working paper
department
of economics

SEMIPARAMETRIC ESTIMATION OF DURATION AND
COMPETING RISK MODELS

Aaron Han, Harvard

Jerry Hausman, MIT

No. 450 November 1986

massachusetts
institute of
technology
50 memorial drive
cambridge, mass. 02139

SEMIPARAMETRIC ESTIMATION OF DURATION AND COMPETING RISK MODELS

Aaron Han, Harvard

Jerry Hausman, MIT

No. 450 November 1986

Semiparametric Estimation of Duration and
Competing Risk Models

by
Aaron Han, Harvard
and
Jerry Hausman, MIT

November 24, 1986

We thank Glenn Sueyoshi for extraordinary research assistance and Bruce Meyer for useful discussions. Lawrence Katz kindly provided the data used in this paper. The NSF and NBER provided research funding.

SEMIPARAMETRIC ESTIMATION OF DURATION AND COMPETING RISK MODELS

Introduction

Since Lancaster's (1979) paper on unemployment, duration models have become commonly used in econometrics. Heckman and Singer (1986) give a recent survey. While econometricians have emphasized the presence of unobserved heterogeneity, statisticians have instead emphasized the use of semiparametric models which do not require parametric specification of the baseline hazard. The leading model used has been the Cox (1972) proportional hazard-partial likelihood specification. While the model's semiparametric specification makes it potentially attractive, it has not been widely used in econometrics. Three possible reasons exist: (i) It is a continuous time specification while most duration data in econometrics is discrete where the discreteness may well be important, e.g., unemployment data. (ii) While various ad hoc procedures have been developed to treat tied failure times within the partial likelihood framework, they become cumbersome in the presence of many ties. Econometric data often has very many simultaneous failures, e.g., unemployment data at 26 weeks. (iii) Unobservable heterogeneity cannot be included without the presence of multiple integrals of the same order as the number of individuals in the risk set which makes estimation difficult, if not impossible.

In this paper, we specify and estimate a semiparametric proportional hazards (duration) model. The model specification is semiparametric in the sense that the baseline hazard is nonparametric while the effect of the covariates takes a particular functional form, which is typically linear although it need not be. The underlying hazard model is based on either an ordered probit or ordered logit model where an unknown parameter is estimated for each time interval over which the model is specified. A particular advantage of the specification is that the estimates of the parameters of the covariates are invariant to the length of time intervals which are chosen. Therefore, the grid of time intervals can be made finer as the sample size increases. We also add parametric heterogeneity to the underlying hazard model specification. The heterogeneity enters in extremely covenient form since the resulting model does not require numerical integration in estimation. In the sample of unemployed individuals examined in this paper, the addition of heterogeneity has very little effect on the results. Whether this finding is general to nonparametric baseline hazard specifications or is a particular finding for our sample, awaits future research.

We then consider competing risk models. Here, two or more hazards exist which may cause failure. We first prove an identification theorem which gives conditions under which the competing risks model is identified even if the covariates for the risks are identical. The identification condition basically requires the presence of at least one partly continuous variable among the covariates. This identification result should diminish the considerable confusion in the literature over whether competing risks models can always be specified as independent hazards models. We then specify a semiparametric proportional hazards model which permits unrestricted correlation among the risks. The basis for the model is a multivariate ordered probit specification where separate coefficients are estimated for each baseline hazard in each time interval of observation. Previous competing risks models which allow for interdependence of the risks have unacceptable restrictions on the form of the hazards. Alternatively, previous attempts at generalization of the semiparametric proportional hazards model to the competing risks situation have allowed only for restricted forms of interdependence among the risks.

In Section III, unemployment duration data is analyzed using the semiparametric duration and competing risks specifications. We first consider the effects of unemployment insurance (UI) and sociodemographic characteristics on the duration of unemployment. We find an important effect arising from the exhaustion of UI benefits at either 26 or 39 weeks. We then follow Katz's (1986) research and divide the hazards into either new jobs or recalls and estimate the competing risks model. Unlike Katz, we neither assume independence of the risks nor do we assume a particular functional form for the baseline hazards. Like Katz, we find significantly different baseline hazards for the two types of risks. However, we develop a test procedure for the particular functional form used by Katz and reject his baseline hazard specification along with his finding of monotonic positive duration dependence in the new jobs hazard.

I. Specification and Semiparametric Estimation of Duration

We assume observations on failure times over the discrete periods $t = 0,1,2,\ldots,T$ for individuals $i = 1,\ldots,n$. For now, we assume that the predetermined variables of each individual X_i do not change with time.[1] Our specification begins with the proportional hazards specification of Prentice (1976), see also Kalbfleisch and Prentice (1980) where

---

[1] Our method can be extended to the case of non-constant X's in a straightforward manner. See G. Sueyoshi (1986).

We again find very important effects from the exhaustion of UI at either 26 or 39 weeks of unemployment.

$$(1.1)\quad \lambda_i(\tau) = \lim_{\Delta \to 0} \frac{P(\tau < t_i < \tau + \Delta \mid t_i > \tau)}{\Delta} = \lambda_o(\tau)\exp(-X_i\beta)$$

is specified as

$$(1.2)\quad \log \int_0^{t_i} \lambda_o(\tau)\, d\tau = X_i\beta + \varepsilon_i$$

where ε_i takes an extreme value form.[2] Now let

$$(1.3)\quad \log \int_0^t \lambda_o(\tau)\, d\tau = \ell_t \qquad t = 1,\ldots,T,$$

so that the probability of failure in period t by individual i is

$$(1.4)\quad \int_{\ell_{t-1} - X_i\beta}^{\ell_t - X_i\beta} f(\varepsilon)\, d\varepsilon$$

We treat the baseline hazards, ℓ_t, as constants in each period and estimate them along with the parameters β. The Cox approach instead treats the baseline hazard function as a nuisance function and conditions it out of the estimation procedure by doing multinomial logit estimation on all the survivors (the risk set) during period t. We, instead, estimate all the parameters simultaneously. Moffit (1985) has also

[2]The extreme value assumption follows directly from the proportional hazards model specification. Thus, no additional assumption is made beyond proportional hazard.

estimated the baseline hazard parameters jointly with the regression parameter, appealing to results of Bailey (1984), but his specification does not guarantee that the probabilities lie in the unit interval and it is not clear how heterogeneity can be included in his model.[3]

Letting $y_{it} = 1$ if failure in period t occurs for person i, and $y_{it} = 0$ otherwise, the log likelihood function takes the form

$$(1.5)\quad \log L = \sum_i \sum_t y_{it} \log \int_{\ell_{t-1} - X_i\beta}^{\ell_t - X_i\beta} f(\varepsilon)\, d\varepsilon$$

for ε with an extreme value distribution then the likelihood function is of an "ordered" logit form and for ε with a standard normal distribution, the likelihood function takes the familiar ordered probit form. Both models are extremely easy to estimate since either a closed form for the integral or an accurate partial fraction expansion is known. In some exploratory research we have done, the estimates of the ordered logit and ordered probit models are very similar except in the extreme left tail, after rescaling, as would be expected from experience in discrete choice models. Note that once the nonparametric estimates of the ℓ_t, together with their covariance matrix, are known, then the applied investigator

---

[3]Moffitt's (1985) approach is not strictly correct in the case of a proportional hazards model with covariates present. Our paper's formulation is consistent for arbitrary discrete periods and equivalent to the Kaplan-Meier estimator, both for discrete data and as the time intervals become arbitrarily small. Moffit's model, along with recently proposed models by Kenan (1985) and Ham and Rea (1986), does not have these consistency properties.

can determine whether various parametric forms such as the exponential or Weibull are consistent with the estimates. In fact, any type of parametric approximation to the ℓ_t can be estimated and hypotheses of increasing or decreasing duration dependence can be considered in a much more flexible manner. We use a test procedure of this type in Section III.

An additional advantage of this specification is that the linear form of the proportional hazards model in equation (1.2) allows handling jointly endogenous variables or errors in variables via instrumental variable techniques. These potential problems commonly occur in empirical applications of duration models, e.g., Diamond and Hausman (1984). The parametric hazard approach used in econometrics so far has not permitted a satisfactory treatment of these common econometric problems.

We now introduce heterogeneity into the specification. Note that while, in principle, heterogeneity can be included in the Cox partial likelihood framework, it involves multiple integration of order n_t, where n_t is the number of survivors in period t. However, in the specification of equation (1.2), heterogeneity is straightforward to include and involves only a single additional order of integration in equation (1.4). In fact, for the case of a parametric gamma distribution specification of heterogeneity, a closed form result occurs which involves no numerical integration. The result is quite similar to Lancaster (1979). Assume a gamma distribution with mean one and variance $\sigma^2 = 1/\theta$. Then rewrite equation (1.2) in exponential form as

$$(1.6) \quad \exp\left\{\log \int_0^t \lambda_o(\tau)\, d\tau - X_i\beta\right\} = \exp(\varepsilon_i + w_i)$$

where w_i represents the unobserved heterogeneity and $\eta_i = \exp(w_i)$ is distributed as a gamma random variable. Then we denote

$$(1.7)\quad I_i(t) = \left\{\int_0^t \lambda_o(\tau)\, d\tau\right\} \exp(-X_i\beta)$$

where $I_i(t)$ is the survivor function in the absence of heterogeneity. Now let $s_i = w_i + \varepsilon_i$ so that $I_i(t) = \exp(s_i) = q_i$ say. Upon integrating both sides of this equation with respect to the random variable q, we find

$$(1.8)\quad \int_{I_i(t)}^{\infty} g(q)\, dq = [1 + (1/\theta)I_i(t)]^{-\theta}$$

Thus, we have a straightforward calculation in closed form since for each period t in the likelihood function

$$(1.9)\quad I_i(t) = \exp(-X_i\beta)\exp(\ell_t)$$

The log likelihood then follows directly as the sum of terms for each person i over the hazard term for each period t and estimation is straightforward. Preliminary results indicate that maximum likelihood estimation of with up to 50 ℓ_t coefficients creates no computational problems.

The next question that we intend to consider is whether the gamma heterogeneity specification is sufficiently flexible. Heckman and Singer (1984a,b) in a series of papers have sharply criticized the specification of parametric heterogeneity and have proposed a nonparametric specification of the heterogeneity. However, their empirical results have all been done in the context of a heavily restricted parametric specification of the baseline hazard function.[4] It may well be the case that a nonparametric specification of the hazard function reduces or eliminates the sensitivity of the estimates to a parametric heterogeneity assumption. Such a result would be quite convenient since our model is considerably easier to estimate than the Heckman-Singer model, and more importantly, yields an asymptotically normal estimator so that standard large sample inference can be used.

We now provide theorems which consider the semiparametric single-risk specifications without and then with heterogeneity. The data setup is:

(D1) $t = 1,2,\ldots,T$ discrete periods.

(D2) (t_i, X_i), for $i = 1,\ldots,N$ observations where t_i denotes the failure period for observation i and X_i' is the vector of predetermined variables. We let $y_{it} = 1$ if the ith individual fails in time period t and $y_{it} = 0$ otherwise.

---

[4]Manton, Stallard, and Vaupel (1986) find that the specification of the baseline hazard function is more important in estimation than is specification of the heterogeneity distribution. Meyer (1986) has proven identification of a Heckman-Singer-type of estimator in a model similar to the one developed in this paper. He is currently implementing the estimator to assess the sensitivity of the results to specification of the heterogeneity distribution.

The likelihood function, from equation (1.5) then becomes

$$(1.10) \qquad L = \prod_{i=1}^{N} \prod_{t=1}^{T} \left[\int_{\ell_{t-1}-X_i\beta}^{\ell_t - X_i\beta} f(\varepsilon)\, d\varepsilon \right]^{y_{it}}$$

We do all asymptotics for fixed T, e.g. T = 50 weeks, and as N becomes large. For the specification without heterogeneity we make the following assumptions:

(A1.1) The error distribution F is twice differentiable with the density function f(.) > 0 almost everywhere.

(A1.2) The X_i's are iid with

(i) $P_X(||X_i|| < M) = 1$ for $M < \infty$.

(ii) For any $\beta \varepsilon R^k$, $\beta \neq 0$ and any constant $c \varepsilon R$,

$P_X(X\beta = c) < 1$.

(A1.3) $\theta_o = (\beta_o, \ell_o)$ is an interior point of a compact set $B \times L$.

Under these rather standard regularity conditions, we prove the usual asymptotic properties of the maximum likelihood estimator (MLE):

<u>Theorem 1</u>: Let $\hat{\theta} = (\hat{\beta}, \hat{\ell})$ be the MLE. Then

$$\sqrt{N}\, (\hat{\theta} - \theta_o) \overset{A}{\to} N\left(0, \left[\lim_{N\to\infty} \frac{1}{N} \frac{\partial^2 \log L}{\partial\theta\partial\theta'} \right]^{-1}\right).$$

We now consider the single risk case with heterogeneity. Thus, the only change in the specification is that we assume that the distribution of the error term ε depends on an unknown (finite) parameter vector γ so that $\varepsilon \sim F(\cdot|\gamma)$. Our assumptions then become:

(A2.1) For all $\gamma \varepsilon \Gamma$, $F(\cdot|\gamma)$ is twice differentiable with

(i) The density function $f(\cdot|\gamma) > 0$ almost everywhere.

(ii) $\{\varepsilon | f(\varepsilon|\gamma) = c \cdot f(\frac{\varepsilon - a}{b} |\gamma_o\}$ for all $\gamma \neq \gamma_o$ and for all a, b, c ε R has Lebesgue measure zero.

(A2.2) Same as (A1.2) with the additional assumption that

(iii) At least one component X_h of X has $\beta_{ho} \neq 0$ and positive density in an open neighborhood, conditional on $\tilde{X} = (X_1, \ldots, X_{h-1}, X_{h+1}, \ldots, X_k)$ almost surely for $P_{\tilde{X}}$.

(A2.3) $\theta_o = (\beta_o, \ell_o, \gamma_o)$ is an interior point of a compact set $B \times L \times \Gamma$.

We are then able to prove the usual asymptotic properties of the MLE:

<u>Theorem 2</u>: Let $\hat{\theta} = (\hat{\beta}, \hat{\ell}, \hat{\gamma})$ be the MLE where $\varepsilon \sim F(\cdot|\gamma_o)$. Then

$$\sqrt{N} (\hat{\theta} - \theta_o) \overset{A}{\rightarrow} N \left(0, \left[\lim_{N\to\infty} \frac{1}{N} \frac{\partial^2 \log L}{\partial\theta \, \partial\theta'}\right]^{-1}\right).$$

In this section of the paper we have specified a semiparametric proportional hazards type model which is well suited for discrete data of the type which occurs in econometrics. Intuitively, the specification of the baseline hazard function is a series of dummy variables which requires no prior assumption of a parametric form. The specification bears close resemblance to the Cox (1972) specification. However, rather than conditioning out the baseline hazard, we estimate it jointly with the coefficients of the predetermined variables. This approach allows us to eliminate the problems of treating ties which are quite common in discrete economic data. The approach also allows the introduction of heterogeneity in a straightforward manner without the necessity of multiple integration. For the case of parametric heterogeneity specifications, the resulting likelihood function is straightforward to compute. Indeed, for the case of gamma heterogeneity the resulting likelihood function exists in closed form. The one additional assumption which is required for identification of the semiparametric model with heterogeneity is that at least one predetermined variable is partly continuous. This assumption need not be satisfied in all econometric applications of our specification.

II. Competing Risk Models

We next consider competing risks models. Competing risks models occur when failure can arise from two or more sources, e.g., a spell of umemployment can end with either a new job or withdrawal from the labor force as estimated in Diamond and Hausman (1984). Katz (1986) has considered the case where unemployment ends with either a recall to the previous job or a new job, and he finds significantly different behavior with respect to the two risks. Cox and Oakes (1984) give a recent survey of these models.

The proportional hazards model has been extended to the bivariate case in ways which allow only quite restricted patterns of interdependence between the two risks, e.g., Clayton and Cuzick (1985). Applied studies have either assumed a very restricted parametric form, e.g., Diamond and Hausman (1984), or have assumed independence between the two risks which is also quite unsatisfactory, e.g., Katz (1986). Lastly, considerable confusion exists over whether the dependent competing risks models are even identified, with the common claim made that independent competing risks specifications are adequate since any general model can be put into this form (an argument reminiscent of the recursivity debate in simultaneous equations). We prove here that identification does exist under quite weak regularity conditions. We then specify and estimate a bivariate ordered probit model with nonparametric baseline hazard specifications and unrestricted patterns of interdependence between the stochastic terms in the model.

The competing risks model can be placed into a latent variable framework similar to the probit model which is familiar to econometricians. Let n_k for $k = 1,...,K$ denote K competing risks. Introduce the latent random variable $Y_k^* \geq 0$ for $k = 1,...,K$, which would be the length of the period before failure if the particular risk were the only risk present. Denote the distribution function of Y_k^* by $F_k(x) = pr(Y_k^* \leq x)$. However, Y_k^* is a latent variable in general since it cannot necessarily be observed. Instead, only the minimum Y_k^* of the theoretical lifetimes is observed:

(2.1) $Y = \min (Y_1^*, \ldots, Y_K^*) = \min_k Y_k^*$

The available information includes the actual outcomes and the fact that if Y is greater than c, then the survivor distribution function, $\bar{F}_y(c) = 1 - F_y(c)$ together with the probability density function gives the conditional hazard rate function $\bar{f}_y(c) = f_y(c)/\bar{F}_y(c)$ for the observable variable Y. We now specialize to the case of K = 2 for notational simplicity. The amount of time until one of the two events occurs is min Y_k^* with only the smaller of Y_1^* and Y_2^* being observed, although neither may be observed if censorship occurs. The latent variable model can then be written as

(2.2)
$$Y_1^* = X_1\beta_1 + \varepsilon_1$$
$$Y_2^* = X_2\beta_2 + \varepsilon_2$$

Katz (1986) assumes that the stochastic disturbances, ε_1 and ε_2 are independent and thus treats equation (2.2) as two standard duration models which were discussed above. Diamond and Hausman (1984) allowed for dependence by assuming that Y_1^* and Y_2^* were distributed as bivariate log normal random variables. The unfortunate consequence of this assumption is to put very strong and non-testable parametric assumptions on the form of the hazard functions. We remove these parametric assumptions through semiparametric estimation of the hazard functions while retaining dependence among the stochastic disturbances in the model.

However, we must first consider the question of whether the competing risks model is identified for an arbitrary multivariate distribution function for equation (2.2). Cox (1959) and Tsiastis (1975) have published non-identifiability results, and a brief review is given in Cox and Oakes (1984). However, this previous work proceeded mostly in the absence of covariates which are typically present in econometric applications. In fact, it is relatively straightforward to demonstrate that if X_1 and X_2 do not have identical variables than the model is identified in the sense that an observationally equivalent independent competing risks model does not exist. However, in many econometric applications such as the unemployment problem $X_1 = X_2$ so that this convenient identifying assumption does not exist.

However, in Theorems 3 and 4 we prove identification of the bivariate competing risks model so long as at least one covariate is partly continuous and certain other regularity conditions are maintained, even if $X_1 = X_2$. Thus, we have solved the long-standing identification problem for competing risks models, at least for many econometric and statistical applications with covariates present.

We first consider the semiparametric specification of a bivariate competing risks model using the notation from the duration model specification considered earlier. We again assume the presence of discrete data with underlying "true" failure times

$$(2.3) \qquad \ell^1_{t_1} = -\log \int_0^{t_1} \lambda^1_o(s)\, ds = X\beta_1 + \varepsilon_1$$

$$\ell^2_{t_2} = -\log \int_0^{t_2} \lambda^2_o(s)\, ds = X\beta_2 + \varepsilon_2$$

Suppose that the failure type is of type 1 so that $t_1 = \min(t_1, t_2)$.

The probability of this outcome is

$$(2.4) \qquad \int_{\ell^1_{t-1} - X\beta_1}^{\ell^1_t - X\beta_1} \int_{m(\varepsilon_1)}^{\infty} f(\varepsilon_1, \varepsilon_2)\, d\varepsilon_2\, d\varepsilon_1$$

where $m(\varepsilon_1)$ is such that the implied failure time of type 2 is greater than the implied failure time of type 1 for a given ε_1. Given a realization of ε^*_1 and assuming linearity, we solve for the implied failure time,

$$(2.5) \qquad X\beta_1 + \varepsilon^*_1 = -\log \int_0^{t^*} \lambda^1_o(s)\, ds$$

where $t^* \varepsilon (t-1,t)$.[5]

Thus, we find

$$(2.6) \qquad X\beta_1 + \varepsilon^*_1 = \ell^1_{t-1} + \kappa_1(\ell^1_t - \ell^1_{t-1})$$

where κ_1 is defined by

$$(2.7) \qquad \ell^1_{t^*} = \ell^1_{t-1} + \kappa_1(\ell^1_t - \ell^1_{t-1})$$

We then use equation (2.6) to solve for κ_1.

$$(2.8) \qquad \kappa_1 = (X\beta_1 + \varepsilon^*{}_1 - \ell^1_{t-1})/(\ell^1_t - \ell^1_{t-1})$$

and we then solve for the support of ε_1 such that t_2 exceeds t^*

$$(2.9) \qquad \varepsilon^*_2 \geq \ell^2_{t-1} - X\beta_2 + \left[\frac{\ell^2_t - \ell^2_{t-1}}{\ell^1_t - \ell^1_{t-1}}\right]\left[\varepsilon^*_1 - (\ell^1_{t-1} - X\beta_1)\right] = m(\varepsilon^*_1)$$

We have thus solved for $m(\varepsilon_1)$ in equation (2.4) so that the probability of failure type 1 in period t is

---

[5]The linearity assumption is tested later to assure that it is accurate enough for estimation purposes.

$$(2.10)\quad \int_{\ell^1_{t-1}-X\beta_1}^{\ell^1_t-X\beta_1} \int_{\ell^2_{t-1}-X\beta_2+\left[\frac{\ell^2_t-\ell^2_{t-1}}{\ell^1_t-\ell^1_{t-1}}\right]\left[\varepsilon_1-(\ell^1_{t-1}-X\beta_1)\right]}^{\infty} f(\varepsilon_1,\varepsilon_2)\, d\varepsilon_2\, d\varepsilon_1$$

Thus, the specification estimates two sets of ℓ_t's which gives a semi-parametric version of the respective hazards. We specify the density function $f(\varepsilon_1,\varepsilon_2)$ to be possibly correlated permitting dependence among the stochastic disturbances. It is important to note that the parametric assumption of f does not impose parametric forms on the cause specific hazards as it did in the previous Diamond-Hausman (1984) specification.

We now specify the log likelihood function which corresponds to this specification of the competing risks model. The data setup is:

(D1) $t = 1,2,\ldots,T$ discrete periods

(D2) $(t_1, d_i, X_{1i}, X_{2i})$ for $i = 1,2,\ldots,N$ where t_i is the period of failure, $d_i = 0$ denotes failure by the first risk while $d_i = 1$ denotes failure by the second risk, X'_{1i} is the vector of covariates for the first risk while X'_{2i} is the vector of covariates for the second risk. We let $y_{it} = 1$ if the ith individual fails at time period t, and $y_{it} = 0$ otherwise. The log likelihood function is

(2.11)

$$\log L = \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it}\Big[(1-d_i)\log \int_{\ell^1_{t-1}-X_{1i}\beta_1}^{\ell^1_t - X_{1i}\beta_1} \int_{[\ell^2_t-X_{2i}\beta_2]+[(\varepsilon_1-(\ell^1_t-X_{1i}\beta_1))\lambda_t]}^{\infty} f(\varepsilon_1, \varepsilon_2)\, d\varepsilon_1\, d\varepsilon_2$$

$$+ d_i \log \int_{\ell^2_{t-1}-X_{2i}\beta_2}^{\ell^2_t - X_{2i}\beta_2} \int_{[\ell^1_t-X_{1i}\beta_1] + [(\varepsilon_2-(\ell^2_2-X_{2i}\beta_2))/\lambda_t]}^{\infty} f(\varepsilon_1, \varepsilon_2)\, d\varepsilon_1\, d\varepsilon_2 \Big]$$

$$-\infty < \ell^1_1 < \ell^1_2 < \ldots < \ell^1_T < \infty,$$

where

$$-\infty < \ell^2_1 < \ell^2_2 < \ldots < \ell^2_T < \infty,$$

and $\lambda_t = (\ell^2_t - \ell^2_{t-1})/(\ell^1_t - \ell^1_{t-1})$ for $t = 2,\ldots,T-1$ with

$\lambda_1 = \lambda_T = 1$

(A3.1) The error distribution F is twice differentiable with the density function $f(.) > 0$ almost everywhere.

(A3.2) $X = (X_1,X_2)$ are $k = (k_1 + k_2)$ vector iid random variables such that for any $\beta_1 \varepsilon R^{k_1}$, $\beta_2 \varepsilon R^{k_2}$ and any c_1, $c_2 \varepsilon R$,

$$P_{X_1}(X_1\beta_1 = c_1) < 1 \text{ and } P_{X_2}(X_2\beta_2 = c_2) < 1$$

(Note that this assumption includes the case of $X_1 = X_2$).

(A3.3) $\theta_o = (\beta_{1o}, \beta_{2o}, \ell^1_o, \ell^2_o)$ is an interior point of a compact set $B_1 \times B_2 \times L^1 \times L^2$.

We then prove the usual properties of the MLE:

<u>Theorem 3</u>: Let $\hat{\theta} = (\hat{\beta}_1, \hat{\beta}_2, \hat{\ell}^1, \hat{\ell}^2)$ be the MLE with error distribution $F(\varepsilon_1, \varepsilon_2)$. Then

$$\sqrt{N}\,(\hat{\theta} - \theta_o) \overset{A}{\to} N\left(0, \left[\lim_{N\to\infty} \frac{1}{N} \frac{\partial^2 \log L}{\partial\theta\, \partial\theta'}\right]^{-1}\right).$$

We now move to the more common case where the error distribution depends on a finite vector of unknown parameters, e.g., the standard bivariate normal distribution with unknown correlation coefficient. We now give a set of conditions for identification where we require at least one continuous variable in both X_1 and X_2. Also, if $X_1 = X_2$, we require additional assumptions on the distribution function F. Our set of conditions is

(A4.1) For all $\gamma\varepsilon\Gamma$, $F(\cdot,\cdot|\gamma)$ is twice differentiable with

(i) A density function $f(\cdot,\cdot|\gamma) > 0$ almost everywhere.

(ii) $\left\{(\varepsilon_1,\varepsilon_2)\,|\,f(\varepsilon_1,\varepsilon_2|\gamma) = c\cdot f\left(\frac{\varepsilon_1 - a_1}{b_1}, \frac{\varepsilon_2 - a_2}{b_2}\,|\gamma_o\right)\right\}$ for $\gamma \neq \gamma_o$ and all constants a, b, c's in εR have Lebesgue measure (R^2) zero.

(A4.2) $X = (X_1,X_2)$ are iid $k = k_1 + k_2$ vector random variables such that

(i) For any $\beta_1 \varepsilon R^{k_1}$, $\beta_2 \varepsilon R^{k_2}$ and any constants c_1, $c_2 \varepsilon R$

$$P_{X_1}(X_1\beta_1 = c_1) < 1,\ P_{X_2}(X_2\beta_2 = c_2) < 1\ .$$

(ii) At least one component X_{1h_1} of X_1 has $\beta_{1h_1} \neq 0$ and has positive density in an open neighborhood conditional on $(\tilde{X}_1, X_2) = (X_{11}, \ldots, X_{1h_1-1}, X_{1h_1+1}, \ldots, X_{k_1}, X_{21}, \ldots, X_{2k_2})$, almost surely $P_{(\tilde{X}_1, X_2)}$.

(iii) The analogous condition holds for at least one component X_{2h_2} of X_2

(A4.3) θ_o is an interior point of a compact set, $B_1 \times B_2 \times L^1 \times L^2 \times \Gamma$.

For the case of $X_1 = X_2$, the following modifications in the assumptions are made:

(A4.1') (i) $\{\varepsilon | F_1(\varepsilon, \frac{\varepsilon - a_1}{b_1} | \gamma) + F_2(\varepsilon, \frac{\varepsilon - a_1}{b_1} | \gamma) \cdot c_1 =$

$F_1(\frac{\varepsilon - a_2}{b_2}, \frac{\varepsilon - a_3}{b_3} | \gamma_o) \cdot c_2 + F_2(\frac{\varepsilon - a_2}{b_2}, \frac{\varepsilon - a_3}{b_3} | \gamma_o) \cdot c_3\}$

for $\gamma \neq \gamma_o$ and all constants a, b, c's in R has Lebesgue measure (R) zero where F_1 and F_2 denote the partial derivatives.

(A4.2') (ii) and (iii)

At least one component X_{1h} of X_1 has $\beta_{1h} \neq 0$ and has positive density in an open neighborhood conditional on $\tilde{X}_1 = (X_{11}, \ldots, X_{1h-1}, X_{1h+1}, \ldots, X_{1k_1})$ almost surely $P_{\tilde{X}_1}$.

Thus the additional required assumption allows us to sort out the marginal distributions of F(.,.). We then have our last theorem on the properties of the MLE:

<u>Theorem 4</u>: Let $\hat{\theta} = (\hat{\beta}_1, \hat{\beta}_2, \hat{\ell}^1, \hat{\ell}^2, \hat{\gamma})$ be the MLE with the error distribution $F(\varepsilon_1, \varepsilon_2 | \gamma_o)$. Then

$$\sqrt{N}\,(\hat{\theta} - \theta_o) \overset{A}{\to} N\left(0, \left[\lim_{N\to\infty} \frac{1}{N} \frac{\partial^2 \log L}{\partial\theta\, \partial\theta'}\right]^{-1}\right).$$

Lastly in this section we consider a Monte Carlo type example to ascertain whether the linearity approximation used in equations (2.4)-(2.11) is sufficiently accurate for estimation purposes. Our sample consists of 1055 observations each with the same two predetermined variables for each of the two risks. The predetermined variables were taken from the data set employed in Section III of this paper. We assume a bivariate normal distribution with each variance set equal to one and correlation coefficient ρ. It is easy to show that the model satisfies the assumptions in (A4). The model is estimated over 20 periods for each risk so that 11 and 12 parameters are estimated for the respective baseline hazard in addition to the unknown parameters for the covariates. The results of one estimation are:

Table 2.1: Test of Linearity Approximation
(Standard Errors)

	Risk 1 Actual	Risk 1 Estimated	Risk 2 Actual	Risk 2 Estimated
β_{01}	0.10	0.08 (0.02)	0.15	0.14 (0.02)
β_{02}	-1.00	-1.00 (.09)	-.50	-.51 (.23)
ρ	0.60	0.72 (0.66)		

Overall, the various samples which we estimated over gave similar results. The estimates of the parameters for the predetermined exogenous variables are quite accurate. ρ is usually estimated accurately but the reported precision is not high. The parameters of the baseline hazards (not shown here) are always estimated very accurately with the average asymptotic "t statistic" about 3. As expected, the precision of the parameters of the baseline hazard decreases in the right hand tails as fewer observations remain in the extreme tails. Overall, we conclude that maximum likelihood estimation of the bivariate competing risks model using the linearity assumption leads to quite accurate parameter estimates. We now proceed to use of the single and competing risks models on unemployment data.

III. Results

Our data are derived from a sample created from the Panel Study of Income Dynamics (PSID) by Katz (1986). Katz emphasizes in his study that recalls from unemployment to a previous job should be treated differently than new jobs. Thus, he formulates a competing risks model where recalls and new jobs are treated separately, rather than being combined into a single risk model as most of the previous unemployment duration literature had assumed. By separating the overall reemployment hazard into these two parts, Katz is able to test the implications of job search models without the potential biases which can arise from the convolution of the two risks. Indeed, his results find strong positive duration dependence in the "new job finding rate" when the recall rate is separated into another risk. He also finds evidence of duration effects for unemployment insurance (UI) near the point of exhaustion of UI benefits.

However, the econometric specification used by Katz for the competing risk model is quite restrictive along two dimensions. First, he assumes that the hazards for recalls and for new jobs are independent. Second, for each of the baseline hazard specifications in the independent proportional hazards models for recalls and new jobs, Katz uses a modification of the one parameter Weibull specification, first used in the econometric literature by Lancaster (1979). Given his emphasis on the estimated shape of the hazard functions, this restrictive specification is very unappealing. Our proposed method permits non-parametric estimation of the baseline hazards removing these restriction while also allowing stochastic dependence between the two hazards.

The sample is taken from waves 14 and 15 of the Panel Study of Income Dynamics (PSID). We adopt Katz's data definitions, in particular whether a particular unemployment spell ends by recall, a new job, or is censored by the date of the interview. The unemployment spells occured in either 1980 or 1981. Potential sampling problems for unemployment spells in the PSID are discussed in Katz (1986). He concludes that these potential problems are extremely minor in affecting his results.

The data set consists of 1055 observations. The sample is limited to heads of households between the ages of 20 and 65. Variable definitions and means are given in Table 3.1. Note that our sample is identical to the Katz (1986) sample with nonwhites being oversampled because of the sample frame of the PSID. Recalls are the most important way in which unemployment ends in the sample: 57% of the spells end in recall.

Of the remaining spells, 23% end in a new job while the remaining 20% of unemployment spells are censored by the interview date. The basic outcomes are given in Table 3.2. Note the increases in exit from unemployment at 26 and 39 weeks which are the exhaustion points of UI benefits.

We now present Kaplan-Meier estimates of the hazard functions for the different exits from unemployment. The Kaplan-Meier estimator is the nonparametric hazard estimated by the number of exits from unemployment divided by the population still in unemployment in that period. Thus, it is the sample analogue of the theoretical hazard without controlling for observed and unobserved differences across individuals. We first present the sample hazard function for the single risk case in Figure 3.1. Here both recalls and new jobs are grouped together as exits from unemployment. Note the prominent spikes at the UI exhaustion points of 26 weeks and 39 weeks. In Figure 3.2 the sample hazards are estimated separately for recalls and for new jobs. Note that the hazard for recalls is decreasing throughout much of its range while the new job hazard is much closer to being constant over most of its range, although it has both increasing and decreasing portions over its range. In Figures 3.3 and 3.4 we divide the sample into two parts based on eligibility for UI. For non-UI recipients the shapes of the hazard functions are now largely missing the spikes at 26 and 39 weeks which appeared in the total sample. However, these spikes are again quite prominent in the UI eligible sample. The interpretation of the increased hazard at UI exhaustion points is more problematic for recalls than it is for new job exists from unemployment. Katz (1986) discusses briefly reasons for the observed spikes in the recall hazards.

We now turn to estimation of the semiparametric single risk model from Section I. We estimate the single risk model without heterogeneity of equation (1.4). We report estimates where f(t) is based on either the normal (ordered probit) or extreme value (ordered logit) distribution functions to determine the sensitivity of the results to choice of this distribution. We then reestimate the model allowing for parametric heterogeneity using the one parameter gamma function as in equation (1.8). In addition to 17 predetermined variables, we also estimate 40 weekly values of ℓ_t for the baseline hazard. Tests using different censoring points for the baseline hazard demonstrated that the results are insensitive to the end point of the baseline hazard, as would be expected given its nonparametric specification.

The estimates are given in Table 3.3. UI has the expected effect of leading to longer spells of unemployment. The sociodemographic and industry variables are all estimated precisely. After rescaling, the estimated hazard functions are quite similar. Thus, we infer that the choice between a normal and an extreme value distribution function matters only in the extreme tails of the distribution. When we allow for gamma heterogeneity, the variance of the distribution is estimated to be 1.23. While both the asymptotic t statistic and the LR statistic indicate that heterogeneity improves the model fit significantly over the no heterogeneity extreme value model, it is interesting to note that the probit model fits the data better than either of the extreme value specifications, with or without heterogeneity. We also infer that allowing for parametric heterogeneity has only a minor effect on the estimated hazard functions. These results are demonstrated graphically in Figure 3.5 where the three estimated hazard functions are quite similar to each other. The extreme value distribution model, with and without heterogeneity, is compared in Table 3.4 where the cumulative distribution functions and asymptotic standard errors are presented from an average across all 1055 individuals using the parameter estimates from Table 3.5. Note that the estimates of the distribution functions are virtually identical and well within one standard error at each week. We thus conclude tentatively, at least in this one sample, that the results are much less sensitive to specification of heterogeneity when a nonparametric specification is used for the baseline hazard function. Thus, when we consider competing risks specifications we do not attempt to include separate distributions for unobserved heterogeneity in our models.[6]

We now turn to estimation of the competing risks models where new jobs and recalls are distinguished as two separate ways to exit from unemployment. We use the bivariate specification of equations (2.3) and (2.4) where the distribution function is assumed to be joint standard normal. Therefore, the only unknown parameter of the distribution function is ρ, the correlation coefficient. The log likelihood function is given in equation (2.11) from which we do maximum likelihhod estimation. We specify the model in two different ways: the first specification again allows for 40 parameters for each of the two competing risks

---

[6] Inclusion of heterogeneity in the competing risks specification of equation (2.11) is straightforward, at least in principle. However, estimation would require numerical evaluation of integrals which would be quite time consuming in practice.

while the second specification adds four additional parameters to allow for interactions at 26 and 39 weeks for UI recipients to allow for the effects of UI exhaustion. We present only the results for the second specification since the previous investigation of the Kaplin-Meier estimates demonstrated that the UI interaction variables are quite important.

The parameter estimates for the semiparametric competing risks model are given in Table 3.5. Note that the presence of UI has an important effect for both new jobs and for recall. As we would expect, the effect of UI is larger on exit from unemployment to new jobs than it is on exit from unemployment to recall. Race and marital status also have important effects with the estimated directions as expected. Lastly, note that the UI interaction effects at both 26 and 39 weeks are estimated to be quite large and significant. Thus, we again conclude that the effect of UI exhaustion has important effects on exit probabilities from unemployment.

We next consider the question of the importance of Katz's (1985) assumption of independence of the hazards and the assumption of a one parameter Weibull specification for the baseline hazard. The estimated ρ is .057 so that the assumption of stochastic independence would not be rejected with usual significance levels. However, the assumption of the Weibull distribution fares less well. In Table 3.6 we present the estimated cumulative distribution functions which correspond to the baseline hazards. We do not find that they have a monotonic downward or upward hazard form which is implied by the specification of a one parameter Weibull family. When we graph the baseline hazard estimates from the estimates of the semiparametric hazard model in Figure 3.6 for individuals who receive UI and Figure 3.7 for individuals who do not receive UI, we note that the monotonic upward or downward shapes are not present. In particular, the new job baseline hazards are initially rising followed by a decline which is followed by another rise. They appear to be far from monotonic, especially for individuals who receive UI. The Weibull specification requires either increasing or decreasing duration dependence; the only question is which direction the duration dependence will go. Our estimates indicate that the requirement of either monotonic increasing or decreasing duration dependence is violated in the data. Thus, we conclude that the Weibull specification is too simple for the PSID unemployment data. The semiparametric estimates indicate that Katz's (1986) finding of "strong positive duration dependence in the new job hazard" for the UI sample appears to arise from the Weibull specification rather than actual individual behavior.

We now turn to a formal test of the Weibull specification. The basis of the test is to determine whether a Weibull specification is consistent with our nonparametric baseline hazard estimates. To do the test, we employ minimum χ^2 type tests. First denote the semiparametric hazard estimates as $\hat{\ell}^1$ and $\hat{\ell}^2$. From our previous theorems we know that

$$(4.1) \quad \begin{vmatrix} \hat{\ell}^1 \\ \hat{\ell}^2 \end{vmatrix} \overset{A}{\sim} N(\ell, \Omega)$$

where Ω corresponds to the lower right hand block of the inverse of the Fisher information matrix. The Weibull specification can be written

$$(4.2) \quad \ell^j = \delta_{j1} t^{\delta_{j2} - 1} = g_j(\delta_j) \text{ for } \delta_j > 0 \text{ for } j = 1, 2$$

To estimate the unknown δ_j's we use minimum chi square estimation

$$(4.3) \quad \min_{\delta_1, \delta_2} W = \begin{vmatrix} \hat{\ell}^1 - g_1(\delta_1) \\ \hat{\ell}^2 - g_2(\delta_2) \end{vmatrix}' \hat{\Omega}^{-1} \begin{vmatrix} \hat{\ell}^1 - g_1(\delta_1) \\ \hat{\ell}^2 - g_2(\delta_2) \end{vmatrix}$$

The estimated δ_j's for the UI and no UI samples are given in Table 3.7. The value of W is then distributed under the null hypothesis of the Weibull specification as asympototic χ^2 with $k_1 + k_2 - 4$ degrees of freedom where k_i is the number of estimated baseline hazard parameters. Since W is estimated to be 98.954 which is well above its expected value of 63 under the null hypothesis, we reject the Weibull specification. In Figure 3.8 and 3.9, we graph the estimated semiparametric baseline hazards together with the Weibull estimates of the baseline hazards.[7] The difference between them is quite striking.

---

[7]Our estimated shapes of the Weibull hazard specifications are quite similar to the Katz (1986) estimates. In particular, the new job hazard for UI recipients is upward sloping which is Katz's main empirical finding.

We recommend this general approach to estimation and testing of baseline hazard functions. Semiparametric estimation puts no restrictions on the shape of the underlying hazard function in discrete data. Given the semiparametric hazard estimates, the econometrician can then test any particular functional hazard specification given our approach. Difficult analysis of residuals is eliminated since the semiparametric hazard estimates will use all the discrete data information. Furthermore, the estimation of particular functional forms of alternative hazard specifications is quite straightforward since it requires only generalized least squares type of estimation rather than repeated remaximization of the likelihood function under each different baseline hazard specification. Tests of duration dependence are thus easy to carry out without undue parametric restrictions.

V. Conclusions

Our approach begins with the Cox proportional hazard model which has been widely used in econometrics. We take account of the discrete nature of much of econometric duration data, and we make use of the typically very large samples which occur in econometrics. Both situations differ from biostatistics where data is often recorded continuously and samples are often quite small. We generalize the single risk Cox model to allow for both nonparametric estimation of the baseline hazard and for parametric heterogeneity. Our findings indicate that much applied econometric work has probably used excessively restricted specifications of the hazard functions. The one or two parameter specifications which are most often used seem too simple for the actual data. On the other hand, our preliminary findings indicate that the addition of heterogeneity has only a minor effect on the results. Attention to unobserved heterogeneity may be less important when nonparametric hazard specifications are used.

We then extend the semiparametric specification approach to competing risks models. We first prove that the competing risks model is identified even if both risks have identical predetermined variables so long as at least one variable is partly continuous. We then specify a semiparametric model which allows for unrestricted correlation across the stochastic disturbances in the competing risks. Lastly, we develop an estimation method which appears to work well in a Monte Carlo example and on actual data.

Simple models of labor market behavior often lead to predictions of monotonic hazard functions for duration of unemployment. Our procedure permits a very flexible approach to estimation of such models and tests of the monotonicity hypotheses. Our results from a sample drawn from the PSID tend to reject the monotonic hazard predictions. We also find, along with previous authors, strong evidence of important UI exhaustion effects. These findings point to the need for more realistic models of labor market behavior for the unemployed.

APPENDIX

AI. Single Risk Model

$t = 1,\ldots,T$ discrete periods

$(t_i, X_i)\ i = 1,\ldots,N$ observations

$$\log L = \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it} \log \left[\int_{\ell_{t-1} - X_i\beta}^{\ell_t - X_i\beta} f(\varepsilon)\, d\varepsilon \right]$$

Theorem 1.1

Let $\hat{\theta} = (\hat{\beta}, \hat{\ell})$ be the MLE. Then under assumptions (A1.1), (A1.2), and (A1.3),

$$\sqrt{N}\,(\hat{\theta} - \theta_o) \overset{A}{\rightarrow} N\left(0, \left[\lim_{N\to\infty} \frac{1}{N} \frac{\partial^2 \log L}{\partial\theta\, \partial\theta'}\right]^{-1}\right).$$

Proof

Identification

Let $\theta \neq \theta_o$.

If $\beta = \beta_o$, $\ell \neq \ell_o$, then there is some t s.t.

$$\int_{\ell_{t-1} - X\beta_o}^{\ell_t - X\beta_o} f(\varepsilon)\, d\varepsilon \neq \int_{\ell_{ot-1} - X\beta_o}^{\ell_{ot} - X\beta_o} f(\varepsilon)\, d\varepsilon \quad \text{a.s. } P_X.$$

Thus the model is identified.

If $\beta \neq \beta_o$, then by Assumption (A1.2.ii),

$$P_X\,(X(\beta - \beta_o) = c) < 1.$$

This implies that there exists some constant c_o s.t.

$$P_X\ (X(\beta - \beta_o) < c_o) > 0$$

$$P_X\ (X(\beta - \beta_o) > c_o) > 0.$$

This in turn implies that for any ℓ,

$$P_X\Big[\int_{\ell_{t-1} - X\beta}^{\ell_t - X\beta} f(\varepsilon)\, d\varepsilon = \int_{\ell_{ot-1} - X\beta_o}^{\ell_{ot} - X\beta_o} f(\varepsilon)\, d\varepsilon\Big] < 1, \quad t = 1,\ldots,T$$

GED.

Theorem 2

Let $\hat{\theta} = (\hat{\beta}, \hat{\ell}, \hat{\gamma})$ be the MLE where

$$\varepsilon \sim F(.|\gamma_o).$$

Then under assumptions (A2.1) (A2.2), and (A2.3),

$$\sqrt{N}\ \ (\hat{\theta} - \theta_o) \overset{A}{\to} N\ (0, \Big[\lim_{N\to\infty} \frac{1}{N} \frac{\partial^2 \log L}{\partial\theta\ \partial\theta'}\Big]^{-1}.$$

Proof

Identification

Case of $\gamma = \gamma_o$: Same as in Theorem 1.

Case of $\gamma \neq \gamma_o$: Suppose (ℓ, β) are the values of the parameters such that for all t,

$F(\ell_t + X\beta|\gamma) = F(\ell_{to} + X'\beta_o|\gamma_o)$ a.s. P_X.

Clearly, this implies that for t = 1,

(A.1) $F(\ell_1 + X\beta|\gamma) = F(\ell_{10} + X\beta_o|\gamma_o)$ a.s. P_X.

Note, however, that under Assumption (A2.2.iii), (A.1) implies

$$\text{(A.2)} \quad \lim_{\Delta\to 0} \frac{F(\ell_1 + \widetilde{X}\widetilde{\beta} + \beta_h (X_h + \Delta)|\gamma) - F(\ell_1 + \widetilde{X}\widetilde{\beta} + \beta_h X_h|\gamma)}{\Delta}$$

$$= \lim_{\Delta\to 0} \frac{F(\ell_{10} + \widetilde{X}\widetilde{\beta}_o + \beta_{ho}(X_h + \Delta)|\gamma_o) - F(\ell_{10} + \widetilde{X}\widetilde{\beta}_o + \beta_{ho} X_h|\gamma_o)}{\Delta}$$

for all X_h in an open neighborhood in R. And this in turn implies

$$\text{(A.3)} \quad \frac{1}{\beta_h} f(\ell_1 + \widetilde{X}\widetilde{\beta} + \beta_h X_h|\gamma) = \frac{1}{\beta_{ho}} f(\ell_{10} + \widetilde{X}\widetilde{\beta}_o + \beta_{ho} X_h|\gamma_o)$$

This contradicts Assumption (A2.1.ii) since equation (A3) implies that the set

$$\{\varepsilon|f(\varepsilon|\gamma) = c\cdot f(\tfrac{\varepsilon - a}{b}|\gamma_o)\}$$

has a positive Lebesgue measure for some c, a, b ε R. QED.

AII. Competing Risks Model

t = 1,...T — discrete periods

(t_i, X_i, d_i) i = 1,...,N — observations

$$\log L = \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it}\Big[(1 - d_i) \log \int_{\ell^1_{t-1} - X^1_i\beta^1}^{\ell^1_t - X^1_i\beta^1} \int_{[\ell^2_t - X^1_i\beta^2] + [(\varepsilon^1 - (\ell^1_t - X^1_i\beta^1))\lambda_t]}^{\infty} f(\varepsilon_1, \varepsilon_2) d\varepsilon_2 \, d\varepsilon_1$$

$$+ (d_i)\log \int_{\ell^2_{t-1} - X^2_i\beta^2}^{\ell^2_t - X^2_i\beta^2} \int_{[\ell^1_t - X^1_i\beta^1] + [(\varepsilon^2 - (\ell^2_t - X^2_i\beta^2))/\lambda_t]}^{\infty} f(\varepsilon_1, \varepsilon_2) d\varepsilon_1 \, d\varepsilon_2 \Big]$$

where $-\infty < \ell_1^1 < \ell_2^1 < \ldots < \ell_{50}^1 <, \infty,$

$$-\infty < \ell_1^2 < \ell_2^2 < \ldots < \ell_{50}^2 < \infty \;,$$

$$\lambda_1 = \lambda_{50} = 1$$

$$\lambda_t = \frac{\ell_t^2 - \ell_{t-1}^2}{\ell_t^1 - \ell_{t-1}^1} \;, \; t = 2,\ldots,49.$$

Theorem 3

Let $\hat{\theta} = (\hat{\beta}^1, \hat{\beta}^2, \hat{\ell}^1, \hat{\ell}^2)$ be MLE with error distribution $F(\varepsilon_1, \varepsilon_2)$.

Then under assumptions (A3.1), (A3.2) (A3.3),

$$\sqrt{N}\,(\hat{\theta} - \theta_o) \overset{A}{\to} N\Big(0, \Big[\lim_{N\to\infty} \frac{1}{N} \frac{\partial^2 \log L}{\partial\theta\, \partial\theta'}\Big]^{-1}\Big).$$

Proof

Identification

Without loss of generality, we will show that for t = 1, the parameters $(\beta_o^1, \beta_o^2, \ell_{o1}^1, \ell_{o2}^2)$ are identified.

Let $(\beta^1, \beta^2, \ell_1^1, \ell_1^2) \neq (\beta_o^1\;, \beta_o^2\;, \ell_{o1}^1, \ell_{o2}^2)$.

Then by Assumption (A3.2), we have

$$P_{X^1}(\ell_1^1 - X^1\beta^1 \neq \ell_{o2}^1 - X^1\beta_o^1) > 0$$

or $\quad P_{X^2}(\ell_1^1 - X^2\beta^2 \neq \ell_{o1}^2 - X^2\beta_o^2) > 0,$

or both. Suppose

$$P_{X^1}\;(\ell_1^1 - X^1\beta^1 > \ell_{o1}^1 - X^1\beta_o^1) > 0.$$

This result implies

$$\ell_1^2 - X^2\beta^2 > \ell_{o1}^2 - X^2\beta_o^2) > 0.$$

such that

$$\int_{-\infty}^{\ell_1^1 - X^1\beta^1} \int_{[\ell_1^2 - X^2\beta^2] + [(\varepsilon_1 - (\ell_1^1 - X^1\beta^1\;))]}^{\infty} f(\varepsilon_1,\varepsilon_2)\, d\varepsilon_2\, d\varepsilon_1$$

$$= \int_{-\infty}^{\ell^1_{o1} - X^1\beta^1_o} \int_{[\ell^2_{o1} - X^2\beta^2_o] + [(\varepsilon_1 - (\ell^1_{o1} - X^1\beta^1_o))]}^{\infty} f(\varepsilon_1,\varepsilon_2)\, d\varepsilon_2\, d\varepsilon_1$$

However, we have then

$$\int_{-\infty}^{\ell^2_{o1} - X^2\beta^2} \int_{\ell^1_1 - X^1\beta^1 + (\varepsilon_2 - (\ell^2_1 - X^2\beta^2))}^{\infty} f(\varepsilon_1,\varepsilon_2)\, d\varepsilon_1\, d\varepsilon_2$$

$$> \int_{-\infty}^{\ell^2_{o1} - X^2\beta^2_o} \int_{\ell^1_{o1} - X^1\beta^1_o + (\varepsilon_2 - (\ell^2_{o1} - X^2\beta^2))}^{\infty} f(\varepsilon_1,\varepsilon_2)\, d\varepsilon_1\, d\varepsilon_2 \quad .$$

Similar arguments apply for other inequalities, and thus the parameters $(\beta^1_o, \beta^2_o, \ell^1_{o1}, \ell^2_{o2})$ are identified.

Theorem 4

Let $\hat{\theta} = (\hat{\beta}^1, \hat{\beta}^2, \hat{\ell}^1, \hat{\ell}^2, \hat{\gamma})$ be the MLE with the error distribution, $F(\varepsilon^1, \varepsilon^2 | \gamma_o)$

Then under assumptions (A4.1), (A4.2), (A4.3),

$$\sqrt{N}\,(\hat{\theta} - \theta_o) \to N\left(0, \left[\lim \frac{1}{N} \frac{\partial^2 \log L}{\partial\theta\partial\theta'}\right]^{-1}\right).$$

Proof

Let $\gamma \neq \gamma_o$. Suppose there exist $(\beta^1, \beta^2, \ell^1, \ell^2)$ such that

$$\text{(A.4)} \quad P_t^1(X^1, X^2) \equiv \int_{\ell_{t-1}^1 - X^1\beta^1}^{\ell_t^1 - X^1\beta^1} \int_{[\ell_t^2 - X^2\beta^2] + [(\varepsilon_1 - (\ell_t^1 - X^1\beta^1))\lambda_t]}^{\infty} f(\varepsilon_1, \varepsilon_2)\, d\varepsilon_1\, d\varepsilon_2$$

$$= P_{ot}^1(X^1, X^2),$$

$$\text{(A.5)} \quad P_t^2(X^1, X^2) \equiv \int_{\ell_{t-1}^2 - X^2\beta^2}^{\ell_t^2 - X^2\beta^2} \int_{[\ell_t^1 - X^1\beta^1] + [(\varepsilon_2 - (\ell_t^2 - X^2\beta^2))/\lambda_t]}^{\infty} f(\varepsilon_1, \varepsilon_2)\, d\varepsilon_1\, d\varepsilon_2$$

$$= P_{ot}^2(X^1, X^2)$$

for all t, almost surely $P_{(X_1,\, X_2)}$. Looking at t = 1 without loss of generality, we then have under Assumptions (A4.3.ii) and (A4.3.iii),

$$\lim_{\Delta^1, \Delta^2 \to 0} \frac{P_1^1(X^1 + e_{h_1} \cdot \Delta^1, X^2 + e_{h_2} \cdot \Delta^2) - P_1^1(X^1, X^2)}{\Delta^1 \cdot \Delta^2}$$

$$+ \frac{P_1^2(X^1 + e_{h_1} \Delta^1, X^2 + e_{h_2} \cdot \Delta^2) - P_1^1(X^1, X^2)}{\Delta^1 \cdot \Delta^2}$$

$$= \lim_{\Delta^1, \Delta^2 \to 0} \frac{P_{o1}^1(X^1 + e_{h_1} \cdot \Delta^1, X^2 + e_{h_2} \cdot \Delta^2) - P_{o1}^1(X^1, X^2)}{\Delta^1 \cdot \Delta^2}$$

$$+ \frac{P_{o1}^2(X^1 + e_{h_1} \cdot \Delta 1, X^2 + e_{h_2} \cdot \Delta^2) - P_{o1}^1(X^1, X^2)}{\Delta^1 \cdot \Delta^2}$$

for a set of (X^1, X^2) with positive Lebesgue measure, where $e_h = (0,\ldots,0,1,0,\ldots,0)$ is the hth basis vector. This, however, implies

$$\frac{f(\ell_1^1 - X^1\beta^1,\ \ell_1^2 - X^2\beta^2 \mid \gamma)}{\beta_{h_1}^1\ \beta_{h_2}^2} = \frac{f(\ell_{o1}^1 - X^1\beta_o^1, \ell_{o1}^2 - X^2\beta_o^2 \mid \gamma_o)}{\beta_{oh_1}^1\ \beta_{oh_2}^2}$$

for a set of (X^1, X^2) with positive Lebesgue measure and contradicts Assumption (4.1.ii).

<u>Case of $X^1 = X^2$</u> (Assumptions (A4.1'), (A4.2'), (A4.3))

Let $\gamma \neq \gamma_o$ and suppose then exists $(\beta^1, \beta^2, \ell^1, \ell^2)$ such that (A.4) and (A.5) hold.

Again looking at t = 1 without loss of generality, we have under Assumptions (A4.2'.ii) and (A4.2'.iii),

$$\lim_{\Delta\to 0} \frac{P1_1(X^1 + e_h \cdot \Delta,\ X^1 + e_h \cdot \Delta) - P_1^1(X^1, X^1)}{\Delta}$$

$$+ \frac{P_{o1}^2(X^1 + e_h \cdot \Delta,\ X^1 + e_h \cdot \Delta) - P_{o1}^2(X^1, X^1)}{\Delta}$$

$$+ \frac{P_1^2(X^1 + e_h \cdot \Delta,\ X^1 + e_h \cdot \Delta) - P_1^2(X^1, X^1)}{\Delta}$$

$$= \lim_{\Delta\to 0} \frac{P_{o1}^1(X^1 + e_h \cdot \Delta,\ X^1 + e_h \cdot \Delta) - P_{o1}^1(X^1, X^1)}{\Delta}$$

for a set (X^1) with positive Lebesgue measure. This however, implies

$$\frac{F_1(\ell_1^1 - X^1\beta^1, \ell_1^2 - X^1\beta^2 \mid \gamma_o)}{\beta_h^1} + \frac{F_2(\ell_1^1 - X^1\beta^1, \ell_1^2 - X^1\beta^2 \mid \gamma_o)}{\beta_h^2}$$

$$= \frac{F_1(\ell_o^1 - X^1\beta_o^1, \ell_o^2 - X^1\beta_o^2 \mid \gamma_o)}{\beta_{oh}^1} + \frac{F_2(\ell_{o1}^1 - X^1\beta_o^1, \ell_o^2 - X^2\beta_o^2 \mid \gamma_o)}{\beta_{oh}^2}$$

for a set (X^1) with positive Lebesgue measure, and contradicts Assumption (4.1'.ii)

Table 3.1: Variable Definitions and Means-Standard Deviations of the PSID Layoff Unemployment Spell Sample (n = 1055)

Variable	Description	Mean Standard Deviation
Duration	= observed spell duration in weeks	17.335
		(22.447)
Age	= age of individual in years	33.154
		(10.607)
Sex	= 1 if female	0.167
		(0.373)
Education	= years of schooling	11.341
		(2.170)
Dependents	= number of dependents	3.038
		(1.640)
Race	= 1 if nonwhite	0.506
		(0.500)
UI	= 1 if worker received UI during spell	0.636
		(0.481)
Married	= 1 if married	0.632
		(0.482)
	Industry Dummy Variables (at onset of spell)	
Equipment	= 1 if in transportation equipment	0.118
Durables	= 1 if in other durable goods manufacturing	0.123
Trade	= 1 if in wholesale or retail trade	0.103
Transportation	= 1 if in transportation or public utilities	0.080
Mining	= 1 if in mining or argiculture	0.034
Service	= 1 if in services	0.172
Construction	= 1 if in construction	0.180

Occupation Dummy Variables (at onset of spell)

Laborer	= 1 if laborer or operative	0.508
Craft	= 1 if craftman or kindred worker	0.228
Clerical	= 1 if clerical, services or sales worker	0.186
Manager	= 1 if manager	0.045
Professional	= 1 if professional or technical worker	0.039

Source: Authors' calculation from PSID sample and Katz (1986).

Table 3.2: Failure Times for the PSID Layoff Unemployment Spell Sample

Weeks	New Job	Recall	Censored	Total
1	10	93	0	103
2	8	118	0	126
3	8	55	0	63
4	23	58	0	81
5	3	18	0	21
6	11	26	0	37
7	1	6	0	7
8	22	38	0	60
9	6	13	1	20
10	7	10	0	17
11	4	4	0	8
12	13	32	1	46
13	10	19	9	38
14	0	9	2	11
15	4	14	2	20
16	10	9	3	22
17	8	7	18	33
18	5	2	6	13
19	2	0	3	5
20	9	12	4	25
21	3	1	7	11
22	5	7	9	21
23	1	0	2	3
24	7	10	4	21
25	2	1	2	5
26	18	15	21	54
27	0	2	1	3
28	0	2	0	2
29	1	0	1	2
30	9	4	9	22
31	0	0	3	3
32	1	0	1	2
33	1	0	0	1
34	2	1	3	6
35	2	0	8	10
36	2	1	0	3
37	0	1	2	3
38	1	0	0	1
39	5	4	7	16
40	4	1	1	6
41	1	0	0	1
42	0	0	2	7
43	1	4	2	7
44	0	0	0	0
45	1	0	0	1

Table 3.2: Failure Times for the PSID Layoff Unemployment Spell Sample Cont'd

Weeks	New Job	Recall	Censored	Total
46	0	0	0	0
47	0	0	2	2
48	0	0	1	1
49	1	0	1	2
50	1	1	0	2
51	0	0	0	0
52	4	0	23	27
53	1	0	0	1
54	0	0	0	0
55	0	0	2	2
56	1	0	0	1
57	0	0	1	1
58	0	0	0	0
59	0	0	0	0
60	1	0	1	2
61	0	0	2	2
62	0	0	0	0
63	0	0	0	0
64	0	0	0	0
65	0	0	1	1
66	1	0	1	2
67	0	1	1	2
68	0	0	0	0
69	0	1	0	1
70	4	3	33	40
Totals	245	603	207	1055

Table 3.3 Parameter Estimates - Single Risk Model

Variable (Standard Error)	Normal (Probit)	Extreme Value	Extreme w/ Heterogeneity
Age	-0.012 (0.003)	-0.012 (0.003)	-0.022 (0.007)
Sex	0.133 (0.102)	0.208 (0.117)	0.172 (0.187)
Education	0.008 (0.018)	0.004 (0.019	0.022 (0.032)
Race	0.330 (0.072)	0.348 (0.080)	0.580 (0.150)
Married	-0.113 (0.086)	-0.130 (0.098)	-0.204 (0.157)
UI	0.029 (0.013)	0.026 (0.014)	0.055 (0.025)
Craft	-0.046 (0.093	-0.031 (0.096)	-0.128 (0.170)
Clerical	-0.074 (0.118)	-0.110 (0.139)	-0.100 (0.207)
Professional	-0.256 (0.222)	-0.373 (0.241)	-0.414 (0.375)
Manager	-0.168 (0.189)	-0.236 (0.209)	-0.263) (0.331)
Equipment	0.140 (0.119)	0.100 (0.125)	0.200 (0.219)
Durables	0.168 (0.114)	0.109 (0.120)	0.299 (0.212)
Trade	0.489 (0.138)	0.403 (0.162)	0.887 (0.266)
Transporation	0.464 (0.140)	0.395 (0.152)	0.847 (0.269)

Table 3.3 Parameter Estimates - Single Risk Model
Cont'd

Variable (Standard Error)	Normal (Probit)	Extreme Value	Extreme w/ Heterogeneity
Mining	0.017	-0.020	-0.004
	(0.182)	(0.182)	(0.340)
Service	0.522	0.442	0.908
	(0.134)	(0.153)	(0.257)
Construction	0.344	0.177	0.671
	(0.121)	(0.127)	(0.232)
σ^2	-	-	1.23
			(0.38)
Log Likelihood	-2956.071	-2959.911	-2956.280
Obs	1055		

Table 3.4: CDFs for the Extreme Value and Heterogeneity Distributions

Average Across Complete Sample: 1055 Individuals

Weeks	Extreme Value CDF	Extreme Value Std. Error	Heterogeneity CDF	Heterogeneity Std. Error
1	0.0976	0.0081	0.0976	0.0075
2	0.2172	0.0115	0.2169	0.0109
3	0.2771	0.0128	0.2768	0.0123
4	0.3539	0.0137	0.3535	0.0134
5	0.3738	0.0140	0.3733	0.0138
6	0.4088	0.0144	0.4083	0.0145
7	0.4155	0.0145	0.4149	0.0146
8	0.4724	0.0150	0.4718	0.0155
9	0.4905	0.0151	0.4898	0.0158
10	0.5067	0.0151	0.5060	0.0161
11	0.5143	0.0151	0.5136	0.0162
12	0.5571	0.0152	0.5564	0.0169
13	0.5847	0.0152	0.5841	0.0173
14	0.5934	0.0152	0.5929	0.0174
15	0.6109	0.0152	0.6104	0.0176
16	0.6295	0.0151	0.6290	0.0176
17	0.6441	0.0149	0.6437	0.0177
18	0.6513	0.0150	0.6510	0.0179
19	0.6534	0.0150	0.6531	0.0180
20	0.6757	0.0149	0.6754	0.0181
21	0.6800	0.0148	0.6797	0.0180
22	0.6931	0.0148	0.6929	0.0181
23	0.6943	0.0148	0.6940	0.0182
24	0.7137	0.0146	0.7135	0.0181
25	0.7172	0.0146	0.7170	0.0181
26	0.7561	0.0146	0.7560	0.0186
27	0.7587	0.0146	0.7586	0.0186
28	0.7614	0.0146	0.7613	0.0187
29	0.7627	0.0146	0.7626	0.0186
30	0.7800	0.0146	0.7799	0.0188
31	0.7870	0.0145	0.7870	0.0187
32	0.7974	0.0145	0.7974	0.0187
33	0.8217	0.0146	0.8216	0.0187
34	0.8325	0.0148	0.8323	0.0188
35	0.8381	0.0148	0.8378	0.0187
36	0.8477	0.0149	0.8474	0.0188
37	0.8504	0.0151	0.8501	0.0190
38	0.8533	0.0153	0.8529	0.0191
39	0.8566	0.0153	0.8562	0.0189
40	0.8634	0.0156	0.8630	0.0189

Note: Reported standard errors are asymptotic standard errors.

Competing Risks

Table 3.5 Parameter Estimates - UI Interaction Model

	New Job	Recall
Age	0.0099 (0.0037)	-0.0145 (0.0087)
Ed	-0.0798 (0.0309	0.0364 (0.0195)
Race	0.3588 (0.1243)	0.2661 (0.0801)
Deps	0.0546 (0.0364)	-0.0023 (0.0265)
UI	0.2451 (0.1051)	0.0871 (0.0779)
Marry	-0.1534 (0.1525)	-0.2646 (0.0938)
UI Interaction (26 weeks)	-1.924 (.478)	-2.530 (.479)
UI Interaction (39 weeks)	-3.555 (1.930)	-3.526 (1.131)
ρ	0.057 (1.18)	
Log LF	-3286.364	
Observations	1055	

Note: Asymptotic standard error in parentheses.

Table 3.6 Estimated CDFS

	No UI Interactions		UI Interactions	
	New Job	Recall	New Job	Recall
1	0.0044	0.0773	0.0044	0.0733
2	0.0094	0.1780	0.0094	0.1780
3	0.0154	0.2299	0.0154	0.2299
4	0.0355	0.2869	0.0355	0.3869
5	0.0385	0.3051	0.0385	0.3051
6	0.0499	0.3319	0.0499	0.3319
7	0.0510	0.3382	0.0510	0.3382
8	0.0767	0.3792	0.0767	0.3792
9	0.0843	0.3937	0.0843	0.3937
10	0.0935	0.4050	0.0935	0.4050
11	0.0989	0.4096	0.0989	0.4096
12	0.1174	0.4473	0.1174	0.4473
13	0.1328	0.4703	0.1328	0.4703
14	0.1328	0.4817	0.1328	0.4817
15	0.1396	0.4994	0.1396	0.4994
16	0.1575	0.5111	0.1575	0.5111
17	0.1723	0.5205	0.1723	0.5205
18	0.1823	0.5234	0.1823	0.5234
19	0.1864	0.5234	0.1864	0.5234
20	0.2056	0.5417	0.2056	0.5417
21	0.2122	0.5432	0.2122	0.5432
22	0.2238	0.5547	0.2238	0.5547
23	0.2262	0.5547	0.2262	0.5547
24	0.2438	0.5723	0.2438	0.5723
25	0.2491	0.5741	0.2491	0.5741
26*	0.2730	0.5841	0.3235	0.6149
27	0.2730	0.5886	0.3235	0.6193
28	0.2730	0.5932	0.3235	0.6238
29	0.2763	0.5932	0.3271	0.6238
30	0.3068	0.6026	0.3598	0.6330
31	0.3068	0.6026	0.3598	0.6330
32	0.3107	0.6026	0.3639	0.6330
33	0.3145	0.6026	0.3680	0.6330
34	0.3223	0.6053	0.3763	0.6357
35	0.3304	0.6053	0.3848	0.6357
36	0.3391	0.6084	0.3939	0.6386
37	0.3391	0.6114	0.3939	0.6416
38	0.3435	0.6114	0.3987	0.6416
39*	0.3613	0.6183	0.4284	0.6591
40	0.3815	0.6218	0.4495	0.6625

Table 3.7 Weibull Specification Test

$$t_i(\delta_{i1}, \delta_{i2}) = \delta_{i1} t^{\delta_{i2}-1}$$

Non-UI Recipients

	New Job	Recall
δ_1	0.007	0.077
	(0.005)	(0.007)
δ_2	1.253	0.596
	(0.251)	(0.051)

$\chi^2 = 98.954$

UI Recipients

	New Job	Recall
δ_1	0.005	0.076
	(0.005)	(0.010)
δ_2	1.446	0.645
	(0.345)	(0.068)

$\chi^2 = 117.620$

Figure 3.1 - Sample Hazard Rates for Re-Employment

Single Risk Model

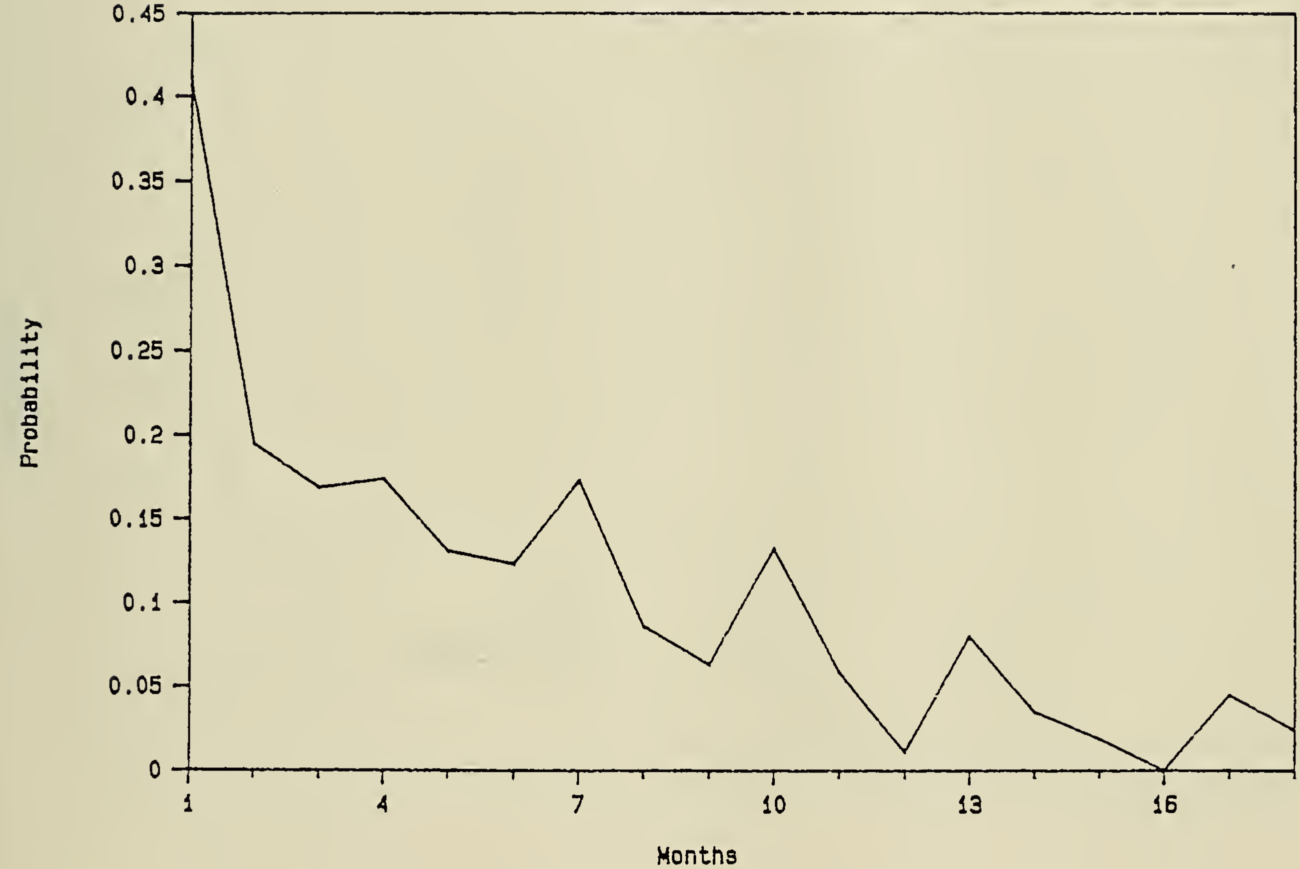


Figure 3.2 - Sample Hazard Rates for Re-Employment
Dual Risk Model

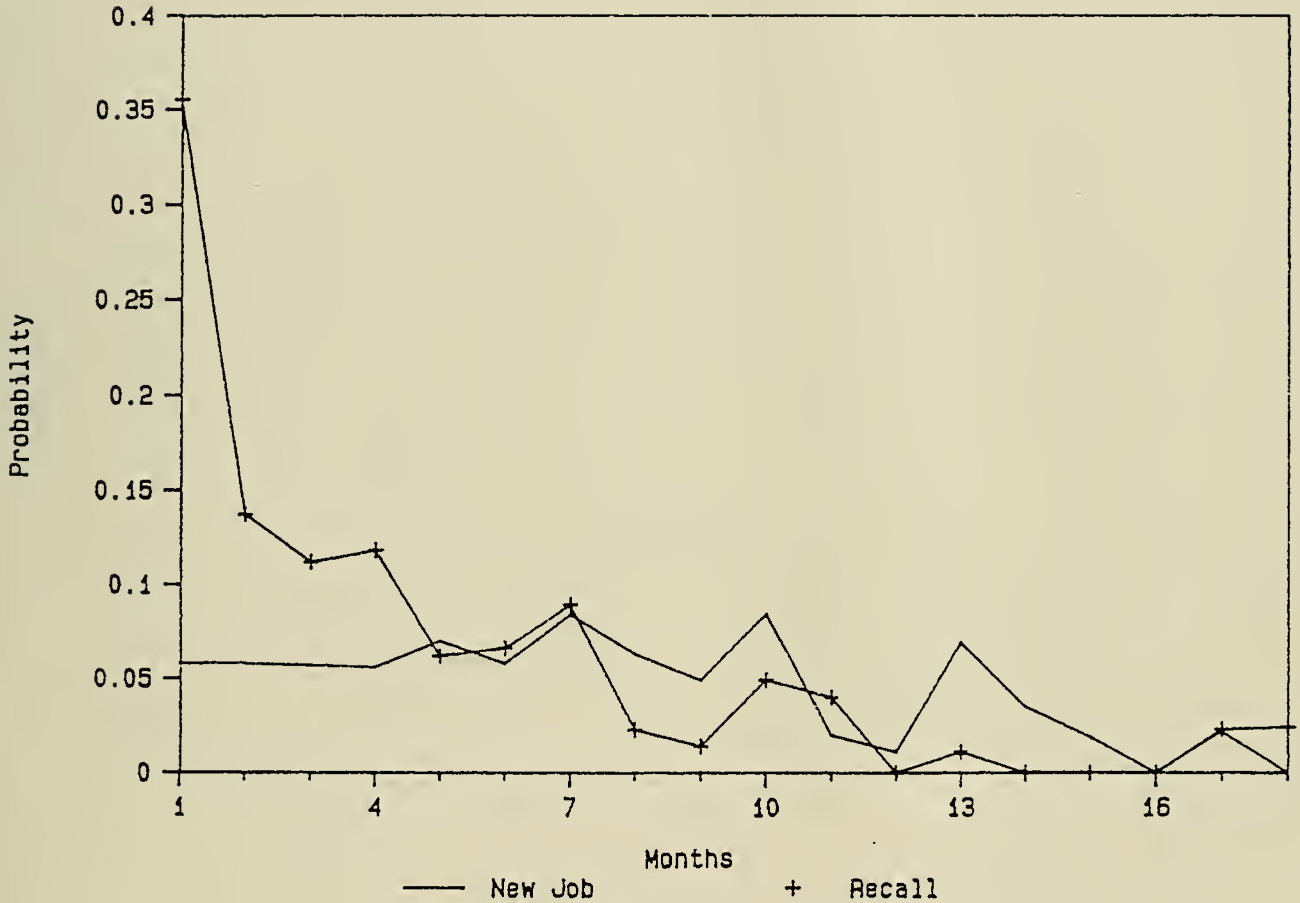


Figure 3.3 - Sample Hazard Rates for Re-Employment

Dual Risk Model - Individuals Receive UI

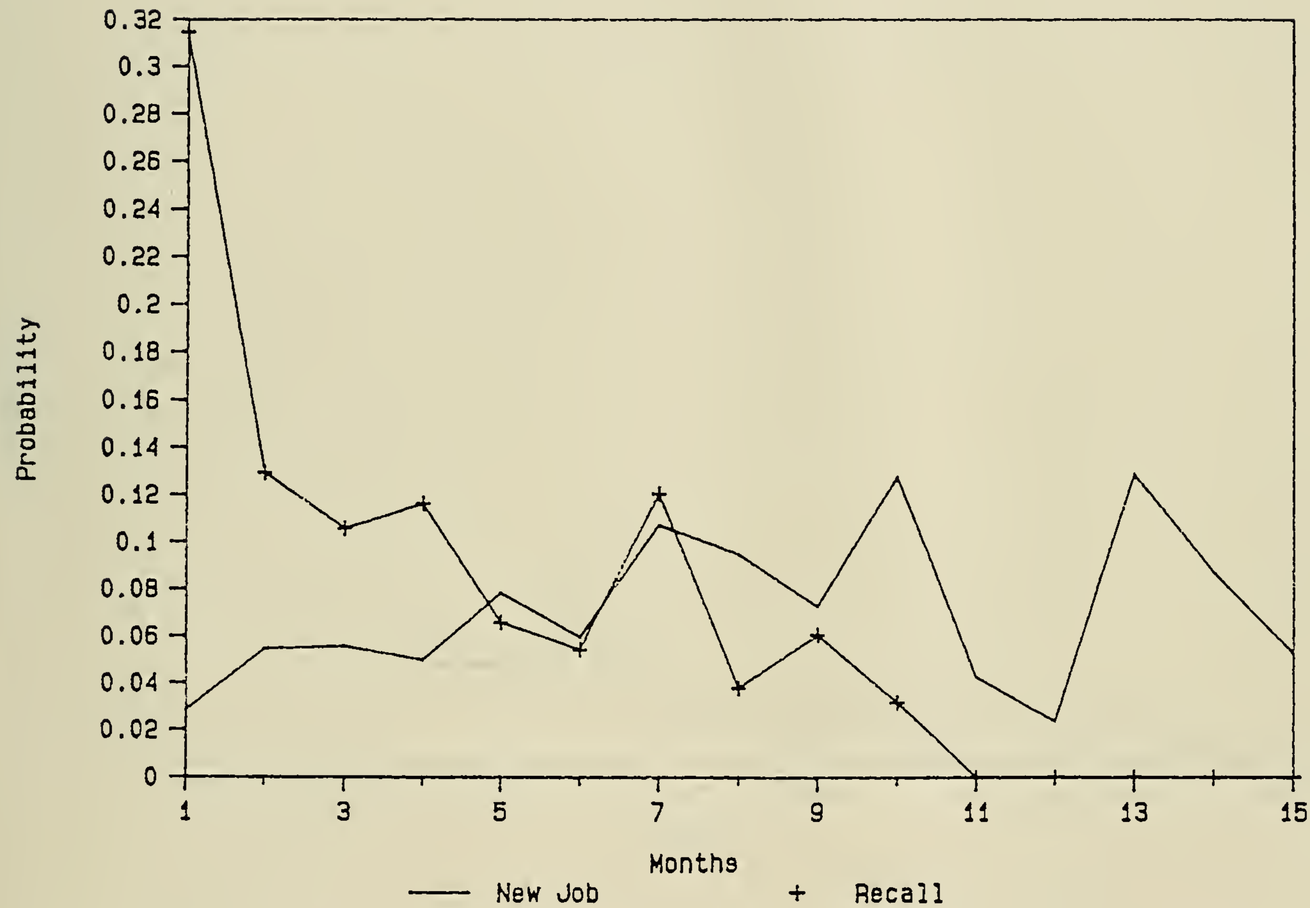


Figure 3.4 - Sample Hazard Rates for Re-Employment

Dual Risk Model - Individuals Do Not Receive UI

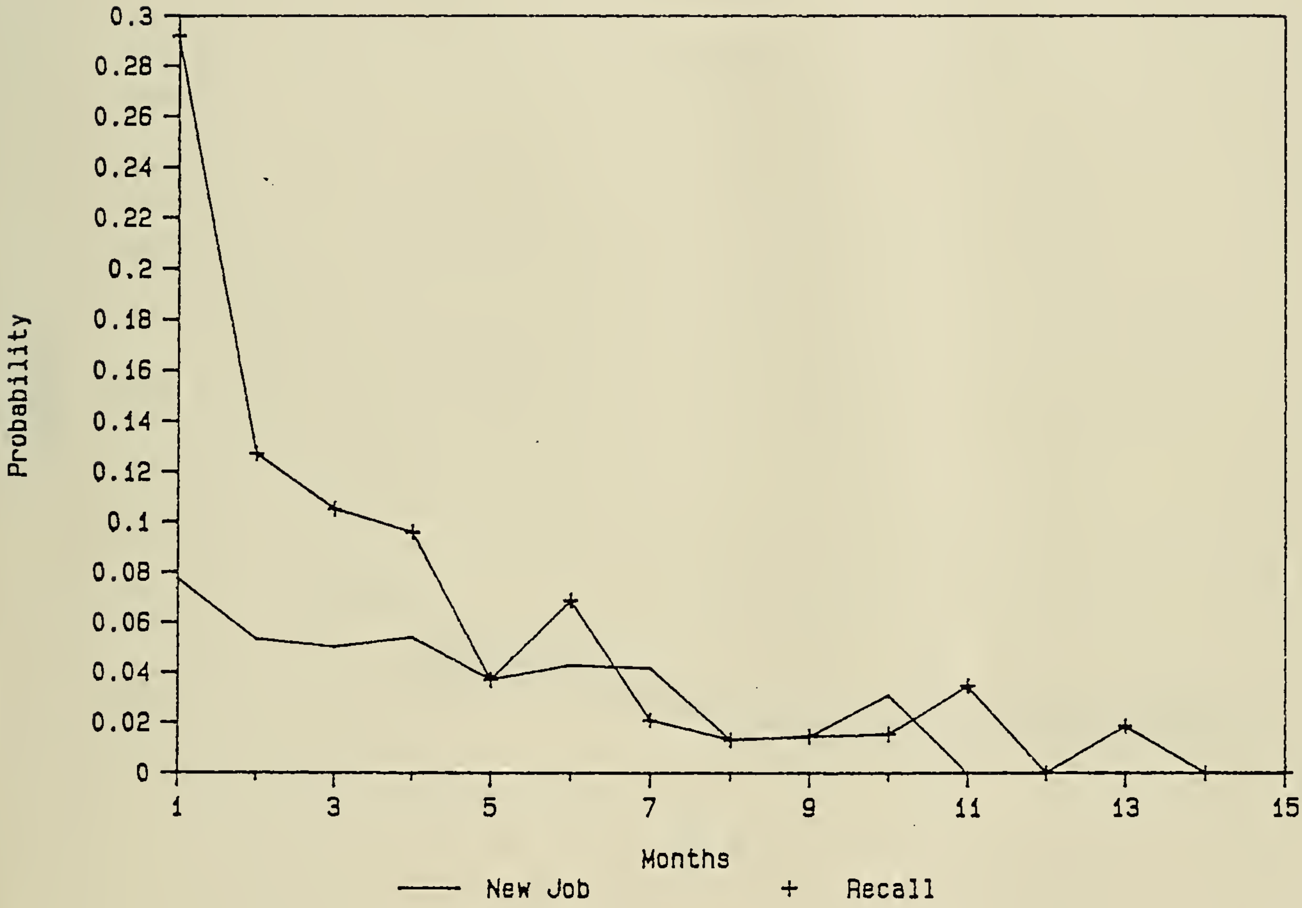


Figure 3.5 - Estimated Hazard Rates for Re-Employment

Single Risk, Various Models

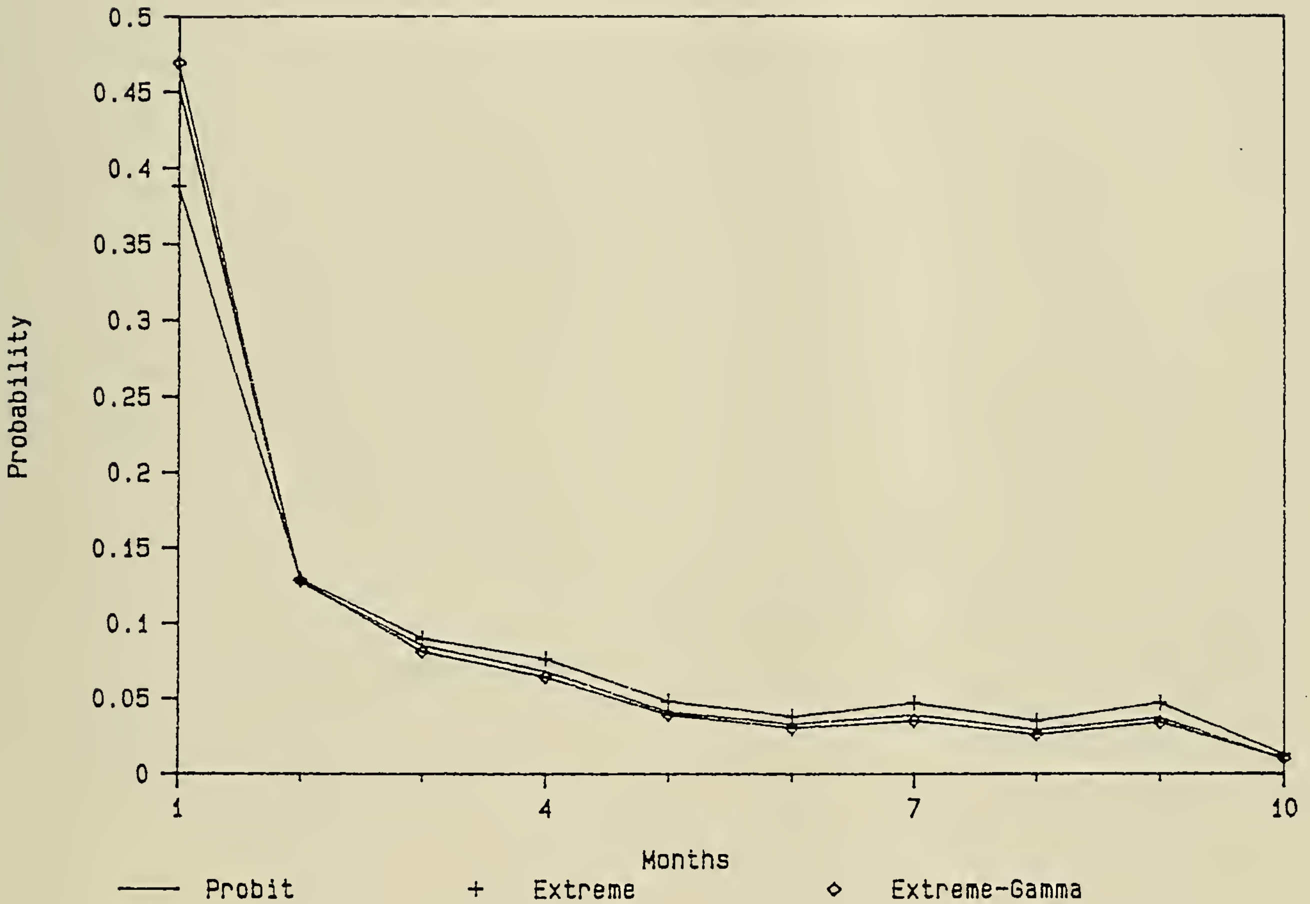


Figure 3.6 - Estimated Hazard Rates for Re-Employment

Dual Risk Model - Individuals Receive UI

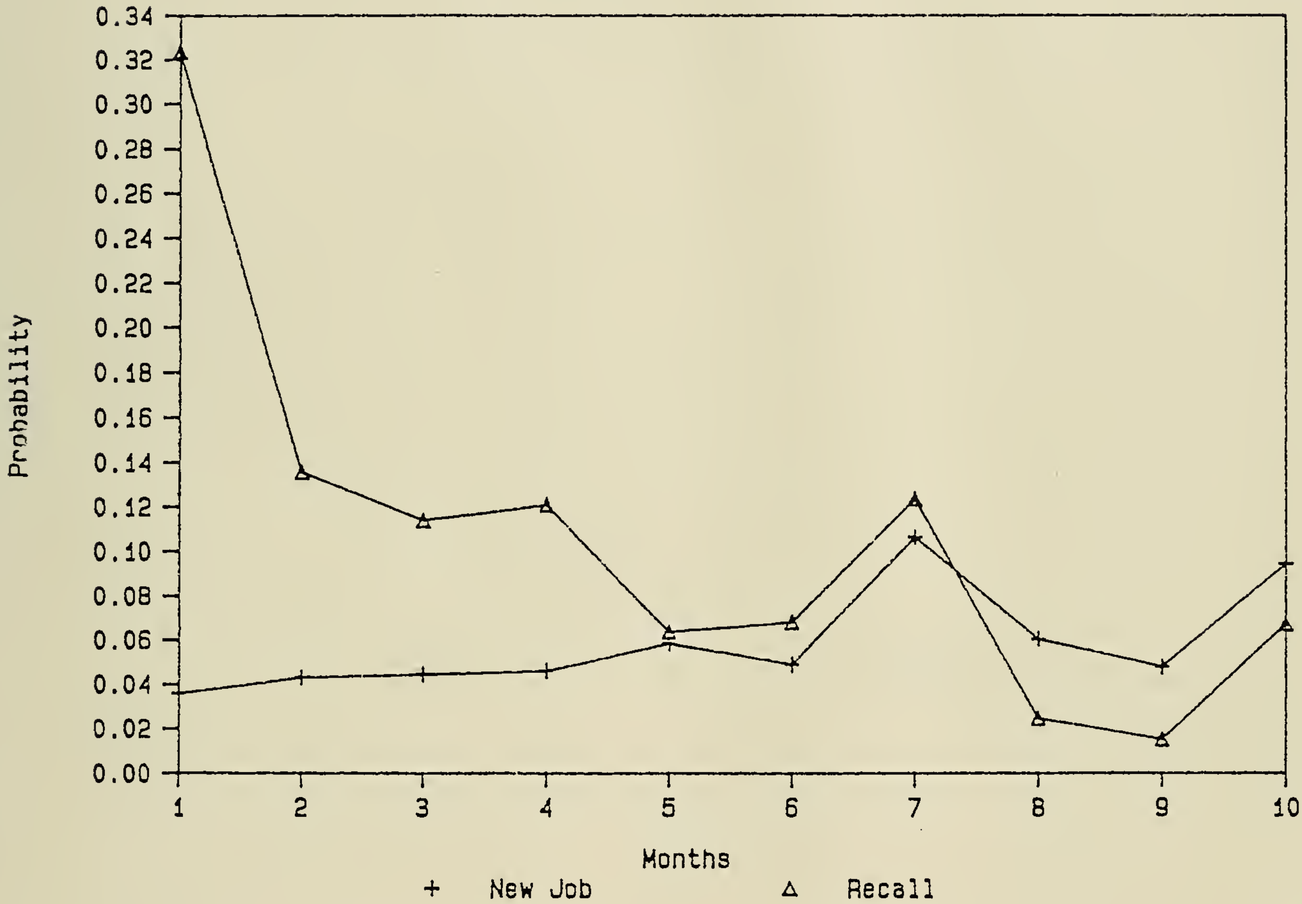


Figure 3.7 - Estimated Hazard Rates for Re-Employment

Dual Risk Model - Individuals Do Not Receive UI

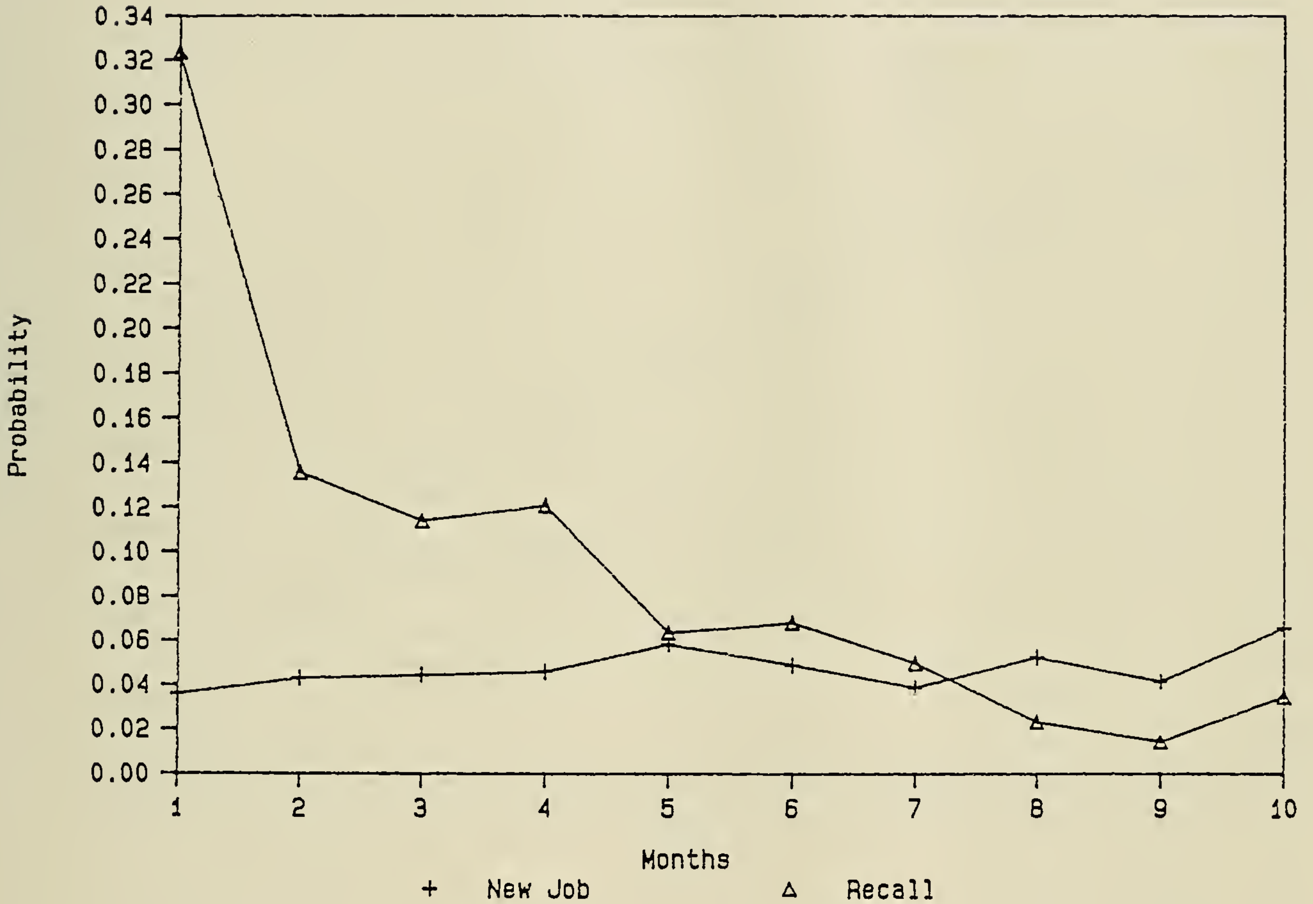


Figure 3.8 - Comparison of Semiparametric Baseline Hazard and Fitted Two-Parameter Weibull Hazard for New Job Duration - Individuals Receive UI

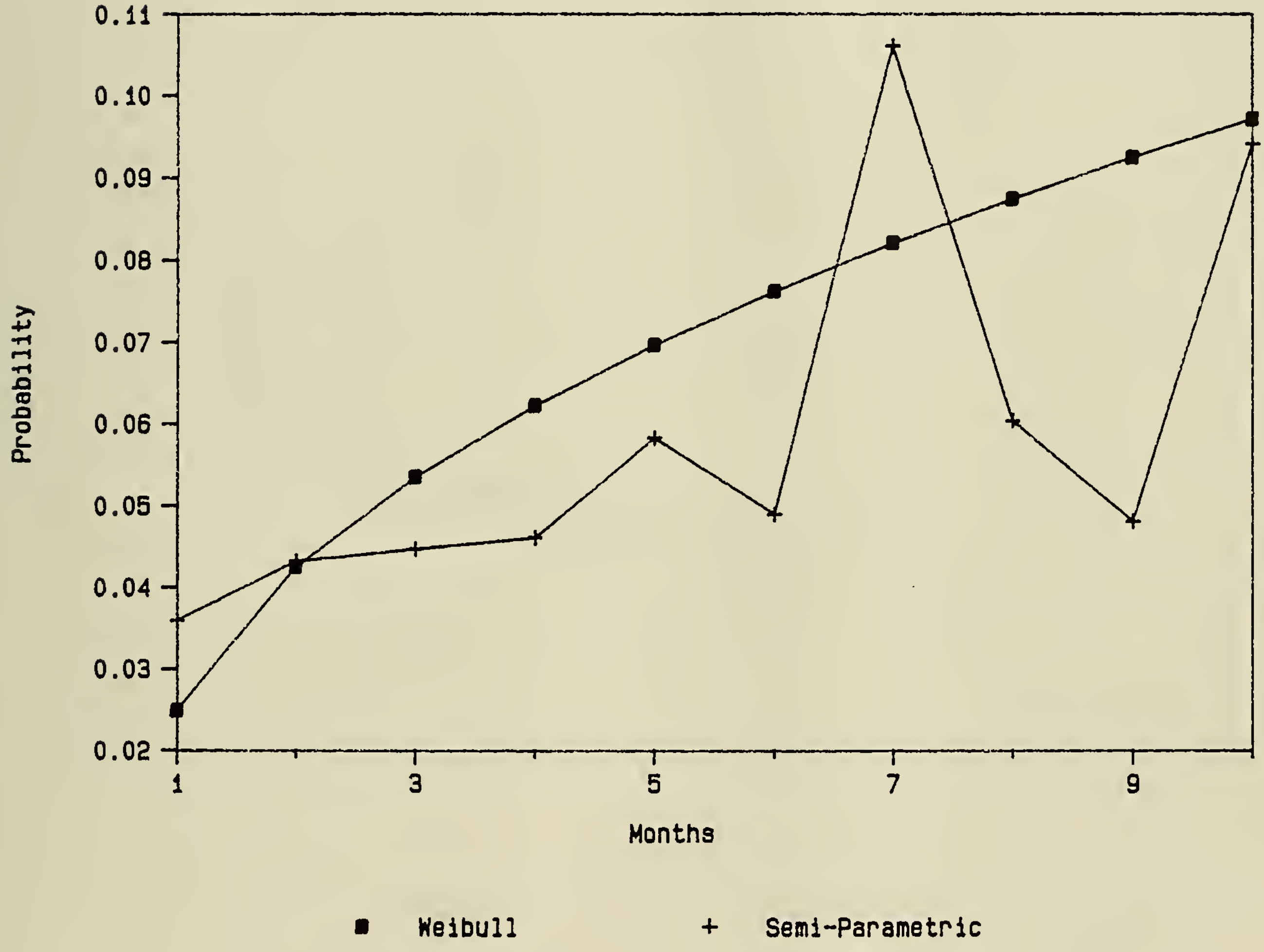
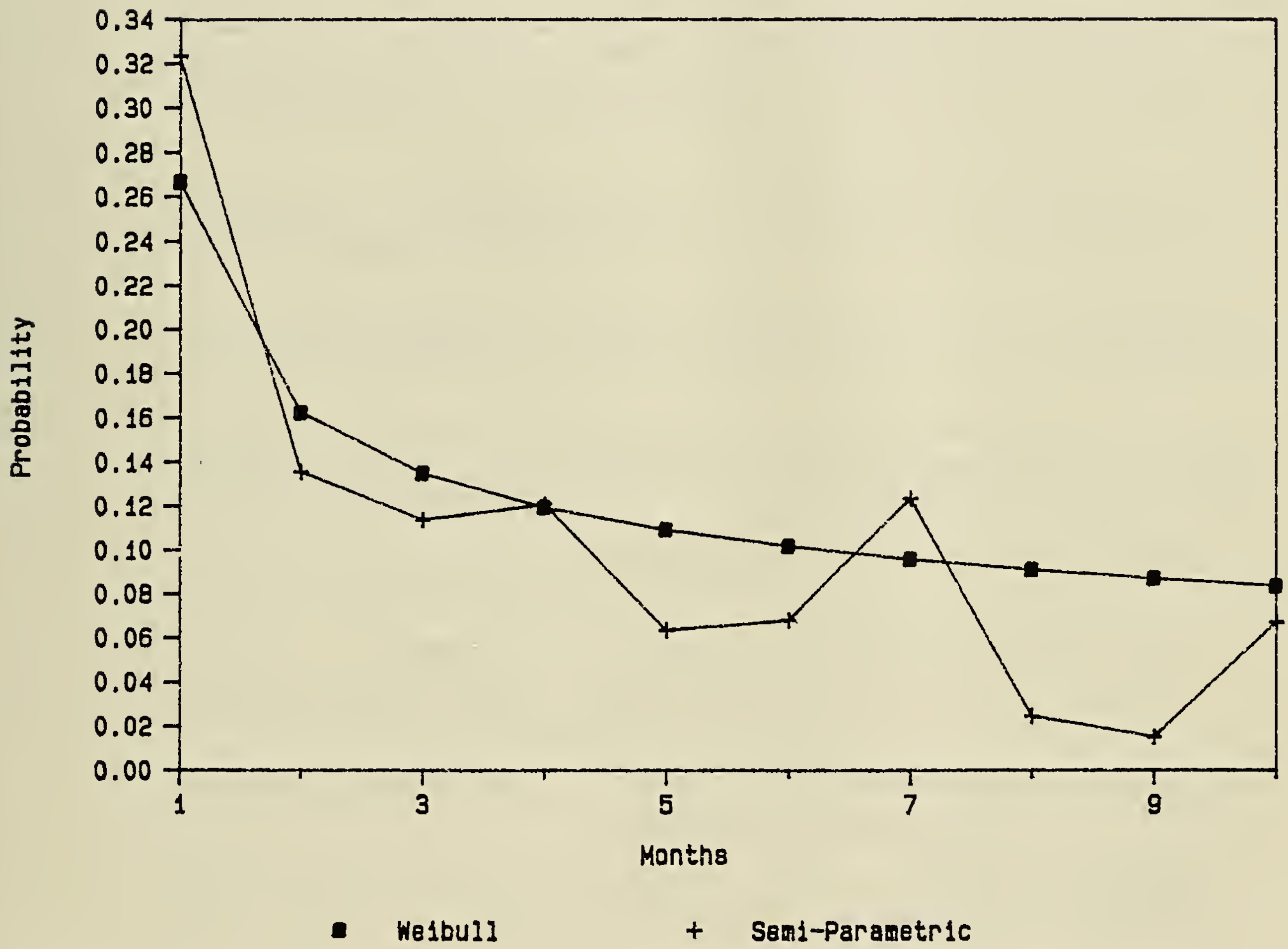


Figure 3.9 - Comparison of Semiparametric Baseline Hazard and Fitted Two-Parameter Weibull Hazard for Recall Duration - Individuals Receive UI



REFERENCES

1. Clayton, D. and Cuzick, J. (1985), "Multivariate Generalizations of the Proportional Hazards Model", Journal of Royal Statistical Society, Series A, 148, 82-117.

2. Cox, D.R. (1972), "Regression Models and Life-Tables", Journal of Royal Statistical Society, Series B, 26, 186-220.

3. Cox, D.R. (1975),"Partial Likelihood", Biometrika, 62, 269-276.

4. Cox, D.R. and Oakes, D. (1984), Analysis of Survival Data, Chapman and Hall.

5. Diamond, P. A., Hausman, J. A. (1984), "The Retirement and Unemployment Behavior of Older Men", in Aaron, H., and Burtless, G. eds. Retirement and Economic Behavior. (The Brookings Institute, Washington, DC).

6. Elber, C. and Ridder, G. (1882), "True and Spurious Duration Dependence" The Identifiability of the Proportional Hazards Model", Review of Economic Studies, 49, 402-411.

7. Ham, J.C. and Rea, S.A. (1986), "Unemployment Insurance and Male Unemployment Duration in Canada", Working Paper No. 212, Industrial Relation Section, Princeton University.

8. Heckman, J. and Singer, B. (1986), "Econometric Analysis of Longitudinal Data", in Grilichis, Z., and Intriligator, M., eds, The Handbook of Econometrics, Vol. 3, (Amsterdam: North Holland).

9. Heckman, J. and Singer, B. (1984a), "Econometric Duration Analysis", Journal of Econometrics, 24, 63-132.

10. Heckman, J. and Singer, B. (1984b), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data", Econometrica, 52, 271-320.

11. Heckman, J. and Singer, B. (1984c), "The Identifiability of the Proportional Hazard Model", Review of Economic Studies, 51, 321-341.

12. Kalbfleisch, J.D. and Prentice, R.L. (1973), "Marginal Likelihood Based on Cox's Regression and Life Model", Biometrika, 60, 267-278.

13. Kalbfleisch, J.D., Prentice, R.L. (1980), Statistical Analysis of Failure Time Data, (Wiley, New York).

14. Katz, L. (1986), "Layoffs, Recall and the Duration of Unemployment", Working Paper No. 1825, National Bureau of Economic Research, Cambridge, MA.

References (Cont'd)

15. Kenan, J. (1985), "The Duration of Contract Strikes in U.S. Manufacturing", Journal of Econometrics, 28, 1-24.

16. Lancaster, T. (1979), "Econometric Methods for the Duration of Unemployment", Econometrica, 47, 939-956.

17. Lancaster, T. and Nickell, S. (1980), "The Analysis of Reemployment Probabilities for the Unemployed", Journal of the Royal Statistical Society, Series A, 143, 141-165.

18. Manton, K.G., Stallard, E., and Vaupel, J.W. (1986), "Alternative Models for the Heterogeneity of Mortality Risks Among the Aged", Journal of Americal Statistical Society, 81, 635-644.

19. Meyer, B. (1986), "Semiparametric Estimation of Duration Models", MIT Mimeo.

20. Moffit, R. (1985), "Unemployment Insurance and the Distribution of Unemployment Spells", Journal of Econometrics, 28, 85-101.

21. Prentice, R.L. (1976), "A Generalization of the Probit and Logit Models for Dose Response Curves, Biometrics, 32, 761-768.

22. Sueyoshi, G. (1986), "Partial Retirement and Full Retirement: A Competing Risks Model of Exit from Full Time Employment".