CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

Discussion Paper Series

CZECH GENDER SEGREGATION AND WAGE GAP WITH ANTI-DISCRIMINATION LEGISLATION

Štěpán JURAJDA

Discussion Paper No. 2004 – 125

April 2004

P.O. Box 882, Politických vězňů 7, 111 21 Praha 1, Czech Republic
http://www.cerge-ei.cz

Czech Gender Segregation and Wage Gap with Anti-Discrimination Legislation

Štěpán Jurajda
CERGE-EI
CEPR, IZA, and WDI

March 5, 2004

Abstract

Using 2002 matched employer-employee data we provide evidence on the size and structure of the Czech gender wage gap after the 2000 introduction of western-type anti-discrimination policies. Despite the new legislation, within-firm sex segregation continues to account for about one third of the gender wage gap while its unexplained portion remains just under two thirds.
JEL Classification: J3, J7, P3

Acknowledgements Partial support provided by the U.S. Agency for International Developments under its Task Order No. 803: Improvement of Economic Policy Through Think Tank Partnership.
Address CERGE-EI, POB 882, Politických vězňů 7, Prague, Czech Republic. E-mail: stepan.jurajda@cerge-ei.cz
CERGE-EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic.

1. Introduction

"Harmonization" of legislation is one of the pre-requisites of accession to the European Union. Hence, in 2000 the Czech Republic enacted the standard set of anti-discrimination policies including the equal pay and equal employment opportunity clauses. Each of these anti-discrimination policies affects a different source of the overall pay difference between men and women. The equal pay provision targets wage differences within job cells, where a job cell is defined as a group of workers with the same occupation within the same firm. The equal employment opportunity clauses affect all forms of segregation, where segregation stands for discriminatory hiring, firing, and promotion practices that result in high concentration of women in low-paying occupations, firms, or job cells.

To guide the implementation of the new policies and to measure their effect, this note decomposes the Czech gender wage gap in the enterprise sector from the first quarter of 2002 into its components corresponding to specific anti-discrimination policies. In particular, we break down the total wage gap into three parts: (i) the gender wage difference attributable to differences in productive characteristics of men and women, (ii) the pay difference one can explain using measures of female segregation and so attribute to potential violations of the equal employment opportunity clauses, and (iii) the unexplained part of the gap, which represents an upper limit on the violation of the equal pay act.

Furthermore, in point (ii), we are able to distinguish among the effects of three types of gender segregation: across occupations, firms, and job cells. The ability to observe within-firm gender segregation comes from the use of matched employer-employee data.[1] We extend the results of Jurajda (2003), who compares wage gaps across the Czech Republic and Slovakia in 1998, before the introduction of anti-discrimination policies (and who offers a brief survey of the existing literature). Our empirical results, based on a sample of over 800 thousand workers and 2000 firms, suggest a minor change occurred in the structure of the gender pay gap between 1998 and 2002.

[1] See Groshen (1991) and Bayard et al. (2003) for gender wage gap analysis using such data from the U.S.

2. Data

The data consists of a national employer survey, the Information System on Average Earnings (ISAE), in which participating firms report hourly wages of all of their employees except top management. The stratified sampling is based on the country firm register and covers only firms employing more than 10 workers; the budgetary sector of health, education, and public administration is not included. The data cover about one third of the entire Czech enterprise employment.

The data are drawn directly from companies' personnel databases, minimizing the extent of reporting error typically present in survey data. The wage measure is a quarterly average used for social security purposes.[2] The data also include a number of firm and personal characteristics.[3]

The sample is not perfectly representative of the population of firms. Thus, to recover population statistics, weights reflecting the sampling procedure were calculated by dividing the population frequency of firms within strata cells by the corresponding sample frequency. The re-weighted data form the basis for the analysis. There are 805,767 worker wage records available from 2,240 enterprises from the first quarter of 2002.

3. Econometric Approach

A large literature on wage discrimination follows Oaxaca (1973) in decomposing the overall mean wage difference between the advantaged (men) and disadvantaged (women) into two parts: the first reflecting the difference in average productive endowments of individuals in each group and the second part due to the differences in coefficients. Following this approach, one first estimates logarithmic wage regressions separately for each gender, controlling for productive characteristics.

[2] It is defined as total cash compensation including bonuses divided by hours worked in a given quarter. The wage data is equivalent to social-security wage records available from many developed economies (i.e. Germany or the U.S.).

[3] The only problem is education, which is missing for 8% of the workers. The incidence of missing education appears to be non-systematic; this part of data is therefore excluded from the analysis.

The decomposition technique then relies on the fact that the fitted regressions pass through sample means. A general form of the mean wage decomposition is as follows:

$$\overline{\ln w_m} - \overline{\ln w_f} = (\overline{X_m} - \overline{X_f})'\widetilde{\beta} + [\overline{X_m}'(\widehat{\beta_m} - \widetilde{\beta}) + \overline{X_f}'(\widetilde{\beta} - \widehat{\beta_f})], \tag{1}$$

where f denotes females and m denotes males, $\overline{\ln w_s}$ is the gender-specific mean of the natural logarithm of hourly wage, $\overline{X_s}$ represents the respective vectors of mean values of explanatory variables for men and women with $s \in \{f, m\}$. Finally, $\widehat{\beta_m}$ and $\widehat{\beta_f}$ are the corresponding vectors of estimated coefficients and $\widetilde{\beta}$ represents a counter-factual non-discriminatory wage structure.

The first term on the right hand side of equation 1 represents that part of the total logarithmic wage difference which stems from the difference in average productive characteristics across gender. The second term originates in the differences in gender-specific coefficients from the non-discriminatory wage structure and is often interpreted as reflecting wage discrimination.

In this note we follow Groshen (1991) and present the wage gap decompositions in a particularly simple form: we estimate pooled regressions based on both male and female data and consider the female dummy coefficient as our estimate of the unexplained portion of the gap. We note that using the pooled-data regression to approximate the non-discriminatory wage structure has been advocated in the literature (e.g., Oaxaca and Ransom, 1994). Further, our simplified approach conserves space and, in Jurajda (2003), it led to results virtually identical to those based on the gender-specific regressions. We therefore estimate log-wage regressions of the following form:

$$\ln w_{ij} = \alpha F_{ij} + X_{ij}^{'}\beta + P_{ij}^{'}\gamma + \epsilon_{ij}, \text{ with } i = 1, ...N_j, \text{ and } j = 1, ..., J, \tag{2}$$

where w_{ij} denotes the hourly wage of the i-th worker in the j-th firm, and where $F_{ij} = 1$ if the worker is female and equals 0 otherwise. J denotes the number of firms in the sample and N_j is the number of workers in the j-th firm. Following Groshen (1991), the effect of gender segregation on wages is captured by conditioning on the "femaleness" of occupations, firms, and job cells, where "femaleness" is measured by the percent of females (P) in a given group of employees. The elements

of the P_{ij} vector are the fraction of female employment in the ij-th worker's occupation, firm, and job cell.

Finally, to capture firm-level clustering of unobservables we use a panel data version of the Huber/White variance estimator:

$$\widehat{V}(\widehat{\delta}) = (Z'Z)^{-1}\left(\sum_{j} Z_j'\widehat{\epsilon}_j\widehat{\epsilon}_j'Z_j\right)(Z'Z)^{-1}, \tag{3}$$

where $\widehat{\epsilon}_j = \ln w_j - Z_j\widehat{\delta}$ is the column vector of estimated error terms for employees of the j-th establishment, δ is the collection of parameters, $\delta' = (\alpha', \beta', \gamma')$, and where the matrix of regressors Z is defined accordingly.

4. Results

Table 1 presents the estimates of the female dummy from a sequence of specifications. In column (1), the female dummy estimate is reported based on a regression where no other explanatory variables are used. This "raw" overall logarithmic wage gap, representing the female wage disadvantage as a fraction of the average male wage, is -0.28. The next question is how much of the wage gap can be accounted for by gender differences in workers' productive characteristics and in the type of their employer. In column (2), we therefore condition on workers' age and education as well as on a quadratic in firm size (employment) and a set of dummy variables reflecting firms' 2-digit industrial classification, ownership type, and region of location. This reduces the female dummy estimate by about 4 percentage points. Fully controlling for firm identity, in column (3), by using a set of firm fixed effects, further decreases the estimates by 1.4 percentage point. Yet, over 80% of the "raw" gap remains unexplained.

Columns (4) to (7) gradually introduce the gender segregation measures. While occupational segregation appears significant in explaining the wage gap in column (4), conditioning on the "femaleness" of occupations lowers the unexplained portion of the pay gap by only about 2 percentage points. Job-cell segregation appears as the most powerful regressor in terms of reducing the unex-

plained wage gap in column (5): the introduction of job-cell female concentration lowers the female dummy coefficient by 4 percentage points. Further, both occupational and job-cell segregation coefficients are highly statistically significant.

The ISAE sample covers 27 2-digit occupations and 541 occupational categories at the 4-digit level. In column (6), we explore the sensitivity of the estimates with respect to using more detailed occupational groups. While there is a small decrease in the female dummy coefficient, the segregation parameters remain quite stable. We therefore proceed with the 2-digit categorization. Finally, in column (7) we control for all three forms of segregation simultaneously. The firm-level "femaleness" coefficient is negative but not statistically significant, while the two within-firm segregation coefficients remain large and significant.

In the left panel of Table 2, the estimates from column (7) of Table 1 are used to calculate a Oaxaca-Blinder decomposition. A clear picture emerges: Almost 60 percent of the wag gap remains unexplained after controlling for detailed worker and firm characteristics and gender segregation. Summing up the three segregation-related wage-gap contributions, we see that about a third of the wage gap is linked to gender segregation. Almost all of the segregation takes place within firms. The effects of other explanatory variables are small and cancel each other.

Our 2002 results can be compared to those from 1998 presented in Jurajda (2003). How does the structure of the wage gap as of two years before the introduction of the anti-discriminination legislation compare to the situation as of two years later? The size of the ISAE sample increased substantially between 1998 and 2002. Unfortunately, we are not able to condition our before/after comparison on a firm panel sub-sample; hence, we rely on industry, ownership and firm size controls to remove the effect of the changing sample structure. It is interesting to note that the share of women out of total sample employment remained constant between 1998 and 2002 at 39%, despite the increase in the sample size from 578 to 805 thousand workers.

The unconditional log wage gap in the enterprise sector in 1998 was -0.297, somewhat above the 2002 level of -0.282. More importantly, given the change in the firm sample, the decrease in the

size of the gap is about 2.5 percentage points after we control for worker and firm characteristics.[4] The right panel of Table 2 offers the relative contribution of different types of segregation to the overall wage gap in 1998.[5] The change in the occupation and job-cell contributions over time is small, but there is a substantial decrease in the wage-gap contribution of firm-level segregation driven by the drop in the parameter estimate. Consequently, the overall contribution of all types of segregation towards explaining the gender wage gap drops from almost 50% in 1998 to about one third in 2002.

5. Conclusion

This note sheds light on the channels through which gender segregation and an individual's gender affect the overall gender wage gap in the Czech enterprise sector and how this relationship changed with the introduction of new anti-discrimination policies in 2000. An important caveat is that the estimated 'pure' wage effect of the individual's gender is likely to be affected by the lack of information on the actual length of labor market experience.

Our results are suggestive of little change in the structure of the wage gap, except for a decrease in the effect of firm-level gender segregation. The drop in the firm-level segregation contribution may be due to the new anti-discrimination legislation. As of 2002, there has been a few court trials concerning unequal hiring practices (CHC, 2002). On the other hand, we continue to find that despite the new legislation almost two thirds of the gender wage gap remains unexplained and segregation still represents a major source of the gap. Segregation now affects gender wage differences by working within firms so that an implementation of the anti-discrimination policies aiming to equalize wages in occupations across firms would have little effect.

---

[4] The gap drops from -0.266 in 1998 to -0.240 in 2002, see column (2) of Table 1.

[5] We note that the gender differences in the (unconditional) means of female concentration by occupation, job cell and firm changed somewhat between 1998 and 2003, but the there is no clear pattern and we hesitate to draw conclusions given the changing size of the sample.

Bibliography

Bayard, K., Hellerstein, J., Neumark D. and K. Troske (2003) "New Evidence on gender Segregation and gender Differences in Wages from Matched Employee-Employer Data," *Journal of Labor Economics*, 21(4), 887-922.

CHC (Czech Helsinki Committee) (2002), *Enforcing the policy of equal opportunity on the Czech labor market (Prosazování politiky rovných příležitostí na trhu práce v ČR)*, CHC, Prague.

Groshen, E.L. (1991) "The Structure of the Female/Male Wage Differential: Is it Who You Are, What You Do, or Where You Work?" *Journal of Human Resources*, 26 (3), 457-72.

Jurajda, Š. (2003) "Gender Wage Gap and Segregation in Enterprises and the Public Sector in Late Transition Countries," *Journal of Comparative Economics*, 31 (2), 199-222.

Oaxaca, R. (1973) "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14(3), 693-709.

Oaxaca, R. and M.R. Ransom (1994) "On Discrimination and the Decomposition of Wage Differentials," *Journal of Econometrics*, 61, 5-21.

Table 1: Estimated Log Wage Differentials by Gender, and Percent Female in Occupation, Firm, and Job Cell: WLS Regressions in Czech Enterprises

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.282	-0.240	-0.226	-0.207	-0.165	-0.143	-0.165
	(0.011)	(0.007)	(0.001)	(0.008)	(0.004)	(0.005)	(0.004)
% female in occupation				-0.149	-0.080	-0.060	-0.084
				(0.015)	(0.019)	(0.019)	(0.02)
% female in job cell					-0.120	-0.129	-0.108
					(0.017)	(0.014)	(0.014)
% female in firm							-0.034
							(0.041)
R-squared	0.080	0.475	0.636	0.479	0.481	0.482	0.481
Worker and firm controls*	No	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects			Firm				
Occupational classification				2-digit	2-digit	4-digit	2-digit

* Worker's education, age and age squared; firms' employment, its square, ownership, two-digit industry, and region.

Standard errors allowing for clustering at firm level in parenteses.

Table 2: Wage Gap Decompositions

	Coefficient estimate	Mean difference women - men	Relative contribution to wage gap	Coefficient estimate	Mean difference women - men	Relative contribution to wage gap
	2002			1998		
	total log wage gap = -0.282			total log wage gap = -0.297		
Variable	(1)	(2)	(1)x(2)/(total gap)	(3)	(4)	(3)x(4)/(total gap)
Female	-0.165	1.00	0.59	-0.189	1.000	0.64
% female in occupation	-0.084	0.29	0.09	-0.104	0.328	0.12
% female in job cell	-0.108	0.57	0.22	-0.104	0.512	0.18
% female in firm	-0.034	0.27	0.03	-0.237	0.236	0.19

Note: Column (1) comes from column (7) of Table 1. The 1998 results are taken from Jurajda (2003).