Censored Quantile Regressions and the Length of Unemployment Periods in West Germany

Elke Lüdemann*

Ralf A. Wilke†

Xuan Zhang‡

April 2005

*Mannheim University and Zentrum für Europäische Wirtschaftsforschung (ZEW Mannheim). E–mail: elke.luedemann@gmx.de

†ZEW Mannheim – Corresponding author: Ralf Wilke, ZEW Mannheim, P.O. Box 10 34 43, 68034 Mannheim, Germany. E–mail: wilke@zew.de.

‡Mannheim University and ZEW Mannheim. E–mail: xuanz98@rumms.uni-mannheim.de

We gratefully acknowledge financial support by the German Research Foundation (DFG) through the research project "Microeconometric modelling of unemployment durations under consideration of the macroeconomic situation". We thank Andreas Ammermüller, Wolfgang Franz, Guillaume Horny, and two anonymous referees for their comments. We are indebted to Bernd Fitzenberger for his stimulating ideas and for the joint work on the project. We also thank a number of employees in the employment offices at Duisburg, Freiburg, Gelsenkirchen, Mannheim, and Oberhausen for many useful hints and information. All errors are our sole responsibility.

Abstract

In this paper, we estimate the effect of different macro and micro variables on the distribution of unemployment duration in West Germany using censored quantile regressions. We analyze unemployment periods of more than 91,000 observations from the years 1981 to 1997 drawn from the IAB employment subsample. The latter is an administrative data set that is representative with respect to the socially insured workforce. Surprisingly, we find that the educational degree and variables indicating the macroeconomic environment such as the unemployment rate have a weak effect only. On the other hand, variables reflecting the (un-)employment history of an individual such as the length of tenure, recall to the same employer in the past, recent unemployment and the position in the population income distribution before unemployment have the strongest effects on unemployment duration. We conclude that work history variables are the ones most suitable in characterizing the unemployment duration of an individual. From a methodological point of view, it is interesting that some regression coefficients have a different sign depending on the quantiles of the unemployment duration distribution. This clearly is a violation of the classical proportional hazard assumption which is very common in unemployment duration analysis.
Keywords: censored quantile regression, unemployment duration, administrative data
JEL: C24, J64

1 Introduction

During the past 25 years, the unemployment rate in West Germany has risen dramatically. Starting at 4.8% in 1981 it almost doubled to 9.5% in 1997, its highest level ever[1].
This paper looks at administrative data on unemployment duration in the period from 1981 to 1997. Applying censored quantile regressions to unemployment duration analysis, it is our purpose to examine the following key issues of interest for labour market policies: Did the increase in the unemployment rate in the 1980s and 1990s come along with a general elongation of unemployment periods? And, more importantly, which micro and macro variables had the strongest effects on the distribution of unemployment duration?

So far, a large number of papers on unemployment duration analysis of West German data have been published. Most of these papers analyze the effect of the German unemployment compensation system on the duration of unemployment. In particular, the reform of this system during the 1980s was subject to many analyses. During this reform, the

[1]The unemployment rate in 2005 is only higher due to a change in the official statistics.

entitlement to unemployment compensation was extended for the older unemployed (aged 42 and older) (for detailed analyses see e.g. Hunt (1995), Hujer and Schneider (1996), and recently Plaßmann (2002) and Fitzenberger and Wilke (2004)). The results of these papers (except for Fitzenberger and Wilke, 2004) are likely affected by the early retirement of elderly workers. For a comprehensive study of incidence and duration of unemployment with regard to early retirement in West Germany see Fitzenberger and Wilke (2004). However, estimation results may differ depending on whether unemployed who are in fact early retired are included in the analysis or not. As opposed to most of the papers to date, we restrict our analysis to the unemployed aged 26 to 41, since the latter are not directly affected by the labor market reforms undertaken in West Germany in the period considered or by the early retirement issue[2]. It is not our purpose to evaluate a specific policy measure or reform, but to obtain a better understanding of the determinants of the distribution of unemployment duration in general, which we believe is helpful for the design of labor market policies.

Moreover, the present paper differs from the papers to date on unemployment duration analysis for West Germany in two important respects:

First, most of the papers to date are based on the German Socio-Economic Panel (GSOEP), which is survey data (e.g. Hunt (1995), Steiner (2001) or Lauer (2003)). Unemployment duration data drawn from surveys have several drawbacks compared to administrative data. This is due to the limited sample size, the imperfect recall of the interviewed individuals and due to misleading replies. See for example Schräpler (2002) who analyzes non-response behavior in the GSOEP and Jürges (2004) who finds considerable inconsistencies in the unemployment information. To circumvent these limitations, we use German administrative data in this paper. However, it should be noted that there are also downsides to using administrative data such as the limited number of variables or possible measurement errors in the exploratory variables. In addition, the sample we use, namely the IABS employment subsample, is representative only with respect to the socially secured workforce as opposed to the GSOEP which is representative with respect to the full population. These differences may account for differing estimation results.

Second and most importantly, the econometric model used in the present paper is different from the papers to date. In fact, the present paper is the first one applying quantile

[2]For the same reason Hunt (1995) excludes older unemployed (aged > 58) from her analysis and she estimates a competing risks model with exits to employment and retirement for the remaining population. Our data does not provide exact information about retirement times. For this reason it is impossible to estimate a competing risks model and therefore we decided to choose a lower upper bound for the age restriction.

regressions to German unemployment data. In particular, we use censored quantile regressions as suggested by Koenker and Bilias (2001). In contrast, in the majority of the papers to date, single spell proportional hazard models have been used. It is well known that estimation results of single spell proportional hazard models that account for unobserved heterogeneity (mixed proportional hazard model) are sensitive to marginal changes in the model specification (see van den Berg (2001) for a survey). At the same time, duration models that do not account for unobserved heterogeneity are expected to be inconsistent. In contrast, we use quantile regression methods because they are robust with respect to an additive error distribution. Moreover, in contrast to mean value methods and many typical duration models [3], they allow us to examine whether the effect of a regressor varies or even changes its sign over the quantiles of the unemployment duration distribution. Wilke (2005) explores whether there are disproportional changes over the duration time and over the calendar time and he finds some indications for this since in some cases the survivor functions cross. A violation of the proportionality property would mean that the proportional hazard specification imposes an empirically unsupported restriction on the nature of the effect of covariates. Our estimation results show that this is indeed the case for some regressors since the estimated coefficients change their sign over the quantiles. Using quantile regression to analyze survival times offers therefore a valuable complement to traditional proportional hazard modeling and helps to reduce the risk of misspecification in parametric duration models. However, it comes at the cost that one cannot allow for time varying regressors and that it is not yet possible to account for several exit states and unobserved heterogeneity. Further research is necessary on these issues.

However, the use of quantile regression techniques in unemployment duration analysis is particularly valuable for the design of labour market policies since the survival times themselves are often of fundamental importance and the quantile regression coefficients are interpretable as direct regression effects on the survival times (or on their transformation). Furthermore, quantile regression methods allow us to consider how the impact of the regressors vary for different parts of the distribution of unemployment duration. Thus, one advantage is that we get hints as to how specific policy measures such as those targeting at a reduction of long-term unemployment should be designed.

So far, only few studies have worked with German administrative data and have not applied proportional hazard models or related models (e.g. Fahrmeir et al. (2003) use semiparametric splines, and Fitzenberger and Wilke (2004) and Wilke (2005) use purely

[3] We do not mean duration models with time varying covariates.

nonparametric methods). Therefore, the question arises whether the results of the studies change when a robust and flexible estimation method such as censored quantile regressions is applied to a large administrative data set.

Furthermore, we believe that our results can provide interesting contributions to the current policy discussion of the so-called "Hartz"-reforms, which have been implemented in Germany as of 2003. The latter comprise far-reaching institutional and legal reforms of the German unemployment compensation system, aiming at an activation of the unemployed and a reduction of unemployment duration (see Hartz, 2002). After presenting our estimation results, we will discuss the relevance of our findings to this current policy discussion.

The remainder of the paper is organized as follows: In section 2, the data set and the relevant German institutional framework is described in detail. In section 3, the econometric model is set out, which is followed by the description and discussion of the estimation results in section 4. Section 5 concludes.

2 Data and Institutions

The analysis is based on German administrative data containing spell information of employment and un-/nonemployment trajectories of about $500,000$ individuals from West Germany. [4] More specifically, we use the IAB employment subsample 1981-1997 -regional file- [5] for our analysis, from which we draw a specific subsample described later. The IAB employment subsample is representative with respect to the socially insured working population. However, it does not contain periods of self-employment and of employment as life-time civil servant (Beamte). The data provides daily information about the beginning and the end dates of socially secured employment as well as unemployment provided that any form of unemployment compensation from the federal employment office (BA) is received.

Until 2004, the German unemployment compensation has consisted of unemployment benefits (Arbeitslosengeld, ALG), unemployment assistance (Arbeitslosenhilfe, ALHi) and maintenance payments during further training (Unterhaltsgeld, UHG). Note that periods of further training are therefore counted as unemployment. Other labor market programmes such as subsidized employment are not identifiable in the data. During the years 1981 to 1997, about 65% to 75 % of all registered unemployed have drawn ALG or ALHi.

In the period we analyze, an employee qualified for unemployment benefits after having

[4] In this analysis an individual is said to be West German if the last employment period before unemployment was in West Germany.

[5] For a general description of the data see Bender et al. (2000).

been in socially secured employment for at least 12 months during the past three years. The monthly amount of unemployment benefits was between 60% and 68% of the previous net monthly wage. After having exhausted the maximum entitlements for unemployment benefits or in case of not being entitled, an unemployed person could draw unemployment assistance, which was means-tested and in case of entitlement also related to the previous net wage (53% - 58% in the period under consideration). Unemployment assistance was provided for an unlimited period but the entitlements were regularly checked. If an unemployed participated in the meantime in further training measures he received specific payments (UHG) during this period. The payment scheme for UHG was related to the one of ALG. For a more detailed description of the German unemployment compensation system see Hunt (1995) or Plaßmann (2002). However, there is no information about the amount of unemployment compensation received in our data set. We only have the more general information if unemployment compensation is drawn or not. However, we indirectly control for the unemployment compensation by using variables computed from the work history of the individuals such as the position in the population income distribution before unemployment.

Moreover, registered unemployment is not recorded in the IAB employment subsample and therefore one cannot precisely distinguish between unemployment and nonemployment periods because unemployment periods without receipt of unemployment compensation from the BA are not observed.

For our analysis, we adopt the definition of "Nonemployment" as introduced by Fitzenberger and Wilke (2004). Nonemployment is any period after an employment period, in which an individual is not (socially secured) employed and receives at least for one day some kind of unemployment compensation from the federal employment office. The latter condition ensures that at least a part of each nonemployment period overlaps with unemployment and rules out purely out-of-the-labor-market periods. At the same time, this results in a sample selection by excluding all unemployment spells without the receipt of unemployment compensation. In fact, the same data is used as in Wilke (2005), but he considers only four years (1981, 1985, 1990 and 1995). Due to the limitations of the underlying register information, our sample of unemployment periods is not representative for all unemployment periods. Moreover, this imperfect proxy of unemployment may influence the estimation results and the latter may deviate from what we would have obtained had the true unemployment period been observed. Using this definition of nonemployment, unemployed which are not entitled for compensation payments from the BA are not considered. However, out-of-the-labor-market-periods may be included in the analysis. Conditioning on employment before unemployment and on the receipt of transfer payments from the BA, we have a preselection of unemployment periods. Moreover, it should be noted that for some groups, the length of unemployment periods is systematically upward biased. This is in particular the case for individuals who are likely to drop out of the labor force for some period, e.g. females in motherhood. Furthermore, there are right-censored nonemployment spells in the data, if the last observed spell of an individual is the receipt of unemployment compensation. We account for right censoring by using censored quantile regressions, a method which will be described in the following section.

We restrict our analysis to unemployment spells starting between 1981 and 1995 of individuals aged 26 to 41 during this period. This restriction is chosen in order to obtain a relatively homogeneous subsample: all individuals are entitled to draw unemployment benefits for at most 12 months, they are too young to be affected by the early retirement issue (Fitzenberger and Wilke, 2004) and too old to be affected by policy measures against youth unemployment.

After selecting observations according to the criteria mentioned above, our sample contains 91,035 observations. For descriptive statistics of the variables used, see Table 5. Unconditional nonparametric quantile functions estimated by the Kaplan-Meier estimator for four calender years are given in figure 1. The quantile functions are higher in 1981 and 1995 than in 1985 and 1990. These descriptive figures do suggest that there is no direct relationship between the unemployment rate and the length of unemployment periods. We shall investigate this in more detail in section 4. Note that the systematic censoring at the end of 1997 affects the results for the top quantiles in 1995.

3 Econometric Model

Quantile regression (QR) is gradually evolving into a comprehensive approach to the statistical duration analysis based on methods to model the quantiles of the response variable conditional on the covariates. Just as classical linear regression methods are usually used to estimate a general class of models for conditional mean functions, quantile regression methods offer a mechanism for estimating models for the conditional median function and the full range of other conditional quantile functions. In contrast to mean value methods and standard proportional hazard models such as the Cox model and the Accelerated Failure Time model QR allow us to obtain different effects of the covariates at different points of the conditional unemployment duration distribution. The advantages of QR based duration analysis are summarized in Koenker and Geling (2001). Koenker and Bilias (2001) and Koenker and Xiao (2002) discuss applications to unemployment duration models and some general problems of inference based on the quantile regression process.

3.1 Quantile regression model

The quantile regression model, first introduced by Koenker and Bassett (1978), can be viewed as a location model. Let y denote the unemployment duration. We model the conditional quantile functions of the logarithm of unemployment duration as linear in the observed covariates, x,

$$lny_i = x_i'\beta^\theta + u_i^\theta, \quad i = 1, ..., N \tag{1}$$

with

$$Quant_\theta(lny_i|x_i) = x_i'\beta^\theta, \tag{2}$$

where x_i is a $k \times 1$ vector of covariates with $x_{1i} \equiv 1$ for all i and β^θ is a vector of regression parameters. The term $Quant_\theta(lny_i|x_i)$ denotes the θth conditional quantile of lny given x. Here u is defined by $u_\theta \equiv lny - x'\beta_\theta$, so that $Quant_\theta(u_\theta|x) = 0$, or alternatively $F_{u_\theta}(0|x) = \theta$. Koenker and Bilias (2001) describe the link between quantile regression and the transformation model and stress a general formulation of treatment effects introduced by Lehmann (1974). The simplest formulation of quantile regression is the two-sample treatment-control model,

$$Quant_\theta(lny|x) = \beta_1^\theta + \beta_2^\theta x \tag{3}$$

with $x = 1$ for treatment and $x = 0$ for the control group. The QR framework is flexible enough to allow for, say, $\beta_2^{0.2} > 0$ but $\beta_2^{0.8} = 0$ - the treatment being effective on left tail but not on the right tail of the duration distribution. If the treatment is continuous, as "age", for example, we assume that the treatment effect, β_2^θ, of changing x from x_0 to $x_0 + 1$ is the same as the treatment effect of changing x from x_1 to $x_1 + 1$.

Another important property of the quantile regression model is that, for any monotone function, $h(\cdot)$,

$$Quant_\theta(h(y)|x) = h(Quant_\theta(y|x)). \tag{4}$$

This equivariance to monotone transformations of the quantile regression model allows us to write, in particular, the family of conditional quantile functions for the untransformed duration y as

$$Quant_\theta(y|x) = exp(x'\beta_\theta). \tag{5}$$

3.2 Censored quantile regression - Estimation

When there is no censoring, the quantile regression coefficients, β^{θ}, can be estimated for given $\theta \in (0,1)$ by the methods introduced by Koenker and Basset (1978). Powell (1984, 1986) developed censored quantile regressions (CQR's) as a robust extension to the censored regression problem (for a recent discussion of censored quantile regression see Fitzenberger, 1997). Consider the sample (lny_i, x_i, yc_i), $i = 1, ..., N$, where yc_i denotes the upper threshold for lny_i ($yc_i = lny_i$ when an observation is censored and $yc_i = +\infty$ when it is not censored), i.e. $lny_i \leq yc_i$ for all i. The quantile regression estimator of β^{θ} is a solution to

$$min \frac{1}{N} \sum_{i=1}^{N} \rho_{\theta}(lny_i - min(x_i'\beta^{\theta}, yc_i)) \tag{6}$$

with,

$$\rho_{\theta}(u) = \begin{cases} \theta \cdot |u| & \text{for } u \geq 0 \\ (1-\theta) \cdot |u| & \text{for } u < 0. \end{cases} \tag{7}$$

The min operator censors $x_i'\beta$ at the larger threshold yc_i from above, i.e. the expression becomes $x_i'\beta$ if $x_i'\beta$ is less than yc_i, or else it becomes yc_i. Under certain regularity conditions, Powell (1984, 1986) showed that the CQR estimator $\widehat{\beta}^{\theta}$ is $\sqrt{N}$-consistent and asymptotically normally distributed. In contrast to quantile regression without censoring the distance function (6) to be minimized is not convex. There are a number of procedures suggested in the literature to calculate the CQR-estimator (Buchinsky, 1998, Fitzenberger, 1997, and Fitzenberger and Winker, 2001). In principle, the distribution of Powell's regression quantile estimators for censored model can be approximated by using the bootstrap aproximation, see e.g. Hahn (1995), Buchinsky (1995) and Fitzenberger (1997,1998). Bilias et al. (2000) suggest a reliable simplified version of the bootstrap. They showed that the method suffices asymptotically to estimate a quantile regression without censoring in the resamples based only on those observations for which the fitted quantile is not censored, i.e. $x_i'\widehat{\beta}^{\theta} < yc_i$. We apply the algorithm BRCENS suggested by Fitzenberger (1997) for the sample estimation problem by using the (censored) LAD procedure in TSP 4.5. 500 resamples are drawn by iid resampling of the entire vector of the logarithm of unemployment duration, regressor and censoring values. The approach suggested by Bilias et al.(2000) is used for the estimation in the resamples and the standard deviation of the coefficient estimates across the resamples is taken as the bootstrap standard error estimate [6].

[6] We do not correct the standard errors for multiple spells in the sample. About 65% of the individuals appear once, 20% appear twice and 15% from three to five times. Therefore, there could be a violation of the assumption that the standard errors are independent across observations.

3.3 Marginal effects

According to Machado and Mata (2000), the marginal effect of "x_j" on the conditional quantile function in our analysis is defined as

$$\gamma_j(\theta, \overline{x}) = \frac{\partial Quant_\theta(y|\overline{x})}{\partial x_j} = exp(\overline{x}'\beta^\theta)\beta_j^\theta, \quad j = 1, ..., k, \tag{8}$$

where $\overline{x}$ denotes the vector of the regressors' sample means and y is untransformed unemployment duration. The marginal effect of each regressor, say of "tenure", measures the change in the unemployment duration which, *ceteris paribus*, would keep an unemployment duration in the same quantile when "tenure" increases by a marginal unit.

3.4 CQR vs. the Cox Model

Compared to quantile regressions the proportional hazard models can account for competing risks, for time varying covariates, and for unobserved heterogeneity in a straightforward way (Wooldridge, 2002, chapter 20), whereas quantile regressions regarding the estimation of competing risks models as well as the models with time-varying coefficients have so far not been considered in the literature. Though a method of estimating quantile regression with unobserved heterogeneity has not yet been developed, it is easy to show, by means of simple simulations in limited samples, that unobserved heterogeneity effectively induces a location shift of the duration distribution while leaving the shape unchanged[7]. Quantile regression conditioning just on the observed covariates therefore yields meaningful results even in the presence of random effects.

However, estimation of a proportional hazard model comes at the cost of the proportional hazard assumption which imposes a restriction on the behavior of the effect of covariates. The (Cox-) proportional hazard (PH) model does not provide a direct analogue of the regression quantile, β_θ, since conditional quantiles under the Cox model are not linear in x. However, Koenker and Geling (2001) suggest a local measure of the marginal effects of various covariates in the Cox model on the conditional quantile at θ. The quantile function for the survival time T in the Cox model is $Q_\theta(T|x) = S_0^{-1}((1-\theta)^{1/\eta(x)})$, where $\eta(x) = e^{-x'\beta}$ and $S_0(t)$ denotes the baseline survival function. Thus the marginal effect in the Cox model is

$$\frac{\partial Q_\theta(T|x)}{\partial x_j} = \frac{(1-\theta)log(1-\theta)\eta(x)}{S_0'(Q_\theta(T|x))}\beta_j, \quad j = 1, ..., k.$$

[7]Simulation results which confirm this property in typical finite sample situations are available on request or can be found in Zhang (2004).

Because the baseline hazard rate $\lambda_0(t) \geq 0$, the sign of the coefficient β_j in the proportional hazard model determines the sign of the marginal effect over the entire distribution. Therefore, a proportional hazard model does not permit behavior where the sign of the effect may change with the size of the response[8]. Quantile regressions are capable of providing a more complete statistical analysis as they can distinguish between differential effects across conditional quantiles and as they allow for consistent estimation of the censored regression model under far less distributional assumptions than commonly required.

4 Estimation Results

Our model includes the following regressors:

- indicators for three periods, 1983 to 1987, 1988 to 1991 and 1992 to 1995, with the reference period 1981 to 1982
- the annual aggregate unemployment rate for West-Germany computed from the social security records (source: IAB Nuremberg)
- an indicator for whether the person became unemployed during the winter months (November to February)
- indicators for female, married and married female in the period 1988 to 1995
- an indicator for "no German citizenship"
- Indicators for apprenticeship and university degree and no apprenticeship in the period 1992 to 1995
- a person's age enters the model as a quadratic
- five quintiles ($0-20\%$, $20-40\%$, $40-60\%$, $60-80\%$, $80-100\%$) of the location of the previous wage in the population income distribution
- the tenure (in days) in the last job before unemployment
- an indicator for whether the person received any form of unemployment compensation (ALG, ALHi, UHG) within the last year before becoming unemployed (LED-spell)

[8]See Koenker and Geling (2001) and Portnoy (2003) for more details

- an indicator for whether the person was recalled by the same employer in the previous period of unemployment
- indicators for agricultural and technical profession
- indicators for employee and part-time worker

The set of regressors is selected according to preliminary estimations with several sets of regressors. Based on Wilke's (2005) nonparametric evidence we include calender time dependent dummy variables which absorb the main evolutions and account for structural breaks in the decades under consideration. This allows us to estimate a pooled model for all calender years. In Figure 2 we present a concise visual representation of the results from the estimation of the model. Each plot depicts one coefficient in the quantile regression model. The solid line represents the point estimates, $\{\beta_j^\theta, j = 1, ..., 25\}$, with the two dashed lines representing a 90% confidence interval for the respective coefficient. In the first panel of the figure the intercept of the model may be interpreted as the estimated conditional quantile function of the log unemployment durations of the control sample and all the other coefficients are simply location and scale shifts of this function. After the log transformation of durations, a location-scale shift would imply that the covariate exerts a time-varying percentage change in the durations. In the following, we focus on some main effects on the macro and on the micro level. We report estimated Cox coefficients in Table 6 (Appendix) for the sake of comparison.

4.1 Calendar time and Macroeconomic Situation

Year and unemployment rate During the years 1981 to 1982, the German economy was characterized by a high, but stable GDP growth rate and a relatively low, but sharply rising unemployment rate. In the period 1983 to 1987, the German unemployment rate remained at a constant high level of about 9%, whereas the GDP growth rate was comparable to that of the years 1981-1982. During the years 1988-1991, the German reunification which took place in 1990 had a strong influence on the economy, bringing about a boom. Hence, there was a low unemployment rate and a high growth rate. In contrast, the German economy during the years 1992-1995 was characterized by a relatively high unemployment rate induced by an economic recession.

In the estimation results the period 1983 to 1987 is associated with a quite uniform effect over the whole range of the distribution of about 93% ($= e^{-0.07}$). Beyond this period the negative effects become stronger in the lower tail and then gradually returned to a null effect (in the last period 1992 to 1995) in the upper tail of the distribution. It is interesting that the unemployment rate of the year when a person became unemployed exerts an estimated detrimental effect at the lower quantiles. However, at higher quantiles (beyond the quantile $\theta = 0.6$), it becomes a significant force for early reemployment. This suggests that the proportional hazard assumption is violated for this regressor[9]. The joint influence of the annual unemployment rate and the period indicators could be measured with the estimated coefficients, given by $exp(\beta^{\theta}_{year} + unemp. * \beta^{\theta}_{unemp.})$, where *unemp.* denotes the unemployment rate and *year* denotes period indicators.

Table 1 presents the relative combined effect of the macroeconomic situation for the selected years, 1985, 1990 and 1995. The year 1981 is chosen as a reference category. In 1990, the good general economic situation led to shorter unemployment duration in all quantiles. Interestingly, the unemployment duration in 1995 tended to be weakly lower than in 1981, although the unemployment rate had risen sharply in the meantime. We observe that there is an effect of the business cycle on the length of unemployment duration, particularly at the lower quantiles. But similar to Wilke (2005) we do not observe that a doubling in the unemployment rate led to a shift in the distribution of unemployment duration to the right for the population under consideration.

Table 1: Relative effect of the calendar time relative to 1981.

Year	$\theta = 0.2$	$\theta = 0.5$	$\theta = 0.8$	Unemployment Rate*	GDP Growth Rate
1981	100%	100%	100%	4.8%	0.1%
1985	99%	98%	79%	8.1%	2.2%
1990	81%	87%	83%	5.9%	5.7%
1995	93%	100%	89%**	8.2%	1.7%***

* West-Germany; source: IAB Nuremberg

** read this with caution due to the censoring of the available data at the end of 1997

*** caution: GDP growth rate for East and West Germany (Gesamtdeutschland)

Winter-season For many quantiles, the duration of unemployment is shorter for individuals who become unemployed in winter. This effect is stronger at the higher quantiles of the

[9]Interestingly, the Cox model predicts that unemployment periods become shorter with an increase in the unemployment rate (see table 6).

distribution. This can be explained by the fact that the proportion of long-term unemployed is smaller among those who become unemployed during the winter months and coincides with the fact that by definition of seasonal unemployment a larger fraction is reemployed after a fixed period. Short unemployment periods are longer because temporary lay-offs last for a minimum period by definition. For this reason, the coefficient is positive for the bottom quantiles, which points to another violation of the proportional hazard assumption. The Cox model predicts a clear shortening effect of this variable (see table 6). Fahrmeier et al. (2003) also find strong seasonal effects but their results are not directly comparable with our results as we have also included a recall variable in our set of regressors. The recall variable is highly correlated with seasonal unemployment in agriculture and in the construction sector (see Wilke, 2005).

4.2 Sociodemographic factors

Gender and marital status The estimated coefficient for females appears increasing across the horizontal line, though barely achieving 10% significance for this effect. Married persons are 21% ($= 1 - e^{-0.24}$) to 25% ($= 1 - e^{-0.29}$) quicker than unmarried persons to exit unemployment. The effect of married women is highly significant positive.

Table 2: Effect of gender and marital status

	$\theta = 0.2$	$\theta = 0.5$	$\theta = 0.8$
Unmarried men	100%	100%	100%
Married men ($= exp(\beta_8^\theta)$)	78%	77%	74%
Unmarried women 1981–1987 ($= exp(\beta_7^\theta)$)	95%	97%	103%
Unmarried women 1988–1995 ($= exp(\beta_7^\theta + \beta_{10}^\theta)$)	82%	85%	83%
Married women 1981–1987 ($= exp(\beta_7^\theta + \beta_8^\theta + \beta_9^\theta)$)	124%	128%	143%
Married women 1988–1995 ($= exp(\beta_7^\theta + \beta_8^\theta + \beta_9^\theta + \beta_{10}^\theta + \beta_{11}^\theta)$)	104%	100%	93%

The joint effect of gender, marital status and the calendar time relative to unmarried men is contained in Table 2. Married men show the shortest unemployment duration of all groups considered. Unmarried women, in contrast, experience about the same unemployment duration as unmarried men in the period 1981 to 1987. Yet, in the years 1988 to 1995, the unemployment duration of unmarried women is shorter, compared to the period before and compared to the duration of unmarried men. Married women in the years 1981 to 1987 are unemployed significantly longer than unmarried men. One possible explanation for the shortened unemployment duration of married as well as unmarried women is the reform of parental leave benefits which was introduced in Germany in 1986. Since then, the length of entitlement to parental leave benefits has been extended gradually (see Table 3 for an overview). This may have forced fewer women in motherhood to register as unemployed.

Table 3: Entitlement to parental leave benefits*

Year	1986	1988	1989	1990	1992
Entitlement	10 months	12 months	15 months	18 months	36 months

*(Source: Weber, 2004)

Citizenship Holding the influence of the other variables in the model constant, employees without German citizenship tend to be unemployed significantly longer than their German colleagues. The effect is stronger for the lower quantiles than for the higher ones. Longer unemployment periods for non German citizens are also found by Fahrmeir et al. (2003) and Wilke (2005).

Education Individuals with a completed apprenticeship exhibit significantly shorter unemployment duration than the reference category which is non-skilled workers. For those with a university degree, we observe an advantage in the lowest and in the highest quantiles only. For a more detailed analysis about the effect of education on unemployment duration see Lauer (2003). She applies a duration model with unobserved heterogeneity to the GSOEP. In general, the GSOEP based studies point to a clearer education pattern as found in this paper. However, our findings that the effect of education seems to be rather limited is in accordance with the results of Fahrmeir et al. (2003) using similar data and a rather different estimation technique. Nonparametric evidence alone (see Wilke, 2005) also suggests the reversed education pattern found in this paper.

Table 4: Effect of education in 1992-1995 relative to no completed apprenticeship

Education	$\theta = 0.2$	$\theta = 0.5$	$\theta = 0.8$
Completed Apprenticeship ($= exp(\beta^{\theta}_{13})/exp(\beta^{\theta}_{15})$)	74%	74%	72%
University Degree ($= exp(\beta^{\theta}_{14})/exp(\beta^{\theta}_{15})$)	85%	88%	82%

Moreover, we observe a significantly positive interaction "no apprenticeship *1992-1995" , which is in line with the descriptive finding of Wilke (2005) for this variable. Over the course of the years, completing an apprenticeship has become more important; this is particularly the case during the mid nineties recession (see Table 4). This may be explained by the fact that, during these years, many jobs for low-skilled workers have been transferred to countries with lower wage levels.

Age The regressor age enters the quantile regression model with a linear and a quadratic term and we found a concave functional relationship between age and unemployment duration. Figure 3 (a) contains the effect of age (in days) on the unemployment duration relative to a 26-year old person . At the 0.8 quantile, for example, a 27-year old person is unemployed about 10 days longer than a 26-year old person. The findings of Fahrmeir et al. (2003) also suggest that the age of the unemployed does not have a strong effect in the age group under consideration.

As denoted earlier, the marginal effect of age on unemployment duration is defined as the derivative of the conditional quantile function with respect to age. In Figure 3 (b), the marginal effect of age on unemployment duration is evaluated for the youngest age (26 years), the sample mean age (32.3 years) and the oldest age (41 years) and for the sample means of all other regressors. Whereas the marginal effect of age on unemployment duration is positive for the younger unemployed, it is negative for older people.

4.3 Individual employment history

Wage quintile We included the nominal variable wage quintile, as the continuous wage variable contained in the data is censored. For each year, we computed the quintile of the earnings distribution of all full-time employed. We then determined in which quintile the unemployed was located when the unemployment spell started. In general, our observation is that the higher the previous wage, the shorter the duration of unemployment. This effect is increasing over the quantiles of the unemployment duration.

Figure 4(a) shows the marginal effect of the wage quintile in days evaluated at the sample mean of all regressors. We observe strong effects at the higher quantiles. This result reflects that the opportunity costs of not working are higher for individuals with higher pre-income earnings and suggests that individuals with high pre-unemployment earnings possess higher abilities and therefore face a much lower risk of long-term unemployment. Note that the level of unemployment compensation in Germany is generally related to the magnitude of the former income[10]. The reduction in unemployment compensation after the exhaustion

[10] Unemployment assistance is also means tested.

of ALG - which is at the latest after 12 months in our sample - is in many cases higher for individuals with higher pre-unemployment earnings. This is because individuals with low pre-unemployment earnings usually obtain unemployment compensation at the level of social benefits from the very beginning of the unemployment duration. For this reason the exhaustion of ALG can be considered as a treatment for particularly the individuals coming from the higher quantiles of the earnings distribution. When we compute the conditional quantile functions at wage quintile 5 and at the sample mean of the other regressors such that predicted duration is 365 days we obtain that this is in between quantiles 0.7 and 0.8. Interestingly, the marginal effect of the wage quintile sharply increases at these quantiles. The treatment of reducing the benefit levels after 12 months may therefore have an impact on the wage quintile coefficient for the higher quantiles. However, further investigations with data containing more information about the receipt of unemployment compensation would be highly interesting. This would allow us to investigate further whether the wage quintile coefficient would be less negative for upper quantiles in a world without ALG.

To make clear how important the level of the previous wage is, we computed the effect of a transition from the lowest wage quintile to a higher one (see Figure 4(b)). This effect is strongest at the 0.8 quantile: There, the difference in unemployment duration from the highest to the lowest wage quintile amounts to about 240 days. This implies that the share of extreme long term unemployment is much higher for individuals with low pre-unemployment income.

Tenure Tenure is associated with a modest but significant increase in unemployment duration. This is more evident for the lower quantiles than for the higher ones.

Figure 5 shows the marginal effect of tenure on the duration (in days). This is the derivative of the conditional quantile with respect to tenure. It should be interpreted as the increase in unemployment duration (in days) which would keep an individual with an additional day of tenure at the same quantile. This marginal effect of tenure is highest at the 0.5 to 0.7 quantiles. We find plausible explanations for this observation: first, unemployed with a long foregoing employment duration may not be used to the situation of being unemployed and therefore there is a higher risk of dropping in long-term unemployment. Another explanation might be that wage expectations are too high due to a continuous wage increase in the former job. Since this reservation wage level is not easily reached, the unemployed waits longer for better job offers. A third explanation is that some long-term employed enter unemployment for personal reasons (e.g. health problems).

Unemployment compensation within the last year Those who have received any form of unemployment compensation (ALG, ALHi, UHG) within the last year before becoming unemployed show significantly shorter periods of unemployment. The effect is stronger for the upper quantiles. If we compare this effect to tenure in the first twelve months of duration time we observe that the two variables balance out and therefore the tenure variable has a strong effect only on unemployed with long-term employment before unemployment.

Recall Employees who have had an unemployment spell before and had been re-employed by their previous employer (recall) also tend to be unemployed significantly shorter. Again, the effect is stronger for the upper quantiles. At the 0.8 quantile, the unemployment duration is only 41% ($= e^{-0.89}$) of that of people without recall. The strong effect of the recall variable is already investigated by Plaßmann (2002) and by Wilke (2005).

5 Summary and Discussion

It is the purpose of this paper to analyze the effect of various micro- and macro variables on individual unemployment duration in West-Germany using censored quantile regression. This is actually the first study analyzing unemployment duration by means of censored quantile regressions.

Moreover, in contrast to most of the former studies on unemployment duration in Germany, our analysis is not based on the GSOEP and therefore possible disadvantages of survey data can be avoided. Instead, we used a subsample of the IAB employment subsample (1981-1997) - regional file - which is administrative data containing information about employment and unemployment periods of socially insured employees and unemployed provided that any form of unemployment compensation is received. The analysis is restricted to unemployed persons aged 26 to 41. In contrast, most of the former studies have included older persons and therefore the effects of regressors on unemployment duration are likely to be confounded with the effects of the reform of the German unemployment compensation system which took place in the 1980s.

Our main result is that, although the unemployment rate has doubled in the observed period, we only find relatively weak effects of the macroeconomic situation on the distribution of unemployment duration.

Surprisingly, we do not find a general elongation of unemployment periods of the considered group of unemployed during the past decades. Our results support the observation of Fitzenberger and Wilke (2004) that the general increase in average unemployment duration for West Germany can be ascribed to the older unemployed (which are excluded in our analysis). Fitzenberger and Wilke (2004) also conclude that this increase of unemployment duration of the older unemployed is linked to the issue of early retirement, the legal basis of which has been gradually abolished. These findings are interesting in view of the current labor market reforms ("Hartz"-reforms) in Germany. The latter aim at activating and reducing unemployment duration for *all* individuals, independent of their age. In contrast, our results suggest that it might be useful to primarily focus on the group of the older unemployed.

Moreover, we find that the educational degree or the profession of the unemployed have a rather limited explanatory degree for the length of unemployment duration. This is important for the design of ALMP. We observe that the individual (un-)employment history, such as the level of pre-unemployment earnings, had a stronger effect on the unemployment duration than sociodemographic variables and the macroeconomic situation. Individuals who had been unemployed before and who were recalled by their former employer exhibit significantly shorter unemployment duration. Moreover, long-term unemployment can be explained better by the individual employment history. Interestingly, long work experience without periods of unemployment increases the probability of long-term unemployment. Another interesting finding is that the unemployment duration of females has shortened during the period under consideration. This may be partly attributed to the introduction of parental leave benefits in 1986. In the 1980s, married females have had the longest unemployment duration.

Although we do not have information about the level of unemployment compensation received, we find that the effect of the regressor "wage quintile" becomes much stronger at the higher quantiles of the unemployment duration distribution where the entitlement to unemployment benefits (ALG) has expired. This may indicate an effect of the unemployment insurance system but it may also be due to better unobserved abilities of unemployed with high pre-unemployment earnings. Further research is therefore necessary on this issue.

From a methodological point of view we have argued that quantile regression offers a constructive complement to existing statistical methods of duration analysis. On the one hand, quantile regression have still three important limitations: they do not allow for time varying regressors, they are not yet extended to a competing risk framework and unobserved heterogeneity is not explicitly modelled. On the other hand, the censored quantile regression estimator enables the accommodation of incomplete duration data. It is a more flexible approach than the conventional proportional hazard models or the accelerated failure time model in the sense that it does not restrict the impact of the covariates - in terms of magnitude or sign - at different points of the distribution. Our analyses suggest that the proportional hazard assumption is violated for some of the regressors. This is why we conclude that quantile regression techniques seem more appropriate for the analysis of unemployment duration.

References

[1] Bender, S., Haas, A. and Klose, C. (2000) The IAB Employment Subsample 1975–1995. Schmollers Jahrbuch, 120, 649–662.

[2] Bilias, Y., Chen, S., and Z. Ying (2000). Simple Resampling Methods for Censored Regression Quantiles. *Journal of Econometrics*, Vol. 99, 373-386.

[3] Buchinsky, M. (1998) Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. *Journal of Human Resources*, 33, 88–126.

[4] Fahrmeir, L., S. Lang, J. Wolff and S. Bender (2003). Semiparametric Bayesian Time-Space Analysis of Unemployment Duration. *Allgemeines Statistisches Archiv*, 87, 281–307

[5] Fitzenberger, B. (1997) A Guide to Censored Quantile Regressions. in *Handbook of Statistics*, eds. G.S. Maddala and C.R. Rao. 15, 405-437.

[6] Fitzenberger, B. (1998) The Moving Blocks Bootstrap and Robust Inference for Linear Least Squares and Quantile Regressions. *Journal of Econometrics*, 82, 235-287.

[7] Fitzenberger, B. and Wilke, R.A. (2004). Unemployment Durations in West-Germany Before and After the Reform of the Unemployment Compensation System during the 1980s. *ZEW Discussion Paper* 04-24.

[8] Hahn, J. (1995). Bootstrapping Quantile Regression Estimators. *Econometric Theory*, 11, 105–121.

[9] Hartz, P. 2002. Moderne Dienstleistungen am Arbeitsmarkt. Vorschläge der Kommission zum Abbau der Arbeitslosigkeit und zur Umstrukturierung der Bundesanstalt für Arbeit. Federal Ministry of Economics and Labour, Berlin.

[10] Hujer, R. und Schneider, H. (1996). Institutionelle und strukturelle Determinanten der Arbeitslosigkeit in Westdeutschland: Eine mikroökonomische Analyse mit Paneldaten. In: B. Gahlen, H. Hesse, H.J. Ramser, eds. *Arbeitslosigkeit und Möglichkeiten ihrer Überwindung*, Wirtschaftswissenschaftliches Seminar Ottenbeuren, 25, J.C.B. Mohr, Tübingen, 53–76.

[11] Hunt, J. (1995). The effect of the Unemployment Compensation on Unemployment Duration in Germany. *Journal of Labor Economics*. Vol.13.1, 88–120.

[12] IAB Employment Subsample 1975-1997 - regional file -, Handbook, IAB Nürnberg (in German)

[13] Jürges, H. (2004). Objects in the mirror are closer than they appear: Unemployment, retrospective error, and life satisfaction. *unpublished manuscript*, Institute for the Economics of Aging, Mannheim

[14] Koenker, R. and Basset, G. (1978). Regression Quantiles. *Econometrica*, Vol. 46, 33–50.

[15] Koenker, R. and Bilias, Y. (2001). Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments. *Empirical Economics*, Vol. 26, 199–220.

[16] Koenker, R. and Geling, O. (2001). Reapprising Medfly Longevity: A Quantile Regression Survival Analysis. *Journal of the American Statistical Association*. Vol. 96, No. 454, 458–468

[17] Koenker, R. and Xiao (2002). Inference on the Quantile Rregression Process. *Econometrica*, Vol. 70, 1583–1612.

[18] Lauer, C. (2003). Education and Unemployment: A French-German Comparison. *ZEW Discussion Paper* No. 03-34.

[19] Lehmann, E. (1974). Nonparametrics: Statistical Methods based on Ranks. *Holden-Day*: San Francisco.

[20] Machado, J. and Mata, J. (2000). Box-Cox Quantile Regression and the Distribution of firm Sizes. *Journal of Applied Econometrics*, Vol. 15, 253–274.

[21] Plaßmann, G. (2002). Der Einfluss der Arbeitslosenversicherung auf die Arbeitslosigkeit in Deutschland. *Beiträge zur Arbeitsmarkt– und Berufsforschung*, 255, Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB) Nürnberg.

[22] Portnoy, S. (2003). Censored Regression Quantiles. *Journal of the American Statistical Association*, Vol. 98, 1001–1012.

[23] Powell, J.L. (1984). Least Absolute Deviations for the Censored Regression Model. *Journal of Econometrics*, Vol. 25, 303–325.

[24] Powell, J.L. (1986). Censored Regression Quantiles. *Journal of Econometrics*, Vol. 32, 143–155.

[25] Schräpler, J.P (2002). Respondant Behavior in Panel Studies - A Case Study for Income-Nonresponse by Means of the German Socio-Economic Panel. *DIW Discussion Paper*. No. 299

[26] Steiner, V. (2001). Unemployment Persistence in the West-German Labor Market: Negative Duration Dependence or Sorting? *Oxford Bulletin of Economics and Statistics*. No. 63.1, 91–113

[27] Van den Berg, G. (2001). Duration Models: Specification, Identification and Multiple Durations. in *Handbook of Econometrics*, eds. J.Heckman and E.Leamer, Vol. 5, North-Holland

[28] Weber, A. (2004). Warum kehren junge Mütter auf den Arbeitsmarkt zurück? Eine Verweildaueranalyse für Deutschland. *ZEW Discussion Paper* No. 04-08

[29] Wilke, R.A. (2005). New Estimates of the Risk and the Duration of Unemployment in West-Germany. forthcoming in: *Schmollers Jahrbuch*

[30] Wooldridge, J.M. (2002). Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge, Massachusetts.

[31] Zhang, X. (2004). Neue Entwicklungen in der Analyse von Verweildauermodellen mit Quantilsregressionen als Alternative zum konventionellen Modell der proportionalen Hazardrate. Diploma thesis, Mannheim University.

Appendix

Table 5: Descriptive Statistics

Variable	Mean	Median	Std.Dev.	Minimum	Maximum
Unemployment Duration (days)	425.74	180	650.30	1	6206
Age	32.30	32	4.60	26	41
Tenure (days)	1032.54	543	1192.00	1	5843

Censored	yes	12.53%
	no	87.47%
Recall	yes	17.65%
	no	82.35%
Unemployment compensation within the last year	yes	42.00%
	no	58.00%
Gender	female	36.55%
	male	63.45%
Marital status	married	53.74%
	unmarried	46.26%
Citizenship	German	88.37%
	Other	11.63%

Education		
	unskilled	29.60%
	skilled	64.68%
	university degree	5.73%

Profession Group		
	agriculture	3.29%
	mining	0.26%
	production	45.34%
	technical professions	3.60%
	service professions	47.46%
	other	0.06%

Table 6: Estimated coefficients of a Cox model

Variable	**Coefficient**	**Standard Error**
Period (83-87)	0.079**	0.014
Period (88-91)	0.111**	0.014
Period (92-95)	0.006	0.016
Unemployment rate	0.012**	0.005
Winter season	0.211**	0.007
Female	0.008	0.017
Marital status	0.223**	0.009
Married female	-0.453**	0.021
Unmarried female in (88-95)	0.111**	0.021
Married female in (88-95)	0.107**	0.025
Citizenship	-0.146**	0.012
Skilled	0.180**	0.010
University degree	0.060**	0.019
Unskilled in period (92-95)	-0.051**	0.018
Age	-0.038**	0.013
Age^2	0.000*	0.000
Wage quintile	0.069**	0.004
Tenure	0.000**	0.000
LED-spell	0.118**	0.010
Recall	0.494**	0.010
Agriculture profession	0.129**	0.020
Technical profession	-0.031	0.020
Employee	-0.180**	0.010
Part time	-0.063**	0.015

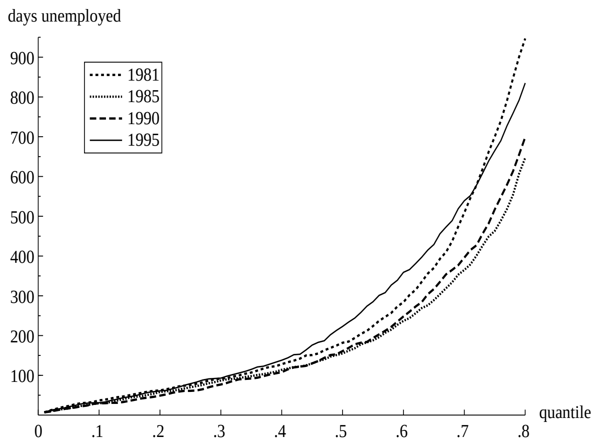


Figure 1: Unconditional nonparametric quantile functions for four calender years.

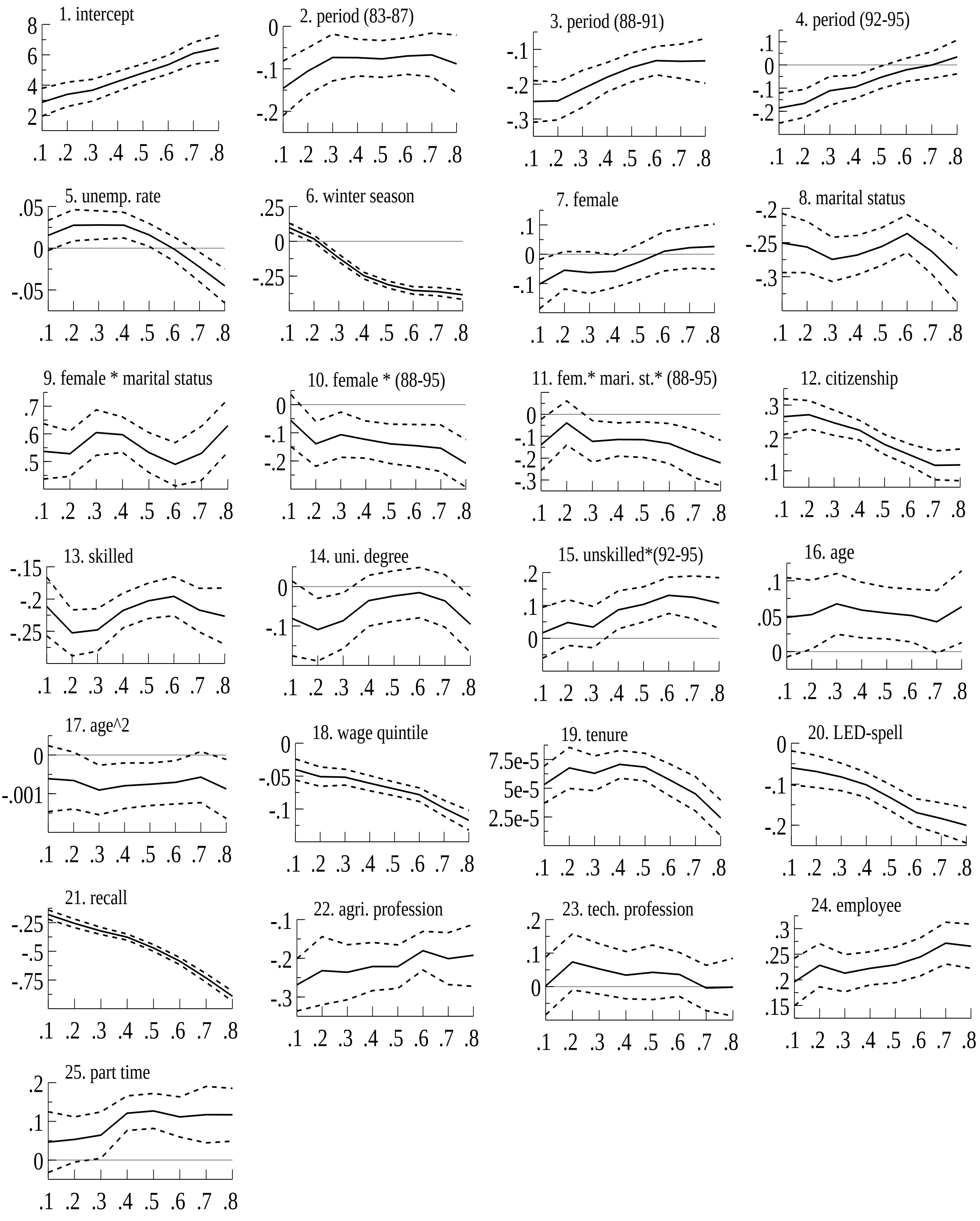
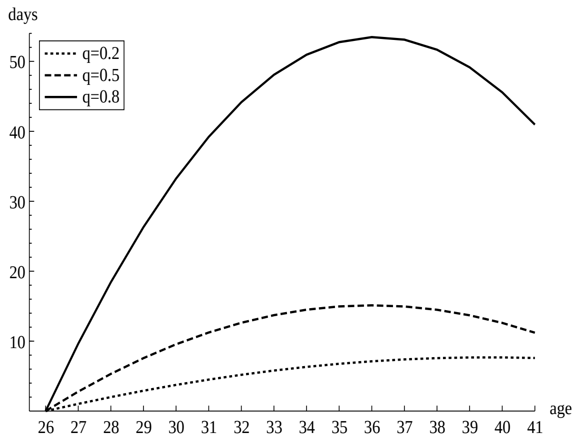
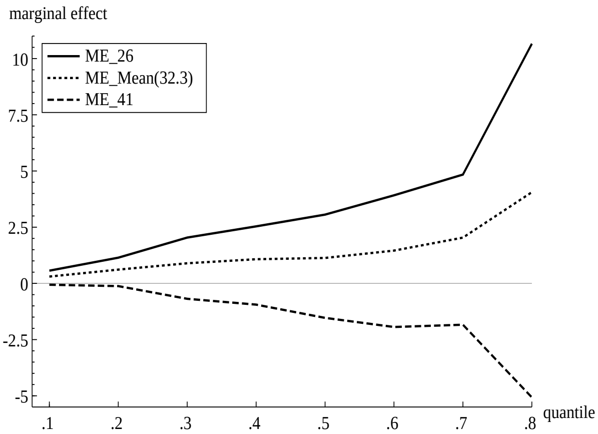


Figure 2: Estimated quantile regression coefficients $\beta_j^\theta, j = 1, ..., 25$, with 90% bootstrap confidence bands for quantile $\theta = 0.1, 0.2, ..., 0.8$

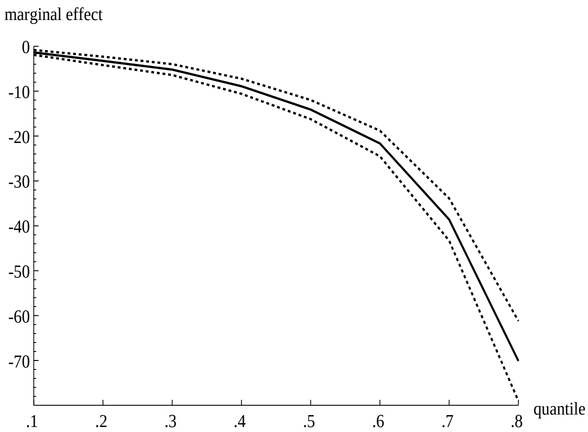


(a) Effect of age relative to a 26-years old

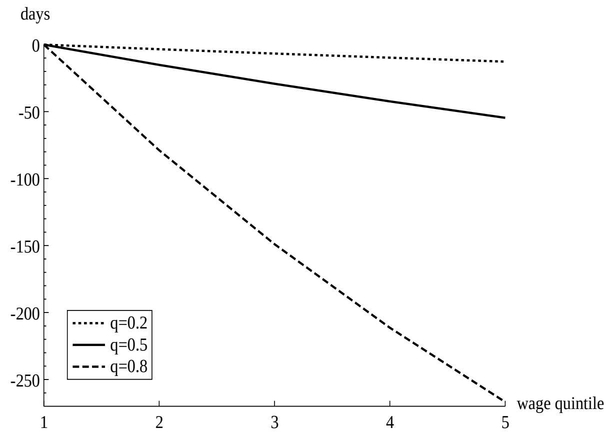


(b) Estimated marginal effect of age

Figure 3: Age



(a) Estimated marginal effect of the wage quintile



(b) Effect of the wage quintile relative to the lowest wage quintile

Figure 4: Wage quintile

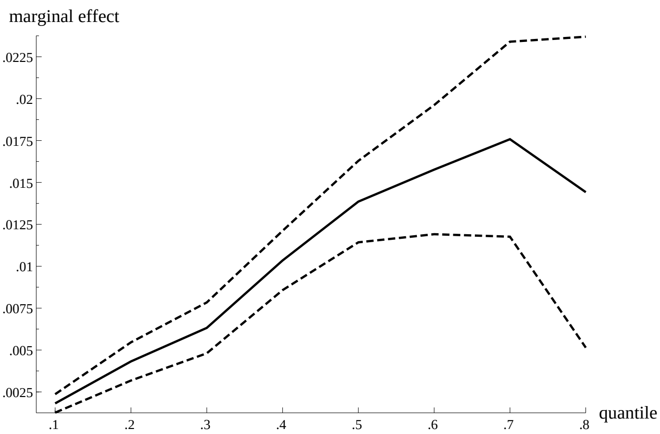


Figure 5: Estimated marginal effect of tenure (in days) with 90% bootstrap confidence bands