DISCUSSION PAPER SERIES

IZA DP No. 6447

Robust Estimation of Wage Dispersion with Censored Data: An Application to Occupational Earnings Risk and Risk Attitudes

Daniel Pollmann
Thomas Dohmen
Franz Palm

March 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Robust Estimation of Wage Dispersion with Censored Data: An Application to Occupational Earnings Risk and Risk Attitudes

Daniel Pollmann
Harvard University
and ROA, Maastricht University

Thomas Dohmen
ROA, Maastricht University
and IZA

Franz Palm
Maastricht University

Discussion Paper No. 6447
March 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 6447
March 2012

ABSTRACT

Robust Estimation of Wage Dispersion with Censored Data: An Application to Occupational Earnings Risk and Risk Attitudes

We present a semiparametric method to estimate group-level dispersion, which is particularly effective in the presence of censored data. We apply this procedure to obtain measures of occupation-specific wage dispersion using top-coded administrative wage data from the German IAB Employment Sample (IABS). We then relate these robust measures of earnings risk to the risk attitudes of individuals working in these occupations. We find that willingness to take risk is positively correlated with the wage dispersion of an individual's occupation.

JEL Classification: C14, C21, C24, J24, J31, D01, D81

Keywords: dispersion estimation, earnings risk, censoring, quantile regression, occupational choice, sorting, risk preferences, SOEP, IABS

Corresponding author:

Thomas Dohmen
ROA, Maastricht University
Tongersestraat 53
6211 LM Maastricht
The Netherlands
E-mail: t.dohmen@maastrichtuniversity.nl

1 Introduction

Important economic issues often center on the shape of distributions. Examples include questions relating to income inequality, the shape of wage offer distributions, or the riskiness of returns to financial assets. In various settings, empirical labor economists have been interested in measures of wage dispersion. More than often, such measures have to be estimated from censored data. For example, the March Current Population Survey (CPS), which contains survey responses on weekly earnings top-coded for anonymization purposes, has been used in several studies. Researchers have frequently dealt with this problem by multiplying top-coded earnings by a factor of 1.3 to 1.5 (e.g., Katz and Murphy, 1992; Juhn, Murphy, and Pierce, 1993). Other studies have relied on distributional assumptions to impute censored earnings in their data (e.g., Dustmann, Ludsteck, and Schönberg, 2009). Closely related, moments can typically be recovered if the shape of the distribution and the censoring rule are known. In many settings, however, the shape of the wage distribution is unknown and possibly itself of interest, and estimation methods that require parametric assumptions typically yield inconsistent estimates when these are violated.[1] More advanced semiparametric methods have been used for social security earnings records matched to the CPS, which suffer from a much higher degree of censoring due to a legal contribution limit (e.g., Chay and Honoré, 1998; Hu, 2002).

We present a measure of group-level dispersion that can be straightforwardly obtained from quantile regression (QR). Our method does not require parametric assumptions on the error terms and is as such consistent under heteroskedasticity and non-normality even for censored data. In addition, by using this simple-to-compute method, which is based on group coefficient estimates at different quantiles rather than residuals, we can avoid dealing with censored residuals. Our semiparametric approach allows to estimate differential patterns of dispersion across occupations. We are thus able to adequately characterize the entire conditional wage distribution while explicitly incorporating the dispersion effect of covariates.

We then demonstrate the usefulness of the estimation procedure in an application in which we relate the estimated occupation-specific wage dispersion as a measure of occupation-specific earnings risk to the risk attitudes of individuals working in these occupations. In order to estimate the occupation-specific cross-sectional earnings risk, we rely on administrative wage data from the IAB Employment Sample (IABS) that contains wage information censored at the statutory limit for social security contributions. The IABS offers great sample size, such that

[1] For example, the Tobin-Amemiya maximum likelihood estimator (Tobin, 1958; Amemiya, 1973) and the two-step Heckit approach (Heckman, 1976, 1979) are inconsistent under deviations from homoskedasticity (e.g., Maddala and Nelson, 1975; Hurd, 1979; Arabmazar and Schmidt, 1981; Brown and Moffitt, 1983; Donald, 1995) and normality (e.g., Arabmazar and Schmidt, 1982; Goldberger, 1983; Paarsch, 1984). The biases these authors derive depend on the exact set-up and the degree of censoring, but can be substantial. While the overall degree of censoring in our application is not that extreme compared to their settings, it is within certain occupation cells. The simulation study of Vijverberg (1987) for the case of non-normality shows that the estimated error variance is often seriously biased, which may trouble our dispersion analysis. Powell (1986b) carries out simulations for heteroskedasticity as well as non-normality and finds that "failure of the homoskedasticity assumption may have more serious consequences than failure of normality in censored regression models." Due to its unbounded influence function, the Tobit estimator is particularly sensitive to outliers and abnormally long tails (e.g., Chay and Honoré, 1998; Wilhelm, 2008). Schulhofer-Wohl (2011) estimates a panel Tobit model, in which he allows heteroskedasticity at the cross-sectional though not longitudinal level. The normality assumption remains crucial, particularly in light of the high degree of censoring and its volatility over time ("30 to 60 percent of observations on prime-age male workers are censored in each year", Schulhofer-Wohl, 2011, p. 931).

we are able to work with more precise occupation definitions than previous studies and reduce the effect of aggregation on variation. We then match the estimated wage dispersion measure of occupations to individuals in the German Socio-Economic Panel Study (SOEP) working in these occupations. The SOEP provides us with survey information on risk attitudes and other individual and household characteristics. Consistent with previous studies (e.g., Bonin et al., 2007; Fouarge, Kriechel, and Dohmen, 2011) that have assessed the relation between occupational earnings risk and risk preferences, we find evidence of a statistically significant correlation between our measure of occupational earnings risk and the risk attitudes of individuals working in a particular occupation: Those who state to be more willing to take risks are more likely to work in occupations with higher cross-sectional wage dispersion.

Our empirical application is related to a large literature that investigates the relationships between risk preferences and occupational choice. Early studies (e.g., Bellante and Link, 1981) have assessed how risk preferences affect the choice between private sector and public sector employment. The typical finding in this strand of the literature is that higher levels of individual risk aversion significantly increase the probability of working in the public sector (e.g., Guiso and Paiella, 2005; Fuchs-Schündeln and Schündeln, 2005; Dohmen and Falk, 2010). A second class of studies has focused on the relationship between risk preferences and the probability of self-employment, which is considered to be more risky than dependent employment. Using data for different countries and employing different measures of risk attitudes, these studies consistently find that a higher propensity to take risks increases the probability of being self-employed (e.g., Cramer et al., 2002; Guiso and Paiella, 2005; Ekelund et al., 2005; Caliendo, Fossen, and Kritikos, 2009; Dohmen et al., 2011; Beauchamp, Cesarini, and Johannesson, 2011).

Most closely related to our empirical application are studies that have related proxies of risk aversion or direct measures of risk attitudes to occupational earnings risk. Saks and Shore (2005), for example, use data from the National Postsecondary Student Aid Survey and find that, as expected under decreasing absolute risk aversion utility, individuals with higher parental wealth more frequently choose college majors leading into occupations with greater conditional earnings variation (see also King, 1974), as estimated on data from the Panel Study of Income Dynamics (PSID) and the Baccalaureate & Beyond survey.[2] Bonin et al. (2007) and Fouarge, Kriechel, and Dohmen (2011) use direct measures of risk attitudes and relate them to an explicit statistic for the riskiness of occupations, the occupation-specific standard deviation of the residuals from a Mincer wage regression. They find a significant positive relationship between this cross-sectional earnings risk measure and individuals' stated willingness to take risks. While Bonin et al. carry out all estimation on data from the SOEP, Fouarge, Kriechel, and Dohmen compute the occupation-specific cross-sectional earnings risk based on administrative wage data from Statistics Netherlands (CBS) and relate it to the self-reported risk attitudes of respondents to the ROA School Leavers Survey, which is based on the SOEP questions on willingness to take risks. Schulhofer-Wohl (2011) uses responses to the question on risky jobs in the Health and Retirement

[2]There is a related literature on the relationship between risk preferences and educational choice (e.g., Belzil and Hansen, 2004; Belzil and Leonardi, 2007, 2009; Chen, 2008; Shaw, 1996). Theoretical predictions about the relationship between risk preferences and educational choice are less clear cut as education may be considered a risky investment (Levhari and Weiss, 1974), but also shield against unemployment (Mincer, 1991; Nickell and Bell, 1996).

Study to relate them to the amount of income risk experienced by individuals, estimated based on data from matched social security earnings records. Schulhofer-Wohl classifies individuals into a low and a high risk tolerance group and finds that the latter carry significantly more of both aggregate and idiosyncratic risk.

Our main contribution to this strand of the literature is the introduction of a robust measure of occupation-specific earnings risk that does not rely on parametric assumptions for the error terms and yields consistent estimates for occupation-specific wage dispersion even if homoskedasticity and normality assumptions are violated. Moreover, the earnings risk measure we propose in the paper can even be estimated in the presence of censored wage information. This can be of great advantage in empirical work as administrative wage data are often top-coded. Importantly, Monte Carlo simulations show that our method for the estimation of wage dispersion is particularly effective compared to conventional approaches.

In our application, we find that individuals with greater stated willingness to take risks work in occupations with higher cross-sectional wage dispersion. After estimating risk profiles of occupations on the IABS data, we match them to individuals in the SOEP working in these occupations. The SOEP provides us with survey information on risk preferences and other individual and household characteristics. The IABS on the other hand offers great sample size, such that we are able to work with more precise occupation definitions than previous studies and reduce the effect of aggregation on variation.

The organization of the paper is as follows. In section 2, we briefly discuss QR and present our method for estimating dispersion in more detail. In addition, we describe a particularly useful estimation algorithm for censored data, the 3-step censored quantile regression (CQR) estimator by Chernozhukov and Hong (2002), which we use in our application on risk preferences and occupational sorting in section 3. Section 4 concludes.

2 Estimation of group-level dispersion

Our method for the estimation of dispersion is not based on residuals, but rather on the difference of coefficient estimates at particular quantiles. As such, it is in the spirit of the heteroskedasticity test of Koenker and Bassett (1982), which carries out a Wald test on the differences of coefficient estimates at different quantiles. Specifically, we first estimate the entire model by (C)QR at different quantiles, such as the 10th, 25th, 50th, 75th, and 90th percentile, including dummy variables for the groups which are to be compared. In our application, for instance, we include dummies for all occupations. We then consider the differences of the coefficient estimates for these dummies at two particular quantiles, such as the 10th and 90th percentile ("10-90 spread"), and compare their values across occupations. Our approach is not only computationally simple, but it also controls for the dispersion effect of covariates, and thereby filters out the (possibly) heteroskedastic effect of, for example, education and tenure in our application.

To introduce notation and build intuition, we briefly summarize quantile regression in section 2.1 before introducing our dispersion measure in section 2.2. Section 2.3 discusses a particularly simple estimator for censored quantile regression used in our application.

2.1 Quantile regression

Quantile regression (QR), introduced by Koenker and Bassett (1978) as a generalization of median regression, allows us to parsimoniously describe the entire conditional wage distribution by estimating conditional quantile functions (CQF) $Q_\tau(Y_i|X_i)$.[3] We denote the conditional τ-quantile of Y given X as

$$Q_\tau(Y_i|X_i) \equiv q_\tau(X_i) \equiv F^{-1}_{Y_i|X_i}(\tau) = \inf_{r\in\mathbb{R}}\{r : F_{Y_i|X_i}(r) > \tau\}. \tag{1}$$

For a linear quantile model, $q_\tau(X_i) = X_i'\beta(\tau)$, and

$$Y_i = X_i'\beta(\tau) + u_i(\tau). \tag{2}$$

While specifying a parametric model of the conditional quantiles, we are agnostic about the error distribution in our semiparametric framework. All we rely on is a conditional quantile restriction, stipulating that the conditional quantile of the error is equal to a constant. We assume that X_i always includes a constant term for the intercept, which affords us the following normalization:

$$Q_\tau(u_i(\tau)|X_i) = 0. \tag{3}$$

Estimation typically proceeds by characterizing conditional quantiles as the solution to a particular expected loss minimization problem, in the context of which it is useful to define the "check" (or weighted absolute loss) function. For $\tau \in (0, 1)$,

$$\rho_\tau(u) \equiv \tau\mathbf{1}(u \geq 0)u + (1-\tau)\mathbf{1}(u < 0)(-u) \tag{4}$$

$$= [\tau - \mathbf{1}(u < 0)]u. \tag{5}$$

In the case of our linear quantile model, $q_\tau(X_i) = X_i'\beta(\tau)$,

$$\beta(\tau) = \arg\min_b \mathbb{E}[\rho_\tau(Y_i - X_i'b)|X_i]. \tag{6}$$

We can define the QR estimator as its sample equivalent and the optimal predictor minimizing the realized loss:

$$\hat{\beta}(\tau) = \arg\min_b \sum_{i=1}^{N} \rho_\tau(Y_i - X_i'b). \tag{7}$$

Asymptotic normality and consistency of the QR estimator can be shown (Bassett and Koenker, 1978).

[3] An excellent non-technical introduction with illustrative examples and an overview of applications can be found in Koenker and Hallock (2001). Buchinsky (1998) summarizes a range of points relevant to the empirical researcher.

2.2 Differences of quantile coefficients

2.2.1 Location-scale model

To build intuition on the meaning of our dispersion measure, we begin with the illustrative yet likely simplistic special case of a location-scale model. Suppose X_i includes a constant, such that $X_{i1} = 1$, as well as additional covariates. Y_i is dependent on X_i in mean and through a re-scaling of variances (the disturbances u_i are *iid*):[4]

$$Y_i = X_i'\beta + \sigma(X_i)u_i \tag{8}$$

$$q_\tau(X_i) = X_i'\beta + \sigma(X_i)F^{-1}(\tau). \tag{9}$$

Suppose $\sigma(\cdot)$ is linear with $\sigma(X_i) = X_i'\zeta$. Then, $\beta(\tau) = \beta + F^{-1}(\tau)\zeta$, and the quantiles have both a location effect through $F^{-1}(\tau)\zeta_1$ and a scale effect through $F^{-1}(\tau)\zeta_{j\neq 1}$. To strike the link to our application, suppose that in addition to a vector $\tilde{X}_i \in \mathbb{R}^p$ of a constant and $p-1$ covariates, X_i contains a vector $\dot{X}_i \in \mathbb{R}^{m-1}$ of dummies for $m-1$ occupation groups such that $X_i = (\tilde{X}_i', \dot{X}_i')'$; in this case, $\zeta_{p+k-1} = \dot{\sigma}_k$ denotes the occupation-specific scale effect of occupation group $k \in \{2, \ldots, m\}$ relative to the reference occupation group.[5] Then, for consistent estimators $\hat{\beta}(\tau_1)$ and $\hat{\beta}(\tau_2)$, $\hat{\beta}_{p+k-1}(\tau_1) - \hat{\beta}_{p+k-1}(\tau_2) \xrightarrow{p} \dot{\sigma}_k[F^{-1}(\tau_1) - F^{-1}(\tau_2)]$ by Slutsky's theorem. Hence, we can consistently estimate $\dot{\sigma}_k$ up to scale.

Chamberlain (1994, p. 186) discusses comparable normal-location models, but considers them inadequate for characterizing the conditional wage distribution. In particular, they imply constant covariate slopes across the quantiles, which are at odds with the quantile patterns of industry wage effects he finds. Moreover, and closest to our application, Chamberlain presents differential patterns across industries and relates these to industry-specific residual dispersion. Close inspection of the industry coefficients in Chamberlain (1994, table 5.4) at different quantiles reveals that even the location-scale model may be too restrictive. A statistical test of the location-scale hypothesis can be based on a Khmaladze transformation, but is only available for uncensored data. Applying a human capital model including occupation dummies to self-reported earnings in the SOEP, the Khmaladze test (Koenker and Xiao, 2002) rejects the location-scale hypothesis at the 1% level.

2.2.2 General (separable) dispersion model

We turn now to a more general model which we believe to more adequately characterize the German wage distribution. The above location-scale model allows some flexibility with respect to heteroskedasticity, but it still imposes a great deal of homogeneity on the shape of the error distribution. In particular, $\frac{\hat{\beta}_{p+k-1}(\tau_1)-\hat{\beta}_{p+k-1}(\tau_2)}{\hat{\beta}_{p+l-1}(\tau_1)-\hat{\beta}_{p+l-1}(\tau_2)} \xrightarrow{p} \frac{\dot{\sigma}_k}{\dot{\sigma}_l}$ for any $\tau_1, \tau_2 \in (0,1)$ and $k, l \in \{2, \ldots, m\}$. For instance, differential asymmetry and tail behavior are precluded.

Consider therefore a more general linear model in which each occupation $k \in \{1, \ldots, m\}$ has its own conditional distribution function, such that $u_i \sim F_k(\cdot\,; X_i)$ if individual i works in occupation k. Suppose further that for each k and at every quantile τ, $F_k^{-1}(\cdot\,; X_i)$ is linearly

[4] For instance, the data may be conditionally normal with variance depending on the covariates.

[5] We omit occupation 1 and assume that its effect is subsumed in the general constant.

separable into a general component $\xi(\tilde{X}_i;\tau)$ linear in $\tilde{X}_i$ and an occupation-specific component $\eta_k(\tau) = \dot{X}_i{}'\eta(\tau)$.[6] The CQF is then given by:

$$q_\tau(X_i) = X_i'\beta + \xi(\tilde{X}_i;\tau) + \eta_k(\tau) \tag{10}$$

$$= X_i'\beta + \tilde{X}_i{}'\lambda(\tau) + \dot{X}_i{}'\eta(\tau). \tag{11}$$

As a result, $\beta(\tau) = \beta + (\lambda(\tau)', \eta(\tau)')'$, and hence, for consistent estimators $\hat{\beta}(\tau_1)$ and $\hat{\beta}(\tau_2)$, $\hat{\beta}(\tau_1) - \hat{\beta}(\tau_2) \xrightarrow{p} (\lambda(\tau_1)' - \lambda(\tau_2)', \eta(\tau_1)' - \eta(\tau_2)')'$. In this way, we can measure the effect of our set of occupation groups on all different quantiles while controlling for the dispersion effect of the additional covariates. This gives us a measure of the conditional dispersion effect of each occupation. While the linear specification of the conditional quantiles may appear restrictive, a linear quantile model is frequently only intended as a reduced-form approximation, such as for the minimum distance (MD) estimators in Buchinsky (1994, p. 409) and Chamberlain (1994, p. 181).[7]

To evaluate the practical merits of the approach, we carry out a set of Monte Carlo simulations (section B in the appendix). We find that compared to conventional approaches, our method is particularly effective for the estimation of dispersion in the presence of interaction effects in variance. For a moderate degree of censoring of 10%, very similar to that in our application, the 10-90 spread obtained from CQR does just as well as when we leave the data uncensored.

2.3 Censored quantile regression

A particular feature of conditional quantiles not shared by conditional expectations is equivariance to monotone transformations. For any non-decreasing function $g(\cdot)$,

$$Q_\tau[g(Y_i)|X_i] = g[Q_\tau(Y_i|X_i)]. \tag{12}$$

As a result, QR is particularly suited for censoring problems. In addition, it does not require the restrictive assumptions of parametric censored estimators. In the case of top-coding, we observe $Y_i = \min(C_i, Y_i^*)$, where Y_i^* is the latent true value of the process of interest and C_i is some upper limit, of which we assume Y_i^* to be conditionally independent. Since for any $C_i \in \mathbb{R}$, $\min(C_i, \cdot)$ is a non-decreasing function, we have (Powell, 1986a):

$$Q_\tau(Y_i^*|X_i) = X_i'\beta(\tau) \Rightarrow Q_\tau(Y_i|X_i) = \min[C_i, X_i'\beta(\tau)]. \tag{13}$$

The censored quantile regression (CQR) estimator follows trivially as the minimizing argu-

[6] This does not state that $\eta_k(\tau)$ is indeed the τ-quantile of the occupation effect; rather, it is the occupation effect at the τ-quantile. Therefore, $\eta_k(\tau)$ need not be increasing; rather, an occupation l may experience lower variance than the reference occupation, in which case $\eta_l(\tau)$ is decreasing.

[7] Formally, Chamberlain (1994, p. 181) recognized that the QR estimator provides a linear approximation to the CQF, albeit of a less "transparent" nature than in the OLS and MD case. Angrist, Chernozhukov, and Fernández-Val (2006) show that QR minimizes a weighted mean-squared error loss function for specification error, implicitly providing a weighted MD approximation to the true nonlinear CQF. Applying the framework to wage regressions with a focus on the education variable, they find QR to provide a useful approximation to the conditional wage distribution.

ment of the Powell objective function (Powell, 1986a):

$$\hat{\beta}_{CQR}(\tau) = \arg\min_b \sum_{i=1}^{N} \rho_\tau[Y_i - \min(C_i, X_i'b)]. \tag{14}$$

As in the uncensored case, the QR-based estimator is consistent under general non-normal distributions and heteroskedasticity (Powell, 1984, 1986a). Buchinsky (1994) gives a well-known application to changes in the US wage structure. For estimation, he presents his iterative linear programming algorithm (ILPA), which iteratively performs QR on observations with predictions in the uncensored region, based on the previous iteration. Convergence is achieved if two subsequent sets of observations are the same; while this need not occur, convergence guarantees local optimality. Another alternative, the BRCENS algorithm, is proposed in Fitzenberger (1997). Unfortunately, both algorithms have less than reliable convergence properties with respect to the Powell estimator (14), particularly in large samples and for high dimensionality, as in our application.

Chernozhukov and Hong (2002) present the three-step CQR method, which avoids a great deal of problems by selecting a more "benign" sample based on an initial regression of the probability of censoring, and subsequently works with standard QR.[8,9]

Step 1. Let $\kappa_i = \mathbf{1}(Y_i \neq C_i)$; that is, κ_i is an indicator of non-censoring (with censoring point C_i). We consider a parametric (e.g., probit or logit) model for the probability of non-censoring:

$$\kappa_i = p(X_i'\gamma) + \epsilon_i. \tag{15}$$

In general, model (15) will be misspecified and therefore inconsistent for the true propensity score $h(X_i, C_i)$. However, it is only used as an auxiliary regression to select an initial sample J_0 with propensity score $h(X_i, C_i) > \tau$, necessary for consistent estimation of quantile τ.[10] To ensure this, we do not base our selection on the condition that $p(X_i'\hat{\gamma}) > \tau$, but rather that $p(X_i'\hat{\gamma}) > \tau + k$, where k is a trimming constant strictly between 0 and $1 - \tau$. Since we do not necessarily have to select the largest subset J_0, there is some freedom in choosing k. For this, we write J_0 as a function of k, $J_0(k) = \{i : p(X_i'\gamma) > \tau + k\}$. The approach taken here, following Chernozhukov and Hong, is to choose the trimming constant k such that

$$\#J_0(k)/\#J_0(0) = 90\%. \tag{16}$$

This means that we discard 10% of those observations with a probability estimate higher than τ. The authors provide the sufficient condition that $p(X_i'\gamma_0) - k$ (for $\gamma_0 = \operatorname{plim}\hat{\gamma}$) be a lower bound on $h(X_i, C_i)$.

[8] Applications include Melly (2005) on wage inequality, Kowalski (2009) on medical expenditure, and Schmillen and Möller (2012) on lifetime unemployment.

[9] The estimators of Buchinsky and Hahn (1998) and Khan and Powell (2001) similarly carry out a first-stage selection, but are impractical for high dimensionality and large data sets.

[10] Note the deviation from Chernozhukov and Hong, who select a sample with propensity score $h(X_i, C_i) > 1 - \tau$. This is an important difference between left- and right-censoring.

Step 2. We obtain the initial (inefficient) estimator $\hat{\beta}^0(\tau)$ by standard QR on the sample J_0:

$$\hat{\beta}^0(\tau) = \arg\min_b \sum_{i \in J_0(k)} \rho_\tau(Y_i - X_i'b). \tag{17}$$

Next, we select $J_1 = \{i : X_i'\hat{\beta}^0(\tau) < C_i - \delta_N\}$, where δ_N is a small number such that as the sample size $N \to \infty$, $\sqrt{N} \cdot \delta_N \to \infty$ and $\delta_N \searrow 0$. We choose δ_N similarly to k, but with a lower percentage of discarded observations of 3%. The aim of this step is to include all observations $\{X_i : X_i'\beta(\tau) < C_i\}$ to build up the efficiency of the next step.

Step 3. We obtain the three-step estimator $\hat{\beta}^1(\tau)$ by running (17) with J_1 instead of J_0.

Step 4. (optional). Step 3 may be repeated a finite number of times on a sample $J_l = \{i : X_i'\hat{\beta}^{l-1}(\tau) < C_i - \delta_N\}$, yielding estimates $\hat{\beta}^l(\tau)$ for $l \in \{2, \dots\}$.

The remaining conditions discussed in Powell (1984, pp. 310-12) and Chernozhukov and Hong (2002, p. 876), most notably independence of observations and the conditional quantile restriction (3), ensure consistency and asymptotic normality. Due to distributional equivalence with the Powell estimator, the estimator inherits its efficiency properties. In contrast to the *iid* case, the variance-covariance matrix now depends on the error density at the τ-quantile, which may vary across the heteroskedastic observations.

One may wonder how much more restrictive three-step CQR is in comparison with Powell's canonical CQR model. For instance, does the parametric first-step classification model invoke additional assumptions which are at odds with, for example, the heteroskedasticity in our wage data? Since the initial model only provides an incorrect lower bound for the true propensity score, no such assumptions are required; neither a particular distribution nor conditional homoskedasticity nor a location-scale submodel are imposed.

In practice, Chernozhukov and Hong report good finite-sample properties in a variety of situations. Their results are rather insensitive to the choice of probability model in step 1; we therefore use a probit model in our application. Additional iterations in step 4 are akin to Buchinsky's ILPA method. Their usefulness at least partially depends on the dimensionality of the regression. When estimating a human capital model on the IABS data including only standard controls — years of education, a cubic experience term, and a quadratic tenure term, but no occupation dummies — step 4 generally does not bring about large improvements in terms of the Powell objective function (14). However, in our main specification, we include dummies to account for the effects of 130 different occupations. In this case, step 4 still turns out to yield substantial improvements. Additional iterations generally lead to only quite small or minuscule improvements, or even an increase in the objective function. In our application, we therefore allow for three additional iterations in step 4, and select the estimates corresponding to the lowest value of the Powell function.

3 The relationship between risk attitudes and occupational earnings risk

3.1 SOEP and IABS data

The German Socio-Economic Panel Study (SOEP) (Wagner, Frick, and Schupp, 2007) contains detailed information on household and individual characteristics, in particular on occupation and risk preferences. We estimate the earnings risk of occupations using a large administrative data set, the IAB Employment Sample (IABS), that contains information on wages and human capital variables collected for social insurance purposes. We restrict the sample to West German men in full-time employment between 25 and 54 years of age. Following Bonin et al. (2007), we discard observations for individuals with wages in the bottom and top percentile of the unconditional wage distribution from the SOEP sample. Since our primary interest lies in occupational choice within dependent employment, we do not consider the self-employed. As the IABS does not include tenured civil servants, we omit them from the SOEP sample as well. We exclude part-time workers, apprentices, and workers in retirement, as well as those employed in military service or alternative civilian service. Occupation identifiers are based on the German KldB coding.[11]

3.1.1 Risk attitudes

Recent waves of the SOEP contain self-reported measures of risk aversion in general and in specific contexts. In particular, individuals were asked to state their willingness to take risks on an eleven-point scale. The general risk question was included in 2004, 2006, 2008, and 2009.[12] Dohmen et al. (2011) confirm the validity of the general risk question by analyzing its predictive power for decisions in a lottery choice experiment with real money at stake, which they administer on a separate sample. In addition, all of the survey measures explain a variety of risky behaviors, including holding stocks, smoking, and participation in active sports. However, measurement error appears to be substantial; we expect to improve on the measurement of risk preferences by utilizing the responses from multiple waves. For each individual, we therefore compute an unweighted average of all available responses, the number of which varies between one and four.

[11] Both SOEP and IABS report occupations using the German *Klassifizierung der Berufe* ("Classification of occupations"), albeit in different versions: While the SOEP employs the 1992 version (KldB 92), the coding in the IABS is based on the earlier 1988 version (KldB 88). Specifically, the IABS uses the 3-digit level of the KldB 88 (*Berufsbezeichnungen*) identifying 328 different occupations; in our IABS R04 file, these are aggregated to 130 different occupations for anonymity purposes (Drews, 2008, pp. 79-86). In the SOEP, occupations are coded according to the 4-digit level of the KldB 92. To match occupations in the SOEP to those in the IABS, we use a cross walk file from the Federal Statistical Office.

[12] The exact wording (translated from German) is as follows: "How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: 'not at all willing to take risks' and the value 10 means: 'very willing to take risks'."

3.1.2 Riskiness of occupations

The IABS, an anonymized sample from a large administrative data set,[13] includes (gross) wage information as well as other employee characteristics reported by the employer for social security contribution purposes. Since misreporting of wages is subject to severe penalties, measurement error is likely to be minimal. The reporting precision of some of the independent variables in our human capital model may be considerably lower, since they are only collected and reported for statistical purposes, but with no pertinence to social security. The education variable in particular is sometimes missing or inconsistent for different employment spells of one individual. Since measurement error may affect our analysis by introducing possibly systematic noise to our dispersion estimates, we apply an imputation-based correction described in Fitzenberger, Osikominu, and Völter (2006). Tenure with an employer, on the other hand, can be computed with great precision due to the spell nature of the data.

For the entire sample period, a statutory limit was in place on the amount of monthly income subject to social security contributions, leading to top-coding of wages. This ceiling is a matter of federal legislation and is adjusted on an annual basis; for the main year of our analysis, 2004, it was EUR 61,800 in annual income. The degree of censoring remains fairly constant over the sample period at around 10%, but varies considerably by groups of age, education, and occupation.[14]

3.2 Mincerian human capital model

We choose a log-linear wage specification including years of education, a cubic polynomial for (potential) experience, and a quadratic polynomial for tenure. The inclusion of occupation dummies serves two purposes: First of all, it captures occupation-specific effects such as compensating wage differentials. More central to our analysis, the estimation of occupation effects at different quantiles is useful to evaluate wage dispersion at the occupation level. Since we do not intend to estimate returns to schooling or find any associated causal link, but rather the earnings variation observable to an individual, we neglect endogeneity bias arising from self-selection into different education streams and unobserved heterogeneity.

We estimate the model using (1) a Tobit ML approach, and (2)-(6) the three-step CQR estimator[15] at the 10th, 25th, 50th, 75th, and 90th percentile (table 1). We construct 95% confidence intervals using the direct percentile method with 100 bootstrap replications. Controlling for occupation, most of the variables have a fairly constant effect across quantiles; heterogeneous returns to education are likely mainly realized through occupational sorting. While we observe the familiar concave effect of experience, there is no clear-cut evidence that variation follows any particular experience pattern. Interestingly, variation is decreasing in tenure; we would instead expect that as employers acquire more knowledge about their workers, variation increases.

[13] In this paper, we use the IABS R04 version, which is a 2% sample of the German social security records for the period from January 1, 1975, to December 31, 2004. From this, we draw a cross section for June 30 for each year considered in our analyses.

[14] Overall, 10.5% of the observations in our sample have their wage information censored from above. The figure is 55.5% for university graduates aged 45 to 54, but as low as 0.4% for 25 to 34 year-olds without any degree.

[15] We use an extended version of the user-written Stata command *cqiv* (Chernozhukov et al., 2011).

However, the difference is rather small.

Under normality and homoskedasticity, the Tobit estimates should be very close to the CQR estimates at the 50th percentile. However, most of the estimates in column (1) and (4) are many standard errors apart, suggesting that these assumptions are not valid and the Tobit estimates biased. We further perform outer-product-of-the-gradient conditional moment tests (Skeels and Vella, 1999); the test statistics for the null hypotheses of normality and homoskedasticity follow $\chi^2(r)$ distributions with $r = 2$ and $r = 273$, respectively. The test statistic for normality, 4,099.3, is far larger than the theoretical 1% critical value of 9.2, and the same goes for homoskedasticity with a test statistic of 14,768.0 against a 1% critical value of 290.6. We thus reject the hypotheses of both normality and homoskedasticity.

3.3 Estimates for risk aversion and dispersion

We regress the wage dispersion within an individual's occupation, departing from a specification using the 10-90 spread, on stated willingness to take risks and a number of controls (table 2). We estimate a positive effect of willingness to take risks in all specifications, significant at either the 5% or the 10% level; the effect is larger for average risk attitudes, likely due to greater measurement precision: For the single-year and average measure of risk tolerance, we estimate an effect size of .0009 and .0018 standard deviations, respectively, per point increase on the 11-point risk scale (assuming normality for purely expositional purposes). Of the control variables, only marital status, education, and median wage[16] are significant.

Surprisingly, median wage enters negatively; it seems unlikely that this is entirely due to wage compression in the top regions of occupations with high median wage. Instead, observations in high-wage occupations are naturally more likely to be censored. If censoring leads to an underestimation of dispersion, this will be picked up by the coefficient on median wage. The effect of median wage is clearly the most pronounced at the 90th percentile (table 3), which is affected most by censoring; in fact, the unconditional 90th percentile is censored. As long as there is sufficient within-occupation heterogeneity that each of them will contain a number of individuals with uncensored conditional 90th percentile, this will not be a problem given correct specification of the model. Since median wage has a standard deviation of around .3, the magnitude of the distortion is not that large in principle; however, it is a lot larger than the effect of risk attitudes. Reassuringly, their effect is very similar when considering the 25-75 spread instead.

Turning our attention to individual median differences, the results in table 3 do not show any difference between the lower and the upper part of the wage distribution; all quantiles are similarly correlated with risk attitudes. Also here, we find a larger coefficient estimate on the average measure of risk tolerance. The significance of the second-stage estimates just presented is sensitive to the number of iterations used in step 4 of the 3-step CQR routine and thus the precision of our wage regressions; stopping at step 3, the results are slightly less clear.

Due to the nature of our data, which does not measure individuals' risk attitudes before

[16] For a few occupations (electrical engineers, managers/executives, doctors and pharmacists), median earnings are above the censoring threshold, and hence, median wage is censored. Omitting these occupations does not affect the results for overall dispersion and the upper regions, but slightly reduces the significance of the estimates at the lower quantiles.

they make their occupational choice, we cannot establish a causal impact of risk attitudes on occupational sorting. Any causal sorting interpretation of our results rests on some sort of stability assumption with respect to risk attitudes. In particular, systematic differences in risk attitudes of individuals working in different occupations should not entirely result from exposure to occupational risk over the working life.[17] Evidence indicating that risk preferences are rather stable is accumulating. Sahm (2008), for example, shows that risk preferences change only gradually with age but are rank-order stable. Changes in macroeconomic conditions have an impact on measured risk tolerance, but changes in income, wealth or other major events that reduce expected lifetime wealth, such as job displacement or a deterioration in health, do not affect individuals' willingness to take risk. Dohmen et al. (2007) analyze the stability of responses to the general risk question in the SOEP. For two subject pools, one a subset of the SOEP, the other a separate one, they find a test-retest correlation of 0.62 and 0.60, respectively, over a six-week horizon. It is plausible to assume that risk preferences do not change dramatically over such a short time period so that the variation in answers in the test-retest samples can be attributed to measurement error. The correlation between the 2004 and 2006 waves of the SOEP, in comparison, is 0.50, which is not too far below the six-week benchmark; this suggests that risk attitudes constitute an inherent and stable trait. Beauchamp, Cesarini, and Johannesson (2011) support this interpretation, as they find very similar results for Swedish data using the same risk measure as is used in the SOEP.

In our setting, a sorting interpretation also requires that the ranking of occupations with respect to their occupational earnings risk has remained stable. Otherwise, the risk profile estimated on the 2004 cross section may not have been relevant at the time when individuals chose their occupation. In an extreme case, risk attitudes might not have been related to differences in occupation-specific earnings risk when individuals sorted into an occupation. Instead, a wage setting mechanism in which preferences of incumbents shape the occupational earnings risk might be a potential channel through which a correlation between risk preferences and wage dispersion can arise. To address the question whether there have been considerable changes in occupation-specific wage dispersion, we estimate wage dispersion measures for the years 1979, 1984, 1989, 1994, and 1999, and compute Pearson correlation coefficients with occupations as cross-sectional unit (table 4). The correlation coefficient is decreasing in the time span considered, but remains positive and high; it is larger than .65 for any pair of years. This suggests that the relative wage dispersion of occupations has been rather stable in West Germany in the period from 1979 to 2004, and that the risk profiles we estimate from a cross section for 2004 are quite close to those relevant at the point of labor market entry for most of the individuals in the SOEP.

Finally, we cannot rule out that the correlation between risk attitudes and wage dispersion is driven by cognitive abilities rather than risk preferences: There is evidence for a negative relationship between risk aversion and cognitive abilities (e.g., Dohmen et al., 2010), and at the same time, dispersion may be particularly attractive for high-ability individuals.

[17] Fouarge, Kriechel, and Dohmen (2011) elicited risk attitudes among school leavers and graduates around the time that they entered the labor market, reducing the potential for endogeneity problems resulting from exposure to occupation characteristics. They find a significant correlation between occupational earnings risk and risk attitudes.

4 Conclusion

We discuss a particular method to estimate group-level wage dispersion, which is based on semiparametric methods. Specifically, we estimate a human capital model, including dummy variables for each of the groups of interest, at a number of different quantiles; we then take the differences of the dummy coefficients at different quantiles as a measure of dispersion within each group. The method is particularly useful when working with data which is either censored or top-coded, such as administrative data and some survey data, since it is more robust to deviations from homoskedasticity and distributional assumptions than parametric estimators. In addition, it controls for the dispersion effect of covariates, and allows us to estimate the entire conditional wage distribution and its differences across groups. In an application which connects a large German administrative data set, the IAB Employment Sample (IABS), which is subject to censoring due to a legislative contribution limit, and a household survey, we find that individuals with greater willingness to take risks work in occupations with higher cross-sectional wage dispersion.

References

Amemiya, Takeshi. 1973. Regression Analysis when the Dependent Variable is Truncated Normal. *Econometrica* 41:997–1016.

Angrist, Joshua, Victor Chernozhukov, and Iván Fernández-Val. 2006. Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure. *Econometrica* 74:539–63.

Arabmazar, Abbas and Peter Schmidt. 1981. Further evidence on the robustness of the Tobit estimator to heteroskedasticity. *Journal of Econometrics* 17:253–58.

———. 1982. An Investigation of the Robustness of the Tobit Estimator to Non-Normality. *Econometrica* 50:1055–63.

Bassett, Gilbert, Jr. and Roger Koenker. 1978. Asymptotic Theory of Least Absolute Error Regression. *Journal of the American Statistical Association* 73:618–22.

Beauchamp, Jonathan, David Cesarini, and Magnus Johannesson. 2011. The Psychometric Properties of Measures of Economic Risk Preferences. Available on request via `https://files.nyu.edu/dac12/public/`.

Bellante, Don and Albert N. Link. 1981. Are Public Sector Workers More Risk Averse Than Private Sector Workers? *Industrial and Labor Relations Review* 34:408–12.

Belzil, Christian and Jörgen Hansen. 2004. Earnings dispersion, risk aversion and education. In *Accounting for Worker Well-Being*, *Research in Labor Economics*, vol. 23, edited by Solomon W. Polachek. Bingley, UK: Emerald Group Publishing Limited, 335–58.

Belzil, Christian and Marco Leonardi. 2007. Can risk aversion explain schooling attainments? Evidence from Italy. *Labour Economics* 14:957–70.

———. 2009. Risk aversion and schooling decisions. Working Papers hal-00411099, HAL.

Bonin, Holger, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde. 2007. Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. *Labour Economics* 14:926–37.

Brown, Charles and Robert Moffitt. 1983. The Effect of Ignoring Heteroscedasticity on Estimates of the Tobit Model. NBER Technical Working Papers 0027, National Bureau of Economic Research, Inc.

Buchinsky, Moshe. 1994. Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression. *Econometrica* 62:405–58.

———. 1998. Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. *Journal of Human Resources* 33:88–126.

Buchinsky, Moshe and Jinyong Hahn. 1998. An Alternative Estimator for the Censored Quantile Regression Model. *Econometrica* 66:653–71.

Caliendo, Marco, Frank Fossen, and Alexander Kritikos. 2009. Risk attitudes of nascent entrepreneurs: New evidence from an experimentally validated survey. *Small Business Economics* 32:153–67.

Chamberlain, Gary. 1994. Quantile regression, censoring, and the structuring of wages. In *Advances in Econometrics: Sixth World Congress*, vol. 1, edited by Christopher A. Sims, chap. 5. Cambridge, UK: Cambridge University Press, 171–209.

Chay, Kenneth Y. and Bo E. Honoré. 1998. Estimation of Semiparametric Censored Regression Models: An Application to Changes in Black-White Earnings Inequality during the 1960s. *Journal of Human Resources* 33:4–38.

Chen, Stacey H. 2008. Estimating the Variance of Wages in the Presence of Selection and Unobserved Heterogeneity. *The Review of Economics and Statistics* 90:275–89.

Chernozhukov, Victor, Iván Fernández-Val, Sukjin Han, and Amanda E. Kowalski. 2011. Stata Command to Implement CQIV. Available at `http://www.econ.yale.edu/~ak669/research.html`.

Chernozhukov, Victor and Han Hong. 2002. Three-Step Censored Quantile Regression and Extramarital Affairs. *Journal of the American Statistical Association* 97:872–82.

Cramer, Jan S., J. Hartog, Nicole Jonker, and C. Mirjam van Praag. 2002. Low risk aversion encourages the choice for entrepreneurship: an empirical test of a truism. *Journal of Economic Behavior & Organization* 48:29–36.

Dohmen, Thomas and Armin Falk. 2010. You Get What You Pay For: Incentives and Selection in the Education System. *Economic Journal* 120:256–71.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review* 100:1238–60.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2007. The Measurement and Stability of Risk Attitudes.

———. 2011. Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. *Journal of the European Economic Association* 9:522–50.

Donald, Stephen G. 1995. Two-step estimation of heteroskedastic sample selection models. *Journal of Econometrics* 65:347–80.

Drews, Nils. 2008. Das Regionalfile der IAB-Beschäftigtenstichprobe 1975-2004: Handbuch-Version 1.0.3. FDZ Datenreport 02/2008(DE), Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nuremberg, Germany. Available at `http://doku.iab.de/fdz/reporte/2008/DR_02-08.pdf`.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg. 2009. Revisiting the German Wage Structure. *Quarterly Journal of Economics* 124:843–81.

Ekelund, Jesper, Edvard Johansson, Marjo-Riitta Jarvelin, and Dirk Lichtermann. 2005. Self-employment and risk aversion — evidence from psychological test data. *Labour Economics* 12:649–59.

Fitzenberger, Bernd. 1997. Computational Aspects of Censored Quantile Regression. *Lecture Notes-Monograph Series* 31:171–86.

Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter. 2006. Imputation Rules to Improve the Education Variable in the IAB Employment Subsample. *Schmollers Jahrbuch* 126:405–36.

Fouarge, Didier, Ben Kriechel, and Thomas Dohmen. 2011. Occupational sorting of school leavers: The role of economic preferences.

Fuchs-Schündeln, Nicola and Matthias Schündeln. 2005. Precautionary Savings and Self-Selection: Evidence from the German Reunification Experiment. *Quarterly Journal of Economics* 120:1085–1120.

Goldberger, Arthur. 1983. Abnormal selection bias. In *Studies in Econometrics, Time Series, and Multivariate Statistics*, edited by Samuel Karlin, Takeshi Amemiya, and Leo Goodman. New York, NY: Academic Press, 67–85.

Guiso, Luigi and Monica Paiella. 2005. The Role Of Risk Aversion In Predicting Individual Behavior. Temi di discussione (Economic working papers) 546, Bank of Italy, Economic Research Department.

Heckman, James J. 1976. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. In *Annals of Economic and Social Measurement, NBER Chapters*, vol. 5. National Bureau of Economic Research, Inc, 120–37.

———. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47:153–61.

Hu, Luojia. 2002. Estimation of a Censored Dynamic Panel Data Model. *Econometrica* 70:2499–2517.

Hurd, Michael. 1979. Estimation in truncated samples when there is heteroscedasticity. *Journal of Econometrics* 11:247–58.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy* 101:410–42.

Katz, Lawrence F. and Kevin M. Murphy. 1992. Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics* 107:35–78.

Khan, Shakeeb and James L. Powell. 2001. Two-step estimation of semiparametric censored regression models. *Journal of Econometrics* 103:73–110.

King, Allan G. 1974. Occupational Choice, Risk Aversion, and Wealth. *Industrial and Labor Relations Review* 27:586–96.

Koenker, Roger and Gilbert Bassett, Jr. 1978. Regression Quantiles. *Econometrica* 46:33–50.

———. 1982. Robust Tests for Heteroscedasticity Based on Regression Quantiles. *Econometrica* 50:43–61.

Koenker, Roger and Kevin F. Hallock. 2001. Quantile Regression. *Journal of Economic Perspectives* 15:143–56.

Koenker, Roger and Zhijie Xiao. 2002. Inference on the Quantile Regression Process. *Econometrica* 70:1583–1612.

Kowalski, Amanda E. 2009. Censored Quantile Instrumental Variable Estimates of the Price Elasticity of Expenditure on Medical Care. NBER Working Papers 15085, National Bureau of Economic Research, Inc.

Levhari, David and Yoram Weiss. 1974. The Effect of Risk on the Investment in Human Capital. *American Economic Review* 64:950–63.

Maddala, G. S. and Forrest D. Nelson. 1975. Specification Errors in Limited Dependent Variable Models. NBER Working Papers 0096, National Bureau of Economic Research, Inc.

Melly, Blaise. 2005. Decomposition of differences in distribution using quantile regression. *Labour Economics* 12:577–90.

Mincer, Jacob A. 1991. Education and Unemployment. NBER Working Papers 3838, National Bureau of Economic Research, Inc.

Nickell, Stephen and Brian Bell. 1996. Changes in the Distribution of Wages and Unemployment in OECD Countries. *The American Economic Review* 86:302–8.

Paarsch, Harry J. 1984. A Monte Carlo comparison of estimators for censored regression models. *Journal of Econometrics* 24:197–213.

Powell, James L. 1984. Least absolute deviations estimation for the censored regression model. *Journal of Econometrics* 25:303–25.

———. 1986a. Censored regression quantiles. *Journal of Econometrics* 32:143–55.

Powell, James L. 1986b. Symmetrically Trimmed Least Squares Estimation for Tobit Models. *Econometrica* 54:1435–60.

Sahm, Claudia R. 2008. How Much Does Risk Tolerance Change? Finance and Economics Discussion Series 2007-66, Board of Governors of the Federal Reserve System (U.S.).

Saks, Raven E. and Stephen H. Shore. 2005. Risk and Career Choice. *The B.E. Journal of Economic Analysis & Policy* 5.

Schmillen, Achim and Joachim Möller. 2012. Distribution and determinants of lifetime unemployment. *Labour Economics* 19:33–47.

Schulhofer-Wohl, Sam. 2011. Heterogeneity and Tests of Risk Sharing. *Journal of Political Economy* 119:925–58.

Shaw, Kathryn L. 1996. An Empirical Analysis of Risk Aversion and Income Growth. *Journal of Labor Economics* 14:626–53.

Skeels, Christopher L. and Francis Vella. 1999. A Monte Carlo investigation of the sampling behavior of conditional moment tests in Tobit and Probit models. *Journal of Econometrics* 92:275–94.

Tobin, James. 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica* 26:24–36.

Vijverberg, Wim P. M. 1987. Non-normality as Distributional Misspecification in Single-Equation Limited Dependent Variable Models. *Oxford Bulletin of Economics and Statistics* 49:417–30.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp. 2007. The German Socio-Economic Panel Study (SOEP) — Scope, Evolution and Enhancements. *Schmollers Jahrbuch* 127:139–69.

Wilhelm, Mark Ottoni. 2008. Practical Considerations for Choosing Between Tobit and SCLS or CLAD Estimators for Censored Regression Models with an Application to Charitable Giving. *Oxford Bulletin of Economics and Statistics* 70:559–82.

A Results

Table 1: Human capital model

Dependent variable: log of censored daily wage						
	(1)	(2)	(3)	(4)	(5)	(6)
Experience	0.064***	0.058***	0.060***	0.058***	0.057***	0.050***
	(0.002)	(0.006)	(0.005)	(0.003)	(0.004)	(0.005)
Experience^2/100	-0.236***	-0.228***	-0.224***	-0.211***	-0.202***	-0.170***
	(0.008)	(0.031)	(0.024)	(0.016)	(0.020)	(0.027)
Experience^3/1000	0.029***	0.029***	0.028***	0.026***	0.024***	0.020***
	(0.001)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
Tenure	0.025***	0.039***	0.027***	0.020***	0.018***	0.016***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Tenure^2/100	-0.057***	-0.091***	-0.061***	-0.043***	-0.038***	-0.035***
	(0.001)	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)
Years of Education	0.056***	0.049***	0.050***	0.052***	0.052***	0.049***
	(0.000)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
Constant	3.305***	3.004***	3.203***	3.419***	3.666***	3.952***
	(0.012)	(0.050)	(0.037)	(0.026)	(0.036)	(0.068)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	168,863	168,863	168,863	168,863	168,863	168,863

Human capital model estimated using (1) Tobit ML approach, and (2)-(6) the three-step CQR estimator at the 10$^{\text{th}}$, 25$^{\text{th}}$, 50$^{\text{th}}$, 75$^{\text{th}}$, and 90$^{\text{th}}$ percentile. Standard errors of coefficient estimates, computed using direct percentile (bootstrap) method with 100 replications and asymptotic normality, in parentheses; ***/**/* indicate significance at 1%/5%/10% level. Dependent variable is the log of censored daily wage. Occupation dummies based on IABS KldB 88 coding.

Table 2: IABS risk profiles

Dependent variable: 10-90 spread of occupation effects						
	(1)	(2)	(3)	(4)	(5)	(6)
General Risk Attitude	0.003**	0.002*	0.003*			
	(0.002)	(0.001)	(0.002)			
General Risk Attitude (av.)				0.005**	0.005**	0.005**
				(0.002)	(0.002)	(0.002)
Experience		0.000	0.000		0.000	0.000
		(0.000)	(0.000)		(0.000)	(0.000)
Tenure		-0.000	0.000		-0.000	0.000
		(0.000)	(0.000)		(0.000)	(0.000)
Years of Education		0.002	0.010**		0.002	0.010**
		(0.004)	(0.004)		(0.004)	(0.004)
Married and living together		-0.017**	-0.016**		-0.017**	-0.017**
		(0.007)	(0.007)		(0.007)	(0.007)
Body height		0.000	0.001		0.000	0.001
		(0.000)	(0.000)		(0.000)	(0.000)
Public Sector Employment		-0.015	-0.017		-0.014	-0.016
		(0.026)	(0.026)		(0.026)	(0.026)
Median wage (occ.)			-0.122*			-0.122*
			(0.068)			(0.068)
Constant	-0.188***	-0.206**	0.164	-0.198***	-0.221**	0.150
	(0.014)	(0.087)	(0.264)	(0.014)	(0.086)	(0.261)
Observations	2,815	2,740	2,740	2,822	2,747	2,747
R-squared	0.002	0.009	0.051	0.004	0.011	0.054

OLS estimates. Robust standard errors of coefficient estimates, allowing for clustering at the IABS KldB 88 occupation level, in parentheses; ***/**/* indicate significance at 1%/5%/10% level. Dependent variable is the 10-90 spread of occupation dummy estimates from QR, on IABS KldB 88 level. "General Risk Attitude" is the response to the 2004 general risk question, "General Risk Attitude (av.)" is the corresponding average over the 2004, 2006, 2008, and 2009 responses.

Table 3: IABS risk profiles

Dependent variable: median differences of occupation effects								
	10-50		25-50		50-75		50-90	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
General Risk Attitude	0.002*		0.000		0.001**		0.002*	
	(0.001)		(0.000)		(0.000)		(0.001)	
General Risk Attitude (av.)		0.003**		0.001*		0.002***		0.003**
		(0.001)		(0.001)		(0.001)		(0.001)
Experience	0.000	0.000	0.000	0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tenure	-0.000	-0.000	0.000	0.000	0.000**	0.000**	0.000*	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of Education	0.005*	0.005*	0.002*	0.002*	0.003*	0.003*	0.005**	0.005**
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Married and living together	-0.009**	-0.009**	-0.003	-0.004	-0.005*	-0.005*	-0.008**	-0.008**
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Body height	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Public Sector Employment	-0.000	-0.000	-0.009	-0.009	-0.001	-0.001	-0.016	-0.016
	(0.016)	(0.016)	(0.007)	(0.007)	(0.009)	(0.009)	(0.012)	(0.012)
Median wage (occ.)	-0.012	-0.012	-0.009	-0.009	-0.002	-0.002	-0.110***	-0.110***
	(0.048)	(0.048)	(0.022)	(0.022)	(0.022)	(0.022)	(0.032)	(0.032)
Constant	-0.114	-0.123	-0.046	-0.049	-0.109	-0.112	0.278**	0.273**
	(0.209)	(0.208)	(0.096)	(0.095)	(0.084)	(0.083)	(0.122)	(0.121)
Observations	2,740	2,747	2,740	2,747	2,740	2,747	2,740	2,747
R-squared	0.025	0.027	0.020	0.022	0.032	0.034	0.133	0.135

OLS estimates. Robust standard errors of coefficient estimates, allowing for clustering at the IABS KldB 88 occupation level, in parentheses; ***/**/* indicate significance at 1%/5%/10% level. Dependent variable are the median differences of occupation dummy estimates from QR, on IABS KldB 88 level. "General Risk Attitude" is the response to the 2004 general risk question, "General Risk Attitude (av.)" is the corresponding average over the 2004, 2006, 2008, and 2009 responses.

Table 4: Temporal stability of risk profiles

	1979	1984	1989	1994	1999	2004
1979	1.000					
1984	0.818	1.000				
1989	0.768	0.902	1.000			
1994	0.781	0.854	0.874	1.000		
1999	0.654	0.778	0.810	0.899	1.000	
2004	0.655	0.694	0.753	0.896	0.922	1.000

Pearson correlation coefficients of 10-90 spread per occupation across years.

B Monte Carlo evidence on estimation of dispersion

In this section, we review the performance of the estimation method described in section 2.2 for both censored and uncensored data and compare it to a residual-based method. For uncensored and censored data, we use the difference between the coefficient estimates of the group dummies at the 90th and 10th percentile from (C)QR. For uncensored data only, we estimate a conventional OLS regression including group dummies and compute the standard deviation of residuals per group. After each of 1,000 simulations, we compute the correlation of an occupation-specific scale σ_k and the three statistics. The models investigated are stylized versions of the wage distribution setting in our empirical analysis; specifically, we first consider a model with only group-specific scale, and then turn to location-scale models in which a regressor has a heteroskedastic effect. The censored data is derived directly from the uncensored data through right-censoring at the 90th percentile such that in each case, 10% of the data are censored, which is intended to resemble the degree of censoring in the IABS data used in our application.

B.1 Group-specific scale model

The DGP has the following linear representation:

$$Y_{ik} = \beta X_i + c_k + \sigma_k \epsilon_i \tag{18}$$

for $i \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, m\}$, $c_k \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$, $\epsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$.

In our simulation, we set $N = 1,000$, $X_i \overset{\text{iid}}{\sim} \mathcal{U}(0, 1)$, $\beta = 2$, $\sigma = .2$, and $\sigma_k = .1 + .2U_k$, where $U_k \overset{\text{iid}}{\sim} \mathcal{U}(0, 1)$. Individuals are randomly assigned to one of $m = 10$ groups according to a uniform distribution.

Table 5 shows a very similar performance for all three statistics, with a correlation close to unity in each case.

Table 5: Group-specific scale model

	Mean	Std. Dev.	Min.	Max.
10-90 spread	0.975	0.021	0.794	0.999
10-90 spread (cens.)	0.968	0.026	0.729	0.999
Resid. std. dev.	0.984	0.013	0.881	0.999

Correlation of risk measures with true dispersion.

B.2 Linear location-scale model

Leaving all else the same,

$$Y_{ik} = \beta X_i + c_k + (\gamma X_i + \sigma_k)\epsilon_i \tag{19}$$

with $\gamma = .5$. Hence, the independent variable X now exerts a heteroskedastic effect.

Already, the 10-90 spread does slightly better (table 6) for the uncensored data. Notably, it also works just as well when only censored data is available.

Table 6: Linear location-scale model

	Mean	Std. Dev.	Min.	Max.
10-90 spread	0.899	0.072	0.228	0.993
10-90 spread (cens.)	0.891	0.076	0.198	0.994
Resid. std. dev.	0.852	0.107	0.159	0.988

Correlation of risk measures with true dispersion.

B.3 Nonlinear location-scale model

We adapt the DGP such that the scale effect of the independent variable X is now negatively related to the occupation variance:

$$Y_{ik} = \beta X_i + c_k + [(1 - \delta\sigma_k)X_i + \sigma_k]\epsilon_i. \tag{20}$$

For $\delta = 1$,

$$Y_{ik} = \beta X_i + c_k + [X_i + (1 - X_i)\sigma_k]\epsilon_i. \tag{21}$$

As reported in table 7, the statistics based on QR are a lot more robust in this case, since the scale effect of X at the different quantiles is explicitly controlled for. The discrepancy will likely be even larger for more irregular distributions. Also, our method for dispersion estimation works equally well for censored data.

Table 7: Nonlinear location-scale model

	Mean	Std. Dev.	Min.	Max.
10-90 spread	0.819	0.112	0.217	0.991
10-90 spread (cens.)	0.837	0.105	0.262	0.992
Resid. std. dev.	0.682	0.191	-0.347	0.973

Correlation of risk measures with true dispersion.