Bandiera, O., Barankay, I. and Rasul, I.: 2010, Social incentives in the workplace, *The Review of Economic Studies* **77**(2), 417–458.

Bartolucci, C.: 2013, Gender wage gaps reconsidered a structural approach using matched employer-employee data, *Journal of Human Resources* **48**(4), 998–1034.

Bayard, K., Hellerstein, J., Neumark, D. and Troske, K.: 2003, New evidence on sex segregation and sex differences in wages from matched employee-employer data, *Journal of Labor Economics* **21**(4), 887–922.

Bertrand, M.: 2011, New perspectives on gender, *in* D. Card and O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 4b, chapter 17, pp. 1543–1590.

Blackburn, M. L. and Neumark, D.: 1992, Unobserved ability, efficiency wages, and interindustry wage differentials, *Quarterly Journal of Economics* **107**(4), 1421–1435.

Blau, F. D. and Kahn, L. M.: 2016, The gender-wage gap: Extent, trends, and explanations, *Journal of Economic Literature* **forthcoming**.

Boeri, T., Lucifora, C. and Murphy, K. J.: 2013, *Executive remuneration and employee performance-related pay: a transatlantic perspective*, OUP Oxford.

Bulow, J.: 1986, An economic theory of planned obsolescence, *The Quarterly Journal of Economics* pp. 729–750.

Busch-Heizmann, A.: 2014, Supply-side explanations for occupational gender segregation: Adolescents' work values and gender-(a) typical occupational aspirations, *European Sociological Review* p. jcu081.

Buser, T., Niederle, M. and Oosterbeek, H.: 2014, Gender, competitiveness and career choices, *The Quarterly Journal of Economics* **129**(3), 1409–1447.

Campbell, C. M.: 1993, Do firms pay efficiency wages? evidence with data at the firm level, *Journal of labor Economics* **11**(3), 442–470.

Cappelli, P. and Chauvin, K.: 1991, An interplant test of the efficiency wage hypothesis, *The Quarterly Journal of Economics* pp. 769–787.

Card, D., Cardoso, A. R. and Kline, P.: 2016, Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women, *Quarterly Journal of Economics* **131**(2), 633–686.

Chiang, H.-Y. and Ohtake, F.: 2014, Performance-pay and the gender wage gap in japan, *Journal of the Japanese and International Economies* **34**, 71–88.

Chiburis, R. and Lokshin, M.: 2007, Maximum likelihood and two-step estimation of an ordered-probit selection model, *Stata Journal* **7**(2), 167.

Clark, A. E.: 2001, What really matters in a job? hedonic measurement using quit data, *Labour economics* **8**(2), 223–242.

Cotton, J.: 1988, On the decomposition of wage differentials, *The Review of Economics and Statistics* pp. 236–243.

Croson, R. and Gneezy, U.: 2009, Gender differences in preferences, *Journal of Economic literature* **47**(2), 448–474.

Dempster, A. P., Laird, N. M. and Rubin, D. B.: 1977, Maximum likelihood from incomplete data via the em algorithm, *Journal of the Royal Statistical Society. Series B (methodological)* pp. 1–38.

Devaro, J. and Heywood, J. S.: 2016, Performance pay and work-related health problems: A longitudinal study of establishments, *ILR Review* p. 0019793916669148.

Dohmen, T. and Falk, A.: 2011, Performance pay and multidimensional sorting: Productivity, preferences, and gender, *The American Economic Review* **101**(2), 556–590.

Doorley, K. and Sierminska, E.: 2015, Myth or fact? the beauty premium across the wage distribution in germany, *Economics Letters* **129**, 29–34.

Eaton, C. and White, W. D.: 1983, The economy of high wages: an agency problem, *Economica* pp. 175–181.

Fang, T. and Heywood, J. S.: 2010, Immigration, ethnic wage differentials and output pay in canada, *British Journal of Industrial Relations* **48**(1), 109–130.

Flory, J. A., Leibbrandt, A. and List, J. A.: 2015, Do competitive workplaces deter female workers? a large-scale natural field experiment on job-entry decisions, *The Review of Economic Studies* **82**(1), 122–155.

Fortin, N., Lemieux, T. and Firpo, S.: 2011, Decomposition methods in economics, *in* O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 4A, Elsevier, pp. 1–102.

Fortin, N. M.: 2008, The gender wage gap among young adults in the united states the importance of money versus people, *Journal of Human Resources* **43**(4), 884–918.

Gneezy, U., List, J. and Price, M. K.: 2012, Toward an understanding of why people discriminate: Evidence from a series of natural field experiments, *NBER Working Paper 17855*, National Bureau of Economic Research.

Goldin, C.: 1986, Monitoring costs and occupational segregation by sex: A historical analysis, *Journal of Labor Economics* **4**(1), 1–27.

Goldin, C.: 2014, A grand gender convergence: Its last chapter, *The American Economic Review* **104**(4), 1091–1119.

Goldsmith, A. H., Veum, J. R. and Darity Jr, W.: 2000, Working hard for the money? efficiency wages and worker effort, *Journal of Economic Psychology* **21**(4), 351–385.

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* **forthcoming**.

Hartley, M. J.: 1978, Comment, *Journal of the American Statistical Association* **73**(364), 738–741.

Heywood, J. S. and O'Halloran, P. L.: 2005, Racial earnings differentials and performance pay, *Journal of Human Resources* **40**(2), 435–452.

Heywood, J. S. and Parent, D.: 2012, Performance pay and the white-black wage gap, *Journal of Labor Economics* **30**(2), 249–290.

Hovakimian, G. and Titman, S.: 2006, Corporate investment with financial constraints: Sensitivity of investment to funds from voluntary asset sales., *Journal of Money, Credit & Banking* **38**(2).

Ichino, A. and Moretti, E.: 2009, Biological gender differences, absenteeism, and the earnings gap, *American Economic Journal: Applied Economics* **1**(1), 183–218.

Jirjahn, U. and Stephan, G.: 2004, Gender, piece rates and wages: evidence from matched employer–employee data, *Cambridge Journal of Economics* **28**(5), 683–704.

Johansson, P., Karimi, A. and Nilsson, J. P.: 2014, Gender differences in shirking: Monitoring or social preferences? evidence from a field experiment, *IZA Discussion Paper 8133*, IZA.

Jung, S.: 2014, Risk attitudes and shirking on the quality of work under monitoring: Evidence from a real-effort task experiment, *Technical report*, THEMA (THéorie Economique, Modélisation et Applications), Université de Cergy-Pontoise.

Kalleberg, A. L. and Marsden, P. V.: 2013, Changing work values in the united states, 1973–2006, *Social Science Research* **42**(2), 255–270.

Kangasniemi, M. and Kauhanen, A.: 2013, Performance-related pay and gender wage differences, *Applied Economics* **45**(36), 5131–5143.

Knoef, M. and Been, J.: 2015, Estimating a panel data sample selection model with part-time employment: Selection issues in wages over the life-cycle. Mimeo, University of Leiden.

Konings, J. and Walsh, P. P.: 1994, Evidence of efficiency wage payments in uk firm level panel data, *The Economic Journal* **104**(424), 542–555.

Krueger, A. B. and Summers, L. H.: 1988, Efficiency wages and the inter-industry wage structure, *Econometrica: Journal of the Econometric Society* pp. 259–293.

Lee, L.-F.: 1978, Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables, *International Economic Review* **19**(2), 415–433.

Leibbrandt, A. and List, J. A.: 2014, Do women avoid salary negotiations? evidence from a large-scale natural field experiment, *Management Science* **61**(9), 2016–2024.

Lokshin, M. and Sajaia, Z.: 2004, Maximum likelihood estimation of endogenous switching regression models, *Stata Journal* **4**, 282–289.

Long, M. C. and Conger, D.: 2013, Gender sorting across k–12 schools in the united states, *American Journal of Education* **119**(3), 349–372.

Maas, V. S. and Torres-González, R.: 2011, Subjective performance evaluation and gender discrimination, *Journal of Business Ethics* **101**(4), 667–681.

Macpherson, D. A., Prasad, K. and Salmon, T. C.: 2014, Deferred compensation vs. efficiency wages: An experimental test of effort provision and self-selection, *Journal of Economic Behavior & Organization* **102**, 90–107.

Maddala, G.: 1986, Chapter 28 disequilibrium, self-selection, and switching models, *in* Z. Griliches and M. Intriligator (eds), *Handbook of Econometrics*, Vol. 3, pp. 1633 –1688.

Maddala, G. and Nelson, F.: 1975, Switching regression models with exogenous and endogenous regimes, *Proceedindgs of the American Statistical Association* pp. 423–425.

Maddala, G. S.: 1983, *Limited-dependent and qualitative variables in econometrics*, Cambridge University Press.

Mastekaasa, A. and Melsom, A. M.: 2014, Occupational segregation and gender differences in sickness absence: Evidence from 17 european countries, *European Sociological Review* **30**(5), 582–594.

Murphy, K. M. and Topel, R. H.: 1990, Efficiency wages reconsidered: Theory and evidence, *Advances in the Theory and Measurement of Unemployment*, Springer, pp. 204–240.

Neumark, D.: 1988, Employers' discriminatory behavior and the estimation of wage discrimination, *Journal of Human Resources* **23**(3), 279–295.

Neumark, D. and Wascher, W.: 1994a, Employment effects of minimum wages and subminimum wages: Reply to card, katz, and krueger, *Industrial and Labor Relations Review* **47**(3), 497–512.

Neumark, D. and Wascher, W.: 1994b, Minimum wage effects and low-wage labor markets: A disequilibrium approach, *NBER Working Paper 4617*, National Bureau of Economic Research.

Oreffice, S. and Quintana-Domeque, C.: 2016, Beauty, body size and wages: Evidence from a unique data set, *Economics & Human Biology* **22**, 24–34.

Pan, J.: 2015, Gender segregation in occupations: The role of tipping and social interactions, *Journal of Labor Economics* **33**(2), 365–408.

Prendergast, C.: 1999, The provision of incentives in firms, *Journal of Economic Literature* **37**(1), 7–63.

Quandt, R. E.: 1958, The estimation of the parameters of a linear regression system obeying two separate regimes, *Journal of the American Statistical Association* **53**(284), 873–880.

Reimers, C. W.: 1983, Labor market discrimination against hispanic and black men, *The Review of Economics and Statistics* pp. 570–579.

Schmitz, P. W.: 2005, Workplace surveillance, privacy protection, and efficiency wages, *Labour Economics* **12**(6), 727–738.

Schumacher, P. and Morahan-Martin, J.: 2001, Gender, internet and computer attitudes and experiences, *Computers in human behavior* **17**(1), 95–110.

Shapiro, C. and Stiglitz, J. E.: 1984, Equilibrium unemployment as a worker discipline device, *American Economic Review* **74**(3), 433–444.

Shatnawi, D., Oaxaca, R. and Ransom, M.: 2014, Movin'on up: Hierarchical occupational segmentation and gender wage gaps, *The Journal of Economic Inequality* **3**(12), 315–338.

Słoczyński, T.: 2015, The oaxaca–blinder unexplained component as a treatment effects estimator, *Oxford Bulletin of Economics and Statistics* **77**(4), 588–604.

Stelcner, M., Van der Gaag, J. and Vijverberg, W.: 1989, A switching regression model of public-private sector wage differentials in peru: 1985-86, *Journal of Human Resources* **24**(3), 545–559.

Weichselbaumer, D. and Winter-Ebmer, R.: 2005, A meta-analysis of the international gender wage gap, *Journal of Economic Surveys* **19**(3), 479–511.

Weiss, A.: 1980, Job queues and layoffs in labor markets with flexible wages, *Journal of Political Economy* pp. 526–538.

Yellen, J. L. and Ackerlof, G.: 1984, Efficiency wage models of unemployment, *American Economic Review* **74**(2), 200–205.

Appendix A. Appendix

Table A.1: Literature approaches to determining β^*

λ value	Interpretation
$\lambda = 1$	Male coefficients taken as a reference
$\lambda = 0$	Female coefficients taken as a reference
$\lambda = 0.5$	Sample average of both, Reimers (1983)
$\lambda = \%male$	Coefficients weighted by % same gender, Cotton (1988)
$\beta^* = pooled$	Coefficients from a pooled regression, without gender dummy Neumark (1988)
$\beta^* = pooled$	Coefficients from a pooled regression, with gender dummy Fortin (2008)
$\lambda = \%female$	Coefficient weighted % opposite gender, Słoczyński (2015)

Appendix B. Sample design in SES 2006

Table B.2: Sampling technique and enterprises covered in each country

	Sampling technique	Size coverage
General	The sampling procedure for the SES contains typically two stages: first, in which a stratified sample of local units is drawn, and second, based on a simple random sample selection of employees within each of the selected local units.	at least 10+
Czech Republic	Only the first stage is employed. Enterprises listed on the Business Register are selected by industry, size group and region.	10+
Finland	The sample is based on national Structure of Earnings Statistics and contains 25% of employments in national data.	10+
France	A stratified random sample is provided.	10+
Greece	Three-stage sampling is employed: firstly selecting enterprises from the business register, secondly a sample of local units from these enterprises and finally a sample of employees.	10+
Hungary	All employers with more than 50 employees are obliged to report a sample of their employees. For employers with less than 50 employees, a 20% random sample is chosen from the business register of the CSO.	5+
Lithuania	Two-stage sampling is employed, stratified by economic activity, legal form of entity and size class.	1+
Luxembourg	Two-stage sampling is employed.	10+
Norway	Only the first stage is employed, stratified by NACE.	3+
Poland	Two-stage sampling is employed, stratified by NACE, ownership and size class.	1+
Portugal	Sample consists of all private sector employees in local units, stratified by NACE.	10+
Romania	Sample stratified by geographic location, economic activity and size class.	10+
Slovakia	No information has been documented in this regard.	1+
Spain	Sample is based on employees registered on the Social Security Register in October 2006.	1+
Sweden	Two-stage sampling is employed.	10+

Notes The size coverage contains an information on the size of enterprises in the sample. 10+ means that the data has been collected only from firms employing more than ten workers. In SES2006 an inclusion of enterprises with fewer than 10 employees was optional.

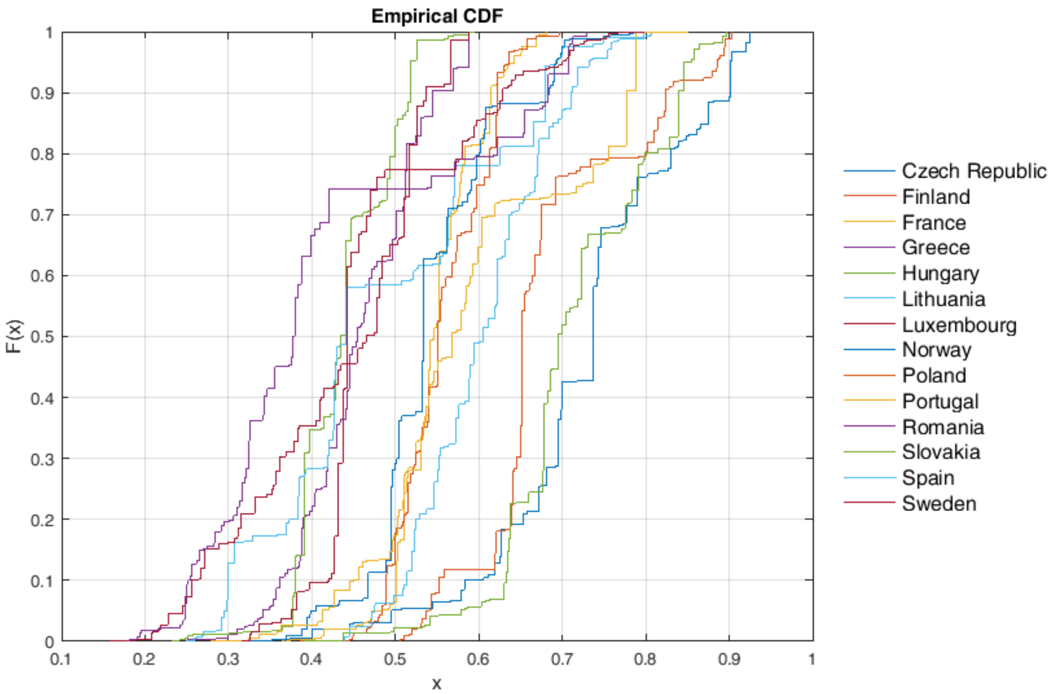
Appendix C. Data

Table C.3: Descriptive statistics (countries in alphabetical order)

		Czech 2006		Finland 2006		France 2006		Greece 2006		Hungary 2006		Lithuania 2006		Luxembourg 2006	
		men	women	men	women	men	women	men	women	men	women	men	women	men	women
Education	primary	8%	13%	16%	13%	15%	14%	28%	20%	11%	15%	8%	4%	27%	25%
	secondary	73%	71%	59%	64%	54%	60%	44%	48%	56%	45%	65%	60%	61%	65%
	tertiary	19%	16%	26%	23%	30%	25%	28%	32%	33%	40%	27%	36%	12%	11%
Occupation	blue collar	50%	26%	41%	15%	29%	12%	47%	20%	31%	16%	54%	26%	50%	22%
	white collar	50%	74%	59%	85%	71%	88%	53%	80%	69%	84%	46%	74%	50%	78%
Age	>25	22%	18%	17%	15%	13%	16%	18%	24%	19%	12%	22%	17%	20%	26%
	25> & <45	26%	24%	26%	21%	28%	30%	34%	35%	31%	24%	25%	25%	34%	36%
	45>	52%	58%	57%	64%	58%	54%	48%	41%	50%	64%	53%	58%	46%	38%
Industry	manufacturing	40%	22%	32%	10%	40%	19%	30%	20%	11%	4%	26%	22%	19%	7%
	services	55%	77%	60%	89%	55%	80%	64%	78%	85%	95%	59%	76%	60%	90%
	agriculture	4%	1%	8%	1%	5%	1%	6%	2%	4%	0%	15%	2%	21%	3%
		Norway 2006		Poland 2006		Portugal 2006		Romania 2006		Slovakia 2006		Spain 2006		Sweden 2006	
		men	women	men	women	men	women	men	women	men	women	men	women	men	women
Education	primary	23%	21%	43%	21%	61%	42%	12%	11%	5%	10%	55%	43%	15%	10%
	secondary	72%	76%	39%	48%	21%	21%	65%	60%	75%	73%	27%	29%	56%	52%
	tertiary	5%	2%	18%	31%	17%	37%	23%	29%	19%	17%	18%	28%	29%	39%
Occupation	blue collar	33%	10%	55%	22%	50%	27%	56%	33%	59%	33%	61%	29%	40%	11%
	white collar	67%	90%	45%	78%	50%	73%	44%	67%	41%	67%	39%	71%	60%	89%
Age	>25	20%	22%	20%	17%	21%	19%	20%	20%	20%	17%	22%	27%	17%	16%
	25> & <45	26%	25%	27%	28%	30%	31%	32%	34%	27%	27%	31%	34%	24%	22%
	45>	54%	53%	53%	55%	49%	51%	48%	46%	54%	55%	47%	39%	59%	62%
Industry	manufacturing	26%	9%	40%	19%	34%	18%	31%	32%	45%	33%	45%	23%	30%	9%
	services	65%	90%	53%	80%	53%	80%	56%	64%	49%	66%	43%	75%	63%	90%
	agriculture	9%	1%	7%	1%	12%	2%	13%	4%	6%	1%	12%	2%	7%	1%

Appendix D. Results

Figure D.1: Estimated probability of sample split – fitted values



Note: figure displays the empirical cumulative distribution functions for the estimated probability of sample split. Each sample split from a separate estimation. Formula for the sample split given by equation (6).

Table D.4: Joint significance test on interactions in the switching equation

	test-stat	p-value
Czech Republic	31317.07	0.000
Finland	471.31	0.000
France	327.22	0.000
Greece	93.70	0.000
Hungary	6777.30	0.000
Lithuania	866.57	0.000
Luxembourg	157.99	0.000
Norway	2990.91	0.000
Poland	2128.48	0.000
Portugal	575.67	0.000
Romania	186.58	0.000
Slovakia	10805.64	0.000
Spain	204.95	0.000
Sweden	735.43	0.000

Notes The appropriate 5% critical value from ${\chi^2}_{(8)}$ distribution is 15.507. The null hypothesis states that the gender specific regressors (male dummy and interactions between male dummy and other explanatory variables) in the switching equation are jointly and simultaneously equal to zero. The null is rejected in all cases.

Table D.5: Wald test for prevalence of two markets

	test-stat	p-value
Czech Republic	606403.84	0.000
Finland	264272.47	0.000
France	11268.91	0.000
Greece	7527.02	0.000
Hungary	381792.76	0.000
Lithuania	-	0.000
Luxembourg	13050.16	0.000
Norway	202222.29	0.000
Poland	55493.37	0.000
Portugal	19287.74	0.000
Romania	95662.03	0.000
Slovakia	175697.56	0.000
Spain	48280.01	0.000
Sweden	84367.11	0.000

Notes The appropriate 5% critical value from ${\chi^2}_{(12)}$ distribution is 21.026. The null hypothesis states that parameters in two markets are the same. The null is rejected in all cases.

Table D.6: Estimated coefficients - Czech Republic

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	4.595	4.598	4.601	3.984	3.987	3.989	4.408	0.646
Gender = man	0.354	0.358	0.363	0.067	0.071	0.074	0.335	0.097
Secondary educ	0.139	0.141	0.143	0.043	0.045	0.046	0.175	0.241
x gender = man	-0.009	-0.005	-0.002	0.091	0.093	0.096	0.031	0.061
Higher educ	0.519	0.522	0.525	0.594	0.596	0.598	0.529	-0.070
x gender = man	0.010	0.014	0.019	0.006	0.009	0.012	0.007	0.013
Blue collar	-0.231	-0.229	-0.227	-0.111	-0.110	-0.109	-0.256	-0.648
x gender = man	-0.048	-0.046	-0.043	-0.003	-0.001	0.002	0.020	0.497
Age <25	-0.066	-0.064	-0.062	-0.064	-0.062	-0.061	-0.055	0.014
x gender = man	-0.131	-0.128	-0.125	-0.042	-0.040	-0.037	-0.159	-0.207
Age >45	0.028	0.029	0.031	0.022	0.023	0.025	0.049	0.127
x gender = man	-0.080	-0.078	-0.076	-0.098	-0.096	-0.094	-0.080	0.013
Manufacturing								0.397
x gender = man								-0.139
Services								-0.378
x gender = man								-0.003

Table D.7: Estimated coefficients - Finland

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	2.574	2.580	2.585	2.127	2.131	2.134	2.442	0.708
Gender = man	0.269	0.277	0.285	0.138	0.143	0.149	0.245	0.150
Secondary educ	0.062	0.066	0.070	0.080	0.083	0.086	0.062	-0.041
x gender = man	-0.008	-0.001	0.005	-0.033	-0.027	-0.022	0.016	0.079
Higher educ	0.421	0.427	0.432	0.233	0.237	0.241	0.374	-0.046
x gender = man	-0.065	-0.058	-0.050	0.025	0.032	0.038	0.010	0.133
Blue collar	-0.135	-0.131	-0.127	-0.093	-0.090	-0.086	-0.133	-0.283
x gender = man	-0.016	-0.011	-0.005	0.042	0.046	0.051	0.015	-0.038
Age <25	-0.155	-0.150	-0.145	-0.069	-0.065	-0.061	-0.142	-0.051
x gender = man	-0.044	-0.038	-0.031	-0.083	-0.079	-0.075	-0.052	0.024
Age >45	0.077	0.080	0.084	0.067	0.070	0.073	0.083	0.029
x gender = man	0.013	0.019	0.024	-0.006	-0.002	0.002	0.016	0.023
Manufacturing								0.511
x gender = man								-0.205
Services								-0.306
x gender = man								-0.189

Table D.8: Estimated coefficients - France

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	2.629	2.648	2.667	2.237	2.249	2.261	2.460	-0.052
Gender = man	0.209	0.228	0.247	0.111	0.124	0.138	0.212	0.100
Secondary educ	0.144	0.157	0.170	0.107	0.116	0.124	0.161	0.100
x gender = man	-0.017	-0.001	0.016	0.040	0.052	0.063	-0.007	-0.156
Higher educ	0.486	0.501	0.516	0.289	0.299	0.309	0.457	0.135
x gender = man	-0.008	0.011	0.030	0.076	0.090	0.104	0.011	-0.186
Blue collar	-0.309	-0.295	-0.281	-0.187	-0.178	-0.169	-0.275	-0.249
x gender = man	-0.003	0.014	0.030	-0.055	-0.044	-0.033	0.005	0.173
Age <25	-0.234	-0.222	-0.210	-0.139	-0.131	-0.122	-0.193	-0.021
x gender = man	-0.087	-0.070	-0.054	-0.224	-0.211	-0.199	-0.083	0.217
Age >45	0.175	0.184	0.193	0.087	0.093	0.100	0.172	0.107
x gender = man	0.054	0.066	0.079	0.017	0.026	0.035	0.056	-0.010
Manufacturing								0.282
x gender = man								-0.085
Services								-0.022
x gender = man								0.033

Table D.9: Estimated coefficients - Greece

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	1.953	1.986	2.020	1.652	1.669	1.686	1.784	-0.415
Gender = man	0.127	0.158	0.188	-0.015	-0.001	0.013	0.093	0.111
Secondary educ	0.167	0.190	0.212	-0.008	0.002	0.013	0.128	0.103
x gender = man	-0.063	-0.036	-0.009	0.033	0.046	0.059	0.011	-0.017
Higher educ	0.456	0.481	0.506	0.150	0.162	0.174	0.405	0.245
x gender = man	0.001	0.032	0.062	0.055	0.070	0.085	0.064	-0.046
Blue collar	-0.238	-0.215	-0.192	-0.081	-0.071	-0.061	-0.190	-0.134
x gender = man	0.122	0.148	0.174	0.046	0.058	0.070	0.116	0.098
Age <25	-0.300	-0.282	-0.264	-0.146	-0.137	-0.128	-0.247	-0.146
x gender = man	-0.054	-0.030	-0.005	-0.033	-0.021	-0.010	-0.026	0.017
Age >45	0.228	0.243	0.259	0.067	0.076	0.084	0.222	0.171
x gender = man	0.063	0.083	0.102	0.040	0.050	0.061	0.088	0.020
Manufacturing								0.022
x gender = man								-0.048
Services								0.173
x gender = man								-0.034

Table D.10: Estimated coefficients - Hungary

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	6.832	6.838	6.843	6.120	6.122	6.125	6.394	-0.418
Gender = man	0.296	0.303	0.311	0.148	0.153	0.157	0.296	-0.134
Secondary educ	0.102	0.107	0.111	0.054	0.056	0.058	0.126	0.037
x gender = man	-0.017	-0.010	-0.002	-0.052	-0.048	-0.044	-0.007	0.012
Higher educ	0.566	0.571	0.576	0.585	0.587	0.590	0.553	-0.091
x gender = man	-0.054	-0.046	-0.038	-0.152	-0.148	-0.144	-0.050	0.203
Blue collar	-0.319	-0.315	-0.310	-0.191	-0.189	-0.186	-0.273	-0.129
x gender = man	-0.097	-0.091	-0.086	-0.087	-0.083	-0.080	-0.138	-0.021
Age <25	-0.127	-0.122	-0.118	-0.143	-0.140	-0.138	-0.100	0.046
x gender = man	-0.088	-0.082	-0.076	0.039	0.043	0.046	-0.040	-0.073
Age >45	0.054	0.057	0.060	0.164	0.166	0.168	0.138	0.032
x gender = man	-0.009	-0.005	0.000	-0.042	-0.039	-0.037	-0.066	-0.049
Manufacturing								0.380
x gender = man								0.297
Services								0.201
x gender = man								0.302

Table D.11: Estimated coefficients - Lithuania

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	2.162	2.188	2.215	1.327	1.331	1.334	1.769	-0.003
Gender = man	0.274	0.306	0.338	0.006	0.014	0.022	0.234	0.397
Secondary educ	0.079	0.103	0.126	-0.006	-0.001	0.004	0.077	0.029
x gender = man	-0.046	-0.016	0.013	0.015	0.022	0.028	0.009	0.060
Higher educ	0.458	0.483	0.508	0.064	0.069	0.074	0.615	0.644
x gender = man	-0.070	-0.038	-0.006	-0.016	-0.008	0.000	-0.094	-0.197
Blue collar	-0.185	-0.173	-0.162	-0.024	-0.022	-0.019	-0.218	-0.352
x gender = man	0.090	0.105	0.121	0.021	0.025	0.029	0.087	0.004
Age <25	-0.089	-0.076	-0.063	-0.008	-0.005	-0.001	-0.099	-0.108
x gender = man	-0.071	-0.053	-0.035	-0.022	-0.017	-0.012	0.017	0.127
Age >45	-0.015	-0.005	0.004	-0.003	0.000	0.003	0.010	0.037
x gender = man	-0.067	-0.053	-0.039	-0.006	-0.002	0.002	-0.045	-0.014
Manufacturing								-0.008
x gender = man								-0.333
Services								-0.213
x gender = man								-0.471

Table D.12: Estimated coefficients - Luxembourg

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	3.042	3.079	3.116	2.313	2.325	2.336	2.622	-0.470
Gender = man	0.132	0.171	0.210	0.184	0.195	0.205	0.218	0.119
Secondary educ	0.173	0.204	0.235	0.061	0.069	0.076	0.255	0.282
x gender = man	-0.104	-0.068	-0.033	-0.010	-0.001	0.007	-0.101	-0.157
Higher educ	0.521	0.562	0.603	0.530	0.542	0.554	0.627	0.224
x gender = man	-0.057	-0.008	0.041	-0.050	-0.035	-0.020	-0.073	-0.173
Blue collar	-0.307	-0.274	-0.241	-0.085	-0.077	-0.069	-0.329	-0.440
x gender = man	-0.103	-0.066	-0.030	0.067	0.076	0.085	-0.016	0.045
Age <25	-0.321	-0.297	-0.273	-0.042	-0.035	-0.027	-0.256	-0.096
x gender = man	-0.041	-0.010	0.021	-0.119	-0.109	-0.099	-0.016	0.028
Age >45	0.145	0.166	0.187	0.003	0.010	0.017	0.122	0.023
x gender = man	0.007	0.033	0.059	0.014	0.023	0.032	0.068	0.115
Manufacturing								0.100
x gender = man								0.211
Services								0.230
x gender = man								0.024

Table D.13: Estimated coefficients - Norway

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	5.202	5.206	5.209	4.833	4.835	4.837	5.019	0.074
Gender = man	0.184	0.188	0.192	0.034	0.036	0.038	0.157	0.410
Secondary educ	0.157	0.160	0.163	0.101	0.102	0.103	0.167	0.165
x gender = man	0.017	0.020	0.024	0.050	0.052	0.054	0.014	-0.150
Higher educ	0.324	0.331	0.338	0.198	0.201	0.205	0.328	0.206
x gender = man	-0.034	-0.026	-0.018	0.107	0.112	0.116	0.026	-0.017
Blue collar	-0.123	-0.120	-0.116	-0.082	-0.080	-0.078	-0.125	-0.253
x gender = man	-0.052	-0.048	-0.043	0.011	0.012	0.014	-0.014	-0.001
Age <25	-0.217	-0.214	-0.211	-0.229	-0.227	-0.226	-0.210	0.075
x gender = man	-0.072	-0.068	-0.064	-0.090	-0.087	-0.085	-0.039	0.186
Age >45	0.037	0.039	0.042	0.045	0.047	0.048	0.043	0.003
x gender = man	0.055	0.059	0.062	0.012	0.014	0.016	0.044	-0.030
Manufacturing								0.235
x gender = man								-0.209
Services								-0.159
x gender = man								-0.328

Table D.14: Estimated coefficients - Poland

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	2.516	2.525	2.533	1.935	1.944	1.952	2.247	-0.117
Gender = man	0.443	0.452	0.461	0.053	0.060	0.067	0.301	0.200
Secondary educ	0.299	0.305	0.311	0.183	0.187	0.192	0.277	0.081
x gender = man	-0.132	-0.124	-0.116	-0.022	-0.016	-0.011	-0.073	-0.041
Higher educ	0.883	0.890	0.897	0.616	0.622	0.627	0.857	0.267
x gender = man	-0.266	-0.257	-0.248	0.013	0.021	0.028	-0.164	-0.118
Blue collar	-0.314	-0.308	-0.301	-0.172	-0.168	-0.163	-0.262	-0.072
x gender = man	0.068	0.076	0.084	0.165	0.171	0.177	0.135	0.013
Age <25	-0.263	-0.258	-0.253	-0.236	-0.232	-0.227	-0.284	-0.102
x gender = man	-0.071	-0.064	-0.057	0.036	0.043	0.049	-0.014	0.012
Age >45	0.088	0.092	0.096	0.136	0.139	0.143	0.122	0.025
x gender = man	-0.058	-0.053	-0.047	-0.112	-0.107	-0.102	-0.064	0.051
Manufacturing								0.162
x gender = man								0.044
Services								0.137
x gender = man								-0.200

Table D.15: Estimated coefficients - Portugal

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	1.622	1.637	1.652	1.090	1.101	1.111	1.382	-0.050
Gender = man	0.423	0.438	0.453	0.202	0.213	0.223	0.412	0.096
Secondary educ	0.333	0.344	0.356	0.213	0.221	0.228	0.370	0.223
x gender = man	-0.052	-0.037	-0.021	-0.066	-0.055	-0.044	-0.070	-0.116
Higher educ	0.855	0.867	0.879	0.578	0.588	0.598	0.996	0.774
x gender = man	-0.229	-0.214	-0.198	-0.035	-0.022	-0.008	-0.210	-0.210
Blue collar	-0.307	-0.296	-0.285	-0.144	-0.137	-0.130	-0.216	0.237
x gender = man	-0.115	-0.101	-0.086	-0.040	-0.031	-0.021	-0.125	-0.262
Age <25	-0.239	-0.228	-0.218	-0.131	-0.123	-0.115	-0.223	-0.129
x gender = man	-0.068	-0.053	-0.038	-0.083	-0.072	-0.061	-0.065	0.011
Age >45	0.278	0.286	0.294	0.058	0.064	0.071	0.240	0.035
x gender = man	-0.055	-0.043	-0.032	0.008	0.018	0.027	-0.019	0.035
Manufacturing								-0.408
x gender = man								0.393
Services								0.041
x gender = man								0.040

Table D.16: Estimated coefficients - Romania

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	0.556	0.572	0.588	0.761	0.764	0.767	1.278	-0.689
Gender = man	-0.007	0.015	0.036	0.022	0.025	0.029	0.038	-0.011
Secondary educ	0.297	0.311	0.325	0.053	0.055	0.057	0.210	0.232
x gender = man	-0.021	-0.003	0.015	-0.010	-0.008	-0.005	0.000	-0.024
Higher educ	1.285	1.301	1.318	0.203	0.207	0.211	0.950	1.087
x gender = man	-0.040	-0.017	0.005	-0.015	-0.010	-0.006	0.020	-0.089
Blue collar	-0.160	-0.150	-0.140	0.011	0.012	0.014	-0.182	-0.012
x gender = man	0.112	0.125	0.138	-0.030	-0.028	-0.026	0.115	0.038
Age <25	-0.274	-0.263	-0.251	-0.032	-0.030	-0.028	-0.183	-0.220
x gender = man	0.007	0.023	0.038	0.004	0.007	0.010	0.008	0.017
Age >45	0.176	0.185	0.194	0.034	0.036	0.037	0.119	0.151
x gender = man	-0.008	0.004	0.017	-0.004	-0.002	0.001	-0.006	0.016
Manufacturing								0.062
x gender = man								0.034
Services								0.000
x gender = man								0.012

Table D.17: Estimated coefficients - Slovakia

	Market 1 CI U	Market 1	Market 1 CI L	Market 2 CI U	Market 2	Market 2 CI L	OLS	Switching model
Constant	4.536	4.543	4.550	3.926	3.930	3.934	4.302	0.113
Gender = man	0.270	0.280	0.290	-0.038	-0.032	-0.026	0.293	0.266
Secondary educ	0.155	0.161	0.166	0.070	0.073	0.076	0.210	0.264
x gender = man	0.039	0.047	0.056	0.091	0.096	0.102	0.062	-0.012
Higher educ	0.657	0.664	0.671	0.478	0.482	0.486	0.691	0.260
x gender = man	-0.051	-0.041	-0.031	-0.124	-0.117	-0.111	-0.001	0.157
Blue collar	-0.208	-0.205	-0.202	-0.107	-0.105	-0.103	-0.192	-0.358
x gender = man	0.039	0.044	0.048	0.082	0.086	0.089	0.066	0.278
Age <25	-0.058	-0.053	-0.049	-0.044	-0.041	-0.038	-0.051	-0.009
x gender = man	-0.116	-0.110	-0.104	-0.009	-0.004	0.000	-0.131	-0.195
Age >45	0.015	0.018	0.022	-0.002	0.000	0.002	0.036	0.111
x gender = man	-0.041	-0.037	-0.032	0.007	0.011	0.015	-0.061	-0.136
Manufacturing								0.463
x gender = man								0.003
Services								-0.024
x gender = man								0.011

Table D.18: Estimated coefficients - Spain

	Market 1 CI U	Market 1	Market 1 CI L	Market 2 CI U	Market 2	Market 2 CI L	OLS	Switching model
Constant	2.065	2.075	2.085	1.528	1.535	1.542	1.887	0.309
Gender = man	0.277	0.286	0.295	0.160	0.167	0.173	0.264	-0.038
Secondary educ	0.124	0.131	0.138	0.033	0.038	0.042	0.126	0.096
x gender = man	0.030	0.038	0.047	0.029	0.035	0.041	0.058	0.070
Higher educ	0.492	0.499	0.507	0.181	0.186	0.192	0.492	0.316
x gender = man	-0.092	-0.082	-0.072	-0.007	0.001	0.009	-0.052	-0.007
Blue collar	-0.157	-0.150	-0.143	-0.027	-0.022	-0.018	-0.145	-0.188
x gender = man	-0.031	-0.022	-0.014	0.031	0.037	0.043	-0.008	-0.022
Age <25	-0.149	-0.142	-0.136	-0.052	-0.047	-0.042	-0.138	-0.093
x gender = man	-0.039	-0.030	-0.021	-0.054	-0.047	-0.041	-0.029	0.027
Age >45	0.135	0.141	0.148	0.004	0.008	0.013	0.134	0.111
x gender = man	0.057	0.064	0.072	0.018	0.024	0.030	0.061	0.019
Manufacturing								-0.007
x gender = man								0.128
Services								-0.182
x gender = man								0.049

Table D.19: Estimated coefficients - Sweden

	Market 1 CI U	Market 1	Market 1 CI L	Market 2 CI U	Market 2	Market 2 CI L	OLS	Switching model
Constant	4.918	4.927	4.935	4.587	4.587	4.587	4.781	0.424
Gender = man	0.216	0.226	0.236	0.081	0.082	0.083	0.185	0.288
Secondary educ	0.053	0.060	0.067	0.018	0.019	0.020	0.046	0.029
x gender = man	0.013	0.022	0.031	-0.012	-0.011	-0.009	0.026	-0.011
Higher educ	0.237	0.244	0.252	0.141	0.142	0.143	0.210	0.016
x gender = man	-0.009	0.001	0.011	-0.033	-0.031	-0.029	0.009	0.029
Blue collar	-0.128	-0.121	-0.115	-0.094	-0.093	-0.091	-0.124	-0.278
x gender = man	-0.156	-0.148	-0.140	0.022	0.024	0.026	-0.051	0.151
Age <25	-0.120	-0.114	-0.108	-0.117	-0.116	-0.114	-0.107	0.097
x gender = man	-0.057	-0.048	-0.040	-0.051	-0.048	-0.046	-0.064	-0.019
Age >45	0.009	0.014	0.018	0.040	0.041	0.043	0.041	0.027
x gender = man	0.067	0.074	0.080	-0.006	-0.004	-0.003	0.038	-0.064
Manufacturing								0.072
x gender = man								-0.435
Services								-0.624
x gender = man								-0.241

Appendix E. LFS

Table E.20: Descriptive statistics: LFS

		Poland 2006		France 2006	
		men	women	men	women
Education	primary	10%	7%	59%	49%
	secondary	74%	66%	16%	19%
	tertiary	16%	27%	26%	32%
Occupation	blue collar	62%	27%	47%	21%
	white collar	38%	73%	53%	79%
Age	>25	2%	3%	3%	3%
	25> & <45	92%	93%	92%	92%
	45>	5%	5%	5%	5%
Industry	manufacturing	37%	22%	11%	1%
	services	47%	76%	61%	87%
	agriculture	16%	2%	28%	12%

Table E.21: Joint significance test on interactions in the switching equation: LFS

	test-stat	p-value
Poland	102.2290	0.000
France	35.3983	0.000

Table E.22: Wald test for prevalence of two markets: LFS

	test-stat	p-value
Poland	842.2278	0.000
France	.	.

Table E.23: Estimated coefficients - France LFS

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	2.278	2.294	2.310	1.493	1.517	1.541	2.084	0.398
Gender = man	0.217	0.232	0.247	0.220	0.256	0.293	0.225	-0.003
Secondary educ	0.097	0.113	0.128	0.107	0.123	0.139	0.092	-0.104
x gender = man	-0.063	-0.041	-0.019	-0.165	-0.121	-0.077	-0.057	0.076
Higher educ	0.402	0.416	0.430	0.394	0.426	0.458	0.379	-0.171
x gender = man	-0.099	-0.078	-0.058	-0.248	-0.200	-0.153	-0.106	0.095
Blue collar	-0.223	-0.209	-0.194	-0.124	-0.088	-0.051	-0.125	0.050
x gender = man	-0.089	-0.070	-0.051	-0.093	-0.045	0.003	-0.048	-0.038
Age <25	-0.348	-0.315	-0.282	-0.338	-0.338	-0.338	-0.268	0.193
x gender = man	-0.046	-0.002	0.042	-0.239	-0.156	-0.072	-0.038	0.077
Age >45	0.096	0.121	0.145	0.098	0.155	0.212	0.124	-0.081
x gender = man	-0.039	-0.005	0.030	-0.087	-0.007	0.072	0.003	0.022
Manufacturing								-0.354
x gender = man								0.233
Services								-0.146
x gender = man								-0.085

Table E.24: Estimated coefficients - Poland LFS

	Market 1			Market 2			OLS	Switching model
	CI U		CI L	CI U		CI L		
Constant	1.788	1.878	1.968	1.152	1.233	1.315	1.612	0.301
Gender = man	-0.110	-0.007	0.095	-1.173	-1.035	-0.896	0.121	1.107
Secondary educ	0.105	0.179	0.253	0.162	0.224	0.286	0.189	-0.112
x gender = man	0.167	0.261	0.355	1.000	1.134	1.268	0.102	-1.198
Higher educ	0.739	0.822	0.905	0.505	0.574	0.643	0.764	0.036
x gender = man	0.027	0.141	0.256	0.980	1.126	1.272	-0.020	-1.353
Blue collar	-0.310	-0.265	-0.219	-0.035	0.000	0.036	-0.170	-0.144
x gender = man	0.017	0.077	0.137	-0.098	-0.048	0.001	0.015	-0.003
Age <25	-0.375	-0.270	-0.166	-0.457	-0.362	-0.267	-0.334	0.021
x gender = man	-0.038	0.111	0.260	0.127	0.259	0.390	0.117	-0.270
Age >45	0.044	0.122	0.200	-0.202	-0.119	-0.037	0.129	0.307
x gender = man	-0.190	-0.087	0.016	0.115	0.224	0.333	-0.031	-0.170
Manufacturing								-0.040
x gender = man								0.332
Services								-0.128
x gender = man								0.172