



UNIVERSITY OF WARSAW

Faculty of Economic Sciences

WORKING PAPERS

No. 21/2014 (138)

KAROLINA GORAUS
JOANNA SIWIŃSKA
JOANNA TYROWICZ
LUCAS VAN DER VELDE

LANGUAGE AND (THE ESTIMATES OF) THE GENDER WAGE GAP

WARSAW 2014



Language and (the Estimates of) the Gender Wage Gap

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

JOANNA SIWIŃSKA
Faculty of Economic Sciences
University of Warsaw
e-mail: siwinska@wne.uw.edu.pl

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

LUCAS VAN DER VELDE
Faculty of Economic Sciences
University of Warsaw
e-mail: lvelde@wne.uw.edu.pl

Abstract

In this paper we link the estimates of the gender wage gap with the gender sensitivity of the language spoken in a given country. We find that nations with more gender neutral languages tend to be characterized by lower estimates of GWG. The results are robust to a number of sensitivity checks.

Keywords:

gender wage gap, language neutrality, meta-analysis

JEL:

C29, J31, J71, Z10

Acknowledgements

Authors express gratitude to Stephen Jarrell, Tom Stanley, Doris Weichselbaumer and Rudolph Winter-Ebdmer for providing their collection of estimates. We thank Wojciech Hardy for collaboration in collecting data. The paper was improved thanks to the comments from Marcin Bielecki, Gianna Giannelli, Saul Estrin, Tomasz Mickiewicz and members of Meta-Analysis of Economics Research Network as well as during seminars at GRAPE.

This research was supported by a grant from the National Science Centre, UMO-2012/05/E/HS4/01510. All opinions expressed are those of the authors and have not been endorsed by NSC or NBP. All remaining errors are ours.

Working Papers contain preliminary research results.

Please consider this when citing the paper.

Please contact the authors to give comments or to obtain revised version.

Any mistakes and the views expressed herein are solely those of the authors.

1. Introduction

”Men imagine that their minds have the command of language, but it often happens that language bears rule over their minds.” - Francis Bacon

Multiple studies seek to identify the sources of cross-country and cross-time variation in the estimates of the gender wage gap (GWG).¹ In one of the most influential analysis, Weichselbaumer and Winter-Ebmer (2007) (henceforth WWE) attempt to link the estimates of GWG to the institutional features of the product and labor markets, arguing that countries with a higher degree of economic freedom have a lower GWG residual than others. In addition, they show that ratification of international conventions supporting equal treatment of men and women tends to be associated with lower subsequent estimates of GWG: ”more competition in an economy means more opportunities for females and subsequently less discrimination”² (p. 273). In a spirit similar to Blau and Kahn (2003), WWE argue that the international conventions triggering change in national policies are the mechanism behind their findings.

While WWE do not directly measure these policies, implicitly they argue that these policies are somehow effective at gradually reducing the estimates of the unexplained GWG component. Similarly, Mandel and Shalev (2009) are optimistic with reference to the role played by the states: it is heterogeneous depending on a social class, but ”welfare states affect the relative pay of men and women” (p. 1900). Also Daly et al. (2006) argue that institutions are important in explaining the relative size of the gender wage gap across countries.

On the other hand, Jurajda (2005) strongly opposes this view, showing that anti-discriminatory policies have little immediate effect on the estimates of GWG, which raises doubt to what extent the time-related correlates offered by WWE are actually indicative. Furthermore, Oostendorp (2009) finds that the openness channel for reducing GWG estimates seems effective only in the case of advanced economies, but not in the case of developing countries. Also, most of the estimates of the GWG used in WWE study have no or weak control over selection into employment. As forcefully argued by Olivetti and Petrongolo (2008), not unequal pay but rather unequal employment stays behind the differences in the estimates of the GWG residuals.

Clearly, the literature lacks consent on the role of institutional features. Some studies are tracing the sources of GWG decrease in policies. Other studies emphasize the role of informal institutions in persistence or even reinforcement of unequal pay between men and women. It is our objective in this paper to shed some more light on this controversy. We test if gender awareness induced by the semantics of a language has explanatory power in the variation of GWG estimates. The idea behind our approach is simple: languages are (usually) exogenous to countries, whereas they form gender awareness already at early stage of childhood. We extend the WWE database until year 2010, and using the original data as well as the extended one, we show that

¹See Blau and Kahn (1996); Stanley and Jarrell (1998); Blau and Kahn (2003); Jarrell and Stanley (2004); Mandel and Semyonov (2005); Weichselbaumer and Winter-Ebmer (2005).

²While more egalitarian societies are likely to have more equal wages, WWE offer an IV approach to identify the actual causal effect.

the language gender intensity can help to explain a variation in the estimates of the unexplained GWG.

2. A role for a language

The idea that language influences the way that people think and act dates back to 1950s and the Sapir-Whorf hypothesis (Whorf, 1956). This view holds that "language can make certain sorts of thought and cognitive process more likely, and more accessible to people" (Carruthers, 2012). Although currently in psychology the rigid hypothesis of linguistic determinism is gradually being dismissed, it is rarely questioned that the language does indeed influence some human behavior (Boroditsky, 2001; Lai and Boroditsky, 2013). The debate as well as the theoretical foundations are laid out in Cameron (1998).

Indeed, language influence on human behavior has already been pursued in economic research. Chen (2013) demonstrates that language influences agents' subjective perception of certain events and affects their choices: the way that languages encode time affects the agents' future oriented behavior, like saving decisions, smoking, diet, etc.

An alternative way in which language differences were introduced into economic thinking is its characterization as transaction costs, e.g. Lang (1986); Melitz (2008). A common example of this view is present in trade models, where language differences represent proxies for the costs of doing business between countries, see Melitz (2008). From a different perspective, Lang (1986) developed a model where the transaction costs imposed by the existence of different languages leads to a segregation within the society. This model has more than one similitude with models of statistical discrimination, where beliefs of productivity led to the same pattern. Moreover, in Lang (1986) when there is a need to cooperate, the transaction costs (learning a new language) fall upon the minority.

Among the language characteristics that can potentially influence thought and behavior, are gender-intensities of different languages. The forms of asymmetric treatment of genders in languages range from grammatical forms, lack of words for female/male professional positions to idiomatic expressions that glorify one gender and deplete the position of another. Combined they amount to gender language bias or language gender intensity.

A number of papers empirically test the proposition that linguistic differences shape women's involvement in different spheres of both public and private environment, including economic outcomes. Bem and Bem (1973) showed that gender-biased job advertisements discouraged both genders from applying for the "opposite-sex" jobs. Mavisakalyan (2011) provides empirical evidence on the strong link between gender systems in languages and women's participation rate in the labor market: gender-intensive languages tend to be related to substantially lower participation rate (both at the country and individual level). Santacreu-Vasut et al. (2014) argue that gender distinctions in grammar affect women's participation in corporate senior management. The effects transcend the sphere of economic activity: Santacreu-Vasut et al. (2013) show that the pervasiveness of gender distinction in language is an important determinant of female quota in politics

3. Method and data

This paper replicates WWE work with two main exceptions. First, we extend the set of explanatory variables to include data on languages. These indicators are described in the data section. Second, in addition to the original WWE dataset, we provide estimates also for the extended dataset. WWE data cover publications prior to 2005. We expand the list of papers included in the study with those written and published between 2005-2010, which roughly doubles the size of the sample.³

3.1. Estimates of GWG

Our primary source of estimates for this meta analysis is the dataset developed by WWE 2007. We also carefully followed their procedure to extend this dataset with papers published between 2005 and 2010. We adopted the same conventions with respect to the language of publication (English) and the search engine (EconLit). We used the same keywords: “(wage* or salar* or earning*) and (discrimination or differen*) and (sex or gender)”. To test if this search was not excessively narrow, we erased one keyword at a time from the first parenthesis, subsequently erasing the logical connectors (“or” “and”).⁴ When more than one version of the paper was available, we selected the published (or the most recent version) version. Finally, we also excluded some not-comparable estimates of the adjusted wage gap (e.g. non-parametric estimates).

This selection resulted in the addition of 1369 new estimates of the adjusted wage gap, which were taken from 117 new studies.⁵ These estimates come from 56 different countries, but the United States, Great Britain and Canada are the countries with the largest number of estimates, in total 46%. This country composition is not surprising, as the search was restricted to papers written in English.

3.2. Language indicator

A comprehensive data source of language grammatical structures is the World Atlas of Linguistic Structures. We employ the gender dummy dataset, based on this source, used also by Santacreu-Vasut et al. (2013, 2014). The set provided by Santacreu-Vasut et al. (2013) is more comprehensive than WALS, which is its additional advantage. The database measures the level of language gender intensity along several axes. The first and most general one is the presence of gender identification at all. Second, the gender system in the language may or may not be based on sexes.

To put these criteria in a perspective, Asian languages have no system of gender identification in the language. English, for example, has gender system but only visible in third person pronouns. Spanish has both noun assignment and pronouns distinction in first person already. Also, there are languages which have genders in a system, but these are not sex-based: e.g. Danish or Swedish in Europe, or Fulu and Nyamwezi-Sukuma in Africa.

³Doris Wichselbaum and Rudold Winter-Ebmer were kind enough to share one version of their data set.

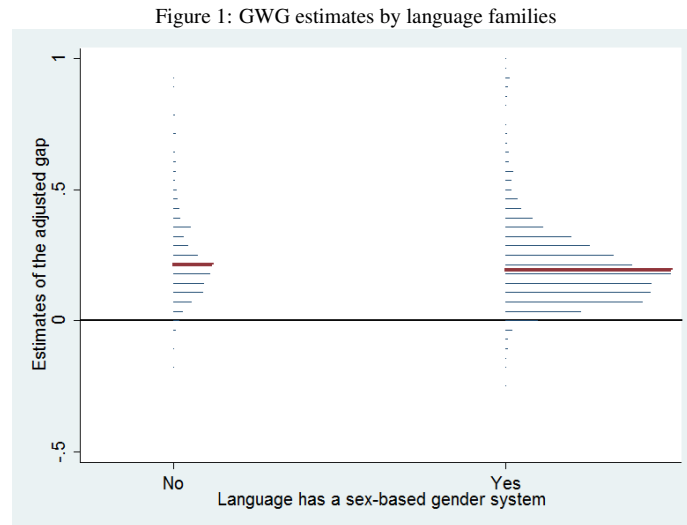
⁴We tried other combinations, but they did not increase the number of results.

⁵The complete list is available upon request.

Of the two indicators, we use the most specific one, i.e. sex-based gender system in the language. There are also other indicators of the gender intensity in the languages. For example, some languages have a gender distinction in pronouns (singular, plural or both), some in verbs, some also in adjectives. Typically these characteristics of languages are coded through the use of dummy variables. One way to measure the intensity of gender distinction in a language would be to add up these dummies. However, WALS has a substantial number of missing data for these categories, mostly because WALS is a collector of studies on languages (and their conclusions) rather than a database of languages characteristics. Thus, we follow the presence of a sex-based gender system as our language gender intensity indicator ⁶.

4. Results

Before we move to a formal meta-regression analysis, we present an intuitive visualisation of how language characteristics may be related to the estimates of GWG, by plotting the estimates against our main language indicator, see Figure 1. Clearly, for languages with and without a sex-based gender system there is a large dispersion in estimates.



Moving to the formal testing, in in column (1) of Table 1 we attempt to replicate WWE estimates on the enlarged sample, which corresponds to column 2 in Table 2 in Weichselbaumer and Winter-Ebmer (2007). Column 2 reports the same estimates but

⁶Notwithstanding the results presented in the next section were found robust to the selection of the gender intensity variable.

for the new sample only⁷ The ratification of ILO conventions, a proxy for countries' commitment to gender equality, appears significant and negative. The results fail to show the same significance for the index of economic freedom and prohibitions to women's work, which may arise from the fact that the estimates from years 2005-2010 substantially more often come from decomposition (rather than an estimate of a gender dummy) and more often control for selection bias, where the two institutional variables are likely to show more power (employment access rather than compensations).

In general, countries for which the main language has a sex-based gender are characterized by higher estimates of the gender wage gap. In column (3) and (5) of Table 1 we show the results once we control for the sex-based gender system in the language. Positive and statistically significant coefficient seems to suggest that the role of language is non-negligible. Estimates of column (3) are estimate-based, whereas our language data are unique for a respective country (and so are the institutional variables used by WWE, but there is time variation as well). Thus, we collapse the estimates for our preferred specifications including language indicators in column (5), so that standard errors were mechanically reduced by the repeated estimates within a language characteristic.⁸ This reduces the sample size substantially, but also increases the point estimate of the sex-based gender system, *reducing* its standard error.

⁷The difference between the used sample and the most comprehensive data set stems from excluding chapters in books (in total: 226 observations). In addition, fertility data are not available for many countries prior to 1970, some countries are also missing the values for the Economic Freedom Index (in total: 166 observations).

⁸In general meta-analyses of GWG cannot fully account for the quality of estimates. In the case of coefficients, usually one would use the standard errors, but in the case of decompositions it is not that simple. Most papers only report the total and the adjusted gap, with only a small fraction of them including e.g. bootstrapped standard errors. By using one observation per paper-country-year we can calculate its accuracy (i.e. the inverse of the standard deviation of the estimates for the same paper-country-year) as weights in the meta-regressions.

Table 1: Results

Variables	(1)	(2)	(3)	(4)	(5)
Language gender intensity			0.035 (0.026)		0.067*** (0.021)
Control variables included in WWE					
GDP pc. (in thou. USD PPP)	-0.007 (0.009)	-0.025 (0.020)	-0.014 (0.009)	-0.012 (0.009)	-0.025** (0.010)
Fertility rate	0.004 (0.011)	0.058*** (0.020)	0.002 (0.012)	0.004 (0.008)	-0.000 (0.008)
Participation rate	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	0.001 (0.001)
Institutional variables included in WWE					
Index of economic freedom	-0.004 (0.009)	0.021 (0.031)	-0.005 (0.009)	-0.006 (0.009)	-0.008 (0.009)
Prohibitions on female work	0.009 (0.012)	0.006 (0.036)	0.004 (0.012)	-0.011 (0.014)	-0.022 (0.014)
Ratifications ILO	-0.026** (0.012)	-0.035* (0.019)	-0.026** (0.012)	0.001 (0.009)	0.001 (0.009)
Constant	0.371*** (0.112)	-0.001 (0.295)	0.349*** (0.116)	0.380*** (0.130)	0.362*** (0.129)
Meta-regression controls (following WWE)					
Sample characteristics	Yes	Yes	Yes	No	No
Methods	Yes	Yes	Yes	No	No
Estimation	Yes	Yes	Yes	No	No
Additional controls					
Regions	Yes	Yes	Yes	Yes	Yes
Years	Yes	Yes	Yes	Yes	Yes
Observations	1,983	896	1,981	509	507
R-squared	0.288	0.381	0.289	0.058	0.098

Notes: OLS in columns (1)-(3), random effects meta regression in columns (4) and (5). Asterisks denote the conventional significance levels. Standard errors between parentheses. Column (4) replicates column (1) for the reduced sample. Column (5) replicates column (3).

We use a meta-regression, which produces an alternative R^2 measure, calculated as $1 - \tau^2/\tau_0^2$, where the numerator is the value of the between studies unexplained variation with all covariates and τ_0^2 the value for the model without any covariates. The coefficients for the control variables included in the estimation available in the Appendix.

In order to ascertain that the results are not an statistical artifact, we conduct several robustness checks, which concern the specification, the error term and the selection of the right statistic from the papers as well as inclusion of other control variables. We thus test the robustness of this estimate by choosing different characteristics of the estimates distribution as well as different set of control variables. The results of this exercise are reported in Table 2. In columns (1) to (3) we test on the collapsed sample whether inclusions of regional dummies, religion controls, and estimation method (regular regression instead of a meta-regression) could affect the estimate on gender language intensity. In columns (4) to (7) rather than using a mean estimate from each paper we use the median, the largest estimate, the smallest estimate and even the estimate with the largest number of covariates. These tests prove that our results are not susceptible to these various checks.

Table 2: Robustness checks: Alternative specifications

Variables	Mean			Median (4)	Min (5)	Max (6)	Max prec. (7)
	(1)	(2)	(3)				
Language gender intensity	0.058*** (0.020)	0.059* (0.033)	0.060*** (0.020)	0.068*** (0.016)	0.056*** (0.019)	0.073*** (0.021)	0.044*** (0.012)
Institutions	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	No	No	No	Yes	Yes	Yes	Yes
Methods	No	No	No	Yes	Yes	Yes	Yes
Estimation	No	No	No	Yes	Yes	Yes	Yes
Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Religion	No	Yes	No	No	No	No	No
Clustered errors	No	No	Yes	No	No	No	No
Observations	506	487	385	653	507	509	1,164
Adjusted gap	.197	.198		.189	.144	.257	.194
R-squared	.056	.039	0.325	.625	.361	.545	.530

Notes: Asterisks denote the conventional significance levels. Standard errors between parentheses. In models (3) and (4) we employed OLS regressions, with robust standard errors clustered in country/year basis. All estimates besides from those in models (3) and (4) were obtained using meta-regression estimations, which means that the R^2 statistics were calculated as in Table 1.

Model (1) has time trend instead of time dummies, model (2) has religion controls, model (3) has no institutional controls and is estimated as unweighted OLS, model (4) is a weighted OLS, where inverse standard deviation of estimates within a paper constitute weights, model (5) has median instead of mean estimate on LHS, model (6) has minimum estimate rather than mean estimate on LHS, model (7) has maximum estimate rather than mean estimate on LHS and model (8) uses the estimate with the largest number of covariates. The complete estimates of variables included but not reported available upon request.

It seems then that sex-based gender distinctions in the language can be translated into unexplained gender differences in pay of a magnitude close to as much as 6 p.p. Our result complements those presented in Santacreu-Vasut et al. (2014) for participation rate and prevalence of women in executive positions.

5. Conclusions

In recent years, literature on the gender wage gap has begun to exploit cross-country variation in the search for the institutional determinants of the differences. In this article, we contribute by bringing one important element to the equation: the language spoken. We hypothesized that in countries where language has a more marked distinction between genders, differences in labor market outcomes will be larger. We tested this hypothesis using meta-regression tools, which allowed us to profit from the large body of literature on the adjusted wage gap already existent. The results seem to confirm the hypothesis.

These results insinuate that indeed GWG may be driven by some deep societal features stemming from such basic social codes as language. This suggests that if reducing GWG was a policy objective, education on gender equality is needed already at early stages of education, when language characteristics are absorbed by children and translated into societal norms.

References

- Bem, S. L., Bem, D. J., 1973. Does sex-biased job advertising “aid and abet” sex discrimination? *Journal of Applied Social Psychology* 3 (1), 6–18.
- Blau, F. D., Kahn, L. M., 1996. Wage structure and gender earnings differentials: An international comparison. *Economica* 63 (250), S29–S62.
- Blau, F. D., Kahn, L. M., 2003. Understanding international differences in the gender pay gap. *Journal of Labor Economics* 21 (1), 106–144.
- Boroditsky, L., 2001. Does language shape thought?: Mandarin and english speakers’ conceptions of time. *Cognitive Psychology* 43 (1), 1–22.
- Cameron, D., 1998. *The Feminist Critique of Language: A Reader*. Psychology Press.
- Carruthers, P., 2012. Language and cognition. In: Margolis E., Samuels R., S. S. (Ed.), *Oxford Handbook of Philosophy of Cognitive Science*. Oxford University Press, Oxford.
- Chen, M. K., 2013. The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. *The American Economic Review* 103 (2), 690–731.
- Daly, A., Meng, X., Kawaguchi, A., Mumford, K., 2006. The gender wage gap in four countries. *The Economic Record* 82 (257), 165–176.
- Jarrell, S. B., Stanley, T. D., 2004. Declining bias and gender wage discrimination? a meta-regression analysis. *The Journal of Human Resources* 39 (3), 828–838.
- Jurajda, S., 2005. Gender segregation and wage gap: An east-west comparison. *Journal of the European Economic Association* 3 (2-3), 598–607.
- Lai, V. T., Boroditsky, L., 2013. The immediate and chronic influence of spatio-temporal metaphors on the mental representations of time in english, mandarin, and mandarin-english speakers. *Cultural Psychology* 4.
- Lang, K., 1986. A language theory of discrimination. *The Quarterly Journal of Economics*, 363–382.
- Mandel, H., Semyonov, M., 2005. Family policies, wage structures, and gender gaps: Sources of earnings inequality in 20 countries. *American Sociological Review* 70 (6), 949–967.
- Mandel, H., Shalev, M., 2009. How welfare states shape the gender pay gap: A theoretical and comparative analysis. *Social Forces* 87 (4), 1873–1911.
- Mavisakalyan, A., 2011. Gender in language and gender in employment. ANUCBE School of.
- Melitz, J., 2008. Language and foreign trade. *European Economic Review* 52 (4), 667–699.

- Olivetti, C., Petrongolo, B., 2008. Unequal pay or unequal employment? a cross-country analysis of gender gaps. *Journal of Labor Economics* 26 (4), 621–654.
- Oostendorp, R. H., 2009. Globalization and the gender wage gap. *The World Bank Economic Review* 23 (1), 141–161.
- Santacreu-Vasut, E., Shenkar, O., Shoham, A., 2014. Linguistic gender marking and its international business ramifications. *Journal of International Business Studies*.
- Santacreu-Vasut, E., Shoham, A., Gay, V., 2013. Do female/male distinctions in language matter? evidence from gender political quotas. *Applied Economics Letters* 20 (5), 495–498.
- Stanley, T., Jarrell, S. B., 1998. Gender wage discrimination bias? a meta-regression analysis. *The Journal of Human Resources* 33 (4), 947–973.
- Weichselbaumer, D., Winter-Ebmer, R., 2005. A meta-analysis of the international gender wage gap. *Journal of Economic Surveys* 19 (3), 479–511.
- Weichselbaumer, D., Winter-Ebmer, R., 2007. The effects of competition and equal treatment laws on gender wage differentials. *Economic Policy* 22 (50), 235–287.

Appendices

Table A1: The estimates from Table 1

Variables	(1)	(2)	(3)
Theme	-0.005 (0.020)	0.054 (0.047)	-0.019 (0.025)
Source	0.024 (0.028)	0.051 (0.045)	0.023 (0.046)
Authors' gender	0.000 (0.018)	0.041 (0.050)	-0.008 (0.023)
Sample characteristics			
New entrants only	-0.055* (0.028)	-0.016 (0.052)	-0.051* (0.028)
Full time only	0.009 (0.014)	-0.001 (0.023)	0.006 (0.018)
Only private	-0.000 (0.015)	-0.024 (0.030)	-0.022 (0.020)
Only public	-0.035* (0.020)	-0.054 (0.046)	-0.052** (0.026)
Only narrowly define occ.	-0.018 (0.022)	0.050 (0.056)	-0.022 (0.019)
Low skill occ.	0.002 (0.039)	-0.075 (0.110)	-0.005 (0.035)
Medium skill occ.	-0.019 (0.030)	-0.010 (0.123)	-0.031 (0.028)
High skill occ.	-0.088*** (0.019)	-0.098 (0.062)	-0.111*** (0.020)
Single only	-0.053* (0.030)	-0.031 (0.204)	-0.063** (0.032)
Married only	-0.011 (0.046)	-0.237* (0.129)	-0.019 (0.055)
Only minorities	0.029 (0.040)	0.031 (0.055)	0.034 (0.033)
Majority only	-0.029 (0.070)	-0.032 (0.093)	-0.039 (0.061)
Method employed			
O-B male component	0.042*** (0.016)	0.122** (0.056)	0.036** (0.016)
Dummy if neum	0.035** (0.015)	0.028 (0.033)	0.032* (0.017)
O-B Reimers	-0.061** (0.030)	-0.034 (0.041)	-0.077** (0.030)
O-B Cotton	0.042** (0.019)	0.062* (0.032)	0.025* (0.014)
Brown decomp.	0.063*** (0.019)		0.047** (0.019)
Dummy variable	0.044** (0.019)	0.071** (0.033)	0.047** (0.021)
Instr. variable	-0.045 (0.031)	0.066 (0.071)	-0.048* (0.027)
Heckman correction	-0.019 (0.016)	-0.018 (0.038)	-0.025 (0.017)
Panel methods	-0.026 (0.024)	-0.014 (0.039)	-0.021 (0.025)

Table A1: The estimates from Table 1

Variables	(1)	(2)	(3)
	Variables (not) included in the estimation		
Not hourly wages	0.015 (0.020)	-0.029 (0.049)	0.018 (0.020)
Constructed hourly w.	-0.016 (0.021)	-0.115 (0.079)	-0.047* (0.027)
Gross wages	0.019 (0.032)	0.113 (0.078)	0.032 (0.043)
Potential experience	0.014 (0.019)	-0.055 (0.062)	0.025 (0.021)
Experience not incl.	0.016 (0.023)	-0.005 (0.055)	0.024 (0.027)
Minority status	-0.000 (0.018)	0.036 (0.036)	0.013 (0.020)
Marital status	0.004 (0.015)	-0.032 (0.034)	-0.015 (0.015)
Children	0.018 (0.015)	0.021 (0.039)	0.009 (0.014)
Child*Married	-0.024 (0.028)	0.015 (0.076)	0.020 (0.029)
Training	0.025 (0.020)	0.022 (0.033)	-0.000 (0.027)
Tenure	0.040*** (0.013)	-0.010 (0.027)	0.041*** (0.014)
Occupation	0.012 (0.012)	0.051* (0.028)	0.027** (0.012)
Industry	0.026* (0.013)	0.053* (0.030)	0.024 (0.015)
Public /private	0.038*** (0.012)	0.035 (0.026)	0.016 (0.014)
Trade union	-0.065** (0.028)	-0.132** (0.065)	-0.040 (0.026)
% Females occ.	0.076*** (0.019)	0.063 (0.038)	0.058*** (0.018)
Full/part time	-0.004 (0.016)	0.016 (0.040)	0.000 (0.018)
Wages not known	-0.041 (0.028)	0.065 (0.073)	-0.002 (0.034)
	Institutional environment		
GDP per cap.	-0.007 (0.009)	-0.025 (0.020)	-0.012 (0.039)
Fertility rate	0.004 (0.011)	0.058*** (0.020)	-0.094* (0.051)
Participation rate	-0.001 (0.001)	-0.001 (0.001)	0.004 (0.003)
Index of economic freedom	-0.004 (0.009)	0.021 (0.031)	-0.006 (0.020)
Limits on women's work	0.009 (0.012)	0.006 (0.036)	-0.004 (0.036)
Sum of ratifications (ILO)	-0.026** (0.012)	-0.035* (0.019)	-0.010 (0.025)
Language has sex-based gender system			-0.343* (0.177)
Constant	0.371*** (0.112)	-0.001 (0.295)	0.479* (0.267)

Table A1: The estimates from Table 1

Variables	(1)	(2)	(3)
Observations	1,983	896	1,981
R-squared	0.288	0.381	0.376

Notes: Asterisks denote the conventional significance levels. Standard errors between parentheses. Estimates in (1) follow the specification by WWE, but for the total sample. In (2) we present the same specification, but restricted to our papers. In (3) we added the language variable to the full sample.



FACULTY OF ECONOMIC SCIENCES
UNIVERSITY OF WARSAW
44/50 DŁUGA ST.
00-241 WARSAW
WWW.WNE.UW.EDU.PL