Gender Wage Gap and the Role of Skills and Tasks: Evidence from the PIAAC Data Set

Michael Christl and Monika Köppl-Turyna

Agenda Austria | Wien

Workshop Arbeitsmarktökonomie 2017

Introduction I

"It is noteworthy that, according to Goldin (2014), the gender wage gap is especially prominent within narrowly defined occupations rather than across the whole range of occupations."

"Autor and Handel (2013) show that tasks in the US vary substantially within certain occupations in the context of gender and race."

Agenda Austria

The goal of the paper I

- We analyze the GWG along the wage distribution in Austria.
- First attempt to include variables such as the skills and tasks of workers, but also the over- or under-qualification as well as the flexibility of work into account for otherwise not included characteristics.
- Control for selection of individuals into workforce.

Literature overview - Austria I

- Grünberger et al. (2009) (Mincer wage regression method (LS)) show that for employees working full time, women earn on average 22% less than men. Controlling for observable characteristics brings the figure down to 12%. They additionally show that more of the GWG can be explained in the lower part of the wage distribution.

- Böheim et al. (2013) (using a similar approach to Machado-Mata decompositon) shows that women earn about 14 % less than men controlling for certain characteristics. They additionally show that the unexplained part of the GWG increases along the wage distribution.

- Grandner and Gstach (2015) (Machado-Mata decomposition on quantile regressions) show that for Austria, the unexplained gender wage gap is about 20% across all income groups. The difference in observable characteristics have almost no explanatory power.

Agenda Austria

Data I

- International Assessment of Adult Competencies (PIAAC) survey conducted by the OECD in 2011/12
- It encompasses 4,810 individual observations, including detailed information about education, skills, income and family background. After filtering out observations with missing data, our sample reduces to about 2,200 observations.
- Dependent variable: Gross hourly wages (with and without bonus payments)
- The greatest advantage of the new dataset, is the possibility of controlling for various skills required and used at work.

Agenda Austria

Data II

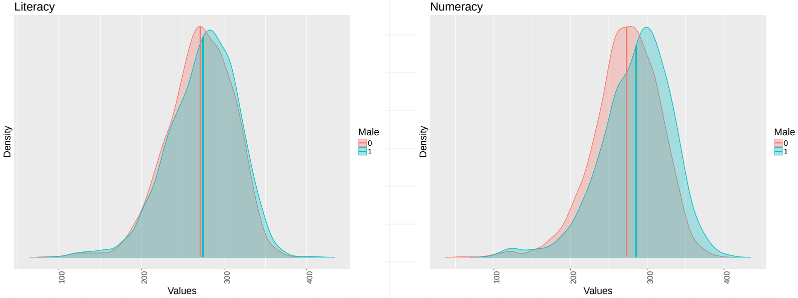
Table : Overview of the independent variables

Type	Variable	Categories
Personal characteristics	Education level	ISCED code
	Relevant job experience	years
	Children	number of
	Age of children	categories
	Out of workforce due to childcare	dummy
	Citizenship	country dummy
	Migration background	migrant (1st or 2nd gen.)
Job characteristics	Employment status	dummy blue white-collar, pub. emp...
	Firm size	1 to 10, 11 to 50, 51 to 250, 251 to 1000
	Participation in on-the-job training	dummy
	Hours worked per week	hours
	Economic sector	NACE (2-digit) sector
	Type of activity	ISCO (2-digit) activity classification
	Work flexibility	none, little, moderate, high, very high

Agenda Austria

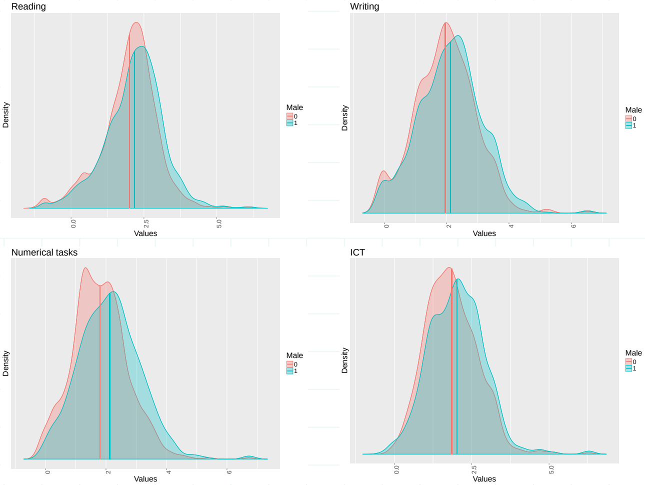
Data III

Figure : Cognitive skills: gender differences in Austria



Data IV

Figure : Work tasks: gender differences in Austria



The Model I

- We estimate the gender wage gap using counterfactual distributions approach by Chernozhukov et al. (2013)
- Let $j = 0, 1$ denote the subpopulation of men ($j = 0$) and the subpopulation of women ($j = 1$).
- Y_j denotes the wages, while X_j is the vector of job-relevant characteristics affecting the wages. Conditional distribution functions $F_{Y_0|X_0}(y|x)$ and $F_{Y_1|X_1}(y|x)$ describe assignment of wages y to individuals with characteristics x.
- If $F_{Y<0|0>}$ and $F_{Y<1|1>}$ are respectively the observed distributions for men and women, we can denote:

$$F_{Y<0|1>}(y) \equiv \int_{\delta_1} F_{Y_0|X_0}(y|x) dF_{X_1}(x), \quad (1)$$

- The decomposition can then be denoted as:

$$F_{Y<1|1>} - F_{Y<0|0>} = [F_{Y<1|1>} - F_{Y<0|1>}] + [F_{Y<0|1>} - F_{Y<0|0>}], \quad (2)$$

Sample selection

- Sample selection is an issue. WHY?
- To correct for sample selection we apply a three-step procedure, following Buchinsky (1998) and Buchinsky (2002)
  - Selection equation using semi-parametric least squares by Klein and Spady (1993) and record the generated single index. Details
  - We regress the outcome variable on the characteristics and a polynomial of the single index series.
  - Once we have obtained consistent estimates of the coefficients and decompose the quantile functions in the manner described above.

Results of the selection equation

Agenda Austria

Results - Wage regression I

- Differencens between gender in returns to skills in the upper part of the wage distribution (but not high)
- Severe differences in returns to tasks between gender:
  - Higher premium for planing, writing and reading tasks for women
  - Higher premium for influencing, numerical and ICT-tasks for men

Results - Gender Wage Gap I

- We use three different specifications:
  - Benchmark model: Similar set of control variables as Böheim et al. (2013), without additional control for individual skills.
  - Skill Task model: We additionally consider skills and tasks
  - Skill Task model including sample-selection: We add skills and tasks and control for sample selection

Agenda Austria

Results - Gender Wage Gap II

Table : Overview of the unexplained differences (in %) - hourly wages

Decil	GWG	unexplained GWG (Male-based)				
		Benchmark	Skills + Tasks	S+T effect (in pp)	Sample selection	Selection effect (in pp)
1	11,47	8,45	1,02	-7,43	1,77	0,75
2	13,89	10,61	3,71	-6,9	3,91	0,2
3	15,99	12,24	5,77	-6,47	6,42	0,65
4	16,62	13,23	8,00	-5,23	8,43	0,43
5	17,48	14,15	9,31	-4,84	10,47	1,16
6	19,24	14,70	10,38	-4,32	11,93	1,55
7	20,62	15,78	11,52	-4,26	12,86	1,34
8	23,31	16,34	13,06	-3,28	14,31	1,25
9	26,55	18,99	15,32	-3,67	16,97	1,65

Agenda Austria

Conclusion I

- We control for skills, tasks, as well as for skill-match.
  - the unexplained part of the gender wage gap decreases by about three to seven percentage points across the entire wage distribution.
- In the lower part of the wage distribution, the unexplained gender wage gap is less than 5% but it increases up to a maximum of about 15%.
- Controlling for sample selection increases the unexplained GWG approximately by 0.5 to 1.5 percentage point.
- We found that we can explain much more of the GWG we thought before. Especially with regard to low-wage earners, we can explain the wage gap almost completely.

References I

Autor, D. H. and M. J. Handel (2013). Putting tasks to the test: Human capital, job tasks, and wages. Journal of Labor Economics 31(S1), S59–S96.

Böheim, R., K. Himpele, H. Mahringer, and C. Zulehner (2013). The distribution of the gender wage gap in Austria: evidence from matched employer-employee data and tax records. Journal for Labour Market Research 46(1), 19–34.

Buchinsky, M. (1998). Recent advances in quantile regression models: a practical guideline for empirical research. Journal of human resources, 88–126.

Buchinsky, M. (2002). Quantile regression with sample selection: Estimating women's return to education in the US. In Economic Applications of Quantile Regression, pp. 87–113. Springer.

Chernozhukov, V., I. Fernández-Val, and B. Melly (2013). Inference on counterfactual distributions. Econometrica 81(6), 2205–2268.

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

Grandner, T. and D. Gstach (2015). Decomposing wage discrimination in Germany and Austria with counterfactual densities. Empirica 42(1), 49–76.

Grünberger, K., C. Zulehner, et al. (2009). Gender-specific wage gaps in Austria. WIFO Monatsberichte (monthly reports) 82(2), 139–150.

Klein, R. W. and R. H. Spady (1993). An efficient semiparametric estimator for binary response models. Econometrica: Journal of the Econometric Society, 387–421.

Agenda Austria

Sample selection in Austria

Austria faces, among comparable European countries, a fairly low labor force participation among women:

- Traditional division of tasks in the family, i.e., females are expected to take over the major part of household and childcare obligations;
- Comparative under-provision of public childcare institutions, in particular for children of age three and less;
- Generosity of the social system and long maternity leaves (of up to three years).

In effect, women are more likely to stay out of the labor force, in particular if they have children, as the opportunity cost of working might be too high.
We therefore include factors affecting the reservation wage, as well as childcare obligations, in the selection equation.

Back to main

Agenda Austria

Sample selection equation

Wage depends linearly on a set of labor market characteristics:

$$y_i^* = \beta_0 \mathsf{X}_{i2} + u_i, \tag{3}$$

whereas y is observed only if it exceeds the reservation wage y^R given by

$$y_i^R = \alpha_0 \mathsf{X}_{i1} + v_i, \tag{4}$$

where (dropping the i index) $\mathsf{X}_2 \subset \mathsf{X}_1$. The observed wage can be written as

$$y = d \cdot y^* = d \cdot (\beta_\theta \mathsf{X}_2 + u_\theta), \tag{5}$$

where $d = I(y^* > y^R)$. The conditional quantile of the observed wage is given by

$$Quant_\theta(y|\mathsf{X}_2) = \beta_\theta \mathsf{X}_2 + h_\theta(\mathsf{X}_1, \gamma_0). \tag{6}$$

For identification, it is necessary to include in equation 4, variables that determine the reservation wage but do not enter the wage equation. In our case, the instruments excluded in the first stage are: age, age squared, dummy for whether a person is living with a spouse/partner and the net replacement rates obtained from the demographic characteristics of the household.

Back to main

Table : Selection equation, dependent variable: Working or not working; probit

	Male	Female
Replacement rate	-0.03***	-0.01***
	(-2.71)	(-3.19)
Age	0.18***	0.09***
	(8.03)	(4.70)
Age^2	-0.00***	-0.00***
	(-9.76)	(-5.61)
Living with partner	0.53***	0.26***
	(4.15)	(2.61)
ISCED 2	-2.18***	-1.34***
	(-6.51)	(-4.36)
ISCED 3	-1.85***	-1.11***
	(-4.94)	(-3.37)
ISCED 4	-1.99***	-0.80**
	(-5.19)	(-2.36)
ISCED 5+6	-1.74***	-0.64*
	(-4.50)	(-1.84)
Migrant	-0.22**	-0.56***
	(-2.01)	(-6.08)
Children	normalized to one	
Observations	1857	1870
χ^2 excl. instr.	174.80	81.69

For comparison, probit model without constant and coefficient for Children normalized to one.
Significance * 0.1, ** 0.05, *** 0.01
t-Statistics in brackets.

Back to main