J Econ Inequal (2016) 14:21–40
DOI 10.1007/s10888-016-9320-z

A distribution-sensitive examination of the gender wage gap in Germany

Ekaterina Selezneva[1] · **Philippe Van Kerm**[2]

Received: 12 March 2014 / Accepted: 6 January 2016 / Published online: 23 January 2016

Abstract This paper provides a new examination of the gender pay gap for Germany based on a family of distribution-sensitive indicators. Wage distributions for men and women do not only differ by a fixed constant; differences are more complex. We show that focusing on the bottom of the wage distribution reveals a larger gender gap. Our distribution-sensitive analysis can also be used to study whether the statistical disadvantage of women in average pay might be 'offset' by lower inequality. Over a broad range of plausible preferences over inequality, we show however that 'inequality-adjusted' estimates of the gap can be up to three times higher than standard inequality-neutral measures in Eastern Germany and up to fifty percent higher in Western Germany. Using preference parameters elicited from a hypothetical risky investment question in our sample, inequality-adjusted gender gap measures turn out to be close to those upper bounds.

Keywords Gender gap · Wage differentials · Wage inequality · Expected utility · Risk aversion · East and West Germany · SOEP · Singh-Maddala distribution · Copula-based selection model

1 Introduction

The gender gap in pay in Germany is one of the largest in Europe. According to recent IAB InfoPlattform briefing, the average gross hourly earnings of women is 22 % lower

✉ Philippe Van Kerm
philippe.vankerm@liser.lu

Ekaterina Selezneva
eka.selez@gmail.com

1 Institut for East and Southeast European Studies, Landshuter Str. 4, 93047 Regensburg, Germany

2 Luxembourg Institute of Socio-Economic Research, 11, Porte des Sciences L-4366, Esch-sur-Alzette, Luxembourg

than men, for an EU average of 16 %.[1] Eurostat's 2011 estimate of Germany's (unadjusted) gender pay gap is third only to Estonia and Austria among 26 European countries.[2] Factors contributing to the gap are generally sought in career breaks, part-time employment and relative concentration of women in low skill and low pay occupations (Al-Farhan 2010a; Antonczyk et al. 2010; Heinze 2010). The size of the gender gap however differs markedly in Western and Eastern Germany: women face a much smaller penalty relative to men in Eastern Germany than in Western Germany (see, e.g., Hunt 2002; Smolny and Kirbach 2011; Kohn and Antonczyk 2013). Smolny and Kirbach (2011) observe that the gender wage gap is one of the few statistics for which there is no convergence to Western levels in the period 1990–2008.

At the same time, Germany recently experienced an increase in overall wage inequality; see Dustmann et al. (2009), Fuchs-Schündeln et al. (2010), and Card et al. (2013). According to Al-Farhan's (2010b) analysis of the German Socio-Economic Panel (SOEP) data, the wage distribution in Germany appeared to stabilize at historically high levels of inequality in the recent ten years, while Antonczyk et al. (2010, 2011) still find increasing wage inequality between 2001 and 2006 using Structure of Earnings Survey data. Observing high levels of wage inequality with a large gender gap in pay is consistent with the demonstration by Blau and Kahn (1992, 1996, 1997) that wage inequality is positively associated to the gender pay gap. As a matter of fact, Al-Farhan (2010a) shows that recent trends in the gender pay gap and in wage inequality in Germany are driven by common underlying factors such as changes in workers' potential experience, occupational positions and firm sizes.

In this context of high gender pay gap statistics and historically high levels of wage inequality, we undertake a new evaluation of the German gender gap in pay. We provide an original 'distribution-sensitive' examination that explicitly accounts for gender differences in wage inequality in addition to differences in mean wage across gender. Comprehensive assessment of wage differentials should be based on comparisons of wage distributions of men and women and not just on differences in average wages. Or, more precisely, assessment should be based on comparisons of the wage distributions of women with the wage distributions that would be observed if women were rewarded as men with identical attributes and productive capacity. As advocated by Dolton and Makepeace (1985), such comparisons should ideally be based on explicit utility functionals defined over those distributions in order to capture all distributional differences between men and women. Evidence of 'glass ceiling' factors which slow women's professional progress and make it difficult for women to reach the very top of the wage distribution emphasizes differences at the high end of the distribution. Alternatively, a 'sticky floors' argument that women are disproportionately maintained in low paid jobs suggests that male-female distribution difference are bigger at the bottom. In sum, as argued by Dolton and Makepeace (1985), it is difficult to claim that distribution differences by gender can be fully summarized by, say, a fixed additive or multiplicative constant. This is more than a mere academic distinction: whether gender gaps are driven by 'sticky floors' or 'glass ceilings' factors calls for different sorts of policy responses.

[1] See http://infosys.iab.de/infoplattform/dokSelect.asp?pkyDokSelect=71 (accessed August 6, 2013).

[2] See http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Gender_pay_gap_statistics (accessed August 6, 2013).

In spite of this, standard indicators of the wage gap generally neglect detailed distribution differences: a typical measure of the gender gap gives "the cents a woman makes on average for every dollar an observationally equivalent man makes on average." Such an indicator controls for gender differences in human capital (and possibly job characteristics) and compares *observationally equivalent* men and women (unlike the raw figures mentioned in the opening sentences of this paper). However it remains limited in that it focuses on comparisons of mean wages of men and women.

Using SOEP data on a sample of individuals aged 25–55 over the period 1999–2008, we calculate new indices of wage differentials for Germany that explicitly incorporate broader distributional concern. We follow the simple strategy proposed in Van Kerm (2013) which involves calculating measures based on well-known 'equally distributed equivalent' wages (or generalised means) as proposed by Atkinson (1970). Those wage gap measures are parameterised by a sensitivity parameter which allows analysts to shift focus on different parts of the wage distribution. For example, one can estimate 'bottom sensitive' indicators that are mostly driven by differences at the lower end of the wage distributions of men and women. Alternatively it is possible to calculate 'top sensitive' indicators which are driven by differences at the upper end. The standard indicator based on comparisons of mean wages is only one special case thereof. Of course, fixing a particular parameter guiding the degree of top- or bottom-sensitivity of the new 'distribution sensitive' index is no less arbitrary than the standard measures focused on mean wages. To address this concern, we calculate indices over a broad range of parameter values – from extreme top sensitivity to extreme bottom sensitivity – and focus discussion on the *range of variation* of the resulting indices and show how misguided can possibly be the standard wage gap indicators.

To preview our results, we find that standard, 'neutral' measures are (close to) the lower bounds of the range of variation of distribution-sensitive measures. This means that focusing on (conditional) mean wage differences identifies a lower bound for the wage gap. It potentially underestimates the magnitude of the wage gap by up to three-fold in Eastern Germany and up to fifty percent in Western Germany. The much larger impact of accounting for wage distribution differences by gender in Eastern Germany challenges current evidence that the gender gap is smaller in the East: taking male-female inequality differences into account very much reduces the contrast between Eastern and Western Germany.

Additionally, we take advantage of a special module on risk attitudes collected in the 2004 wave of the survey (Dohmen et al. 2005, 2011) to estimate individual-level coefficients of relative risk aversion from the sample itself. To the extent that individual risk aversion estimates can approximate preferences over wage inequality, these values can be plugged in our wage gap measure to provide an indicator that reflects preferences elicited from the respondents themselves. We find that those measures turn out to be close to the upper bounds in our range of distribution sensitive indicators.

We estimate the measures with and without correction for endogenous labour market participation using a copula-based selection model (Smith 2003). Accounting for endogenous selection does not change our main results however.

The paper proceeds as follows. Section 2 describes the distribution-sensitive measure of wage differentials and estimation methods. Section 3 provides information on the sample and variables of interest and describes construction of the individual measures of risk aversion from answers to the SOEP hypothetical risky investment question. Results are presented in Section 4. Section 5 concludes.

2 Methods

2.1 Distribution-sensitive measures of wage differentials

A typical indicator of the wage gap gives the "cents a woman makes on average for every dollar an observationally equivalent man makes on average" (see, e.g., Jenkins 1994):

$$\Delta_1 = \exp\left[\int_{\Xi} \left[\log\left(\mu_x^w\right) - \log\left(\mu_x^m\right)\right] h^w(x)dx\right] \quad (1)$$

where (i) μ_x^w is the average wage of a woman with characteristics x (e.g., education, experience, etc.), μ_x^m is the average wage of a man with the same characteristics x and (ii) $h^w(x)$ is the (multivariate) probability density function of characteristics x among women. Δ_1 is usually estimated from wage regressions of the form $\log(w_i) = x_i\beta^s + u_i$ $(i = 1, \ldots, N^s)$ for $s \in \{m, w\}$, and plugging regression coefficients to obtain the ubiquitous Oaxaca-Blinder measure of wage differences

$$\Delta_1^{\text{OB}} = \exp\left[\int_{\Xi} x\left(\beta^w - \beta^m\right) h^w(x)dx\right]\theta \quad (2)$$

$$= \exp\left[\bar{x}^w\left(\beta^w - \beta^m\right)\right]\theta \quad (3)$$

where $\bar{x}^w$ is the vector of average women characteristics (see, e.g., Fortin et al. 2011).[3,4]

The anatomy of Eq. 1 reveals a double averaging: first by considering the differences in average wage of men and women conditionally on characteristics $x(\mu_x^w$ and $\mu_x^m)$ and, second, by averaging these differences over the characteristics of women ($h^w(x)$).

This approach, albeit empirically convenient, has been criticised for putting narrow focus on mean differences and discarding normatively relevant and empirically important distributional differences. Most notably, Jenkins (1994) and del Río et al. (2011) claim that wage discrimination measures should be sensitive to variations among women in the size of the wage disadvantage experienced; that is, they criticise the second averaging over $h^w(x)$ inherent to Δ_1. They argue, for example, that a situation in which all women experience a wage disadvantage of 10 cents per dollar should be evaluated differently from a situation in which half of women face a disadvantage of 20 cents and half of women face no disadvantage. Jenkins (1994) and del Río et al. (2011) develop methods to address this concern. Del Río et al. (2011) further discuss how individual-level disadvantage can be captured by quantile regression.

Van Kerm (2013) is instead concerned with the first averaging in Eq. 1 and proposes to use a straightforward generalisation of Δ_1 where the conditional means are substituted by parameterised power means, also known as 'equally distributed equivalent' wage standards (see, e.g., Foster and Székely 2008):

$$\Delta_2(\epsilon) = \exp\left[\int_{\Xi} \left[\log\left(C(F_x^w;\epsilon)\right) - \log\left(C(F_x^m;\epsilon)\right)\right] h^w(x)dx\right] \quad (4)$$

[3] θ reflects differences in residual variance in the wage regressions for men and women (see Blackburn 2007, Van Kerm 2013).

[4] Women are taken as group of interest and men as reference group here. The measure captures the disadvantage of women relative to men, what Jenkins (1994) calls 'discrimination against women'. Other choices are obviously available: as Fortin et al. (2011) argue, this question does not have any unambiguous econometric solution.

where $C(F_x^s; \epsilon)$ is a (conditional) power mean of order $(1 - \epsilon)$

$$C(F; \epsilon) = \left(\int_0^\infty y^{1-\epsilon} dF(y) \right)^{\frac{1}{1-\epsilon}}$$

for $\epsilon \neq 1$ and $C(F; 1) = \exp\left[\int_0^\infty \ln(y) dF(y)\right]$. The basic idea is to substitute one particular 'wage standard' representative of the conditional distribution of wages for women (or men) with characteristics x – the conditional mean μ_x^s – by a family of alternative 'wage standards' parameterised by ϵ, $C(F_x^s; \epsilon)$. Varying ϵ allows shifting the relative importance given to low or high wages in the representative wage standard: low (negative) ϵ gives greater weight to the top of the wage distribution whereas increasing ϵ shifts weight to low wages. At the limit $\epsilon \longrightarrow -\infty$ (resp. $\epsilon \longrightarrow \infty$), the representative wage standard is the highest (resp. smallest) possible wage obtained by workers of characteristics x. The pivotal case $\epsilon = 0$ leads to the standard arithmetic mean as a special case.

The normative significance of adopting $\Delta_2(\epsilon)$ stems from the fact that $C(F_x^s; \epsilon)$ is also the certainty equivalent (or equally distributed equivalent) for the outcome described by the distribution F_x^s under constant relative risk aversion von Neumann-Morgenstern expected utility over the wage distribution (Atkinson 1970). Adopting $C(F_x^s; \epsilon)$ instead of μ_x^s can therefore also be interpreted as accounting for gender differences in wage inequality in the assessment of the wage gap. It is easy to see that adopting $\epsilon > 0$ leads to $C(F; \epsilon) \leq C(F; 0)$: dispersion in the wage distribution is penalized. The 'ϵ-sensitive wage standard' representative of unequally distributed wages is smaller than the arithmetic mean and the more unequally distributed is the distribution, the lower is the standard $C(F; \epsilon)$ relative to $C(F; 0)$. On the contrary, adopting $\epsilon < 0$ leads to $C(F; \epsilon) \geq C(F; 0)$: dispersion in the wage distribution is rewarded and the 'ϵ-sensitive wage standard' representative of unequally distributed wages gets larger, the more unequally distributed are wages. So this simple framework makes it straightforward to incorporate inequality adjustments in a statistical assessment of the gender wage gap. When women with characteristics x face wage distributions F_x^w with greater inequality than observationally equivalent men F_x^m, this will tend to inflate the wage gap index $\Delta_2(\epsilon)$ for any positive ϵ. In reverse, lower inequality among women distributions may compensate for lower wages and thereby mitigate standard estimates of the wage gap. If, on the contrary, one considers that greater inequality should be positively rewarded (that is, assume preference for risk or inequality) and therefore choose a negative ϵ, inequality may mitigate the gender wage gap if there is more inequality in the wage distributions of women than of observationally equivalent men.

Inequality comparisons between men and women are made here over *conditional* wage distributions F_x^w and F_x^m. A related branch of the literature focuses on comparisons of quantiles of the *unconditional* wage distributions of men and women; see, among others, Juhn et al. (1993), Machado and Mata (2005), Melly (2005), and Firpo et al. (2009) on methods and Antonczyk et al. (2010), Heinze (2010) or Al-Farhan (2010a) for recent applications to the gender wage gap in Germany. This approach also departs from a narrow focus on mean wages and, by comparing quantiles of the wage distributions, flexibly identifies differences in pay at, say, the bottom or the top of the wage distributions. The focus is however on understanding and decomposing differences in the *unconditional* wage distribution of men and women (disentangling composition effects from wage structure effects).[5] Our aim instead is to assess the impact and relevance of introducing normative considerations of inequality in aggregate assessments of the gender wage gap and we base this on comparisons of *conditional* wage distributions as is the classic Oaxaca-Blinder indicators.

So, Eq. 4 defines a class of 'ϵ-sensitive' wage gap measures indexed by a parameter of inequality aversion. To avoid merely substituting one (possibly arbitrary) value judgment by another one, we will report estimates for ϵ in the range $[-4, 10]$ covering a wide range of positions from inequality aversion – or 'bottom sensitivity' – to inequality preference – or 'top sensitivity' – . Variations in the statistic for different ϵ will inform us upon the degree to which overall (conditional) wage distributions differ between women and observationally equivalent men in our sample and therefore how much misguided can be assessments based solely on Δ_1.

In principle, a distinct ϵ could also be specified for each women in the sample, or for each configuration of characteristics, to reflect heterogeneity in individual-level attitudes towards inequality or risk (Van Kerm 2013). This is only a straightforward extension of Eq. 4 as

$$\Delta_2(\tilde{\epsilon}) = \exp\left[\int_{\Xi}\int_{\epsilon}\left[\log\left(C(F_x^w; e)\right) - \log\left(C(F_x^m; e)\right)\right] g_x^w(e) h^w(x) de\, dx\right] \tag{5}$$

where g_x^w is the density distribution of individual-level values for the parameter ϵ (over the domain ϵ of possible risk-aversion parameters) among women of characteristics x. More intuitively re-expressed in discrete notation, this leads to

$$\Delta_2(\tilde{\epsilon}) = \exp\left[\frac{1}{N^w}\sum_{i=1}^{N^w}\left[\log\left(C(F_{x_i}^w; e_i)\right) - \log\left(C(F_{x_i}^m; e_i)\right)\right]\right]. \tag{6}$$

If individual parameters e_i adequately reflect individual preferences over inequality, the measure resulting from such calculations would incorporate population preferences directly into the evaluation of the wage gap. It would not require the analyst to set one particular ϵ or to look at ranges of variation of the index. Individual preferences over inequality are however difficult to quantify. The SOEP dataset contains individual-level measures of risk attitudes in a special module in the 2004 wave of the survey. Responses to the small set of questions asked were validated in an incentive compatible field experiment with representative subjects (Dohmen et al. 2011). To benchmark our estimation results with constant ϵ for all women, we will tentatively exploit individual-level parameter estimates of risk aversion that can be derived from this module under the assumption of constant relative risk aversion utilities; see Section 3.2 for details. Under the assumption that risk aversion parameters so derived can also reflect individual positions regarding wage inequality, we will estimate $\Delta_2(\tilde{\epsilon})$ with heterogeneous, individual-level e_i parameters. The degree to which the gap estimate will differ from indices with constant parameters will depend both on the size of the estimated risk-aversion parameters and on the association between individual-level risk aversion parameters and differences in wage inequality across gender.

2.2 Estimation

Calculation of $\Delta_2(\epsilon)$ requires estimates of conditional wage distributions F_x^m and F_x^w for all x observed in the sample. Several alternative estimators can be chosen from, such as quantile

[5] The unconditional wage distribution for women is $F^w(y) = \int F_x^w(y)h_w(x)dx$ and counterfactual unconditional distribution that would be observed if women were paid as men is $F^*(y) = \int F_x^m(y)h_w(x)dx$ which can be inverted to consider unconditional quantiles; see, e.g., Albrecht et al. (2003), Millimet and Wang (2006), Arulampalam et al. (2007), and Christofides et al. (2013) for applications to the gender wage gap.

regression, non-parametric kernel estimation, 'distribution regression'; see, e.g., Hall et al. (1999) for non-parametric approaches or Rothe (2010) and Chernozhukov et al. (2013) for recent general discussions. We follow Biewen and Jenkins (2005) and Van Kerm (2013) and adopt a fully parametric approach. Wages are specified as Singh-Maddala distributed conditionally on covariates, with the three parameters of the distribution allowed to vary log-linearly with covariates:

$$F_x^s(y) = 1 - \left[1 + \left(\frac{y}{b^s(x)}\right)^{a^s(x)}\right]^{-q^s(x)} \tag{7}$$

where $b^s(x) = \exp(x\theta_b^s)$ is a scale parameter, $a^s(x) = \exp(x\theta_a^s)$ is a shape parameter modifying both tails and $q^s(x) = \exp(x\theta_q^s)$ is a shape parameter modifying the upper tail (Singh and Maddala 1976). Power means for the Singh-Maddala distribution have convenient closed-form expressions. This makes estimation of $\Delta_2(\epsilon)$ easy from coefficient estimates of model (7):

$$C(\hat{F}_x^s;\epsilon) = \hat{b}^s(x)\left(\frac{\Gamma\left(1+(1-\epsilon)/\hat{a}^s(x)\right)\Gamma\left(\hat{q}^s(x)-(1-\epsilon)/\hat{a}^s(x)\right)}{\Gamma\left(\hat{q}^s(x)\right)}\right)^{\frac{1}{1-\epsilon}} \tag{8}$$

where $\Gamma(\cdot)$ is the Gamma function and $(\hat{\theta}_a^s, \hat{\theta}_b^s, \hat{\theta}_q^s)$ are (say, maximum likelihood) estimates of the Singh-Maddala parameters (Kleiber and Kotz 2003). Such a specification, albeit parametric, is flexible and allows for broad variations in degrees of skewness and kurtosis of the wage distributions. It can also deal with the heavy tail typical of earnings data.

A prime advantage of the parametric approach over quantile regression or non-parametric procedures is that it can be easily adapted to handle endogenous labour market participation.[6] As is well-known, endogenous selection may lead to an over-estimation of the wage distributions when estimated only from observed wages. A differential effect of sample selection between women and men might appear, as the former group is more likely to include workers with lower earnings potential and/or higher reservation wages. See Hunt (2002) for a discussion of this in the context of East Germany. Van Kerm (2013) shows how a Singh-Maddala distribution for wages can be combined with the standard Heckman-selection-type normality assumption (Heckman 1979) to correct for endogenous participation using a copula-based selection model (Smith 2003).[7] Say z denotes participation for a given agent with wage y observed if $z = 1$. Let z^* be a latent propensity to participate in the labour market with $z = 1$ if $z^* > 0$ and $z = 0$ otherwise. Assuming the pair (y, z^*) is jointly distributed with cumulative distribution H_x and expressing H_x using its copula, the marginal Singh-Maddala distribution for y and an assumed distribution for the latent z^* (denoted G_x) leads to

$$H_x^s(y, z^*) = \Psi^s(F_x^s(y), G_x^s(z^*)) \tag{9}$$

As in standard selectivity-corrected regression models, G_x^s can be assumed normal with mean $x\delta^s$ and unit variance. Following Van Kerm (2013), we take Ψ^s to be a Clayton copula:

$$\Psi(u, v; \theta_C^s) = \left(u^{-\theta_C^s} + v^{-\theta_C^s} - 1\right)^{-1/\theta_C^s} \tag{10}$$

[6] Huber and Melly (2011) discuss conditions for consistency of quantile regression with sample selection.

[7] See Pigini (2014) for a recent survey of selection models with copula-based dependence specifications and semi-parametric approaches. See Picchio and Mussida (2011) for an alternative strategy.

where θ_C^s is an association parameter to be estimated. Joint estimation of θ_a^s, θ_b^s, θ_q^s, δ^s and θ_C^s, for example via maximum likelihood, leads to estimates of the Singh-Maddala distribution appropriately corrected for endogenous labour market participation. See Christofides et al. (2013) for a recent application of this model.

For inference on estimates of $\Delta_2(\epsilon)$, we assess sampling variability on the basis of bootstrap resampling. We implement the repeated half-sample bootstrap algorithm of Saigo et al. (2001) with a resampling that takes into account the repetition of individuals in the pooled longitudinal dataset as well as the sampling dependence between observations implied by the survey design (stratification and clustering).

3 Data

3.1 Sample definition and labour market variables

The German Socio-Economic Panel (SOEP) is a nationally representative longitudinal survey on living conditions in Germany collected annually by the Deutsches Institut fur Wirtschaftsforschung (DIW, Berlin). Multiple topics are covered and include, e.g., income and employment, housing, health, educational achievements (see Wagner et al. 1993, 2007). SOEP data are available yearly since 1984 (since 1990 for East Germany). For our analysis we pool ten waves of data covering the period 1999–2008 in order to achieve sample sizes that allow accurate estimation of the conditional wage distributions required for calculations of the distribution-sensitive indices. We cover a period in which inequality has been documented to be historically high in Germany, but do not pool waves of data collected after the onset of the Great Recession to avoid mixing data potentially driven by different labour market forces. Using those rounds of SOEP data also allows us to exploit special modules on risk attitudes collected in 2004 and on wealth and asset holdings collected in 2002 and 2007 and match those to individuals observed at work in any of the waves 1999–2008.

To limit the influence on wages of transitions into and out of the labour market, we restrict our analysis to respondents aged between 25 and 55. We consider the wages of individuals working in the private or public sector at least 15 hours per week in a regular (full-time or part-time) job. This excludes workers on vocational training, in sheltered workshops, or reporting marginal or irregular part-time employment. Self-employed workers are excluded as their wage rate is ill-defined. We exclude workers in the agricultural sector. Individuals with hourly wage below €4 are considered as out of regular employment. Cross-section sample weights are used throughout the analysis to correct for unequal sampling probabilities in the SOEP sampling design.

Gross hourly wage in the main job is calculated as gross earnings received in the month preceding interview divided by 4.32 and then by average weekly hours worked. Gross monthly earnings exclude additional payments such as holiday money or back-pay, while include money earned for overtime. Hours of work include information on average hours worked normally during a week, including overtime. We inflate all wages to 2008 prices using the national consumer price index.[8] To prevent outlying data from driving our estimates, abnormally large wages are excluded: observations with hourly wages above €70 are discarded. We also recode as missing any hourly wage from a person reported as normally working more than 65 hours per week.

[8]Data on CPI are taken from the website of the Germany Federal Statistical Office, http://www.destatis.de.

Our final samples consist of 24,029 male observations and 17,330 female observations for Western Germany (with mean real hourly wages of €17.01 and €13.35, respectively), and of 7,047 male observations and 6,831 female observations for Eastern Germany (with mean wages of €11.69 and €10.81 euros, respectively). Descriptive statistics on wages percentiles and means by gender and region along with sample sizes are presented in Table 1. Note the regional contrast in differences between men and women in wages and in the share of the sample observed in our subsample of (salaried) workers with non-missing wage.

The vector of individual characteristics we condition upon when computing our gender gap index contains age (in quadratic form), educational attainment (recorded as five categories from general elementary (or less) to higher, post-secondary education), years of potential labour market experience (computed as age minus years normally required to complete attained education minus six and classified in four groups: below 6, 6–10, 11–20 and 20 or more), and whether respondent was born in Germany. Because we pool data over multiple waves we also include a set of year dummies.[9] We adopt a parsimonious specification to avoid including variables that are potentially strongly gender-determined such as actual years of labour market experience, working hours and occupation or job characteristics (Neal and Johnson 1996).

In models controlling for endogenous labour market participation, we take into account family structure in the selection equation $x\delta^s$ (whether living with a partner and the education-level of the partner, the number of children in the household in various age ranges) and θ_C^s is allowed to vary according to the foreign-born status (in the Western Germany samples only).

Means of all covariates are reported in Table 2 by region and gender and within the full sample and for the subsample of observations with non-missing wage. Mean age is about 41 in all samples. The main contrasts between men and women are in educational achievements and years of potential experience (in particular in Eastern Germany). Employed women also have fewer children (in particular in Western Germany) which signals selective labour market participation.

3.2 Individual measures of risk aversion

The SOEP dataset contains individual-level measures of risk attitudes in a special module collected in the 2004 wave of the survey. One particular question allows us to approximate an individual-level risk aversion parameter ϵ_i under the assumption of CRRA utilities. Individual-level parameters can then be used in Eq. 4 to compute a measure of wage differentials that incorporates heterogeneous individual-level attitudes towards risk, instead of a constant parameter.

After a range of questions on their personal willingness to take risks (in general and in a number of specific contexts), survey respondents were presented an hypothetical investment opportunity. They were asked to report how much of a windfall gain of €100,000 they would be willing to invest in a risky asset. With equal probability, they would lose half of the value of their investment or they would double their investment. Respondents were asked to select a value for the amount invested $k \in \{€0, €20000, €40000, €60000, €80000, €100000\}$. The expected gain after investment of an amount k in the lottery is $G(k) =$

[9] In particular, equations for $b^s(x)$ include all controls (age minus 25 and its square, education dummies, potential predicted wage experience in years, wave dummies). Equations for $a^s(x)$ include the dummy for not being born in Germany, age and age squared and wage dummies. Equations for $q^s(x)$ only include waves dummies (aggregated in three sub-periods).

Table 1 Mean wage and selected percentiles (by gender and region)

	Western Germany		Eastern Germany	
	Men	Women	Men	Women
Mean	17.01	13.35	11.69	10.81
10th percentile	10.03	7.87	6.73	5.49
25th percentile	12.73	9.84	8.23	6.94
50th percentile	15.60	12.79	10.42	9.84
75th percentile	19.84	15.78	14.00	13.89
90th percentile	25.46	19.20	18.23	17.09
Total sample size	31801	35082	10893	11450
Obs. with non-missing wage (in salaried employment)	24029	17330	7047	6831
Share of obs. with non-missing wage	0.76	0.49	0.65	0.60

Pooled sample of respondents aged 25–55 over survey years 1999–2008. Sample weights applied. See text for definitions of hourly wage and of the sub-sample of workers with non-missing wage

$(100000 - k) + \frac{1}{2}\left(\frac{k}{2} + 2k\right) = 100000 + \frac{k}{4}$. We follow Dohmen et al. (2005, 2011) and combine each respondent's investment choice with measures of initial endowments taken from the data to infer values for individual CRRA coefficients.

Under the assumed CRRA utility function, expected utility of respondent i with initial wealth endowments w_i^0 investing an amount k of the windfall €100000 gain (denoted W) is

$$\bar{u}(w_i^0, k; e_i) = \frac{1}{2(1-e_i)}\left(\left(w_i^0 + W - \frac{k}{2}\right)^{1-e_i} + (w_i^0 + W + 2k)^{1-e_i}\right) \tag{11}$$

where e_i is the respondent's CRRA. e_i is unknown but we assume that the investment level k_i selected by the respondent leads to a higher expected utility than any other possible investment k_{-i}:

$$\bar{u}(w_i^0, k_i; e_i) \geq \bar{u}(w_i^0, k_{-i}; e_i). \tag{12}$$

Plausible estimates of risk aversion coefficients for respondent i are given by the range of values of e_i for which inequality (12) is satisfied.[10] The ordinal nature of the set of possible investment choices implies that we are only able to estimate *ranges* of coefficients of risk aversion (Dohmen et al. 2005). We will therefore construct different values for Eq. 4 based on lower bounds and upper bounds for individual risk aversion coefficients. Also, the range of investment options does not allow us to capture any negative risk aversion: the lower bound for risk aversion is set to 0. For plausibility, we cap the maximum value for e_i at 10.

Table 3 shows estimates of the individual measures of risk aversion obtained in our sample with three alternative measures of initial endowments w_i^0: the net worth of

[10]In our application, we solve this problem numerically on a grid of 21 potential values for $e_i \in \{0, 0.5, 1, \ldots, 10\}$.

Table 2 Covariate means for observations with non-missing wage (1) and in the full sample (2) (by gender and region)

	Western Germany				Eastern Germany			
	Men		Women		Men		Women	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Age	41.59	40.63	41.03	40.24	41.01	41.00	41.71	40.59
Foreign born	0.17	0.18	0.16	0.19	0.01	0.01	0.00	0.02
Education:								
General elementary	0.10	0.11	0.10	0.16	0.04	0.06	0.03	0.06
Middle vocational	0.51	0.51	0.51	0.53	0.62	0.62	0.53	0.56
Vocational plus Abitur	0.07	0.08	0.13	0.09	0.06	0.04	0.05	0.04
Higher vocational	0.10	0.10	0.06	0.08	0.10	0.11	0.11	0.07
Higher education	0.22	0.20	0.19	0.15	0.19	0.18	0.28	0.27
Potential experience:								
0–10 years	0.07	0.10	0.12	0.12	0.09	0.10	0.11	0.13
11–20 years	0.29	0.30	0.25	0.30	0.28	0.27	0.22	0.28
21–30 years	0.38	0.34	0.35	0.32	0.37	0.37	0.38	0.33
31 years or more	0.22	0.21	0.23	0.22	0.21	0.20	0.22	0.21
Education of partner:								
No partner	0.30	0.31	0.41	0.29	0.37	0.35	0.28	0.28
General elementary	0.10	0.12	0.05	0.09	0.03	0.03	0.02	0.03
Middle vocational	0.35	0.36	0.30	0.35	0.38	0.38	0.43	0.42
Vocational plus Abitur	0.08	0.06	0.05	0.04	0.03	0.02	0.03	0.02
Higher vocational	0.05	0.06	0.07	0.08	0.05	0.04	0.10	0.09
Higher education	0.11	0.09	0.12	0.14	0.15	0.18	0.14	0.16
Number of children in age range:								
0–1 year old	0.03	0.05	0.01	0.05	0.04	0.04	0.01	0.04
2–4 years old	0.12	0.13	0.06	0.14	0.07	0.09	0.08	0.11
5–7 years old	0.13	0.14	0.09	0.16	0.10	0.08	0.13	0.11
8–10 years old	0.13	0.14	0.12	0.16	0.07	0.08	0.10	0.10
11–12 years old	0.10	0.10	0.08	0.10	0.05	0.07	0.04	0.08
13–15 years old	0.12	0.14	0.10	0.16	0.05	0.13	0.07	0.16
16–18 years old	0.09	0.10	0.09	0.12	0.08	0.13	0.09	0.16

Pooled sample of respondents aged 25–55 over survey years 1999–2008. Sample weights applied. See text for definition of variables

household wealth (total household assets minus debts), annual household disposable income, and neglect of initial endowments altogether (w_i^0 set to zero).[11] Estimates of individual risk aversion span the range of value from 1 to 10. Using annual disposable income

[11] Estimates in Table 3 differ from those of Dohmen et al. (2005) because of our sample restrictions (e.g., to working-age individuals). When computed on the whole set of respondents estimates are similar. Estimates of household net worth are taken from the 2002 and 2007 SOEP wealth modules (averaged over the two years for respondents with both values available).

Table 3 Mean, 5th, 50th and 95th percentiles of estimated individual-level risk aversion parameters by gender and region

	Western Germany				Eastern Germany			
	Men		Women		Men		Women	
	LB	UB	LB	UB	LB	UB	LB	UB
Initial endowment at net worth:								
Mean	5.9	7.9	6.3	8.4	5.6	7.9	6.0	8.4
P5	1.0	1.5	1.5	1.5	1.5	1.5	1.5	1.5
P50	5.5	10.0	6.0	10.0	5.5	10.0	5.5	10.0
P95	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
Initial endowment at annual household income:								
Mean	4.6	7.2	5.0	7.8	4.6	7.5	5.0	8.0
P5	1.0	1.0	1.5	1.5	1.0	1.0	1.5	1.5
P50	6.0	10.0	6.0	10.0	6.0	10.0	6.0	10.0
P95	7.5	10.0	8.0	10.0	7.5	10.0	7.5	10.0
No initial endowment:								
Mean	3.4	6.7	3.7	7.3	3.6	7.1	3.8	7.6
P5	1.0	1.0	1.0	1.5	1.0	1.0	1.0	1.5
P50	5.0	10.0	5.0	10.0	5.0	10.0	5.0	10.0
P95	5.0	10.0	5.0	10.0	5.0	10.0	5.0	10.0

Individual-level coefficients of risk aversion calculated from an hypothetical investment question (see text for details). Pooled sample of respondents aged 25–55 over survey years 1999–2008 which responded to 2004 module on willingness to take risks and for which initial endowments can be estimated. Net worth is estimated from SOEP wealth modules of 2002 and 2007. LB and UB refer to lower and upper bounds on individual risk aversion parameter estimates (see text for details)

as a level of initial endowments seems to lead to a more credible distribution of CRRA coefficients (which mostly range *within* the limits of zero and 10) than using net worth. Using data on wealth as endowment leads to more extreme estimates of risk aversion. Like us if we do not cap the estimates at 10, Dohmen et al. (2005) find estimated parameters of risk aversion up to 20 and suggest such extreme values to be potentially explained by measurement error in wealth. Disposable household income has the double advantage in this respect of being generally more precisely measured than wealth and of being available in SOEP in 2004 when the investment question was asked (wealth data are recorded in 2002 and 2007). It also has the advantage of being a strictly financial measure which can be compared to a windfall gain of € 100,000, whereas net worth data combine non-financial assets and debts that may not be relevant amounts for initial endowments in the present context.

As suggested in Section 2, individual-level CRRA measures can be plugged into Eq. 5 to benchmark the range of gender gap indices with constant $\epsilon \in [-4, 10]$ against a measure of wage differentials based on expressed risk aversion among women. Of course the resulting index is informative only to the extent that individual preferences regarding risk (as elicited by the hypothetical lottery question) are equal – or at least similar – to individual preferences regarding the distribution of income or wage in a group. Evidence (mostly experimental) exists that risk aversion and inequality aversion are indeed correlated (Beck 1994; Carlsson et al. 2005; Ferrer-i Carbonell and Ramos 2010). However, the two are not necessarily perfect substitutes. One must therefore be cautious in interpreting estimates of $\Delta_2(\tilde{\epsilon})$ based on individual-specific risk aversion parameters at face value, and mainly use it in comparison to estimates based on fixed, common ϵ parameters.

4 Results

Before going to the wage gap estimates *per se*, Table 4 summarises differences in the conditional wage distribution estimates between men and women. Numbers reported in the table are averages over women's characteristics of the mean, median and percentile ratios calculated from estimated conditional distributions, F_x^m and F_x^w. It is immediately clear that women have lower wages generally: the mean and median are remarkably lower for women. Estimates of percentile ratios additionally reveal that wages are more dispersed among women, with the exception of selection corrected estimates in Western Germany. Estimates also indicate that bottom-half inequality (the P50/P10 ratio) is larger than top-half inequality (the P90/P50 ratio) and relatively more so among women. Differences in the percentile ratios between men and women appear smaller than differences in mean or median wage levels.

How these differences in conditional distributions translate in our distribution-sensitive wage gap indicators is shown in Fig. 1 and Table 5. Each of the four panels of Fig. 1 shows estimates of $\Delta_2(\epsilon)$ for distribution-sensitivity parameters ϵ ranging from -4 to 10 (point estimates are bracketed by 90 percent percentile bootstrap variability bands). Since $\Delta_2(\epsilon)$ represents the '(inequality-adjusted) cents a women makes for every (inequality-adjusted) euros a man makes', the 'no wage gap' reference is the horizontal line at 1. The larger the

Table 4 Average mean, median and percentile ratios in men and women conditional wage distributions

	Western Germany			Eastern Germany		
	Men	Women	Ratio (W/M)	Men	Women	Ratio (W/M)
No selection correction:						
Mean wage	14.61	11.99	0.82	10.81	9.72	0.91
Median wage	14.14	11.65	0.83	10.21	9.16	0.91
P90/P10 ratio	2.13	2.24	1.05	2.36	2.57	1.09
P50/P10 ratio	1.49	1.55	1.04	1.55	1.64	1.06
P90/P50 ratio	1.43	1.44	1.01	1.53	1.56	1.02
With selection correction:						
Mean wage	15.99	12.52	0.79	11.22	9.79	0.88
Median wage	15.35	12.18	0.80	10.61	9.24	0.88
P90/P10 ratio	2.24	2.22	0.99	2.37	2.57	1.08
P50/P10 ratio	1.52	1.55	1.02	1.56	1.65	1.06
P90/P50 ratio	1.47	1.43	0.98	1.53	1.56	1.02

Calculations from estimates of the conditional wage distributions for men and women, averaged over women's characteristics

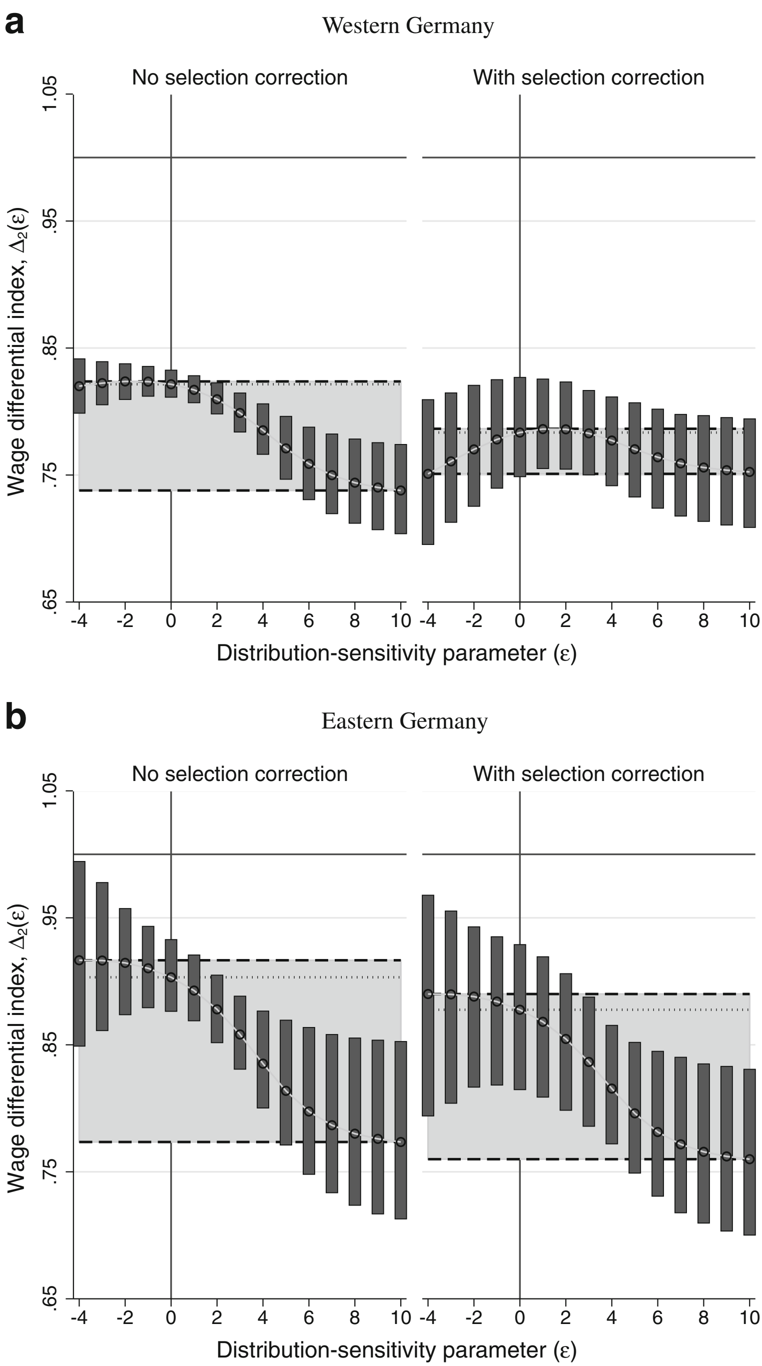

Fig. 1 Distribution-sensitive gender wage gap estimates $\Delta_2(\epsilon)$ for $\epsilon \in [-4, 10]$. *Vertical bars* show 90 % percentile-based bootstrap variability bands (based on 500 repeated half-sample bootstrap replications). The *grayed areas* show the range of variation of $\Delta_2(\epsilon)$ estimates for $\epsilon \in [-4, 10]$. The dotted line identifies $\Delta_2(0)$

Table 5 Summary of distribution-sensitive gender wage gap estimates

	No selection correction		With selection correction	
	Western Germany	Eastern Germany	Western Germany	Eastern Germany
Inequality-neutral gender gap, $\Delta_1 = \Delta_2(0)$:				
	0.82	0.90	0.78	0.88
Range of inequality-adjusted gender gap estimates $\Delta_2(\epsilon)$ for $\epsilon \in [-4, 10]$:				
Minimum	0.82	0.92	0.79	0.89
Maximum	0.74	0.77	0.75	0.76
Inequality-adjusted gender gap estimates with individual-specific risk aversion parameters $\Delta_2(\tilde{\epsilon})$:				
Lower bound parameters	0.77	0.81	0.77	0.80
Upper bound parameters	0.75	0.79	0.76	0.77

Individual-specific risk aversion parameters are based on the household disposable income initial endowment (see Section 3.2 for details)

vertical distance of the estimated indices to 1, the greater the disadvantage of women.[12] The grayed areas show the overall range of variation of $\Delta_2(\epsilon)$ for ϵ in -4 to 10: they show upper and lower bounds attained by the gender gap index once gender differences in inequality are taken into account in this range for ϵ. The bounds need not be attained for identical ϵ in all cases: the upper bound in the gap is attained at the corner $\epsilon = 10$ in the four cases, but the lower bound is attained at $\epsilon = -4$ in Eastern Germany and at $\epsilon = -2$ or $\epsilon = 1$ in Western Germany whether selection corrections are made or not. Note also that $\Delta_2(\epsilon)$ needs not be monotonically related to ϵ. The classic, 'neutral' gender gap index with $\epsilon = 0$ is identified on Fig. 1 by the horizontal dotted line. Table 5 summarizes the main estimates – the neutral baseline $\Delta_2(0)$ and the maximum and minimum values for 'distribution-sensitive' measures in the range covered by ϵ – along with wage gap measures based on estimated individual-level CRRA coefficients estimates as explained in Section 3.2.

The main result is that the classic 'neutral' index turns out to be close to the lower limit of the range inequality-adjusted indices, in particular in Western Germany. $\Delta_2(0)$ evaluates to 0.82 and 0.91 for Western and Eastern Germany respectively if endogenous selection is ignored and to 0.78 and 0.88 respectively if self-selection is taken into account. But the upper bound for the inequality-adjusted index evaluates to 0.73 and 0.78 for Western and Eastern Germany without selection correction and 0.75 and 0.77 after selection correction. The difference from the upper bound to the classic index is therefore relatively large. Women tend to face wage distributions which are unfavourable compared to the distribution of men, with a disadvantage going beyond facing lower means. The most bottom-sensitive is the measure (that is, the higher ϵ), the larger appears to be the wage gap. Although there is no one-to-one connection between our distribution-sensitive indices (which focus on gender differences in conditional distributions) and estimates of gender gaps at different quantiles of the unconditional wage distribution, these findings are coherent with evidence on the latter. For example, Arulampalam et al. (2007), Heinze (2010), and Christofides et al. (2013)

[12]The variant of index $\Delta_2(\epsilon)$ obtained by averaging over men's covariate distribution h^m instead of women's covariate distribution leads to similar estimates. These results are therefore not reported here but are available on request.

generally find larger gender differences at low wages in Germany, with more women being low paid.

In all cases, the gender gap estimates turn out to be smaller in Eastern Germany than in Western Germany. This is consistent with earlier research (see Maier 2007). But the impact of adjusting the estimates for inequality differences between men and women is much larger in Eastern Germany. East-West differences in the upper bounds of inequality-adjusted gender gap estimates are much smaller than appears when looking at the classic, 'inequality-neutral' gap estimate. This result is consistent with recent findings by Kohn and Antonczyk (2013) showing that wage inequality had been rising in East Germany in the aftermath of the German reunification and that this particularly affected women in the lower part of the wage distribution (primarily through changes in industry-specific remuneration patterns). Our results confirm that this resulted in markedly higher wage inequality in the female pay distribution as compared to the male distribution.

Unsurprisingly, accounting for endogenous selection generally leads to larger gender gap estimates. This is true in both regions, but the impact is stronger in Western Germany. The effect of accounting for selection in Western Germany not only appears to affect the male-female difference in the levels of wage (which would translate in a general shift in all estimates of $\Delta_2(\epsilon)$). It also affects the estimates of the overall shape of the conditional wage distributions, and in fact appears to reduce gender differences in wage distributions: the bounds of $\Delta_2(\epsilon)$ are much tightened, although at levels much higher than what would be indicated by $\Delta_2(0)$ with no selection correction.

Table 5 also reveals that estimates based on individual-level risk aversion parameters lead to wage gap indices much higher than the classic Δ_1 measures. They tend to be close to the upper limit of the range of estimates based on $\epsilon \in [-4, 10]$. This is unsurprising since our individual risk aversion parameter estimates tend to be relatively large on average (see Table 3). With those parameter values, wage gap estimates increase by approximately 20 % in Western Germany compared to the inequality-neutral case, while it more than doubles in Eastern Germany. Note also that gender gap indices for Eastern Germany are then much closer to the levels observed in Western Germany.

Finally, Table 6 presents some gender gap estimates for a set of specific subgroups of women and for three values of the ϵ parameter. The estimates presented in the table are simple averages among women of each of the subgroups of their 'individual inequality-adjusted gap' relative to a man of identical characteristics. In other words, it is the average over the $i \in \{1.., N^s\}$ women in each subgroup s of $\left[\log\left(C(F^w_{x_i};\epsilon)\right) - \log\left(C(F^m_{x_i};\epsilon)\right)\right]$.[13]

Partitioning by education level leads to contrasted results between Western and Eastern Germany. In Western Germany, the gender gap is larger for more highly educated women, while the opposite is observed in Eastern Germany. For $\epsilon = 4$, the gap is not even statistically significant for highly educated women in Eastern Germany. The partition into age groups reveals an inverted-U shape with lower gaps observed among middle-aged women. Finally, comparing estimates from the sample observations observed in 2000 and in 2008 show an interesting contrast for different ϵ: the gap has narrowed over time according to $\epsilon = -4$ (especially in Eastern Germany), has remained stable (in the West) or increased moderately (in the East) for $\epsilon = 0$, but it has increased for $\epsilon = 4$ (especially so in the East). This pattern signals an increase in the difference in (conditional) wage inequality between men and women over time, with inequality becoming higher among women.

[13] For any partition into subgroups, $\Delta_2(\epsilon)$ can be expressed as the average of the subgroup indices weighted by the subgroup shares.

Table 6 Inequality-adjusted gender wage gaps for population subgroups

	No selection correction			With selection correction		
	$\epsilon = -4$	$\epsilon = 0$	$\epsilon = 4$	$\epsilon = -4$	$\epsilon = 0$	$\epsilon = 4$
Western Germany						
All	0.82	0.82	0.79	0.75	0.78	0.78
	[0.80,0.84]	[0.81,0.83]	[0.77,0.81]	[0.70,0.81]	[0.75,0.83]	[0.74,0.81]
General elementary ed.	0.86	0.87	0.86	0.77	0.82	0.83
	[0.83,0.89]	[0.85,0.89]	[0.83,0.88]	[0.71,0.84]	[0.78,0.86]	[0.79,0.88]
Middle vocational ed.	0.83	0.81	0.76	0.78	0.79	0.76
	[0.80,0.85]	[0.80,0.83]	[0.73,0.78]	[0.73,0.83]	[0.76,0.83]	[0.72,0.79]
Higher education	0.78	0.79	0.76	0.71	0.75	0.75
	[0.75,0.81]	[0.78,0.81]	[0.74,0.79]	[0.65,0.77]	[0.71,0.80]	[0.72,0.79]
Aged 25–34	0.75	0.76	0.72	0.69	0.73	0.73
	[0.72,0.78]	[0.73,0.78]	[0.70,0.76]	[0.63,0.76]	[0.68,0.78]	[0.68,0.77]
Aged 35–44	0.86	0.86	0.83	0.77	0.80	0.80
	[0.82,0.91]	[0.82,0.90]	[0.79,0.87]	[0.70,0.84]	[0.75,0.86]	[0.75,0.85]
Aged 45–55	0.81	0.81	0.78	0.74	0.77	0.77
	[0.78,0.84]	[0.79,0.84]	[0.75,0.81]	[0.69,0.80]	[0.73,0.81]	[0.73,0.80]
Year 2000	0.79	0.81	0.78	0.72	0.77	0.77
	[0.76,0.83]	[0.80,0.82]	[0.75,0.81]	[0.64,0.79]	[0.74,0.81]	[0.74,0.81]
Year 2008	0.82	0.81	0.77	0.76	0.78	0.76
	[0.79,0.86]	[0.79,0.83]	[0.74,0.81]	[0.71,0.83]	[0.74,0.82]	[0.72,0.81]
Eastern Germany						
All	0.92	0.90	0.84	0.89	0.88	0.82
	[0.85,0.99]	[0.88,0.93]	[0.80,0.88]	[0.79,0.97]	[0.81,0.93]	[0.77,0.87]
General elementary ed.	0.85	0.86	0.84	0.83	0.84	0.82
	[0.77,0.93]	[0.83,0.91]	[0.78,0.91]	[0.72,0.91]	[0.77,0.90]	[0.76,0.89]
Middle vocational ed.	0.88	0.87	0.82	0.86	0.85	0.80
	[0.80,0.97]	[0.84,0.91]	[0.77,0.86]	[0.76,0.94]	[0.79,0.90]	[0.75,0.86]
Higher education	1.00	0.96	0.85	0.97	0.93	0.83
	[0.91,1.10]	[0.92,1.00]	[0.80,0.91]	[0.85,1.07]	[0.85,0.99]	[0.77,0.89]
Aged 25–34	0.88	0.87	0.81	0.85	0.84	0.79
	[0.74,1.05]	[0.75,1.01]	[0.69,0.95]	[0.70,1.03]	[0.72,0.99]	[0.68,0.94]
Aged 35–44	0.88	0.88	0.83	0.85	0.85	0.81
	[0.78,1.02]	[0.78,1.00]	[0.74,0.96]	[0.73,0.99]	[0.75,0.96]	[0.71,0.93]
Aged 45–55	0.90	0.89	0.82	0.88	0.87	0.80
	[0.82,1.01]	[0.84,0.94]	[0.76,0.88]	[0.77,0.98]	[0.81,0.93]	[0.74,0.86]
Year 2000	0.90	0.93	0.88	0.87	0.90	0.87
	[0.82,0.99]	[0.90,0.96]	[0.84,0.93]	[0.75,0.96]	[0.84,0.95]	[0.81,0.92]
Year 2008	0.96	0.90	0.78	0.94	0.87	0.76
	[0.88,1.08]	[0.86,0.94]	[0.72,0.85]	[0.83,1.05]	[0.80,0.93]	[0.70,0.83]

Figures in brackets are 90 % percentile bootstrap confidence intervals based on 500 repeated half-sample bootstrap replications

5 Summary and conclusion

Our analysis examines the size of the gender wage gap in Germany using indicators that fully incorporate (conditional) wage distribution differences between men and women. We provide range estimates for the gender wage gap that cover a broad range of possible distribution-sensitivity parameters, from extreme bottom-sensitive (or inequality averse) to extreme top-sensitive (or inequality loving). Standard indicators of the wage gap insensitive to inequality differences turn out to be close to the lower bound of the calculated range. Estimates of distribution-sensitive measures can be up to three times standard mean gap indicators in Eastern Germany and up to fifty percent higher in Western Germany. Indicators constructed on the basis of individual risk aversion parameters elicited from survey questions on a hypothetical risky investment are close to these upper bound. The situation of women appears to be worse than suggested by classic indicators: they tend to be penalized both by lower mean wages and by unfavourable configurations of higher moments. An argument that women's lower average wages could be compensated by favourable differences in higher moments is clearly not supported by our results, even under extreme preferences towards inequality.

The impact of accounting for inequality is particularly striking in Eastern Germany. While the gender pay gap remains generally smaller than in Western Germany, the inequality-related penalty is large enough in the East that the regional difference in inequality-adjusted indicators becomes largely muted.[14]

The analysis exploits a flexible parametric model of wage distributions that incorporates adjustments for endogenous labour market participation. Reassuringly, the overall picture remains unaffected once selectivity is taken into account, albeit – unsurprisingly – with generally larger gender gap estimates.

The more we increase the sensitivity of the measure to the differences at the bottom of the distributions, the larger appears to be the wage gap. Potential policy actions aiming to close the gender wage gap should therefore be oriented towards 'sticky floors' concern, rather than towards 'glass ceiling' problems; that is, the gap is mostly driven by women having disproportionately low wages compared to men, rather than by women not being able to achieve as high wages as men can at the top. So, policies that contribute to reducing female wage *inequality* by 'raising the floor' for low paid women would be most beneficial to equalise wage distribution across gender. We can speculate from, e.g., Antonczyk et al. (2010) that the sources of higher wage inequality among women (conditionally on human capital characteristics) are to be sought in the greater prevalence of low paid part-time employment, lower actual work experience (conditional on age or 'potential' experience), and possibly greater dispersion of occupational choices (especially towards low-paid jobs).

Acknowledgments This paper was prepared when Van Kerm visited the Institute for East and Southeast European Studies (IOS Regensburg) whose support and hospitality are gratefully acknowledged. The project was also supported by core funding for CEPS/INSTEAD from the Ministry of Higher Education and Research of Luxembourg. The data used in this publication were made available to us by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin. We thank Carlo Coen Castellino, Richard Frensch, Ksenia Gatskova, Jürgen Jerger and Donald Williams for stimulating comments and discussion. Comments from participants to the 2011 New Directions in Welfare

[14]The relevance of this result may be compounded in the face of the stronger preferences for redistribution and/or more egalitarian distributional norms expressed in Eastern Germany in general (Alesina and Fuchs-Schündeln 2007; Kuhn 2013).

congress (OECD Paris), the 2012 biennial EACES conference (Perth) and the 2013 ECINEQ meeting (Bari) are gratefully acknowledged.

References

Al-Farhan, U.F.: Changes in the gender wage gap in Germany during a period of rising wage inequality 1999-2006: Was it discrimination in the returns to human capital? SOEPpapers 293, DIW Berlin, The German Socio-Economic Panel (SOEP) (2010a)

Al-Farhan, U.F.: A detailed decomposition of changes in wage inequality in reunified post-transition Germany 1999-2006, SOEPpapers 269, DIW Berlin, The German Socio-Economic Panel (SOEP) (2010b)

Albrecht, J., Björklund, A., Vroman, S.: Is there a glass ceiling in Sweden? J. Labor Econ. **21**(1), 145–177 (2003)

Alesina, A., Fuchs-Schündeln, N.: Goodbye Lenin (or not?): The effect of communism on people. Am. Econ. Rev. **97**(4), 1507–1528 (2007)

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

Antonczyk, D., Fitzenberger, B., Sommerfeld, K.: Anstieg der lohnungleichheit, rückgang der tarifbindung und polarisierung. Zeitschrift für Arbeitsmarktforschung **44**(1–2), 15–27 (2011)

Arulampalam, W., Booth, A.L., Bryan, M.L.: Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. Ind. Labor Relat. Rev. **60**(2), 163–186 (2007)

Atkinson, A.B.: On the measurement of inequality. J. Econ. Theory **2**, 244–263 (1970)

Beck, J.H.: An experimental test of preferences for the distribution of income and individual risk aversion. East. Econ. J. **20**(2), 131–145 (1994)

Biewen, M., Jenkins, S.P.: Accounting for differences in poverty between the USA, Britain and Germany. Empir. Econ. **30**(2), 331–358 (2005)

Blackburn, M.L.: Estimating wage differentials without logarithms. Labour Econ. **14**(1), 73–98 (2007)

Blau, F.D., Kahn, L.M.: The gender earnings gap: learning from international comparisons. Am. Econ. Rev. **82**(2), 533–38 (1992)

Blau, F.D., Kahn, L.M.: Wage structure and gender earnings differentials: an international comparison. Economica **63**, S29—S62 (1996)

Blau, F.D., Kahn, L.M.: Swimming upstream: trends in the gender wage differential in 1980s. J. Labor Econ. **15**(1), 1–42 (1997)

Card, D., Heining, J., Kline, P.: Workplace heterogeneity and the rise of West German wage inequality. Q. J. Econ. **128**(3), 967–1015 (2013)

Carlsson, F., Daruvala, D., Johansson-Stenman, O.: Are people inequality-averse, or just risk-averse? Economica **72**(287), 375–396 (2005)

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

Christofides, L.N., Polycarpou, A., Vrachimis, K.: Gender wage gaps, 'sticky floors' and 'glass ceilings' in Europe. Labour Econ. **21**, 86–102 (2013)

del Río, C., Gradín, C., Cantó, O.: The measurement of gender wage discrimination: The distributional approach revisited. J. Econ. Inequal. **9**(1), 57–86 (2011)

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G.: Individual risk attitudes: new evidence from a large, representative, experimentally-validated survey, Discussion Papers of DIW Berlin 511, DIW Berlin, German Institute for Economic Research (2005)

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G.: Individual risk attitudes: measurement, determinants and behavioural consequences. J. Eur. Econ. Assoc. **9**(3), 522–550 (2011)

Dolton, P.J., Makepeace, G.: The statistical measurement of discrimination. Econ. Lett. **18**, 391–5 (1985)

Dustmann, C., Ludsteck, J., Schönberg, U.: Revisiting the German wage structure. Q. J. Econ. **124**(2), 843–881 (2009)

Ferrer-i Carbonell, A., Ramos, X.: Inequality aversion and risk attitudes, IZA Discussion Paper 4703. Institute for the Study of Labor (IZA), Bonn, Germany (2010)

Firpo, S., Fortin, N.M., Lemieux, T.: Unconditional quantile regressions. Econometrica **77**(3), 953–973 (2009)

Fortin, N., Lemieux, T., Firpo, S.: Decomposition methods. In: Ashenfelter, O., Card, D. (eds.): Handbook of Labor Economics, vol. 4A, pp. 1–102, North-Holland, Amsterdam (2011)
Foster, J.E., Székely, M.: Is economic growth good for the poor? Tracking low incomes using general means. Int. Econ. Rev. **49**(4), 1143–1172 (2008)
Fuchs-Schündeln, N., Krueger, D., Sommer, M.: Inequality trends for Germany in the last two decades: A tale of two countries. Rev. Econ. Dyn. **13**(1), 103–132 (2010)
Hall, P., Wolff, R.C.L., Yao, Q.: Methods for estimating a conditional distribution function. J. Am. Stat. Assoc. **94**(445), 154–63 (1999)
Heckman, J.J.: Sample selection bias as a specification error. Econometrica **47**(1), 153–161 (1979)
Heinze, A.: Beyond the mean gender wage gap: Decomposition of differences in wage distributions using quantile regression, ZEW Discussion Papers 10-043. Center for European Economic Research, Mannheim (2010)
Huber, M., Melly, B.: Quantile regression in the presence of sample selection, Economics Working Paper 1109, School of Economics and Political Science, Department of Economics, University of St. Gallen (2011)
Hunt, J.: The transition in East Germany: When is a ten-point fall in the gender wage gap bad news? J. Labor Econ. **20**, 148–169 (2002)
Jenkins, S.P.: Earnings discrimination measurement: a distributional approach. J. Econ. **61**, 81–102 (1994)
Juhn, C., Murphy, K.M., Pierce, B.: Wage inequality and the rise in returns to skill. J. Polit. Econ. **101**(3), 410–42 (1993)
Kleiber, C., Kotz, S.: Statistical size distributions in Economics and Actuarial Sciences, Wiley series in probability and statistics. Wiley, New York (2003)
Kohn, K., Antonczyk, D.: The aftermath of reunification. Econ. Transit. **21**(1), 73–110 (2013)
Kuhn, A.: Inequality perceptions, distributional norms, and redistributive preferences in East and West Germany. Ger. Econ. Rev. **14**(4), 4 (2013)
Machado, J.A.F., Mata, J.: Counterfactual decomposition of changes in wage distributions using quantile regression. J. Appl. Econ. **20**, 445–465 (2005)
Maier, F.: The persistence of the gender wage gap in Germany, Discussion Paper 01/2007, Harriet Taylor Mill-Institut für Ökonomie und Geschlechterforschung, Berlin School of Economics (2007)
Melly, B.: Decomposition of differences in distribution using quantile regression. Labour Econ. **12**, 577–90 (2005)
Millimet, D.L., Wang, L.: A distributional analysis of the gender earnings gap in urban China. Contributions to Economic Analysis and Policy **5**(1), 1–48 (2006)
Neal, D.A., Johnson, W.R.: The role of premarket factors in black-white wage differences. The J. Polit. Econ. **104**(5), 869–895 (1996)
Picchio, M., Mussida, C.: Gender wage gap: a semi-parametric approach with sample selection correction. Labour Econ. **18**(5), 564–578 (2011)
Pigini, C.: Bivariate non-normality in the sample selection model. J. Econ. Methodol. **4**(1), 123–144 (2014)
Rothe, C.: Nonparametric estimation of distributional policy effects. J. Econ. **155**(1), 56–70 (2010)
Saigo, H., Shao, J., Sitter, R.R.: A repeated half-sample bootstrap and balanced repeated replications for randomly imputed data. Survey Methodology **27**(2), 189–196 (2001)
Singh, S.K., Maddala, G.S.: A function for size distribution of incomes. Econometrica **44**(5), 963–970 (1976)
Smith, M.D.: Modelling sample selection using Archimedean copulas. Econ. J. **6**(1), 99–123 (2003)
Smolny, W., Kirbach, M.: Wage differentials between East and West Germany: Are they related to the location or to the people? Appl. Econ. Lett. **18**(9), 873–879 (2011)
Van Kerm, P.: Generalized measures of wage differentials. Empir. Econ. **45**(1), 465–482 (2013)
Wagner, G.G., Burkhauser, R.V., Behringer, F.: The English language public use file of the German Socio-Economic Panel Study. J. Hum. Resour. **28**(2), 429–433 (1993)
Wagner, G.G., Frick, J.R., Schupp, J.: The German Socio-Economic Panel Study (SOEP) - scope, evolution and enhancements. Schmollers Jahrbuch J. Appl. Soc. Sci. Stud. **127**(1), 139–169 (2007)