QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 1, 2014, pp. 113 – 124

WAGES INEQUALITIES BETWEEN MEN AND WOMEN: EUROSTAT SES METADATA ANALYSIS APPLYING ECONOMETRIC MODELS[1]

Aleksandra Matuszewska-Janica
Department of Econometrics and Statistics, Warsaw University of Life Sciences
aleksandra_matuszewska@sggw.pl

Abstract: In this paper there are presented the results of investigation of the various factors impact on the level of male and female wages inequality. These factors are as follows: level of wages in employees group in comparison to the national average wages, the proportion of women in the group of employees, women labor market activity in the states, and variables such as the age, job seniority, level of education of the employees, type of employment contract, occupation (ISCO88), branch where the enterprises operate (NACE rev. 1.1), size of the company and collective pay agreement.

Keywords: labor market, the wage gap, the gender wage gap, SES

INTRODUCTION

Eurostat estimated that in 2012 in the EU women earned on average 16.4% less than men. This rate (GPG – Gender Pay Gap) varies e.g. among EU countries, economic sectors. We can also observe that at the GPG rate affect age, education, job seniority of employees and size of enterprise among other. The wage differences between men and women are largely explained on the basis of human capital theory (see e.g. [Haager 2000], [Polachek 2004]) and the discrimination theory (see e.g. [Becker 1971]). This phenomenon has a social dimension as well as economic importance (see e.g. discussion presented in [Klasen 1999], [Seguino 2000], [Blecker and Seguino 2002], [Löfström 2009], [Sinha et al. 2007]).

[1] Work performed within the project that has been funded by the National Science Centre, decision number DEC-2011/01/B/HS4/06346 "Wages Inequalities between Men and Women in Poland in the Framework of the European Union".

Situation of women in the labor market is subject to European Union policy. Prevention of discrimination against women has been included in Strategy for equality between women and men 2010-2015.

The aim of the study is to estimate the impact of various factors on the level of men and women wages inequality in different countries and different groups of employees. Groups of employees are characterized by one of the following features: economic branch, age, occupation, job seniority, size of enterprise, collective pay agreement, type of employment contract. There is observe that on the gender wage differences influence some other factors like: feminization of employees groups, level of wages in employees groups or women activity rate at labor market in individual countries. So such variables also are included into estimated models. For the analysis is employed Eurostat SES metadata.

DATA DESCRIPTION

Analysis is provided upon the European Union Structure of Earnings Survey (SES) data collected in 2006.[2] There are used aggregated data, that Eurostat calls Metadata. SES is a survey conducted in accordance with the Council Regulation No. 530/1999 and the Commission Regulation No. 1916/2000 as amended by Commission Regulation No. 1738/2005. The SES for 2006 is the second of a series of four yearly. The SES is a survey providing information on relationships between the level of remuneration, individual characteristics of employees and their employer (economic branch, age, occupation, job seniority, size of enterprise, collective pay agreement, type of employment contract among others). The statistics of the SES refer to the enterprises with at least 10 employees.

Data on employment and wages are encompassed in the database that contain different characteristics, as is presented in Figure 1.

To measure income inequality is often used *GPG* (Gender Pay Gap) coefficient. *GPG* represents the difference between average gross hourly earnings of male paid employees ($\overline{GHE}_M$) and of female paid employees ($\overline{GHE}_F$) as a percentage of average gross hourly earnings of male paid employees (see Fernandez-Aviles et al. 2010):

$$GPG = \left(1 - \frac{\overline{GHE}_F}{\overline{GHE}_M}\right) \cdot 100 \tag{1}$$

where:

[2] Structure of Earnings Survey has been provided every four year since 2002. There are some differences between metadata because of imported correction in every survey. For example in each survey (2002, 2006 and 2010) was different definition (SES 2006 and SES 2010) or different range (SES 2002 and SES 2006) of economic branches. So in presented analysis is used database from 2006.

$$GPG - 100 = \frac{\overline{GHE}_F}{\overline{GHE}_M} \cdot 100 = HE\_FPCM = GPC \quad (2)$$

is publicized by Eurostat.

Figure 1. Structure of SES database

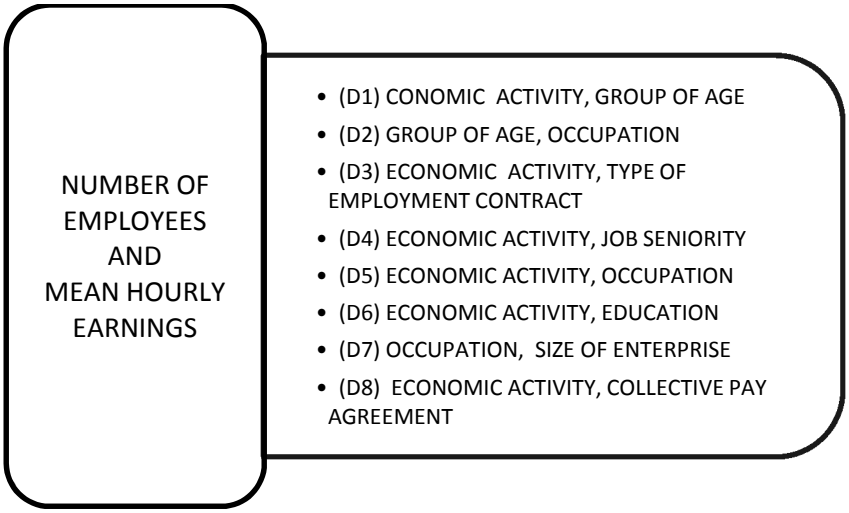
DATA ON EMPLOYMENT AND WAGES					
NUMBER OF EMPLOYEES	HOURLY EARNINGS	MONTHLY EARNINGS	ANNUAL EARNINGS	HOURS PAID	ANNUAL HOLIDAYS

CHARACTERISTICS OF EMPLOYEES/EMPLOYERS IN INDIVIDUAL DATASETS							
SEX; COUNTRY; ECONOMIC ACTIVITY; AGE	SEX; COUNTRY; ECONOMIC ACTIVITY; TYPE OF CONTRACT	SEX; COUNTRY; ECONOMIC ACTIVITY; EDUCATION	SEX; COUNTRY; ECONOMIC ACTIVITY; OCCUPATION	SEX; COUNTRY; ECONOMIC ACTIVITY; JOB SENIORITY	SEX; COUNTRY; OCCUPATION; SIZE OF ENTERPRISE	SEX; COUNTRY; OCCUPATION; AGE	SEX; COUNTRY; ECONOMIC ACTIVITY; COLL. PAY AGREEMENT

Source: own elaboration.

In the SES 2006 we can distinguish eight types of sets of aggregated data. Every data sets contained two types of information: measurable and no measurable. Measurable variable were mean hourly earnings and number of employees given for men, women and total employees. No measurable were variables as follows: sex, country and two others that were different for each data set (see Figure 2).

Figure 2. The variables that differentiate data sets



Source: own elaboration. In parentheses are given dataset names.

METHODOLOGY

In the analysis were estimated one equation econometric models:

$$\ln HE\_FPCM_{ij} = \beta_0 + \beta_1 \ln Wages_{ij} + \beta_2 \ln FEM_{ij} + \\ + \beta_3 \ln Activ_i + \sum_{k=1}^{m} \gamma_k Dummy_{kij} + \varepsilon_{ij} \quad (3)$$

HE_FPCM is share of average gross hourly earnings of female paid employees ($\overline{GHE}_F$) as a percentage of average gross hourly earnings of male paid employees ($\overline{GHE}_M$). In the paper Witkowska (2013) this rate was named as gender pay convergence ratio. *HE_FPCM* is published by Eurostat (see formula 2). Value of *HE_FPCM* equals 100 inform that between men and women wages there are no differences. When *HE_FPCM* is greater than 100 – women earn more than men on average.

In the models was used gender pay convergence coefficients calculated as:

$$\ln HE\_FPCM_{ij} = \ln \frac{\overline{GHE}_{Fij}}{\overline{GHE}_{Mij}} \quad (4)$$

where: $\overline{GHE}_{Fij}$ – average hourly female earnings of employees in *i*-th country and *j*-th group of employees; $\overline{GHE}_{Mij}$ – average hourly male earnings of employees in *i*-th country and *j*-th group of employees; in each model groups of employees refers to one of the employee's or enterprise's characteristic like economic branch, age, occupation, job seniority, size of enterprise, collective pay agreement, type of employment contract.

Variable *Wages* refers to the structure of hourly earnings in selected group of employees in each country. In the models variable *Wages* was calculated as natural logarithm of the ratio of the average wage in the *j*-th group of employees to average wages in the country:

$$\ln Wages_{ij} = \ln \frac{\overline{GHE}_{ij}}{\overline{GHE}_i} \quad (5)$$

where: $\overline{GHE}_{ij}$ – average hourly earnings of employees in *i*-th country and *j*-th group of employees; $\overline{GHE}_i$ – average hourly earnings of employees in *i*-th country. Previous analysis indicated that higher wages ratio is positive associated with gender wag gap (see [Witkowska et al. 2013]).

Feminization rate (*FEM*) is a variable that refers to gender employment structure in selected group of employees. In the models this variable was calculated as:

$$\ln FEM_{ij} = \ln \frac{EF_{ij}}{EF_{ij} + EM_{ij}} \quad (6)$$

where: EF_{ij} – number of employed women in *i*-th country and *j*-th group of employees; EM_{ij} – number of employed men in *i*-th country and *j*-th group of employees.

Activity rate $lnActiv_i$ is a natural logarithm of share of active women in the labor market in the *i*-th country in whole women population in working age 20-64 in *i*-th country. Previous analysis indicated that higher women's activity at the labor market is positive associated with gender wag gap (see [Witkowska et al. 2013]).

Each model contains dummy variable. Every of them refers to one of the employee's or enterprise's characteristic like economic branch, age, occupation, job seniority, size of enterprise, collective pay agreement, type of employment contract. It is defined as: $Dummy_{kij} = 1$ – when the variable concerns *k*-th option in *j*-th group of employees and *i*-th country, $Dummy_{kij} = 0$ – otherwise.

Options of dummy variables are presented in Table 1. In presented models is not investigated country effect.

Table 1. Dummy variables and theirs options

Dummy variable	Options
Age	Y0_29 - less than 30 years Y30_39 - between 30 and 39 years Y40_49 - between 40 and 49 years Y50_59 - between 50 and 59 years Y_GE60 - 60 years and over
Branch (economic sector)	C - Mining and quarrying D - manufacturing E - electricity, gas and water supply F - construction G - wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods H - hotels and restaurants I - transport, storage and communication J - financial intermediation K - real estate, renting and business activities L - public administration and defense; compulsory social security M - education N - health and social work O - other community, social, personal service activities
Size of Enterprise (number of employees)	10_49 - between 10 and 49 50_249 - between 250 and 499 250_499 - between 500 and 999 500_999 - between 50 and 249 gt_1000 - more than 1 000

Source: own elaboration.

Table 1. (cont.) Dummy variables and theirs options

Dummy variable	Options
Occupation	ISCO1 - Legislators, senior officials and managers ISCO2 - Professionals ISCO 3 - Technicians and associate professionals ISCO 4 - Clerks ISCO 5 - Service workers and shop and market sales workers ISCO 7 - Craft and related trades workers ISCO 8 - Plant and machine operators and assemblers ISCO 9 - Elementary occupations
Education (ISCED 1997)	ED0_1 - Pre-primary and primary education - levels 0-1 ED2 - lower secondary education – level 2 ED3_4 Upper secondary and post-secondary non-tertiary education - levels 3-4 ED5A Tertiary education - level 5A ED5B Tertiary education - level 5B ED6 Tertiary education - level 6
Collective pay agreement	NAT - A national level or interconfederal agreement IND - B industry agreement IND1 - C agreement for individual industries in individual regions ENT - D enterprise or single employer agreement UNIT - E agreement applying only to workers in the local unit OTH - F any other type of agreement NONE - N no collective agreement exists
Type of contract	INDEF - Indefinite duration FIX - Fixed term (except apprentice and trainee) APPR – apprentice or trainee
Job seniority	Y_LT1 – less than 1 year Y1_5 – between 1 and 5 years Y6_9 – between 6 and 9 years Y10_14 – between 10 and 14 years Y15_19 – between 15 and 19 years Y20_29 – between 20 and 29 years Y_GE30 – 30 years or more

Source: own elaboration based on Structure of Earnings Survey 2006: Eurostat's arrangements for implementing the Council Regulation 530/1999, the Commission Regulations 1916/2000 and 1738/2005.

RESULTS

In this section were presented eight models. In each model are included three the same (in respect of variable construction) quantitative variables: *lnWages*, *lnFEM* and *lnActiv*. Models differs in dummy variables. Characteristics of each model are presented in the Table 2.

Table 2. Characteristics of data set used for model estimation

Model No.	Dummy variable	Reference option	Data set	Number of obs.
1	SECTOR	C	(D6)	324
2	AGE	Y0_29	(D2)	134
3	OCCUPATION	ISCO9	(D2)	231
4	EDUCATION	ED0_1	(D6)	135
5	SIZE OF ENTERPRISE	10_49	(D7)	135
6	JOB SENIORITY	Y_LT1	(D4)	179
7	COLLECTIVE PAY AGREEMENT	NONE	(D8)	59
8	TYPE OF EMPLOYMENT CONTRACT	INDEF	(D3)	52

Source: own elaboration.

Results of model estimation are presented in the Tables 3, 4 and 5. In model No. 2, where dummy variable represents age effect, we can observe the highest level of adjusted R^2 (0,6141) among all estimated models. The regression results show that there are negative associations between convergence rate (*lnHE_FPCM*) and wage level (*lnWages*), convergence rate (*lnHE_FPCM*) and women activity at the labor market (*lnActiv*). Convergence rate (*lnHE_FPCM*) is also negative associated with feminization rate (*lnFEM*, at the significance level $\alpha = 0.1$). The "age effect" is visible only for the eldest group of employees (only for variable Y_GE60 parameter is significant). The wages difference between men and women that are at least 60 years old is significant higher than wages difference between men and women under 30. In models based on the less aggregated data (see [Witkowska et al. 2013], models number: 3, 4, 7, 10, 14) we can observed that all dummy variable that represented age are significant.

In model number 3 dummy variables designate several groups of occupations. Variable *lnWages* has not significant influence on explained variable (*lnHE_FPCM*). But two other: feminization rate (*lnFEM*) and women activity at the labor market (*lnActiv*) are negative associated with gender pay convergence ratio (*lnHE_FPCM*) For this data set we can observe occupation effect (some dummy variable are significant). The wages difference between men and women with elementary occupations (ISCO9) is significant higher than men's and women's wages difference for employees working as legislators, senior officials and managers (ISCO1), technicians and associate professionals (ISCO3), craft and related trades workers (ISCO7) and plant and machine operators and assemblers (ISCO8).

Table 3. Parameters of estimated models (1)

Model No. 2 Dummy variable: AGE			Model No. 3 Dummy variable: OCCUPATION			Model No. 4 Dummy variable: EDUCATION		
variable	coefficient		variable	coefficient		Variable	coefficient	
const	-0,3159	***	const	-0,2388	***	const	-0,2701	***
lnWages	-0,3065	***	lnWages	0,0727		lnWages	-0,0080	
lnFEM	-0,0927	*	lnFEM	-0,0521	***	lnFEM	0,0374	*
lnActiv	-0,2403	***	lnActiv	-0,1987	***	lnActiv	-0,1919	***
Y30_39	-0,0227		ISCO1	-0,2450	***	ED2	0,0053	
Y40_49	-0,0361		ISCO2	-0,0765		ED3_4	0,0109	
Y50_59	-0,0326		ISCO3	-0,0838	**	ED5A	-0,0367	
Y_GE60	-0,0873	***	ISCO4	0,0131		ED5B	-0,0143	
			ISCO5	0,0009		ED6	0,0254	
			ISCO6	0,0110				
			ISCO7	-0,2151	***			
			ISCO8	-0,1493	***			
R^2 adj.	0,6141		R^2 adj.	0,3880		R^2 adj.	0,4533	
F	31,24	***	F	14,25	***	F	14,89	***

Source: own calculation. *** denotes significance level $\alpha = 0.10$, ** $\alpha = 0.05$ and * $\alpha = 0.1$

In model number 4 dummy variables represents several education groups. In this model variable *lnWages* also has not significant influence on dependent variable (*lnHE_FPCM*). Women's activity rate (*lnActiv*) has significant negative impact on the gender pay convergence ratio (*lnHE_FPCM*). Increasing feminization rate (*lnFEM*) causes increase of *lnHE_FPCM*. In examined data set differences between wages of men and women with the different education level are similar. All dummy variable are not significant. This result is opposite to obtained by [Witkowska et al. 2013], see models number: 1, 9, 12, 16). In these models we observed significant differences between gender wag gaps for employees with elementary and highest level of education.

Job seniority was investigated in model number 6. There we can observe that in groups of employees with greater job seniority gender pay convergence ratio is smaller than for employees shorter than one year. The same results was obtained for less aggregated data (see [Witkowska et al. 2013], models number: 2, 8, 11, 15). In this model variable *lnWages* was not significant, but we observed significant negative influence on gender wages convergence other two variable: *lnFEM* and *lnActiv*.

In models No. 1 and 5 were take onto account such attributes of enterprise like economic branch and size of enterprise. In bigger companies gender wage inequalities are greater than in smaller enterprises (see results obtained for model number 5). Also differences between wages of men and women varies among branches. Gender wage gap is significant smaller (gender wage convergence ratio is higher) in construction (F), hotels and restaurants (H), transport, storage and

communication (I), and education (M) than in mining and quarrying (C). Similar results were obtained for less aggregated data (see [Witkowska et al. 2013]. Research presented in Oi and Idson (1999) indicated that there are significant differences in average salaries according to economic branch and size of enterprise. In bigger enterprises average remuneration are higher. Higher salaries could make for greater gender wage differences (see results presented in [Witkowska et al. 2013]). In model 1 variable *lnWages* was not significant. Other two quantitative variable: *lnFEM* and *lnActiv* have significant negative influence on gender wages convergence. In model 5 only *lnActiv* was significant and it was also negative associated with explained variable *lnHE_FPCM*.

Table 4. Parameters of estimated models (2)

Model No. 6 Dummy variable: JOB SENIORITY			Model No. 1 Dummy variable: BRANCH			Model No. 5 Dummy variable: SIZE OF ENTERPRISE		
variable	coefficient		variable	coefficient		variable	coefficient	
const	-0,301	***	const	-0,370	***	const	-0,265	***
lnWages	0,081		lnWages	-0,043		lnWages	-0,055	
lnFEM	-0,093	**	lnFEM	-0,083	***	lnFEM	-0,039	
lnActiv	-0,308	***	lnActiv	-0,084	**	lnActiv	-0,186	**
Y1_5	-0,069	***	D	-0,031		S50_249	-0,024	
Y6_9	-0,110	***	E	0,058		S250_499	-0,081	***
Y10_14	-0,109	***	F	0,117	***	S500_999	-0,095	***
Y15_19	-0,097	***	G	-0,014		GE1000	-0,075	**
Y20_29	-0,098	***	H	0,105	*			
Y_GE30	-0,088	**	I	0,103	**			
			J	-0,056				
			K	0,045				
			M	0,142	**			
			N	0,079				
			O	0,074				
R^2 adj.	0,3784		R^2 adj.	0,3788		R^2 adj.	0,2189	
F	13,04	***	F	15,07	***	F	6,96	***

Source: own calculation. *** denotes significance level $\alpha = 0.10$, ** $\alpha = 0.05$ and * $\alpha = 0.1$

In model number 7, where was investigated influence of collective pay agreement on gender wage differences, non e dummy variable was significant. We do not observe significant mean of collective pay agreement for gender wage differences at this level of aggregation data. In this model we observe significant negative influence of women activity (*lnActiv*) on the gender wage convergence and positive influence of feminization rate (*lnFEM*) on explained variable *lnHE_FPCM*. In model number 8 only women activity (*lnActiv*) was significant and negative associated with gender wage convergence ratio (*lnHE_FPCM*).

Table 5. Parameters of estimated models (3)

Model No. 7 Dummy variable: COLLECTIVE PAY AGREEMENT		Model No. 8 Dummy variable: TYPE OF EMPLOYMENT CONTACT	
variable	coefficient	variable	coefficient
const	-0,2725 ***	const	-0,3126 ***
lnWages	-0,0400	lnWages	-0,2029
lnFEM	0,1034 **	lnFEM	0,0137
lnActiv	-0,3516 ***	lnActiv	-0,3150 ***
ENT	-0,0181	APPR	0,1102
IND	0,0118	FIX	0,0585
IND1	-0,0069		
NAT	0,0759		
OTH	-0,0583		
UNIT	-0,0458		
R^2 adj.	0,1193	R^2 adj.	0,5066
F	1,87 *	F	11,47 ***

Source: own calculation. *** denotes significance level $\alpha = 0.10$, ** $\alpha = 0.05$ and * $\alpha = 0.1$

SUMMARY

In the states with low women's labor market activity (e.g. Malta, Italy) we can observe smaller wages differences between men and women. On the other hand, in the states with high rate of women's labor market activity gender pay gap is much bigger. Obtained results confirm this. In each model we observe that only one variable has significant impact on gender wage convergence. It is *lnActiv* - women activity at the labor market. This variable is negative associated with explained variable in every case. So we can conclude that higher participation of women in the labor market is connected to greater difference in men and women wages. Women tends to concentrate in low pay jobs, co

Statistical analysis of SES data provided information that in groups of employees with wages that are larger than the average in the state, male and female wage differences are also larger. It was the reason to introduce *lnWages* variable. Wages level (*lnWages*) is significant variable only in model number 2, where dummy variable refers to age. So we can conclude that wages level is not so strong connected to gender wage convergence at this level of aggregation data. For less aggregated observation the level of remuneration in more visible (see models presented in [Witkowska et al. 2013]).

The problem of feminization of occupations is wide discussed in the literature (see e.g. [Anker 1998], [England et al. 2007], [Perales 2010]). Women tends to concentrate to lower paid jobs. So we can suppose that in high feminized jobs male and female wage differences would be higher, and in the opposite

situation - *GPG* would be lower. Feminization rate applied in analyzed models gives different results. In two models number 4 (education) and number 7 (collective pay agreement) is positive associated with gender pay convergence coefficient. In two other models (no. 5 (size of enterprise) and no. 8 (type of employment contract)) this variable has none significant impact on the explained variable *lnHE_FPCM*. In four models (models number 1, 2, 3, 6) feminization rate has negative impact on *lnHE_FPCM*. So we can conclude that feminization of labor market has both negative and positive effects on gender wage differences.

Effects that are represented by dummy variable (economic branch, age, occupation, job seniority, size of enterprise, collective pay agreement, type of employment contract) in not so strong for analyzed data like for less aggregated data. In presented model we can observe significant differences in wage inequalities especially in branches, different size enterprises, for groups of employees with different job seniority, occupation and age. There are not detected differences between groups of employees with different level of education, type of employment contract or collective pay agreement.

Next step of the study will be analysis with the use of low aggregated and more detailed data.

REFERENCES

Anker R. (1998) Gender and Jobs: Sex Segregation of Occupations in the World, International Labor Office.

Becker G.S. (1971) The Economics of Discrimination.

Blecker R., Seguino S. (2002) Macroeconomic Effects of Reducing Gender Wage Inequality in an Export-Oriented, Semi-industrialized Economy, Review of Development Economics, Wiley Blackwell, Vol. 6(1), 103-119.

England P., Allison P.D., Wu Y. (2007) Does bad pay cause occupations to feminize, Does feminization reduce pay, and How can we tell with longitudinal data?, Social Science Research, Vol. 36, 1237-1256.

Fernández-Avilés G., Montero J.M., Witkowska D. (2010) Gender Wage Gap in EU States. Application of Taxonomic and Spatial Methods, [in:] Witkowska D., Nermend K. (ed.), Regional Analysis: Globalization, Integration, Transformation, Wydawnictwo Uniwersytetu Szczecińskiego, Szczecin, 71-95.

Haager D.M. (2000) How do Investments in Human Capital Differentially Affect Gender Income? An Analysis of the Gender Wage Gap, The Park Place Economist, vol. VIII.

Klasen S. (1999) Does Gender Inequality Reduce Growth and Development? Evidence from Cross-Country Regressions, World Bank Working Paper Series, No. 7.

Löfström Å. (2009) Gender equality, economic growth and employment, working paper.

Morrison A., Raju D. Sinha N. (2007) Gender Equality, Poverty and Economic Growth, Policy World Bank Policy Research Working Paper, No. 4349.

Oi W.Y., Idson T.L. (1999) Firm Size and Wages in. Handbook of Income Distribution, vol. 3B, ed. Ashenfelter O., Card D., Elsevier, Amsterdam, 2165-2214.

Perales F. (2010) Occupational Feminization, Specialized Human Capital and Wages: Evidence from the British Labour Market, ISER Working Paper Series: 2010-31.

Polachek S.W. (2004) How the Human Capital Model Explains Why the Gender Wage Gap Narrowed, IZA Bonn Discussion Paper, No. 1102.

Seguino S. (2000) Gender Inequality and Economic Growth: A Cross-Country Analysis, World Development, Elsevier, vol. 28(7), 1211-1230.

Witkowska D. (2013) Gender Disparities in the Labor Market in the EU, International Atlantic Economic Research, Vol. 19(4), 331-254.

Witkowska D., Kompa K., Matuszewska-Janica A. (2013) Gender Earning Differences in the European Union Member States, European Journal of Management, Vol. 13(4), 129-148.