

Latent and Behavioral Responses to Extensions in  
Unemployment Insurance Benefits

Gordon B. Dahl

UC San Diego

[gdahl@ucsd.edu](mailto:gdahl@ucsd.edu)

June 2011

## **Abstract**

An important question facing economists and policymakers is how long individuals would collect unemployment insurance (UI) if it were made available for a longer period of time. This is a difficult task because (i) distributional assumptions can have a large impact past the censoring point (i.e., after UI benefits are exhausted) and (ii) there may be a behavioral response to any change in the maximum allowed benefit duration. To estimate the survival function past the censoring point, I adopt a semiparametric approach which builds on Chen, Dahl, and Khan (2005). I flexibly model the location and scale parameters of an accelerated failure time (AFT) model, without specifying the error term distribution. Using administrative-level data from New Jersey's UI system, the semiparametric estimates predict a relatively flat exit rate from UI past the censoring point in the absence of a behavioral response. In contrast, the parametric Weibull model significantly biases UI exit rates upwards. To estimate the incentive effects past the censoring point of 26 weeks, I take advantage of a unique policy experiment which exogenously increased the maximum duration from 26 to 39 weeks (see Card and Levine, 2000). I find only a modest behavioral response in weeks 26 to 39 due to the extension in benefits. The results suggest the long-term unemployed have a difficult time re-entering the labor force, with UI benefits serving as an important income maintenance program for these workers.

# 1 Introduction

Unemployment insurance (UI) plays an important role in buffering the income shocks of individuals who experience job loss in the United States and many other countries. In the United States, the maximum duration of UI benefits is set at the state level, but is generally 26 weeks during normal economic times. An important question facing economists and policymakers is how long individuals would collect unemployment insurance if it were made available for a longer period of time.

However, many claimants exhaust their UI benefits, resulting in a significant fraction of censored observations. This censoring makes estimation of the complete survival curve and the predicted costs of an extension difficult. Even though distributional assumptions generally make little difference for the non-censored portion of a survival curve (i.e., before UI benefits are exhausted), they can have a large impact past the censoring point. Moreover, there may be a behavioral response to any change in the maximum allowed benefit duration.

Beyond the costs of an extended benefits program, researchers and policymakers are also concerned about UI's effect on job search behavior. Extended benefits might provide additional time to find a better job match, but many worry that such extensions might also reduce claimants incentives to find a job and hence slow re-entry into the labor force. Both costs and disincentive effects are often cited as reasons to limit the length of UI benefits.

In this paper I estimate bounds on both the costs and disincentive effects of extending UI benefits in a flexible way. I find the long-term unemployed are predicted to exit the UI system very slowly over time. I also find the incentive effects to delay exiting UI are minimal. Combining these two results, it appears the long-term unemployed have a difficult time

re-entering the labor force, with UI benefits serving as an important income maintenance program for these workers.

To estimate the survival function past the censoring point, I adopt an approach which builds on Chen, Dahl, and Khan (2005). UI spell duration is modeled in an accelerated failure time (AFT) framework, with flexible modeling of the location and scale parameters as a function of observed characteristics of individuals. The approach is semiparametric in two ways: (i) it does not specify the distribution of the error term, and (ii) it does not specify the functional form for how the  $x$ 's influence median failure time. The key requirements for identification are a median zero error term and a relatively large support for the error term compared to the effect of the independent variables on the location and scale parameters. Since the conditional distribution of the error term can be identified up to scale, I can also estimate other quantiles of the survival function beyond the censoring point.

Semi-parametric estimation in this framework quickly becomes computationally infeasible with more than one or two continuous variables. To avoid this curse of dimensionality, I discretize the data into a set of mutually exclusive and exhaustive cells. These cells are based on demographic information, characteristics of a claimant's previous job, and a measure of local labor market conditions. Since estimates are calculated at the cell-level, I propose a simple way to combine these cell-level estimates to create aggregate survival functions.

I estimate the survival curve for New Jersey's UI program in the late 1990s, using a large dataset of individual-level administrative records. I find that incorrect distributional assumptions can significantly bias the results. The Weibull model is a commonly used parametric model, and a special case of the semiparametric AFT model used in this paper. While both estimates are similar to the Kaplan-Meier estimate below the censoring point,

the Weibull estimate predicts a much quicker drop off in UI claims past the censoring point. A policymaker who based predictions on this incorrect distributional assumption would grossly underestimate the cost of an extension in UI benefits, even in the absence of any behavioral response. While some parametric AFT models based on alternative distributions do not underestimate costs, the point is that it is difficult to know which distribution to use in advance.

I then estimate an upper bound on the behavioral response using a unique policy experiment first explored by Card and Levine (2000). In the late 1990's, New Jersey increased the duration of UI benefits from 26 weeks to 39 weeks, for arguably exogenous reasons unrelated to economic conditions. I compare the actual duration of UI spells from 1 to 39 weeks during the policy experiment to that predicted using the semiparametric estimator and data outside the policy experiment period. Under a monotonicity assumption about selection into New Jersey's extended benefits program and an assumption that behavioral responses do not decrease spell length, the difference between the two survival functions yields an upper bound estimate of the behavioral response. There is little evidence of a behavioral response past the censoring point. My findings suggest that the slow re-entry of the long-term unemployed is not primarily due to a disincentive effect where workers decrease their job search intensity.

The remainder of the paper proceeds as follows. I first briefly discuss the previous literature and outline how to model UI receipt in an AFT framework. In Sections 3 and 4 I discuss the data from New Jersey's UI system and lay out the semiparametric estimation approach. Section 5 discusses results, first at the cell level and then at a more aggregated level. Section 6 estimates an upper bound on the behavioral response and Section 7 concludes.

## 2 Modeling Unemployment Insurance Receipt

Previous research on the determinants of UI spell length has focused on the effect of benefit levels (Anderson and Meyer(1997), Chetty (2008), Ham and Rea(1987), Hunt(1995), Lalive et al. (2006), McCall(1995), and Meyer(1992)) and the maximum allowed benefit duration (Card and Levine(2000), Jurajda and Tannery (2003), Katz and Meyer(1990), Meyer(1990), Moffitt and Nicholson(1982), Moffitt(1985), Ours and Vodopivec (2006), Roed et al (2008), Woodbury and Murray(1997)). This literature generally concludes that both higher benefit levels and higher maximum durations significantly increase the length of an individual's spell of unemployment insurance receipt. The empirical findings mainly rely on variation in benefit amounts and maximum benefit lengths over time, across states, or between individuals for identification.

While such studies have helped researchers partly understand the determinants of UI spell length, one limitation of many studies is the endogeneity of the identifying variation. Time variation is not likely to be exogeneous since benefits are usually extended when labor market conditions are poor (see Blaustein, et al(1993) and Blank and Card (1991)). Individual states also sometimes extend the maximum duration of benefits in response to a slack labor market. These changes are endogeneous since they occur precisely when spell lengths and benefit exhaustions would otherwise be predicted to increase. To combat problems of endogeneity, a small number of researchers have taken advantage of quasi-experimental policy changes (e.g., Card and Levine (2000), Lalive et al. (2006), Meyer (1992), and Ours and Vodopivec (2006)). Of particular relevance is the paper by Card and Levine (2000), since I make use of the same policy experiment for a different purpose, as described later in the

paper.

Prior research has concentrated on the responses of individuals with non-censored spells. For example, many studies document a sharp spike in exit rates immediately prior to benefit exhaustion. But existing studies say little about spell durations past the censoring point, since such predictions would need to be made either using parametric assumptions about the distribution of failure times or functional form assumptions for how covariates affect failure times past the censoring point. In contrast, this paper uses an approach based on Chen, Dahl, and Khan (2005), which allows estimation of the survival function beyond the censoring point with a minimum of parametric and structural assumptions.

I model log-failure time (i.e., when an individual stops collecting UI) as a semiparametric location-scale model. The model of interest is

$$\log t_i^* = \mu(x_i) + \sigma(x_i)\epsilon_i \tag{1}$$

$$t_i = \min(t_i^*, 26) \tag{2}$$

The variable  $t_i$  is the observed number of weeks a claimant collects unemployment insurance benefits, the exogenous vector  $x_i$  contains claimant characteristics, and  $\epsilon_i$  is a homoskedastic and median zero error term. The censoring point is fixed at 26 weeks; the latent number of weeks a claimant would like to collect benefits,  $t_i^*$ , is observed only when a claimant collects benefits for less than this censoring point.

As described in the next section, the approach taken in this paper can estimate the location function  $\mu(x_i)$  as well as the conditional distribution of failure times beyond the censoring point. The reader is cautioned that the estimator does not necessarily predict what will

happen after an increase in maximum duration, since both censored and non-censored individuals may exhibit a behavioral response to a longer time limit. Conceptually, the effect of extending the maximum duration can be separated into the sum of claimants (potentially censored) desired number of weeks under the current time limit plus any behavioral shift associated with the incentive effects of a longer program. In Section 6, I estimate an upper bound on the size of the incentive effect for a temporary increase in UI benefits in New Jersey. I find weak evidence of a small incentive effect, implying the estimates in this setting come close to approximating the total effects. More generally, the approach outlined here would estimate a lower bound on the effect of increasing the UI time limit, as long as time limit increases do not shorten individual spell lengths.

The accelerated failure time (AFT) model adopted here is a common approach to estimating failure time models. Another widely used model, particularly in economics, is the proportional hazards (PH) model.<sup>1</sup> When a Weibull distribution is assumed, the model has both an AFT and a PH interpretation. Although convenient, there is little theoretical justification for an AFT (or PH) model based on a particular distribution, and faulty distributional assumptions can significantly bias the results. For example, Moffitt(1985) reports that estimates of UI spell length for Tobit-type models are very sensitive to the assumed distribution. Any inconsistencies generated from incorrectly specifying the distribution are likely to be further compounded when predicting survival times beyond the censoring point.

Researchers have relaxed distributional assumptions using a variety of semiparametric approaches for censored duration models (e.g., Cox(1972), Powell(1984), Horowitz(1996)). While useful, these semiparametric approaches cannot identify the shape of the survival

---

<sup>1</sup>One advantage of the PH model is that it can more easily accommodate time-varying covariates.



function beyond the point of censoring without further assumptions. A commonly used assumption for a variety of estimators is to specify the functional form of the location function  $\mu(x_i)$ ; however, economic theory provides little guidance on what functional form to assume. The method proposed in this paper does not specify the distribution of the homoskedastic component of the error term or the form of the location function, but still permits estimation of the conditional distribution of failure times beyond the censoring point.

### **3 Data**

The model is estimated using individual-level administrative records from New Jersey's UI system. The baseline dataset is drawn from the 322,907 individuals who received their first payment between June 1, 1996 and October 25, 1997. Of these claimants, 63 percent were eligible for the legislated maximum of 26 weeks of UI benefits. The sample is restricted to claimants between the ages of 18 and 65, with complete demographic information, with no more than one week of partial UI benefits, and who were eligible for 26 weeks of UI. These restrictions result in a usable dataset containing 192,162 observations (see Card and Levine (2000) for further details about this dataset).

New Jersey's UI system is administered at the state level, with benefits being financed by a tax on both workers and firms. These taxes are subject to maximum and minimum rates, and are partially experienced rated. Individuals can collect unemployment benefits if they have a sufficiently long work history and they continue to actively seek work. Benefits are paid weekly and are based on an individual's previous earnings with a maximum benefit of \$362 in 1996. During the period of the data, New Jersey experienced a strong labor market;

for the individuals in this dataset, the median unemployment rate measured at the county level was 5.5 percent. During the baseline period of the data, the maximum duration of UI benefits remained constant at 26 weeks. Of those individuals eligible for 26 weeks in the dataset, 43 percent exhausted their benefits.

A variety of characteristics influence the length of time an individual remains on unemployment insurance. In the analysis, I include demographic information, characteristics of a claimant's previous job, and a measure of local labor market conditions. The curse of dimensionality quickly makes the approach developed in Chen, Dahl, and Khan (2005) computationally infeasible with more than one or two continuous variables in the model, since since the approach requires local polynomial estimation. To make estimation feasible, I discretize the continuous covariates so the data can be grouped into a set of mutually exclusive and exhaustive cells. Many important characteristics such as race, gender, and union status are already discrete. Other characteristics, such as years of schooling, previous weekly earnings, age, or tenure, can be broken up into a few defining categories. Nonparametric quantile regression applied to this type of cell-grouped data is simple and computationally fast. With such cell-grouped data, the estimator discussed in the next section requires only easily computed estimates of the median and other quantiles at the cell level combined with simple algebra.

Table 1 contains summary statistics for different characteristics during the baseline period, including the fraction of claimants who exhaust their UI benefits. A typical UI claimant in New Jersey is male, white, middle-aged, not a union member, and has a high school degree or less. Claimants have varied previous earnings histories and a large fraction of UI claimants have been at their job for less than two years. Exhaustion rates differ widely

across characteristics. For example, the exhaustion rate for whites is 40 percent compared to 53 percent for blacks and the exhaustion rate for union members is 34 percent compared to 45 percent for non-union members. Exhaustion rates are also higher for women, older workers, workers with long tenure, and in counties with high unemployment rates. Claimants from the construction, agriculture, or mining industries also have markedly shorter UI spells on average.

The complete interaction of all the characteristics in Table 1 yields a possible 2,586 non-overlapping cells, 2,225 of which are nonempty. In 1,528 of the cells, the median time on UI benefits is below the censoring point of 26 weeks. These cells contain approximately 69 percent of all observations. The remaining 31 percent of the data are grouped in 697 cells where more than 50 percent of claimants are observed to exhaust benefits.

## 4 Nonparametric Estimation Approach

This paper adapts the estimation approach developed in Chen, Dahl, and Khan (2005), where the estimator is motivated by an identification argument. It is immediate that  $\mu(x)$  is identified for any value of the covariates  $x_0$  such that  $\mu(x_0) < \log(26)$ , i.e., for any  $x_0$  where the median lies in the noncensored region. In other words, if the true median implies less than 50% censoring for a particular value of the covariates, the 0.5 quantile can be identified. Under minimal assumptions, it is possible to show the location function  $\mu(x)$  is identified in the censored region as well (i.e., for values  $x_1$  where  $\mu(x_1) > \log(26)$ ). I outline this proof since it motivates the cell-level estimation technique used in this paper. As the outline will show, the key requirements for identification are a median zero error term

and a relatively large support for the error term compared to the effect of the independent variables on the location and scale parameters.<sup>2</sup>

Consider the AFT model written out in equations (1) and (2). Let  $q_{\alpha_j}$  denote the  $j^{\text{th}}$  quantile of log failure time. Assuming that  $\mu(x)$  and  $\sigma(x)$  are finite for all values of  $x$  and that  $\epsilon$  has support on the entire real line, there exist quantiles  $\alpha_1 < \alpha_2 < 1$  (for the homoskedastic component of the error term  $\epsilon$ ) such that

$$\mu(x) + c_{\alpha_i}\sigma(x) < \log(26) \quad i = 1, 2 \quad (3)$$

Equation (3) follows since there is always some quantile of  $\epsilon$  which will result in a noncensored observation, no matter the value of the covariates  $x$ . Thus the following relationships hold

$$q_{\alpha_1}(x) = \mu(x) + c_{\alpha_1}\sigma(x) \quad (4)$$

$$q_{\alpha_2}(x) = \mu(x) + c_{\alpha_2}\sigma(x) \quad (5)$$

Using these two relationships to substitute out  $\sigma(x)$  and solve for  $\mu(x)$  yields

$$\mu(x) = \frac{c_{\alpha_2}}{\Delta c}q_{\alpha_1}(x) - \frac{c_{\alpha_1}}{\Delta c}q_{\alpha_2}(x) \quad (6)$$

where  $\Delta c = c_{\alpha_2} - c_{\alpha_1}$ . If the fractions  $\frac{c_{\alpha_1}}{\Delta c}$  and  $\frac{c_{\alpha_2}}{\Delta c}$  could be identified, then  $\mu(x)$  could be identified using the above equation. Note that  $\frac{c_{\alpha_2}}{\Delta c} = 1 + \frac{c_{\alpha_1}}{\Delta c}$ , so all that is required is one of these two fractions.

---

<sup>2</sup>For simplicity, a relatively large support for the error term is taken to be support on the entire real line, while the location and scale parameters are assumed to be finite for all values of  $x$ . For details on the technical assumptions required and a formal proof of consistency, see Chen, Dahl, and Khan (2005).

The identification of the location function for values of  $x_0$  where the median is in the non-censored region (i.e.,  $\mu(x_0) < \log(26)$ ) can be used to identify  $\frac{c_{\alpha_1}}{\Delta c}$  and  $\frac{c_{\alpha_2}}{\Delta c}$ . The idea is to algebraically combine the following values of the conditional quantile function at the three distinct quantiles 0.5,  $\alpha_1$ ,  $\alpha_2$ , evaluated at the regressor value  $x_0$ .

$$q_{0.5}(x_0) = \mu(x_0) \tag{7}$$

$$q_{\alpha_1}(x_0) = \mu(x_0) + c_{\alpha_1}\sigma(x_0) \tag{8}$$

$$q_{\alpha_2}(x_0) = \mu(x_0) + c_{\alpha_2}\sigma(x_0) \tag{9}$$

After simple algebra, this enables one to write  $\frac{c_{\alpha_1}}{\Delta c}$  and  $\frac{c_{\alpha_2}}{\Delta c}$  as

$$\frac{c_{\alpha_1}}{\Delta c} = \frac{q_{\alpha_1}(x_0) - q_{0.5}(x_0)}{\Delta q(x_0)} \tag{10}$$

$$\frac{c_{\alpha_2}}{\Delta c} = \frac{q_{\alpha_2}(x_0) - q_{0.5}(x_0)}{\Delta q(x_0)} \tag{11}$$

where  $\Delta q(x_0) = q_{\alpha_2}(x_0) - q_{\alpha_1}(x_0)$ . The ability to substitute out  $\mu(x_0)$  from these expressions relies on a median zero error term, an assumption which is used in equation (7).

This immediately translates into identification of  $\mu(x)$  from the relationship

$$\mu(x) = \frac{q_{\alpha_2}(x_0) - q_{0.5}(x_0)}{\Delta q(x_0)} q_{\alpha_1}(x) - \frac{q_{\alpha_1}(x_0) - q_{0.5}(x_0)}{\Delta q(x_0)} q_{\alpha_2}(x) \tag{12}$$

Equation (12) forms the basis of estimation in this paper. Since the data will be partitioned into cells, the values of the  $x$  variables will uniquely characterize each cell. Cells where the median is uncensored will be used to estimate both  $\frac{c_{\alpha_1}}{\Delta c}$  and  $\frac{c_{\alpha_2}}{\Delta c}$  as described in equations (10) and (11). These estimated coefficients are common to all cells, no matter the value of the covariates  $x$ . These estimated coefficients can therefore be used to get cell-level

estimates of the median function for all cells, including those whose median is censored, using equation (12).

Estimation proceeds in three stages. Let  $\hat{q}_{\alpha_j}$  denote the estimated value of  $q_{\alpha_j}$  and let  $\hat{c}_1$  and  $\hat{c}_2$  denote the estimated values of  $\frac{c_{\alpha_1}}{\Delta c}$  and  $\frac{c_{\alpha_2}}{\Delta c}$ . First, median log failure times are calculated for each cell where the median is not censored. In the second stage, estimates of the unknown constants  $\frac{c_{\alpha_1}}{\Delta c}$  and  $\frac{c_{\alpha_2}}{\Delta c}$  are calculated. This requires first estimating the  $\alpha_1$  and  $\alpha_2$  cell quantiles (two quantiles lower than the median). For each cell where the median is below the censoring point, the expressions  $\hat{q}_{\alpha_1} - \hat{q}_{0.5}$  and  $\hat{q}_{\alpha_2} - \hat{q}_{0.5}$  are divided by  $\hat{q}_{\alpha_2} - \hat{q}_{\alpha_1}$ . The sums of these cell calculations, using only cells where the median is below the censoring point, yield the estimates  $\hat{c}_1$  and  $\hat{c}_2$ . The third stage estimates the median separately for each cell by taking a simple algebraic combination of the estimated constants  $\hat{c}_1$  and  $\hat{c}_2$  (common to all cells, but estimated only using the cells where the median was not censored) and the estimated quantiles  $\hat{q}_{\alpha_1}$  and  $\hat{q}_{\alpha_2}$  (specific to each cell), as suggested by equation (12)

$$\hat{\mu}(x) = \hat{c}_2 \hat{q}_{\alpha_1}(x) - \hat{c}_1 \hat{q}_{\alpha_2}(x). \quad (13)$$

Once medians have been estimated for all cells, other quantiles can be estimated for each cell as well since the conditional distribution of  $\epsilon_i$  can be identified up to scale. The  $\alpha_j$  quantile of a cell is estimated as

$$\hat{q}_{\alpha_j}(x) = \hat{\mu}(x) + \hat{c}_j (\hat{q}_{\alpha_2}(x) - \hat{q}_{\alpha_1}(x)) \quad (14)$$

where the estimator  $\hat{c}_j$  is the average of the expression  $(\hat{q}_{\alpha_j}(x_0) - \hat{q}_{0.5}(x_0)) / (\hat{q}_{\alpha_2}(x_0) - \hat{q}_{\alpha_1}(x_0))$  over cells where both the estimated median and the  $\alpha_j$  quantile lie below the censoring point. It should be noted that the estimated quantiles  $\hat{q}_{\alpha_1}$  and  $\hat{q}_{\alpha_2}$  appearing in

these expressions need not be the same as those used to calculate  $\hat{\mu}$ . They are kept the same largely for convenience; I find that the estimates are not overly sensitive to the choice of these two quantiles. Equation (14) allows estimation of the survival curve for each cell up to the largest  $\alpha_j$  quantile for which  $\hat{c}_j$  can be estimated using other cells which are not heavily censored.

## 5 Results

### 5.1 Estimates at the Cell Level

Estimation parallels the stages described in the previous section. In the first two stages, for each cell where the median lies below the censoring point, I calculate cell medians and obtain cell-level estimates of the two unknown lower quantiles of the error term up to scale. I chose the 30<sup>th</sup> and 40<sup>th</sup> quantiles for these two quantiles below the median. To construct the estimates, the 30<sup>th</sup> and 40<sup>th</sup> quantiles must be less than 26 weeks and the 40<sup>th</sup> quantile cannot equal the median (to avoid division by zero). The first restriction eliminates approximately seven percent of the data and the second restriction eliminates less than one quarter of one percent of the data. Of course, other quantiles could be used as well. Other choices, such as the 17<sup>th</sup> and 33<sup>rd</sup> quantiles, do not alter the general findings.

One gauge of the method's applicability to UI claims is how well the estimate of the location function compares to the observed median in cells with less than 50 percent censoring. Figure 1 plots the observed median versus the estimate of the location function for cells where the observed median is less than 26 weeks. For cells with few observations, the estimate  $\hat{\mu}$

is highly variable and has a large bootstrapped standard error. To simplify the graph, it plots only those cells with more than 30 observations. The observations are generally clustered around the forty-five degree line, although the estimate  $\hat{\mu}$  is somewhat higher in the weeks immediately prior to censoring.<sup>3</sup>

A primary objective of the empirical application is to estimate the location function for cells with more than a 50 percent exhaustion rate. (Note: Table 2 needs to be updated, as it refers to a different way of splitting the data into cells.) Since it is impractical to present estimates for all of these cells, in Table 2 we provide estimates for the subset of cells with more than 250 observations. Since only 13 percent of claimants belong to a union and union members have low exhaustion rates to begin with, all of the cells appearing in Table 2 contain only non-union claimants. Otherwise, there is a rich variety of characteristics defining cells. The estimated medians vary substantially, ranging from approximately 27 weeks (male, black, young, college, middle earnings, short tenure, high unemployment rate) to 40 weeks (female, black, mid-age, high school, high earnings, long tenure, high unemployment rate). The standard errors appearing in Table 2 are estimated using the bootstrap. The bootstrap estimates are based on 400 replications for each cell, with samples equal in size to the number of observations in a cell and drawn with replacement.

Conditional quantiles for all cells can be estimated up to the largest probability  $\alpha_j$  for which sufficiently many cells have  $\alpha_j^{th}$  quantiles observed below the censoring point. In the application, I chose 85 percent as the cutoff probability. There are 17 cells with an observed 85<sup>th</sup> quantile less than 26 weeks, with a total of 2,526 observations in these cells. Using information from these observations, I am able to estimate up to the 85<sup>th</sup> quantile

---

<sup>3</sup>Two outliers are omitted from the graph for visual clarity; these two cells have 34 and 42 observations.



for all cells in the dataset. For lower quantiles, of course, information can be used from more than just these 17 cells. For example, there are 119 cells (18,688 total observations) with an observed 70<sup>th</sup> quantile less than 26 weeks. (Note: Figure 2 needs to be updated, as it refers to a different way of splitting the data into cells. The new cell definitions enable estimation out to the 98<sup>th</sup> quantile. The other figures and tables in the paper utilize the new cell definitions.)

Figure 2 graphs the estimated quantiles of UI receipt for eight cells with a diverse set of characteristics. The varied shapes of the quantile functions suggest the presence of conditional heteroskedasticity. The top four panels show the estimates for four cells with observed medians well below the censoring point. The Kaplan-Meier estimates (with the axes reversed) closely track the estimates obtained using the semiparametric AFT model and generally lie within the bootstrapped 95 percent confidence intervals. As expected, these confidence intervals fan out as the fraction of active claims falls, i.e., for estimates of the higher quantiles. Notice that in panel 1, the observed 85<sup>th</sup> quantile occurs at less than 26 weeks. Intuitively, the structure of the assumptions allows cells with more severe censoring to take advantage of information on the error distribution from cells like those found in panel 1 to estimate quantiles beyond the censoring point.

The bottom four panels in Figure 2 depict estimates for cells with more severe censoring. For example, in panel 5 close to 60 percent of the claimants in this cell exhaust benefits. Even with such severe censoring, however, one can estimate quantile values greater than 26 weeks. Although one could plot up to the 85<sup>th</sup> quantile for each of these cells as done in the top panels, I choose instead to include the estimated quantiles until the point estimate exceeds 52 weeks. This choice was made so the scale of the graphs would more clearly

illustrate the comparison to the Kaplan-Meier estimate and what is happening immediately after the 26 week censoring point. As before, the Kaplan-Meier estimates are similar to the semiparametric estimates.

## 5.2 Aggregate Survival Functions Based on Cell-Level Estimates

While cell-level estimates of medians and other quantiles beyond the censoring point are useful, they do not provide a concise summary of what is happening at a more aggregate level. In this section I describe how to combine cell-level estimates to create aggregate survival functions. The idea behind aggregation is to take a weighted vertical sum of the survival curves of individual cells.

To calculate the aggregate survival function, first notice the failure time distribution for all the data,  $F(\cdot)$ , can be written as a weighted average of the failure time distributions for each cell. The weights are merely the fraction of observations belonging to each cell and do not depend on the failure time. Let  $F_i(\cdot)$  denote the distribution of failure times and  $w_i$  denote the weight for cell  $i$ . The  $p^{th}$  quantile is the number  $t$  such that

$$F(t) = \sum_i w_i F_i(t) = p \quad (15)$$

For each cell, quantile estimates up to the 85<sup>th</sup> quantile have already been obtained. These cell quantiles are just the inverse distribution functions,  $F_i^{-1}(\cdot)$ , evaluated at various probabilities. To calculate the unconditional  $p^{th}$  quantile for the entire population, one needs to find  $p_1, p_2, \dots, p_N$  such that

$$F_1^{-1}(p_1) = F_2^{-1}(p_2) = \dots = F_N^{-1}(p_N) \quad (16)$$

and

$$\sum_i w_i p_i = p \tag{17}$$

where  $N$  indicates the total number of cells the data have been divided into.

A grid method is used to compute probabilities associated with a given value for the inverse distribution function,  $F_i^{-1}(\cdot)$ . For each cell I estimate the quantiles at 1,000 probabilities evenly spaced between 0 and 1. To calculate the aggregate survival function at a specific failure time, for each cell  $i$  we first find the probability  $p_i$  which corresponds as closely as possible to the quantile equaling this failure time. That is, I take a weighted average of probabilities straddling integer failure times, where the weights on the probabilities depend on the distance of the associated quantiles from the integer failure time. I then take the weighted sum of these cell probabilities, where the weights are estimated by the fraction of all observations belonging to cell  $i$ . The overall survival function is evaluated at failure times which take on positive integer values, although it should be noted the survival function could be evaluated at other points as well.

One challenge in the current application is that I am limited to quantile estimates below the chosen cutoff, i.e., the 98<sup>th</sup> quantile. For cells where the 98<sup>th</sup> quantile occurs relatively early, for example at 28 weeks, what probability should be assigned for failure times greater than 28 weeks? The maximum probability for failure times beyond the 98<sup>th</sup> quantile is one, implying all remaining observations fail immediately after the 98<sup>th</sup> quantile. The minimum probability does not increase at all, implying that no observations fail after the 98<sup>th</sup> quantile. These two extreme assumptions provide conservative upper and lower bounds on the cell-specific probabilities associated with failure times beyond the 98<sup>th</sup> quantile in a cell.

Notice that aggregate estimates can be calculated for subsets of the data as well as for the entire dataset. For example, unconditional survival functions can be calculated for different races, men and women, or union and non-union members separately. In this paper I present just one such example, for the subset of data with severe cell-level censoring. Figure 3 displays the aggregate survival functions for claimants in cells with an observed median greater than 26 weeks. For this group of heavily censored claimants, the predicted median spell length is approximately 32 weeks, or 6 weeks longer than the actual censoring point. Hence, this group is of special interest when considering extensions to UI benefits. The figure includes the upper and lower bound estimates of the survival function calculated as described above, as well as the Kaplan-Meier estimate. The standard errors for the Kaplan-Meier estimate are relatively small, so to keep things visually simple confidence intervals for the Kaplan-Meier estimate are excluded from the graph.

For this group, the upper and lower bound estimates of the survival curve are very similar, and any differences are not visually apparent in the graph. This is because this set of individuals exhaust benefits quickly, so the 98<sup>th</sup> quantile generally occurs far later than 52 weeks, the last week shown in the graph. The estimates track the Kaplan-Meier estimate in the non-censored region fairly well. In the figure we have also added the estimated survival curve using a Weibull model, a model which has both a proportional hazards and an accelerated failure time (AFT) interpretation. It should be noted that my approach generalizes parametric AFT models, so if the Weibull model is correct it should yield a similar estimate. The pointwise confidence intervals appearing in the graph for the Weibull estimate are based on standard errors calculated using the delta method.

While the Weibull estimate and the upper and lower bound estimates are similar imme-

diately prior to the censoring point, they have very different shapes to the right of 26 weeks. In particular, the Weibull estimate is noticeably lower for failure times past the censoring point. For claimants belonging to heavily censored cells, the semiparametric approach predicts between roughly 35 percent of claims would still be active at 52 weeks compared to 26 percent for the Weibull model. Put another way, the median residual life (i.e.,  $\text{median}[t - 26 \mid t > 26]$ ) for my model is about 39 weeks compared to only 24 weeks for the Weibull model—a difference of more than 50 percent. These inconsistencies illustrate the costs of making incorrect distributional assumptions.

While not shown in the figure, some parametric AFT models based on alternative distributions do not underestimate costs. For example, assuming a lognormal distribution yields an estimated survival curve which is similar to the semiparametric survival curve. However, the point is that it is difficult to know which distribution to use in advance, especially since various distributions all fit the noncensored region fairly well.

Figure 4 graphs the aggregate survival functions for all claimants in the dataset. To the left of the censoring point, the upper and lower bound estimates are similar to the Kaplan-Meier estimate. As before, the bootstrap estimates for the pointwise confidence intervals are based on 400 replications, with samples equal in size to the total number of observations and drawn with replacement. Immediately prior to the censoring point at 25 weeks, around 42 percent of claims are estimated to be active compared to 41 percent for the Kaplan-Meier estimate. To the right of 26 weeks, the shape of the survival curve flattens out. As the number of weeks increase, the upper and lower bound point estimates start to diverge as a larger fraction of cell-level estimates exceed the 98<sup>th</sup> quantile. Six weeks after the censoring point, between 35.5 (upper bound) and 34.2 (lower bound) percent of claims are

estimated to still be ongoing. Although the bounds are somewhat wider in latter weeks, the estimates suggest a significant fraction of claimants would like to continue collecting UI benefits beyond the legally-specified maximum duration.

The information contained in the estimated survival functions can readily be used to derive estimates of the monetary cost of extending UI benefits in New Jersey. For each cell and each week, cell-level costs are calculated by multiplying the average weekly benefit for individuals in a cell by the number of claims estimated to be ongoing in a cell. Average payouts vary across cells since benefit amounts depend on pre-displacement wages and work history. Aggregate cost estimates are then calculated by summing these cell-level cost estimates.

The first week of UI claims is estimated to cost around 49 million dollars in benefit payouts. By the censoring point of 26 weeks, this aggregate weekly payout falls to approximately 20 million dollars with an estimated cumulative cost of roughly 880 million dollars. Since the survival curve flattens out after the censoring point, predicted costs decline slowly. Increasing the maximum duration by fifty percent to 39 weeks is predicted to cost between an extra 209 (lower bound estimate) and 221 (upper bound estimate) million dollars. The reader is reminded that these estimates represent aggregate costs assuming no behavioral shift associated with the incentive effects of a longer program. Since time limit increases should not decrease the length of time an individual receives UI, these estimates should be interpreted as the minimum cost of extending the maximum allowed duration. Any behavioral response would further add to the costs of extending UI benefits.

## 6 Behavioral Response

In this section, we take advantage of a unique policy experiment to estimate an upper bound on the behavioral response to an increase in maximum duration. We explore the behavioral response to New Jersey's Extended Benefits (NJEB) program, a temporary increase in UI duration first examined in a paper by Card and Levine (2000). In the late 1990's, New Jersey increased the duration of UI benefits from 26 weeks to 39 weeks, for arguably exogenous reasons unrelated to economic conditions (see Card and Levine for details).

I compare the actual duration of UI spells from 1 to 39 weeks during the policy experiment to that predicted using the semiparametric estimator and data outside the policy experiment period. Card and Levine compare hazard functions from 1 to 26 weeks, and the spike immediately prior to exhaustion. The focus of this paper is to compare predicted and actual survival curves from weeks 27 to 39 weeks, in addition to 1 to 26 weeks.

It is important for policy purposes to understand what happens past 26 weeks, since 43% of claimants exhaust benefits when the maximum allowed duration is 26 weeks. What these 43% of claimants will do after 26 weeks if benefits are extended is important for cost predictions. Moreover, the behavioral response of the long-term unemployed (individuals on UI over 26 weeks are classified as "long-term" unemployed by the government) may or may not be similar to the behavioral response for those individuals who exit UI relatively quickly.

One challenge in using the NJEB program is that roughly 30% of claimants never sign up for extended benefits, even though they have exhausted their regular 26 week benefits and are eligible. Card and Levine discuss several potential reasons for this, the most salient

of which relate to the temporary nature of the NJEB program. The results in this section should certainly take the temporary nature of the program into account, as a more permanent change might have larger behavioral effects. It should be recognized, however, that analyzing temporary changes is useful, since many extensions to UI benefits are in fact transitory.

The idea behind estimating the behavioral response is to compare the predicted survival curve when individuals are only eligible for 26 weeks to the actual survival curve when individuals are eligible for 39 weeks. The difference between the two survival functions yields an estimate of the behavioral response, assuming that economic conditions did not change between the two periods. (In fact, economic conditions were remarkable similar during the two periods, and the demographic composition of claimants was very similar.) However, since 30% of claimants do not sign up for extended benefits, I can only estimate an upper bound on any behavioral response.

To identify an upper bound on the behavioral response, I make three assumptions. First, I assume that individuals who self-select into the NJEB program are not positively selected. By this I mean that claimants who signed up for extended benefits would not have had shorter spell lengths (past 26 weeks) compared to those who did not sign up for the program. This assumption seems natural, since claimants who sign up for extended benefits are likely to be those who need them the most.

Second, I assume that increasing the maximum duration from 26 to 39 weeks did not decrease average spell length. This assumption says that the NJEB program did not cause claimants to exit UI more quickly.



Finally, I assume no exits from UI at 26 weeks. We do not know how many claimants exit UI at 26 weeks (eg, because they find a job at 26 weeks) versus how many claimants fail to sign up for the program but would otherwise not have exited at 26 weeks. I make the conservative assumption that no claimants exit at 26 weeks, so that the survival function is flat through 26 weeks. While this assumption is extremely conservative, it delivers an upper bound estimate under the worst case scenario.

Under these three assumptions, the difference between the two survival curves yields an upper bound estimate of the behavioral response. Figure 5 graphs the behavioral response. The two solid lines plot the upper and lower bound semiparametric estimates of the aggregate survival function. As a reminder, the semiparametric survival curve is estimated using data when the maximum allowed duration was 26 weeks. The step function graphs the Kaplan-Meier estimate, assuming random participation in the NJEB program (i.e., random censoring of individuals at 26 weeks) and no exits at 26 weeks. Past 26 weeks, this step function is an upper-bound estimate of the survival curve given the three assumptions outlined above.

Interestingly, there is little evidence of a large behavioral response past the censoring point. The upper bound Kaplan-Meier estimate is near the upper 95 percent confidence interval for the semiparametric estimator's upper bound. Evaluated at each week past the censoring point, there is never a week where the estimated quantiles from the two approaches are significantly different.<sup>4</sup> Figure 5 suggests the slow re-entry of the long-term unemployed is not primarily due to a disincentive effect where workers decrease their job search intensity.

---

<sup>4</sup>A test which aggregates over all weeks would likely find a significant difference, but I have not yet performed such a test.

## 7 Conclusion

Predicting how UI spell length would change in response to an extension in allowable duration is difficult as many claimants have censored spells. In the New Jersey data analyzed in this paper, 43 percent of claimants exhaust their benefits at 26 weeks. How long would these censored claims remain active if UI benefits were made available for a longer time period? The answer would be particularly useful in evaluating the costs and benefits of extending the maximum duration of UI benefits. Also of interest is whether such an extension provides disincentive effects to leave the UI system in a timely manner.

Using a flexible semiparametric approach, the costs of extending UI benefits are estimated to be quite substantial. This is because the shape of the survival curve is relatively flat past the censoring point; few individuals are predicted to exit the program each week after 26 weeks of payments. If time limit increases do not decrease an individual's desired spell length, a lower bound estimate for the cost of increasing the maximum duration can be calculated. The minimum cost estimate for a 50 percent increase in maximum allowed duration (from 26 weeks to 39 weeks) is estimated to raise the amount of money spent on New Jersey's UI program by over 25 percent. If researchers rely on an incorrect parametric model, these costs could be significantly underestimated.

Any behavioral response would further add to the costs of extending the maximum allowed duration. The behavioral response is estimated with a unique policy experiment in New Jersey which exogenously increased the duration of UI benefits from 26 to 39 weeks. The analysis reveals at most a small disincentive effect after 26 weeks when benefits are extended to 39 weeks. The results imply the group of individuals who have been unemployed

for at least 26 weeks have a difficult time finding a job and leaving UI. This suggests an important role for UI as an income maintenance program for the long-term unemployed.

## References

- [1] Anderson, P. and B.D. Meyer (1997) "Unemployment Insurance Take Up Rates and the After-Tax Value of Benefits", *Quarterly Journal of Economics*, 112, 913-938.
- [2] Blank, R. M. and D. Card (1991) "Recent Trends in Insured and Uninsured Unemployment: Is There an Explanation?", *Quarterly Journal of Economics*, 106, 1157-1190.
- [3] Blaustein, S.J., W.J. Cohen and W. Haber (1993) *Unemployment Insurance in the United States: The First Half Century*, W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.
- [4] Buchinsky, M. (1998) "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research." *Journal of Human Resources* 33, 88-126.
- [5] Card, D. and P.B. Levine (2000) "Extended Benefits and the Duration of UI Spells: Evidence from the New Jersey Extended Benefit Program", *Journal of Public Economics*, 78, 107-138.
- [6] Chen, S., G.B. Dahl, and S. Khan (2005) "Nonparametric Identification and Estimation of a Censored Location-Scale Regression Model", *Journal of the American Statistical Association*, 100, 212-221.
- [7] Chen, S. and S. Khan (2000) "Estimation of Censored Regression Models in the Presence of Nonparametric Multiplicative Heteroskedasticity", *Journal of Economet-*

*rics*, 98, 283-316.

- [8] Chetty, R. (2008) "Moral Hazard versus Liquidity and Optimal Unemployment Insurance", *Journal of Political Economy*, 116, 173-234.
- [9] Cox, D.R. (1972) "Regression Models and Life Tables" (with discussion), *Journal of the Royal Statistical Society, Series B*, 34, 187-220.
- [10] Ham, J.C. and S.A. Rea Jr. (1987) "Unemployment Insurance and Male Unemployment Duration in Canada", *Journal of Labor Economics*, 5, 325-353.
- [11] Härdle, W. and O. Linton (1994) "Applied Nonparametric Methods", in Engle, R.F. and D. McFadden (eds.), *Handbook of Econometrics, Vol. 4*, Amsterdam: North-Holland.
- [12] Honoré, B.E., Khan, S. and J.L. Powell (2002), "Quantile Regression under Random Censoring", *Journal of Econometrics*, 109, 67-105.
- [13] Horowitz, J.L. (1996) "Semiparametric Estimation of a Regression Model with an Unknown Transformation of the Dependent Variable", *Econometrica*, 64, 103-137
- [14] Hunt, J. (1995) "The Effect of Unemployment Compensation on Unemployment Duration in Germany", *Journal of Labor Economics* , 13, 88-120.
- [15] Jurajda, S. and F. Tannery. (2003) "Unemployment Durations and Extended Unemployment Benefits in Local Labor Markets", *Industrial and Labor Relations Review*, 56, 324-348.
- [16] Katz, L.F. and B.D. Meyer (1990) "The Impact of the Potential Duration of Unemployment Benefits on the Duration of Unemployment", *Journal of Public Economics*, 105, 973-1002.

- [17] Khan, S. and J.L. Powell (2001) "Two-Step Quantile Estimation of Semiparametric Censored Regression Models ", *Journal of Econometrics*, 103, 73-110.
- [18] Lalive, R., J. van Ours, and J. Zweimuller (2006) "How Changes in Financial Incentives Affect the Duration of Unemployment", *Review of Economic Studies*, 73, 1009-1038.
- [19] Lewbel, A. and O.B. Linton (2002) "Nonparametric Censored and Truncated Regression", *Econometrica*, 70, 765-779.
- [20] McCall, B.P. (1995) "The Impact of Unemployment Insurance Benefit Levels on Reciprocity", *Journal of Business and Economic Statistics*, 12, 189-198.
- [21] Meyer, B.D. (1990) "Unemployment Insurance and Unemployment Spells", *Econometrica*, 58, 757-782.
- [22] Meyer, B.D. (1992) "Quasi-Experimental Evidence on the Effects of Unemployment Insurance from New York State", Unpublished Manuscript.
- [23] Moffitt, R. (1985) "Unemployment Insurance and the Distribution of Unemployment Spells", *Journal of Econometrics*, 28, 85-101.
- [24] Moffitt, R. and W. Nicholson (1982) "The Effect of Unemployment Insurance on Unemployment: The Case of Federal Supplemental Compensation", *Review of Economics and Statistics*, 64, 1-11.
- [25] van Ours, J. C. and M. Vodopivec (2006) "How Shortening the Potential Duration of Unemployment Benefits Affects the Duration of Unemployment: Evidence from a Natural Experiment", *Journal of Labor Economics*, 24, 351-378.
- [26] Powell, J.L. (1994) "Estimation of Semiparametric Models", in Engle, R.F. and

D. McFadden (eds.) , *Handbook of Econometrics, Vol. 4*, Amsterdam: North-Holland.

[27] Roed, K., P. Jensen, and A. Thoursie (2008) “Unemployment Duration and Unemployment Insurance: A Comparative Analysis Based on Scandinavian Micro Data”, *Oxford Economic Papers*, 53, 254-274.

[28] Woodbury, S.A. and A.R. Murray (1997) “The Duration of Benefits”, in O’Leary, C.J., Wandner, S.A. (eds.), *Unemployment Insurance in the United States: Analysis of Policy Issues*, W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.

Table 1. Characteristics of Unemployment Insurance Recipients in New Jersey.

Characteristic	Percent	Percent Exhausting Benefits
Gender		
male	56.7	40.1
female	43.3	47.1
Race		
white (not hispanic)	62.7	39.8
black (not hispanic)	18.0	53.4
hispanic and other	19.3	44.4
Age		
age $\leq$ 35	37.8	40.0
35 < age $\leq$ 50	40.3	43.0
50 < age $\leq$ 65	21.9	48.5
Education		
high school or less	60.3	43.9
some college or more	39.7	41.8
Weekly Earnings at Previous Job		
earnings $\leq$ \$375	33.2	42.1
\$375 < earnings $\leq$ \$625	33.5	47.6
earnings > \$625	33.3	39.5
Tenure at Previous Job		
less than 2 years	43.3	39.6
greater than 2 years	56.7	45.8
Union Status at Previous Job		
union member	12.7	33.5
not a union member	87.3	44.5
County Unemployment Rate		
less than 5.5%	49.6	40.9
greater than or equal to 5.5%	50.4	45.3
Industry		
Construction, Mining, or Agriculture	18.4	15.9
Manufacturing, Wholesale Trade, or Retail Trade	44.0	47.3
Transportation, Services, FIRE, or Public Administration	37.6	49.3
All Observations	100.0	43.1

Notes: Data consists of 192,162 individuals from administrative records of the New Jersey Department of Labor. Sample is restricted to claimants age 18 to 65 who were eligible for 26 weeks of benefits and received their first payment between June 1, 1996 and October 25, 1997. Sample also excludes claimants with missing information on age, earnings, or UI claim characteristics.

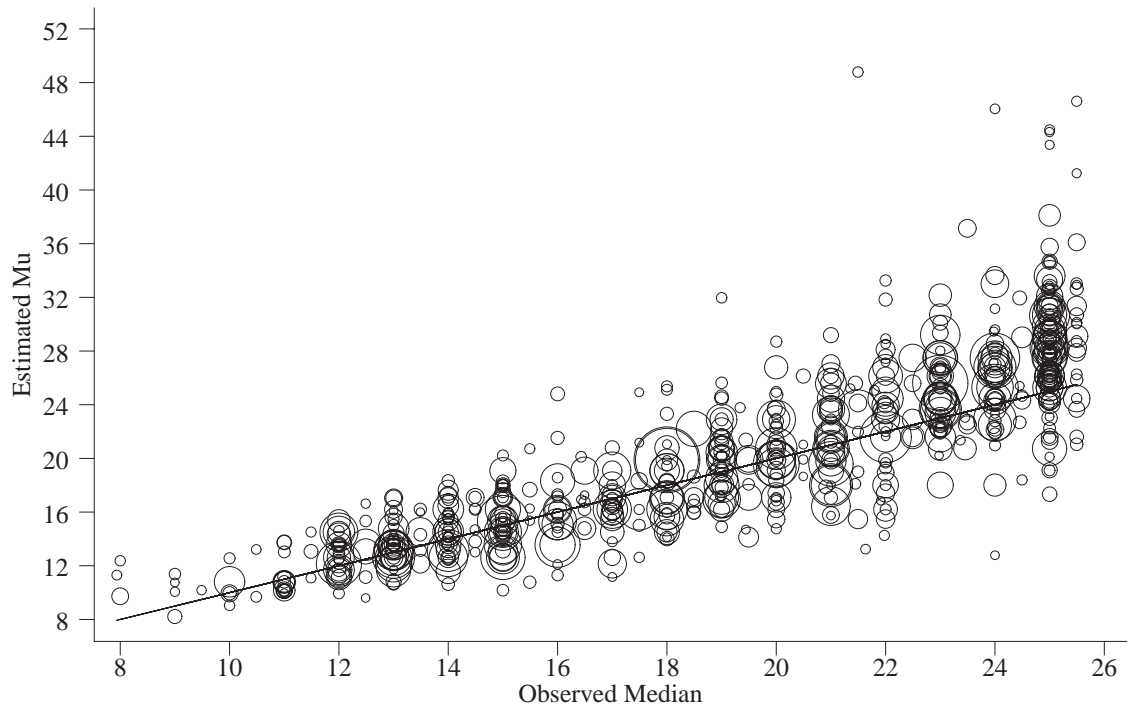
Table 2. Estimated Medians for Cells with More Than a Fifty Percent Exhaustion Rate and More Than 250 Observations.

Gender	Race	Age	Education	Earnings	Tenure	Unemp. Rate	Estimated Median	Std. Error	Obs. in Cell
male	white	young	H.S.	middle	long	low	33.50	1.95	992
male	white	young	H.S.	middle	long	high	32.33	1.92	731
male	white	young	H.S.	middle	short	high	32.72	2.26	477
male	white	young	H.S.	high	long	low	28.74	2.38	304
male	white	mid-age	H.S.	middle	long	low	29.76	1.59	1500
male	white	mid-age	H.S.	mid-age	long	high	35.82	2.49	1135
male	white	mid-age	H.S.	high	long	low	34.31	2.49	819
male	white	mid-age	college	mid-age	long	high	31.25	3.02	712
male	white	older	H.S.	mid-age	short	high	33.50	4.98	258
male	white	older	college	mid-age	long	low	38.47	3.36	607
male	white	older	college	mid-age	long	high	39.70	3.84	353
male	white	older	college	high	long	low	33.50	2.85	857
male	black	young	H.S.	low	long	low	34.52	5.56	253
male	black	young	H.S.	low	short	high	34.31	1.91	1091
male	black	young	college	mid-age	long	high	35.14	2.24	429
male	black	young	college	mid-age	short	high	26.95	2.96	387
male	black	mid-age	H.S.	low	short	high	34.31	2.65	524
male	black	mid-age	college	low	short	high	32.72	4.01	251
male	hispanic	young	H.S.	mid-age	long	high	37.28	3.59	262
male	hispanic	young	college	mid-age	long	high	30.57	2.37	253
female	white	older	H.S.	mid-age	long	low	32.72	2.70	699
female	white	older	H.S.	high	long	low	28.94	1.97	1345
female	white	older	H.S.	high	long	high	32.72	3.63	981
female	white	older	college	mid-age	long	low	31.46	3.35	276
female	white	older	college	mid-age	short	low	35.14	4.95	277
female	white	older	college	high	long	low	31.97	1.79	2114
female	white	older	college	high	long	high	29.76	2.28	1085
female	black	young	H.S.	low	long	high	33.50	2.81	649
female	black	young	H.S.	mid-age	short	high	28.54	1.64	415
female	black	young	college	mid-age	short	high	26.95	4.40	321
female	black	mid-age	H.S.	low	long	high	33.22	2.34	488
female	black	mid-age	H.S.	low	short	low	31.97	4.23	301
female	black	mid-age	H.S.	mid-age	long	low	31.50	2.23	288
female	black	mid-age	H.S.	mid-age	short	high	31.50	2.24	344
female	black	mid-age	H.S.	high	long	high	39.70	4.13	297
female	black	mid-age	college	low	short	high	33.22	3.06	252
female	black	mid-age	college	high	long	high	33.50	4.67	339
female	black	older	H.S.	mid-age	long	high	36.53	4.58	337

Notes: All cells contain non-union claimants. Standard errors are calculated using the bootstrap, based on 400 replications with samples equal in size to the number of observations in a cell and drawn with replacement.

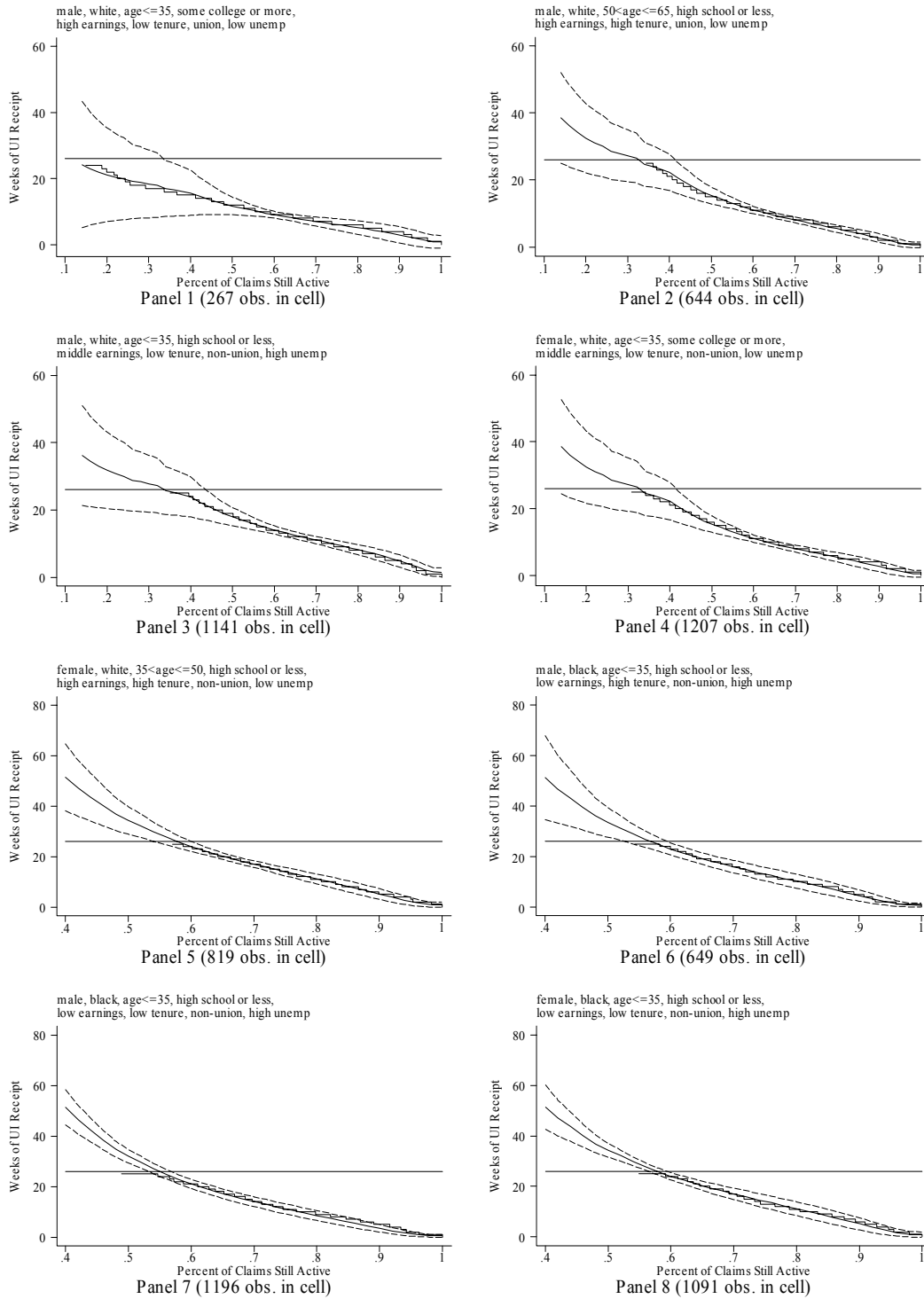


Figure 1. Comparison of the Estimated Location Function  $\hat{\mu}(x)$  to the Observed Median for Cells with Less than Fifty Percent Censoring.



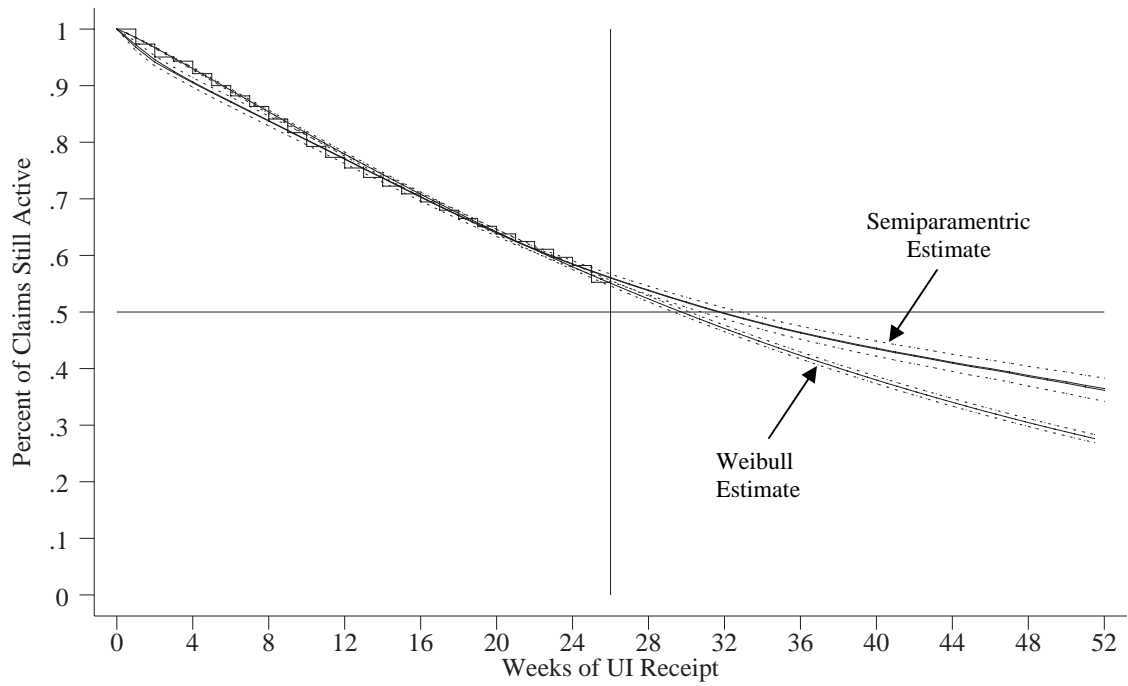
Notes: Solid line indicates the 45 degree line. Circle size is proportional to cell size. For visual clarity, sample restricted to cells with at least 30 observations.

Figure 2. Estimated Quantiles of Unemployment Insurance Receipt for Selected Cells, with a Comparison to the Kaplan-Meier Estimates.



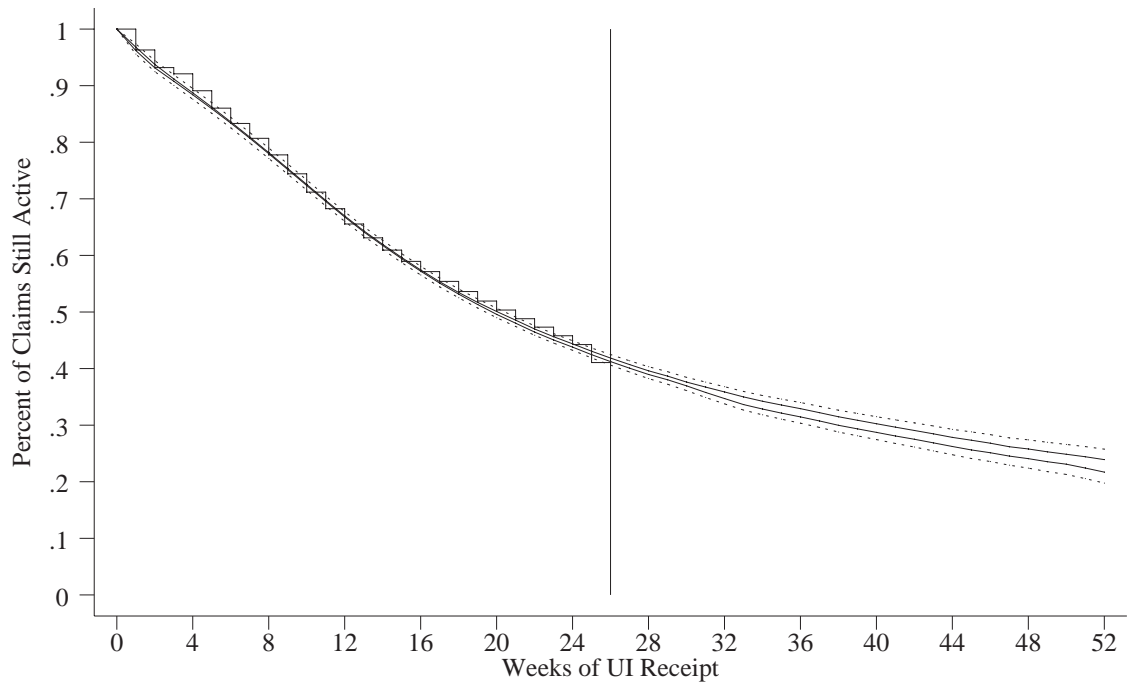
Notes: Smooth lines indicate estimated quantiles and broken lines indicate pointwise 95 percent confidence intervals. The step function is the Kaplan-Meier estimate (with the axes reversed). A line is drawn at 26 weeks to indicate the censoring point. Panels 1-4 plot the estimated quantiles up to the point where 15 percent of claims are predicted to be active. Panels 5-8 plot the estimated quantiles until the point estimate exceeds 52 weeks.

Figure 3. Aggregate Survival Function for Claimants in Cells with More than Fifty Percent Censoring, with a Comparison to the Kaplan-Meier and Weibull Estimates.



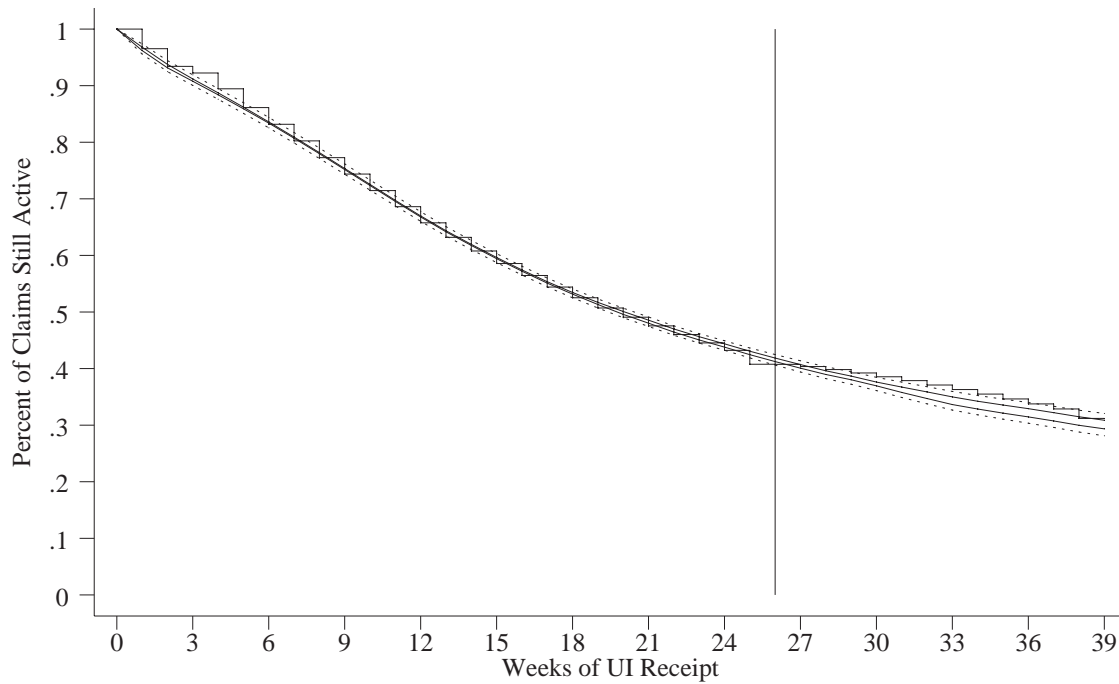
Notes: Broken lines indicate pointwise 95 percent confidence intervals. The step function is the Kaplan-Meier estimate. A vertical line is drawn at 26 weeks to indicate the censoring point.

Figure 4. Upper and Lower Bound Estimates of the Aggregate Survival Function for All Claimants, with a Comparison to the Kaplan-Meier Estimate.



Notes: Two solid lines indicate the upper and lower bound estimates of the aggregate survival function. Dashed lines indicate pointwise 95 percent upper and lower bound confidence intervals for the estimates. The step function is the Kaplan-Meier estimate. A vertical line is drawn at 26 weeks to indicate the censoring point.

Figure 5. Behavioral Response to an Increase in Maximum Allowed Duration from 26 to 39 weeks.



Notes: Two solid lines indicate the upper and lower bound estimates of the aggregate survival function, using data from the period when the maximum allowed duration was 26 weeks. Dashed lines indicate pointwise 95 percent upper and lower bound confidence intervals for the estimates. The step function is the upper bound Kaplan-Meier estimate, using data when the maximum allowed duration was extended to 39 weeks. The upper bound Kaplan-Meier estimate conservatively assumes random participation in the extended benefits program and no exits at 26 weeks. The difference between the solid lines and the step function captures the behavioral response to the extended benefits program. A vertical line is drawn to indicate the censoring point in the period when the maximum allowed duration was 26 weeks.