Has long become longer or short become shorter? Evidence from a censored quantile regression analysis of the changes in the distribution of U.S. unemployment duration *

Juliana Guimarães
Universidade NOVA de Lisboa, Faculdade de Economia
José A.F. Machado
Universidade NOVA de Lisboa, Faculdade de Economia †
Pedro Portugal
Banco de Portugal
Universidade NOVA de Lisboa, Faculdade de Economia

This Version: April 2004

Abstract

There is conflicting evidence regarding the recent evolution of unemployment duration in the U.S. In this study we rely on censored quantile regression methods to analyze the changes in the US unemployment duration distribution. We employed the decomposition method proposed by Machado and Mata (2003)to disentangle the contribution of the changes generated by the covariate distribution and by the conditional distribution and adapted it to a duration analysis framework.The data used in this inquiry are taken from the nationally representative Displaced Worker Survey of 1988 and 1998. We provide evidence that the unemployment duration distribution shifted leftward. The main driving force behind that shift was the sharp leftward move in the unemployment rate distribution. This force was partially counteracted by the ageing of the displaced population, the striking absence of impact from being displaced via a plant shutdown, and the higher sensitivity of unemployment duration to unemployment rates.

KEYWORDS: Quantile Regression, Duration Analysis, Unemployment Duration, Counterfactual Decomposition
JEL CODES: C14, C21, C41, J64

*The authors also gratefully acknowledge the partial financial support from the Fundação para a Ciência e a Tecnologia. The usual disclaimer applies.
†Corresponding author: José A.F. Machado, Faculdade de Economia, Universidade NOVA de Lisboa, Campus de Campolide, 1099-032 Lisboa, Portugal. E-mail: jafm@fe.unl.pt

1 Introduction

There is conflicting evidence regarding the recent evolution of unemployment duration in the U.S. Whereas Abraham and Shimer (2001) argue that mean elapsed duration increased above it expected level during the nineties, Farber (2003) documents an increase in post-displacement reemployment rates for the same period. Abraham and Shimer, using the Current Population Survey (CPS), show that the rate of very long-term unemployment (longer than 26 weeks) increased sizeably relative to the aggregate unemployment rate. This outcome is partially produced by a decrease in the unemployment to employment transition rates. Farber employed the Displaced Worker Survey (DWS) to show that, while rates of job loss were higher than expected, the economic costs of job loss diminished because the transition rates from joblessness into employment increased.

A number of significant demographic changes have been occurring in the US labor market that are likely to impact on the distribution of unemployment duration. On one side, the aging of the baby-boom generation is likely to increase the duration of unemployment. On the other side, the increased attachment to the labor force of women tends to lead to a decrease of transitions into inactivity, making them to stay longer unemployed.

A complete characterization of the shape of the unemployment duration distribution is of interest for o number of reasons. In first place, a high incidence of long-term unemployment generates an unevenly distributed burden of unemployment. In second place, persistent unemployment may generate hysteresis due to human capital depreciation, stigmatization, loss of social networks, or specialization in home production. In third place, unemployment duration affects significantly the prospects of finding a job because less employable individuals dynamically sort themselves into long-term unemployment. In fourth place, the low reemployment rates of long-term unemployed may justify public interventions such as retraining or job search assistance. And, in fifth place, long-term unemployed compete less effectively for a job than short-term unemployed, pre-empting downward pressure on wages.

The data used in this inquiry are taken from the nationally representative Displaced Worker Survey of 1988 and 1998. The DWS is a retrospective survey that has been conducted biennially since 1984. In contains information on the nature of the job lost and subsequent joblessness duration of displaced workers by reason of plant closure, slack work, or abolition of shift or position. The DWS is particularly well suited to study the distributional shape of unemployment duration because, unlike the CPS, it provides information on completed spells of unemployment.

In this study we rely on censored quantile regression methods to analyze the changes in the US unemployment duration distribution. Quantiles seem appropriate to analyze unemployment duration for, at least, two main reasons. The methodology estimates the whole quantile process of duration time conditional on the attributes of interest which constitutes, as does the more traditional hazard function, a complete characterization of the distribution of duration time or, if one wishes, of the survivor function. Therefore quantiles provide a natural way of characterizing important concepts as short or long-term unemployment by focusing on the relevant tails of the duration distribution. For instance, comparison of the quantile regressions for the 20th and for the 80th percentiles (say) may shed important insights on the different determinants of short or long-term unemployment. From a methodological vintage point, it is worth noticing that quantile regression, although certainly not the only way of performing those comparisons, provide a unified and flexible framework for such an analysis. Moreover, quantile regression, as the seminal work of Powell (1986) reveals, is particularly well equipped to perform consistent inferences with censored data, a typical situation in duration studies.

The law of total probability implies that changes over time in the distribution of unemployment duration may result from changes in the distribution of the conditioning variables (e.g., labor force characteristics such as the age distribution) or from changes in the conditional distribution of duration itself (which may be thought of as changes in the way those labor force characteristics impact duration). Machado e Mata (2003) proposed a method (henceforth, M&M decomposition) of disentangling those effects. The method is based on the estimation of marginal distribution of the variable of interest consistent with a conditional distribution estimated by quantile regression as well as with any hypothesized distribution for the covariates. Comparing the marginal distributions implied by different distributions for the covariates one will then able to perform counterfactual exercises and identify the sources of the changes in the distribution of duration over the ten years period.

The paper is organized as follows. Section 2 describes the data set used and provides a careful comparison of the DWS for 1988 and for 1998. In section 3 we discuss the econometric methodology. The basic regression results are presented in section 4. Section 5 uses the M&M decomposition to sort out the forces behind the changes in unemployment duration. Finally, section 6 concludes.

2 Data

2.1 General Description

The data used in this inquiry are taken from the nationally representative, Displaced Worker Supplement to the February 1988 and 1998 Current Population Survey. The dataset - and changes in the survey including the wording of the core displacement question and the recall period over which information on job loss is recorded - are well described elsewhere (see, for example, Kletzer, 1998; Farber, 2003), so that only brief introductory remarks are required here. The DWS has been conducted biennially since 1984. It contains information on the nature of the lost job and subsequent joblessness for workers displaced by reason of plant closure, slack work, or abolition of shift or position. Such data can be supplemented by extensive information on the personal characteristics of the worker contained in the parent CPS. The choice of the 1988 and 1998 surveys was guided by the need to use a comparable framework as much as possible. The 1988 DWS survey was the first to provide information for a single spell of joblessness (until 1986 the recorded jobless duration included multiple spells of joblessness). The 1998 survey is the most recent available survey with adequate data on joblessness duration. But there remain some issues of comparability that will be discussed below.

The DWS has a number of advantages over administrative data. Firstly, unlike the unemployment registry, the DWS survey covers both unemployment benefits recipients and non-recipients. Secondly, because it is retrospective, the information on unemployment duration is not censored at the time of the exhaustion of benefits. And, thirdly, the DWS allows the identification of transitions of displaced workers to another job without any intervening spell of unemployment.

There are inevitably some shortcomings of the DWS data. Thus, retrospective data are subject to recall bias - individuals experiencing displacement in past years may be more likely to understate their jobless duration than are more recent job losers - and respondents are prone to round (to months and quarters) their reported spells of unemployment. Beginning with the 1994 survey, however, the period over which job loss is measured has been reduced from five to three years, which should reduce the recall bias problem.

As mentioned above, since the 1988 survey the measure of unemployment refers to the length of the single spell of joblessness that followed the displacement event and resulted in reemployment. To be sure, the definition still does not require the unemployed individual to be engaged in active search so that this single spell may include intervals of suspended job search/withdrawal but it no longer includes multiple spells of joblessness. A more recent innovation which affects the 1998 survey is that the DWS unemployment data are no longer top coded (at 99 weeks of joblessness). An additional source of right censoring in the data stem from our inclusion (via the CPS) of those individuals who failed to find work after displacement but who were nevertheless economically active as of the survey date.

Although we included those who wanted but never found employment after losing their jobs - as well as those individuals who transitioned directly into reemployment without any intervening spell of joblessness - we excluded individuals who were not economically active at the time of the survey. Further, because the nature of displacement is not well defined for certain individuals and sectors, those employed part time and in agriculture at the point of displacement were also excluded, as were those aged less than 20 years and above 61 years. These restrictions yielded a sample of 2,837 individuals for 1988 and 2762 for 1998.

2.2 Survey Comparisons

There are a number of comparability issues that need to tackled. First, and most importantly, whereas the 1988 survey is a five year retrospective data set of displaced workers based in the question *"In the past five years, that is since January 1983, has ...lost or left a job because of a plant closing , an employer going out of business, a layoff from which...was not recalled, or other similar reason?"*, the 1998 survey is a three year retrospective data set based in the question *During the last three calendar years, that is, from January of 1995 through December of 1997, did (name/you) loose a job, or leave one because a plant or company closed or moved, (your/his/her) position or shift was abolished, insufficient work, or another similar reason?"*. If the response to the job loss core question was positive, the respondent is asked whether the reason for displacement was 1) plant closing, 2) slack work, 3) position shifted or abolished, 4) seasonal job ended, 5) self-employment failed, and 6) other reasons. In line with the CPS definition of job displacement, solely the first three situations will be considered in this study

Whereas the slight change of wording is unlikely to raise significant comparison problems, the reduction of the retrospective period is much more serious. Since there is information on the year of displacement of the worker, one can minimize this problem excising from the 1988 sample the individuals displaced in 1983, 1984, and 1988. But this does not solve the issue. If an individual experienced multiple spells of joblessness (which is likely to occur for a fraction of displaced workers) the interviewer has instructions to record the episode where the worker lost the job with the longest duration. It may well occur that an individual after loosing a long-tenure job during 1983 or 1984 is displaced again during the 1985-1987 period. In this case, this displacement from a short-duration job is not registered. There is a clear implication for distortion of the distribution of job duration, with short job durations being likely to be under sampled in the 1988 survey in comparison with the 1988 survey. But there is no unambiguous implication for the distribution of unemployment duration [1].

Second, even though unemployment rates were falling and labor market conditions were improving over the survey periods, the cyclical conditions were not identical. In fact, the average state unemployment rate at the time of displacement is 1.7 percentage points lower in the 1998 survey in comparison with the 1988 survey. We hope that, by conditioning de unemployment duration distribution on labor market tightness, we will be able isolate the impact of the business cycle.

And third, in both surveys the displaced workers are asked whether they received advance notice of impending redundancy, but in the 1998 survey this question is restricted to written notice where in the 1988 survey the individuals distinguish between informal and written notice. In order to make this variable as comparable as possible we will solely consider notified those workers that received written notice with at least two month advance to the date of displacement.

Descriptive information on the two samples is provided in Table 1. The composition of the 1998 sample differs significantly from that of 1988.

[1] There are, however, a number of checks that can be done. First, one can compare the job duration distribution for the 1983-1984 period with the 1985-1987 period. Second, one can exclude from both samples workers with less than two years of tenure in the pre displacement job. And third, one can use our decomposition methodology to simulate the 1998 unemployment distribution with the 1988 job duration distribution. In all cases we arrive to the conclusion that the issue of multiple spells does not affect significantly the comparison of the two unemployment duration distributions

	Sample Means	
	1985-1987	1995-1997
Age	35.7	38.3
Gender	0.650	0.562
Race	0.869	0.863
Marital status	0.606	0.562
Marital*Gender	0.176	0.222
Schooling	11.6	13.2
Tenure	4.6	4.7
Plant Closing	0.480	0.395
Written Notice	0.054	0.131
Unemp. Rate	7.0	5.3
Unemp. Duration (completed)	12.7	11.2
Proportion censored	0.149	0.092
Number of observations	2837	2762

Table 1: Sample Descriptive Statistics.

- Displaced workers in the nineties are significantly older and better educated than during the eighties, very likely reflecting the ageing of the baby-booming generation (see Figure 1.

- The proportion of female workers among displaced also increased sizeably, probably because labor market participation rates of women in risk of being displaced also increased.

- The likelihood of receiving formal notice of job redundancy more than doubled in the nineties, due, probably, to the introduction of the WARN act that made pre-notification mandatory for mass-layoffs generated by large firms.

- Interestingly, despite the change in the reference period of job displacements (from five to three years), there are no significant changes in the distribution of job duration in the pre displacement job (see Figure 1). It may still happen, however, that workers are now displaced with longer tenure than before.

- Finally, and very importantly, unemployment duration is visibly shorter in the 1995-97 period than during the 1985-87 period. This indication is best understood in the empirical survival functions (Kaplan-Meier estimates) exhibited in Figure 2.

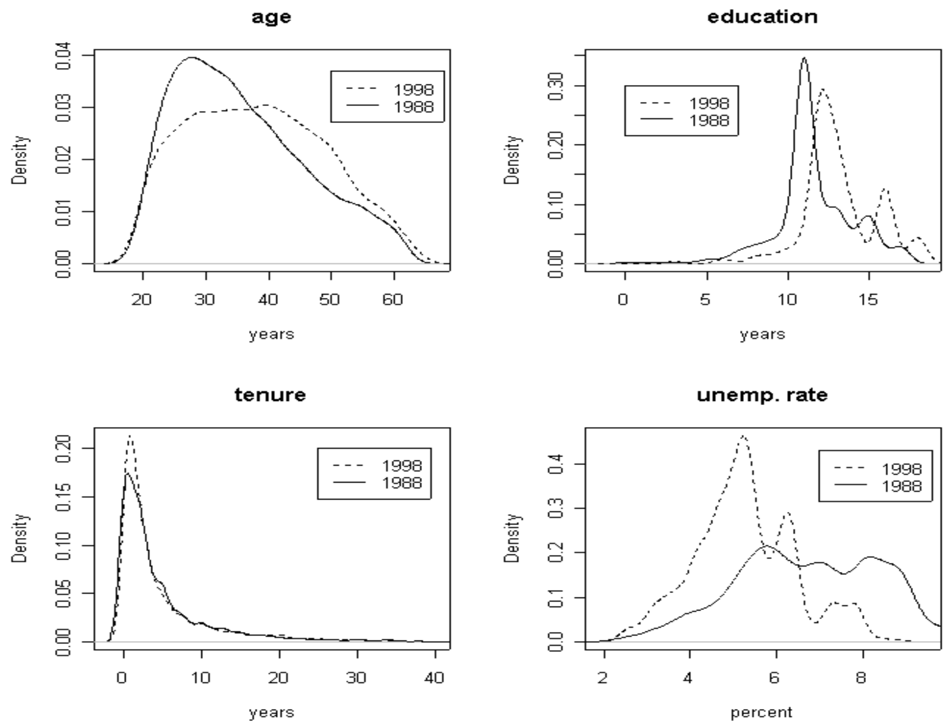

Figure 1: COVARIATES' DENSITIES

3 Exploring the information in Quantile Regression Estimation

3.1 Quantile Regressions in Duration Data Analysis

Models specified in terms of hazard functions undoubtedly dominate the analysis of duration data. Yet, in some instances, regression-type models may prove natural and useful. Regression models for the duration time are typically framed in a strict parametric setting. Let T be the duration of stay in a given state, and x_i ($x_{1i} \equiv 1$) be the vector of covariates for the ith observation. In our application T_i represents the duration of the "most representative" unemployment spell of individual i. A parametric regression model assumes that

$$y(T_i) = x_i'\beta + \sigma\epsilon_i \tag{1}$$

where, β and σ are unknown parameters, $y(\cdot)$ is a transformation function and ϵ is a zero mean and unit variance random variable with density f, not depending

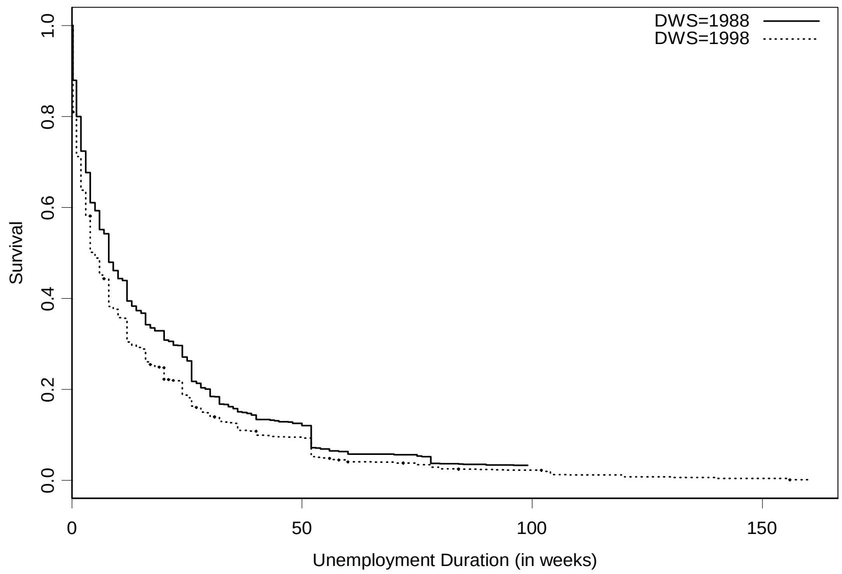

Figure 2: Kaplan-Meier Survival Functions

on x, (e.g., Gaussian, lognormal, smallest extreme value, Weibull or exponential). A leading example of this class is the Accelerated Failure–Time (AFT) model where

$$\log T_i = x_i'\beta + \sigma\epsilon_i \tag{2}$$

and f is left unspecified. The Proportional hazard (PH) model with Weibull baseline also fits in the class, as it is equivalent to the Accelerated Life model with ϵ being the log of a unit Exponential variate.

The set-up above is restrictive in two main ways. First, it assumes a known duration distribution f so that the model may be estimated by maximum likelihood. As is well known, the resulting estimators are "optimal" if the model is correctly specified but lack robustness to departures from the assumed distribution.

Second, and perhaps even more importantly, (1) assumes that only the conditional mean of $y(T)$ depends on the covariates. In technical terms, the distribution of the duration time conditional on the covariates is restricted to the translation family that is, all the heterogeneity in the distribution of duration time for different levels of the covariates is assumed to be captured by mere location shifts (Manski, 1988). To put it plainly, the distributions corresponding to different individuals differ only on location; other distributional attributes such as scale, skewness or tail behavior are deemed independent of the conditioning variables.

Quantile regression (QR) directly addresses these two limitations of a strict parametric approach. Let $F_T(t|X = x)$ denote the conditional distribution of T given $X = x$; for $p \in (0, 1)$, the pth quantile of T is

$$Q_T(p|x) = \inf\{t|F_{T|X}(t|X = x) \geq p\}.$$

We consider statistical models specifying

$$Q_{y(T)}(p|x) = x'\beta(p) \tag{3}$$

where $y(\cdot)$ is a monotone link function, known possibly up to a finite number of parameters $\lambda(p)$, (we shall take $y(\cdot) \equiv \log$) and $\beta(p)$ is a vector of QR parameters, varying from quantile to quantile.

The conditional quantile process – i.e., $Q_y(p \mid x)$ as a function of $p \in (0, 1)$ – provides a full characterization of the conditional unemployment duration in much the same way as ordinary sample quantiles characterize a marginal distribution.

When there is no censoring, the quantile regression coefficients, $\beta(p)$, can be estimated for given $p \in (0, 1)$ by the methods introduced by Koenker and Bassett (1978). Powell (1984, 1986) developed estimators of the QR coefficients for the case of censored data with known, but possibly varying, censoring points, (for a recent discussion of censored quantile regression see Fitzenberger, 1997).

Our sample provides information on complete unemployment durations but there are some incomplete spells (right-censoring). Moreover, to avoid problems with taking logs of very short spells (0 or close to 0 weeks) we, arbitrarily, censored durations inferior to 0.5 at 0.5 weeks. The sample information we consider may thus be represented by $(y_i^\star, x_i)$, $i = 1, \ldots, n$ where $y_i^\star = \min[\max(y_i, l), u_i]$, u_i denotes the upper threshold for y_i ($u_i = \infty$ when observation i is not censored) and l the left-censoring point ($l = \log(0.5)$). The QR estimator minimizes the sample objective function

$$\sum_{i=1}^{n} \rho_p(y_i - \min[u_i, \max(x_i'b, l)])$$

with,

$$\rho_p(z) = \begin{cases} p\,z & \text{for } z \geq 0 \\ (p-1)\,z & \text{for } z < 0. \end{cases}$$

Due to censoring it may not be possible to identify the whole quantile process. Let (p_l, p_u) represent the range of quantiles quantile that can be consistently estimated. Technically, any p in that range must be such that

$$M_n(p) = E\{\frac{1}{n}\sum_{i=1}^{n} I(l+\xi < x_i'\beta(p) < u_i - \xi)x_i x_i'\}$$

is uniformly positive definite in n for some $\xi > 0$ (Fitzenberger (1997), Theorem 2.1).

3.2 From Conditional to Marginal Quantiles

The resampling procedures proposed in Machado and Mata (2003) (henceforth, M&M) provide an easy way of simulating a random sample, $\{T_i^\star, i = 1, \ldots, m\}$, from a conditional distribution of duration times that is consistent with the restrictions imposed on the conditional quantiles by the QR model. The theoretical underpinnings of this procedure are quite simple. On the one hand, the probability integral transformation theorem from elementary statistics implies that one is simulating a sample from the (estimated) conditional distribution of T given $X = x_0$. On the other hand, the results in Bassett and Koenker (1986) establish that under regularity conditions the estimated conditional quantile function is a strongly consistent estimator of the population quantile function, uniformly in p on a compact interval in $(0, 1)$.

For completeness we outline here the procedure:

1. Generate m random draws from a Uniform distribution on (p_l, p_u), $\pi_i,\ i = 1, \ldots, m$;

2. For each π_i estimate the QR model (3), thereby obtaining m vectors $\hat{\beta}(\pi_i)$;

3. For a given value of the covariates, x_0,

$$T_i^\star \equiv \hat{Q}_T(\pi_i|x_0) = g(x_0'\hat{\beta}(\pi_i)) \ \ i = 1, \ldots, m,$$

   is a random sample from the estimated conditional c.d.f. $F_T(t|X = x_0)$ censored at p_l and p_u.

The sample generated by the procedure above is drawn from the conditional distribution. In many instances it is important to integrate out the conditioning covariates. This integration or marginalization can be performed with respect to different joint distributions, $g(x)$, of the covariates. The approach in M&M may be described as follows:

1. As described before, generate $\pi_i,\ i = 1, \ldots, m$ and estimate the corresponding $\hat{\beta}(\pi_i)$;

2. Generate a random sample of size m from a given $g(x)$; let it be denoted by $\{x_i^\star\}$, $i = 1, \ldots m$.

3. Obtain

$$T_i^\star \equiv \hat{Q}_T(\pi_i|x_i^\star) = g(x_i^{\star'}\hat{\beta}(\pi_i)),$$

   which is a random sample from the marginal distributions of durations times implied by the model postulated for the quantile process and by the assumed joint distribution of the covariates.

In the implementation of the method in this paper we made $p_l = 0.20$ and $p_u = 0.95$ and estimated the quantile regression coefficients at equally space intervals of length 0.01. We then draw 1000 ($= m$) of such estimates with replacement. A code in R with the whole procedure is available on request.

When $g(x)$ is an estimate of the actual distribution of the covariates in the population, the resulting sample of durations is drawn from the actual marginal distribution. In this case, $\{x_i^\star\}$ may be obtained by drawing with replacement from the rows of $\mathcal{X}$, the regressors' data matrix.

3.2.1 Counterfactual durations

But, in reality, $g(x)$ may be any distribution of interest. If it is an estimate of the distribution of the covariates in 1988 ($g(x(1988))$), the resulting durations will constitute a simulated sample from the marginal distribution of durations that would have prevailed in 1998 if all covariates had been distributed as in 1988, (assuming, of course, that β were estimated with 1998 data).

Comparing this counterfactual sample with samples of durations from the actual marginals for 1998 and 1988 it is possible to derive Oaxaca type decompositions for the entire distribution rather than just for its mean. Specifically, it is possible to decompose the observed changes in those due to changes in the conditional distribution of durations (the β's) and those stemming from changes in the joint distribution of the covariates.

Other decompositions of interest often involve to isolate the contribution of a single covariate. Suppose we wish to simulate a random sample from the counterfactual distribution of durations that would have prevailed in 1998 if a given covariate, x_k, had been distributed as in 1988 and the other covariates (denoted by x_{-k}) as in 1998. We shall assume that x_k is discrete (or was discretized according, say, to its deciles) with support $\mathcal{S}_k(t) = \{c_1(t), \ldots, c_{L_k}(t)\}$, for $t = 1988, 1998$. Now, the relevant counterfactual distribution of the covariates is $g(x_{-k}(1998)|x_k = c)P(x_k(1988) = c),\ c \in \mathcal{S}_k(1988)$. (For further details on how to implement this decomposition see M&M.)

3.3 Hazard Functions

Model (3) provides a complete characterization of the (conditional) distribution of duration time T or, if one wishes, of the survivor function, (obviously, $Q_T(p|x)$ is the $(1-p)$th quantile of the conditional survivor function). The hazard function,

$$h(t|x) = \frac{f_{T|X}(t|x)}{1 - F_{T|X}(t|x)}$$

provides still another characterization of the same probability distribution. Since it constitutes the most popular frame for duration analysis, it is important to relate it to models for the conditional quantile function (CQF).

Having obtained a simulated random sample, $\{T_i^\star, i = 1, \ldots, m\}$, from the distribution of duration time of interest (conditional, marginal or counterfactual) the usual methods of density estimation and hazard function estimation may be applied. In situations where, due to censoring, the top quantiles cannot be consistently estimated, the estimated function must be adequately rescaled. Specifically, assuming that quantile process is only identified in (p_l, p_u), the results in Silverman (186, p.148) yield,

$$\hat{h}(t|x) = \frac{(p_u - p_l)f^\star(t)}{p_u - F^\star(t)}$$

where $f^\star(t)$ is the usual kernel density smoother of $T_i^\star$,

$$f^\star(t) = \frac{1}{mh}\sum_{i=1}^{m} K(\frac{t - T_i^\star}{h})$$

and the distribution function estimator is,

$$F^\star(t) = \frac{1}{m}\sum_{i=1}^{m} \mathcal{K}(\frac{t - T_i^\star}{h})$$

with

$$\mathcal{K}(u) = \int_0^u K(v)dv.$$

Besides hazard functions, other standard outputs of duration analysis such as survivor function, residual duration and mean duration are also quite easily estimated from a quantile model such as (4). For instance, given an estimate of the quantile function of T, $\hat{Q}_T(p|x)$, the quantile process of the survivor time conditional on x can be estimated by $\hat{Q}_T(1-p|x)$ which, upon "inversion", yields an estimate of the survivor function (see, Bassett and Koenker, 1986). The mean duration conditional on x can be estimated as $\int_0^1 \hat{Q}_T(p|x)dp$ which can be easily computed by Monte-Carlo methods. Likewise, the distribution of the residual duration—i.e., the duration of all those that have survived longer than $Q_T(p^*|x)$, for a given p^*—may be summarized by $\int_{p^*}^1 Q_T(p|x)dp$.

4 Unemployment Duration in 1985-87 and 1995-97

Empirical results for selected quantiles from fitting the QR model are given in Table 2 and 3. For comparison purposes, we also provide the estimates obtained from a Cox proportional hazard model and from an accelerated failure time (AFT) model that employs an extended generalized gamma distribution [2].

Estimation of the censored quantile regression was performed iteratively using Buchinsky's (1994) ILPA procedure [3]. The iterative procedure is quite well known: at each iteration the observations for which $x_i'\hat{\beta}(p) \geq u_i$ or $x_i'\hat{\beta}(p) \leq l$ are discarded; then, the coefficients are re-estimated with the remaining observations until convergence is reached. The quantile estimation uses the Frisch-Newton algorithm (see Koenker and Portnoy, 1997) implemented in the function `rq` in the `quantreg` package for R, Koenker (1991 -). For the estimation of standard errors for the individual coefficients we resort to the bootstrap. Since the "errors" from the QR equation are not necessarily homogeneously distributed, to achieve robustness we resample (y, x, l, u) following the method of Billias et al. (2000).

In general, the regression coefficient estimates are fairly conventional [4]:

- Age reduces escape rates proxying, very likely, the reduced arrival rate of job offers with age.

[2] See Addison and Portugal (1987) for an application of the extended generalized gamma distribution to unemployment duration.

[3] See Fitzenberger (1997) for a discussion of limitations and alternatives.

[4] The continuous regressors were centered at their sample means. Consequently, the intercept estimates the quantile of the distribution of log duration for the "population" corresponding to these mean values and to the reference values of the binary regressors.

- Tenure in the previous also leads to longer unemployment. The effect of the tenure variable most probably captures the elevated reservation wages of long-serving workers.

- Schooling enhances the chances of getting a job. More educated workers might be expected to have higher escape rates because of their greater search efficiency, higher opportunity cost of staying unemployed, and generally better job prospects.

- The result for race is familiar and captures the poorer opportunities facing blacks as a result of both objective and discriminatory factors.

- The familiar (opposing) effects of marital status on reemployment probabilities - positive for males and negative for females - are also obtained. The result for married males presumably picks up a household head effect,and thus likely reflects the higher opportunity cost of unemployment for married males and their greater search intensity.

- Higher state unemployment rates are associated with longer spells of joblessness, reflecting, at the state level, lower arrival rates of job offers.

- Altogether less transparent are the effects of written pre-notification - defined as written notice of at least two months - and job loss by reason of plant closure. It is often argued in the displacement literature that the compositional or labor quality implications of plant closings - all workers are 'canned' when a plant closes its doors rather than a subset of workers (selected by management) in the case of slack work or abolition of shift or position - and the enhanced search facilitated by advance notice should each lead to lower joblessness. This indication is obtained for the 1988 survey but not for the 1998

Comparison across different model specifications - Quantile Regression, Cox Proportional hazard, and Accelerated Failure Time - also reveals broad agreement, at least in terms of sign and statistical significance of the regression coefficients, in particular if we take the highest quantiles as comparators.

The coefficient estimates for lower quantiles (for example, the 20th quantile in Tables 2 and 3), however, disclose some interesting features:

- First, advance notice of displacement exerts a significant influence on joblessness duration at low quantiles in contrast with the small and statisti-

	Quantile Regression				
	20th	50th	80th	AFT	Cox
Age	0.010	0.024	0.016	0.021	-0.014
(in years)	(0.007)	(0.004)	(0.003)	(0.004)	(0.002)
Gender	0.060	0.379	0.219	0.328	-0.224
(male=1)	(0.153)	(0.129)	(0.103)	(0.099)	(0.066)
Race	-0.392	-0.257	-0.352	-0.443	0.307
(white=1)	(0.172)	(0.102)	(0.113)	(0.096)	(0.064)
Marital Status	-0.425	-0.352	-0.236	-0.327	0.209
(married=1)	(0.140)	(0.109)	(0.085)	(0.083)	(0.055)
Married*Gender	0.606	0.845	0.493	0.681	-0.429
(married female=1)	(0.278)	(0.179)	(0.141)	(0.131)	(0.088)
Schooling	-0.109	-0.046	-0.041	-0.065	0.040
(in years)	(0.023)	(0.019)	(0.014)	(0.013)	(0.009)
Tenure	-0.004	0.007	0.020	0.011	-0.008
(in years)	(0.010)	(0.008)	(0.006)	(0.006)	(0.004)
Plant Closing	-0.716	-0.424	-0.183	-0.389	0.219
(Shutdown=1)	(0.120)	(0.082)	(0.068)	(0.064)	(0.041)
Written Notice	-0.260	0.133	0.141	0.064	-0.065
	(0.269)	(0.156)	(0.163)	(0.140)	(0.093)
Unemp.Rate	0.081	0.122	0.116	0.106	-0.071
	(0.024)	(0.021)	(0.016)	(0.016)	(0.011)
Constant	1.233	2.278	3.529	2.039	
	(0.158)	(0.121)	(0.111)	(0.247)	
scale parameter				1.536	
				(0.035)	
shape parameter				0.639	
				(0.070)	

Table 2: Unemployment duration regression results for 1985-1887. The first entry in each cell is the regression coefficient point estimate with continuous regressors centered at their mean; in parenthesis are the standard errors. The standard errors for the QR estimators were computed as the half-length of a 95% bootstrap confidence interval divided by 1.96. (N=2837)

	Quantile Regression				
	20th	50th	80th	AFT	Cox
Age (in years)	0.010 (0.005)	0.023 (0.004)	0.024 (0.004)	0.020 (0.004)	-0.012 (0.002)
Gender (male=1)	0.469 (0.166)	0.160 (0.117)	0.287 (0.104)	0.270 (0.104)	-0.111 (0.061)
Race (white=1)	-0.523 (0.155)	-0.360 (0.114)	-0.288 (0.102)	-0.435 (0.102)	0.225 (0.061)
Marital Status (married=1)	-0.414 (0.122)	-0.391 (0.117)	-0.399 (0.105)	-0.435 (0.097)	0.210 (0.057)
Married*Gender (married female=1)	0.601 (0.191)	0.496 (0.167)	0.529 (0.156)	0.532 (0.139)	-0.257 (0.082)
Schooling (in years)	-0.086 (0.016)	-0.024 (0.018)	-0.036 (0.015)	-0.058 (0.015)	0.033 (0.009)
Tenure (in years)	-0.016 (0.007)	0.022 (0.010)	0.029 (0.007)	0.016 (0.006)	-0.013 (0.004)
Plant Closing (Shutdown=1)	0.081 (0.079)	-0.117 (0.092)	-0.044 (0.080)	-0.063 (0.072)	0.041 (0.042)
Written Notice	-0.312 (0.100)	-0.345 (0.158)	-0.079 (0.156)	-0.244 (0.106)	0.050 (0.061)
Unemp.Rate	0.146 (0.035)	0.184 (0.040)	0.170 (0.028)	0.183 (0.029)	-0.107 (0.018)
Constant	0.084 (0.177)	2.037 (0.137)	3.174 (0.116)	1.102 (0.299)	
scale parameter				1.743 (0.031)	
shape parameter				0.317 (0.076)	

Table 3: UNEMPLOYMENT DURATION REGRESSION RESULTS FOR 1995-97. The first entry in each cell is the regression coefficient point estimate with continuous regressors centered at their mean; in parenthesis are the standard errors. The standard errors for the QR estimators were computed as the half-length of a 95% bootstrap confidence interval divided by 1.96. (N=2762)

cally insignificant effects at higher quantiles. This result is especially clear in the second survey.

- Second, the impact of schooling is much stronger at low quantiles.This result obtains in both surveys.
- Third, again in both periods, the positive impact of tenure on unemployment duration is not present at low quantiles.
- And fourth, for the 1988 survey, the impact of plant shutdown fades away with the duration of unemployment.

Clearly, these effects would not be detected by conventional parametric approaches. Indeed, the results from the estimation of the AFT and Cox models appear to average out the time-varying regression effects.

There are variables such as the unemployment rate, age and race, and marital status that exert a statistically significant influence throughout the entire distribution. More interestingly, covariates such as education and pre-notification are only relevant on the left tail of the duration distribution, that is, for short-term unemployment.

It is worth noting that the variables that have significantly higher effects during the early phase of the unemployment spell very likely reflect the influence of on-the-job search (advance notice of displacement and dislocation by plant closing) or human capital (as captured by schooling). In the latter case it can be argued that larger human capital endowments are associated with greater job opportunities and higher opportunity costs of unemployment that necessarily erode with the progression of the unemployment spell. A number of explanations can be suggested here. Human capital depreciation, unobserved individual heterogeneity correlated with the measures of human capital, or stigmatization would lead to a fading human capital effect on the transition rate out of unemployment.

It has been argued that the beneficial effects of pre-notification accrue via the increase in on-the-job search intensity (Addison and Portugal, 1992). Faced with the prospect of an imminent discharge, the worker will engage in on-the-job search. If successful, he or she will experience a short spell of unemployment. Identically, workers displaced by reason of plant closing — in comparison with workers dismissed due to slack work or position shifted or abolished — benefit from an essentially short-term advantage conveyed by job search assistance and early (and unmistakable) warning of displacement. In essence, both on-the-job search and human capital depreciation point to time varying effects of the covariates and, thus, to non-proportional hazard. These types of effects may be labeled "transient effects" after Cox and Oakes (1984).

Despite broad agreement between the regression coefficient estimates from the two surveys, there are, however, some differences. The most striking change is related with the impact of plant closing which is very strong in the first survey, but vanishes in the second. Also interesting is the notable strengthening of the effect of the unemployment rate variable from the 1988 sample to the 1998 one. Finally, there is some indication that the influence of advance notice is stronger and persists for a longer period in the 1998 survey.

5 Changes in unemployment duration

5.1 Identifying the sources of changes

The law of total probability implies that changes over time in the distribution of unemployment duration may result from changes in the distribution of the conditioning variables or from changes in the conditional distribution of duration itself or both. Figure 3 sorts out these contributions.

Panel (1,2), "Changes in marginal hazards", plots the difference between that marginal hazard function for 1998 and the marginal hazard for 1988. These marginals are those implied by our model for the conditional unemployment duration and by the actual distribution of covariates in each year. They were estimated using the methods of M&M; since these are based on resampling and it is always dangerous base inferences on a single realization, we plot estimates for several (20) samples. For comparison, the first panel plots the changes in the empirical hazards, that is, those estimates from the actual data on unemployment duration. It is immediately apparent that the marginal implied by the model, capture pretty well the actual change in unemployment duration.

The plots in the second row represent counterfactual decompositions of the change in the (marginal) hazard changes. We compare the marginal hazard functions for 1998 with those that would have prevailed if the covariates were distributed as in 1988, the "covariates contribution"; we also compare the marginal hazard and survival functions for 1998 with those that would have prevailed if the conditional quantile function of unemployment duration was as in 1988, the "coefficients contribution". Again several realizations of the estimates were

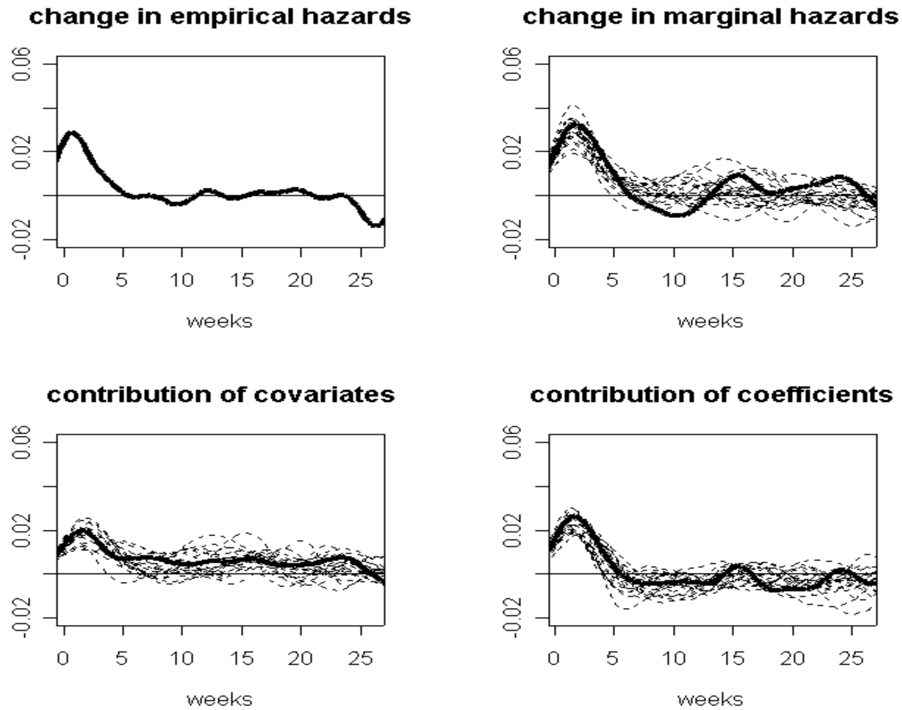

Figure 3: Hazard Changes

plotted. These reveal that overall contribution of the covariates and of the coefficients is roughly evenly splitted. The shift to the left on the unemployment duration distribution owes to significant changes in both the shape of the covariates' distribution and of the conditional distribution. Having said that, it is interesting to note that the contribution of the covariates is more important at low durations and persists beyond 26 weeks, in contrast with the contribution of the conditional distribution which looses steam at around 26 weeks.

	Quantiles		
	20	50	80
Marginal	-1.960;-1.159	-4.402;-2.291	-8.875; -4.057
Cont. Cov's	-1.219;-0.498	-3.308; -1.331	-9.818;-4.267
Cont. Coef's	-1.435;-0.741	-2.525;-0.680	-3.682;0.548

Table 4: Change in the quantiles of the unemployment distribution. 90% intervals estimates (in weeks) of the changes in quantiles (1997"minus" 1987) of the marginal and of the counterfactual distributions (based on 500 replications).

In table 4 assesses more rigoursly this same conclusion. For instance, the second row presents a 95% confidence interval for the difference between a given quantile of the unemployment duration in 1998 and the value that the same quantile would have had if the joint distribution of the covariates was as in 1988. Duration as shifted to the left at all points: smaller durations have become smaller and larger durations. On the middle and left tail, both factors played a role in explaining that shift.On the right tail, however, only the contribution of covariates appears to matter[5]. Which covariates play a significant role will be analyzed in what follows.

5.2 What is behind the changes in conditional and marginal duration?

Table 5 looks more deeply to the contribution of the changes in the covariates. Most of them do not affect significantly the distribution of the unemployment durations. It appears that the decisive factor reshaping the distribution are the changes in the unemployment rates, partially offseted by the changes of the age distribution.

	Quantiles		
	20	50	80
Age	0.068	0.433	1.417
	-0.214;0.383	-0.422;1.269	-0.803;3.645
Unemp. Rate	-0.567	-2.121	-6.368
	-0.967;-0.168	-3.409;-0.972	-9.369;-3.372

Table 5: CONTRIBUTION OF SELECTED COVARIATES TO THE CHANGE IN THE QUANTILES OF THE UNEMPLOYMENT DISTRIBUTION. Mean and 90% interval estimates (in weeks) of the changes in the quantiles (1997"minus" 1987) of the marginal and of the counterfactual distributions (based on 500 replications).

The contribution of the change in the distribution of the unemployment rate

[5]There are alternative ways of evaluating the contributions of coefficients and covariates. In the table below, we take as a reference 1988: the "contribution of covariates" now compares the the counterfactual distribution that would result if the covariates were as in 1998 but the conditional as in 1988 with the marginal for 1988; the "contribution of coefficients" compares the the counterfactual distribution that would result if the conditional was as in 1998 but the covariates were as in 1988 with the marginal for 1988. It is clear that the conclusions dont change; if anything, the contribution of the coefficients appears less significant.

	Quantiles		
	20	50	80
Cont. Cov's	-0.903;-0.049	-2.625;-0.355	-6.836;-2.197
Cont. Coef's	-1.131;-0.255	-2.038;0.246	-2.089;3.572

is felt throughout, but is especially strong for longer durations; the impact at the 80th quantile is estimated in about 6 weeks (or, 27%): if the unemployment rate was as in 1988, that quantile would be 27% bigger. Age, on the other hand, acts chiefly on the right tail but its effect, if any, is to increase unemployment duration, a result in line with the findings in Abraham and Shimer (2001).

We turn now to the individual components of the contribution of conditional distribution. Figure 4 shows the changes three such distributions. The panels in the first row clearly reveal the leftward shift in the unemployment distribution gauged either by the survivor or by the hazard function. When one controls for the "reference population" (that is, when the conditional distributions are evaluated at the same values of the conditioning variables), as it is done in the two bottom rows, the shift is less pronounced and also less clear cut. For instance, when both conditional functions are evaluated at the 1988 sample averages of the continuous covariates (second row), the direction of the change is much less transparent since the hazards cross several times. However, it is arguable that even here there is a reduction of shorter durations say, up to five weeks.

Since we modeled conditional distributions by quantile regressions, analyzing individual contributions to conditional distribution is tantamount to analyze the changes in the quantile regression coefficients (see Figure 6).

Since the quantile regressions were estimated with centered continuous regressors (see footnote 4), the change in the intercepts just mirrors the shift of the survivor function to the left depicted in the frame (1,1) of Figure 4. Again one sees that the reduction was more pronounced for shorter durations. But one already knows that part of that shift owes to changes in mean values of covariates such as unemployment rate and age.

Other significant changes in the coefficient contribution are exhibited by the plant closing and unemployment rates coefficients. In both cases the changes in the coefficients work in the direction of an increase in the unemployment duration. Changes in the gender effects at low quantiles and of advance notice effects at the median are also noticeable, although their overall impact in the unemployment duration distribution is rather muted.

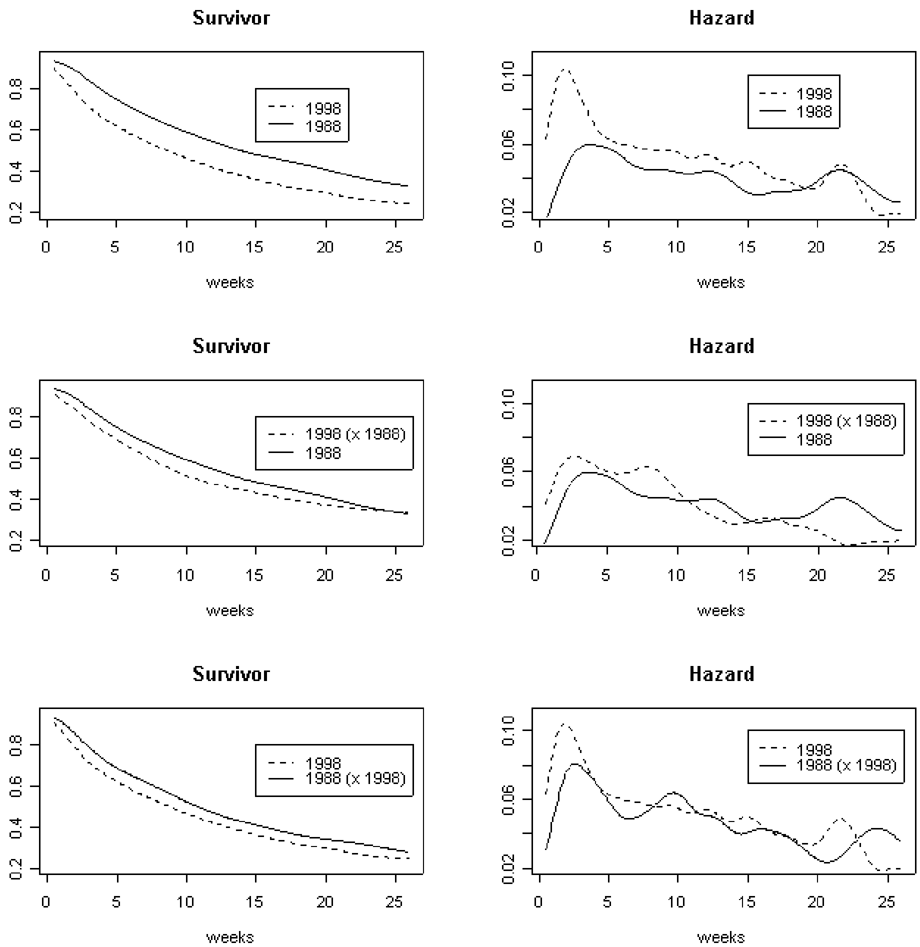

Figure 4: CONDITIONAL SURVIVOR AND HAZARD FUNCTIONS. Evaluated at baseline population of white married males, with average schooling and tenure, displaced without notice and by reasions other than plant closing in a state with average unemployment. The second and third rows use the 1988 and 1998 sample means, respectively.

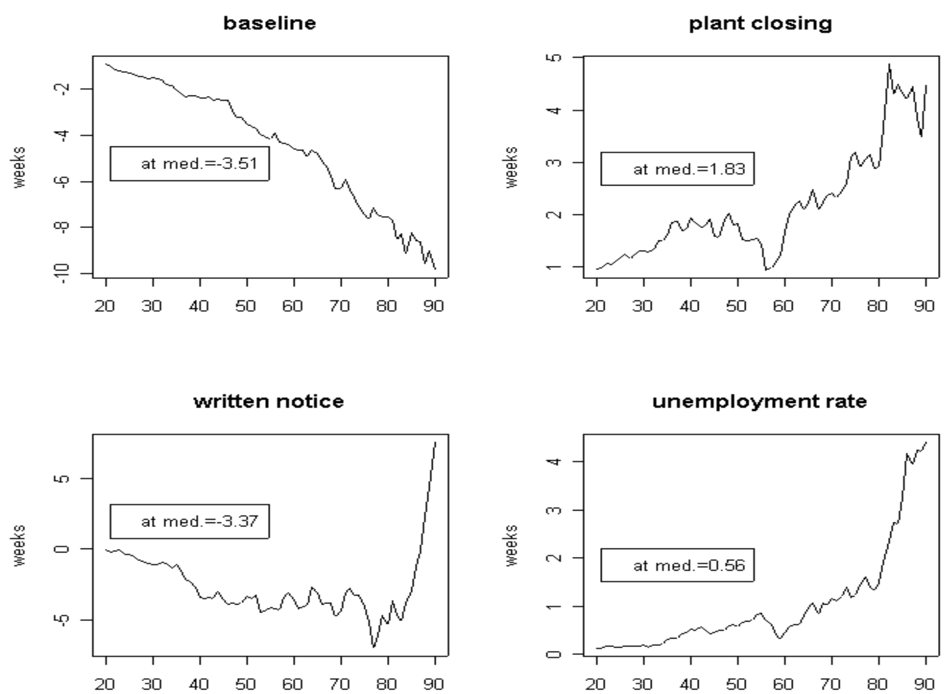

Figure 5: Impact of changes in QR coefficientsBaseline as in Figure (4). The "unemployment rate" impact refers to a 1 p.p. increase.

6 Conclusions

Comparing the DWS survey from 1988 with the one collected in 1998, there are noticeable changes in the shape of the unemployment duration distribution. We provided suggestive evidence that the unemployment duration distribution shifted leftward, leading to significantly higher hazard rates at the early phase of the unemployment spell.

We employed the decomposition method proposed by Machado and Mata (2003)to disentangle the contribution of the changes generated by covariates' distribution and by the conditional distribution and adapted it to a duration analysis framework.

Proceeding this way we arrived to the conclusion that the "coefficient contribution" and the "covariate contribution" worked in roughly equal parts in reshaping the unemployment duration distribution. The main driving force behind the "covariate contributions" was the sharp leftward move in the unemployment rate distribution, whereas the main factor behind the "coefficient contribution" was the sharp decline in the intercept. Those forces were partially counteracted by the ageing of the displaced population, the striking absence of impact from being displaced via a plant shutdown, and the higher sensitivity of unemployment duration to unemployment rates.

The indication that a significant part of the change in the unemployment duration distribution is unrelated with observed characteristics of the displaced workers is unfortunate, in the sense that it makes more difficult to pin down the routes of this leftward movement of the unemployment distribution. A possible interpretation of this result is that it is generated by upward trend of job to job transitions in the U.S. labor market indicated by Farber (2003). Our evidence is not necessarily inconsistent with the one provided by Abraham and Shimer (2001) as well. Indeed, the impact of changes in covariates leads to visibly lower hazard rates for longer term unemployed, possibly generated by the ageing of the displaced worker population pointed by Abraham and Shimer.

A note of caution is also in order. This results can not be generalized to the whole U.S. labor market since they rely solely on the joblessness experience of displaced workers. For example, the unemployment experience of job market incomers and reentrants or job quitters was not contemplated. A longer time frame may also prove to be necessary in order to circumvent outcomes that may be cycle idiosyncratic.

Finally, the use of a censored quantile regression model provided a flexible and thorough representation of unemployment duration distribution, and enabled a natural operational distinction between short- and long-term unemployment. The Machado and Mata (2003) decomposition method proved to be a useful analytical tool to study duration data.

References

Abraham, K. and R. Shimer (2001). Changes in unemployment duration and labor force attachment. NBER Working Paper No. 8513.

Addison, J. and P. Portugal (1987). On the distributional shape of unemployment duration. *Review of Economics and Statistics*, 68, 520-526.

Addison, J. and P. Portugal (1992). Advance notice and unemployment: New evidence from the 1988 Displaced Worker Survey. *Industrial and Labor Relations Review*, 45, 645-664.

Bilias, Y., Chen, S. and Z. Ying (2000). Simple resampling methods for censored regression quantiles. *Journal of Econometrics*, 99, 373-386.

Bassett, G. and R. Koenker (1986). Strong consistency of regression quantiles and related empirical processes. *Econometric Theory*, 2, 191-201.

Buchinsky, M. (1994). Changes in the U.S. wage structure 1963-1987: Application of quantile regression. *Econometrica*, 62, 405-458.

Cox, D. and D. Oakes (1984). *Analysis of survival data.* Chapman

Farber, H. (2003). Mobility and stability: the dynamics of job change in labor markets. in O. Ashenfelter Maddala and D. Card, eds. *Handbook of Labor Economics, vol. 3B*, North-Holland.

Fitzenberger, B. (1997). A guide to censored quantile regression. in G. Maddala and C. Rao, eds. *Handbook of Statistics, vol. 15*, Elsevier Science B. V.

Koenker, R. (1991-). Quantreg: A quantile regression package for R. http://cran.r project.org.

Koenker, R. and G.S. Bassett Jr. (1978). Regression Quantiles. *Econometrica*, 46, 33-50.

Koenker, R. and S. Portnoy (1997). THe Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators. *Statist, Sci.*, 12, 279-300.

Machado, J. and J. Mata (2003). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, forthcoming.

Manski, C. (1988). *Analog estimation methods in econometrics.* Chapman and Hall.

Powell, J. (1984). Least absolute deviations for the censored regression model. *Journal of Econometrics*, 25, 303-325.

Powell, J. (1986). Censored regression quantiles. *Journal of Econometrics*, 32, 143-155.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall.

Appendix

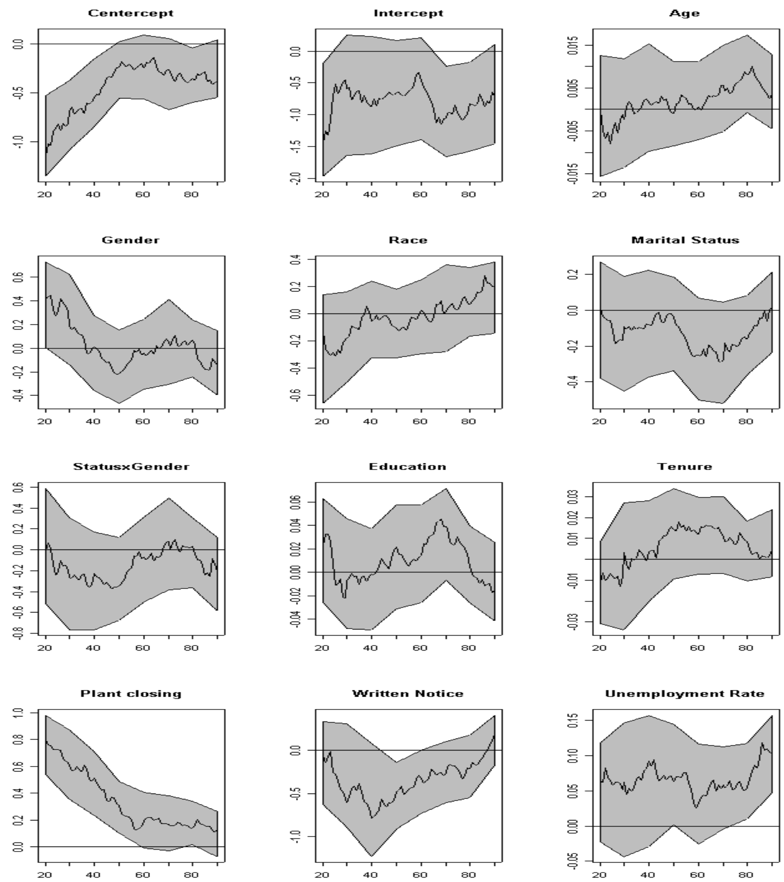

Figure 6: DIFFERENCE OF QR COEFFICIENTS 1998 minus 1988; shaded region represents 90% bootstrap confidence intervals for the deciles.