JOURNAL OF APPLIED ECONOMETRICS, VOL. 5, 1–28 (1990)

FLEXIBLE PARAMETRIC ESTIMATION OF DURATION AND COMPETING RISK MODELS

AARON HAN
Department of Economics, Harvard University, Cambridge, MA 02138, USA

AND

JERRY A. HAUSMAN
Department of Economics MIT., Cambridge MA 02139, USA

SUMMARY

In this paper we specify and estimate a flexible parametric proportional hazards model. The model specification is flexibly parametric in the sense that the baseline hazard is non parametric while the effect of the covariates takes a particular functional form. We also add parametric heterogeneity to the underlying hazard model specification. We specify a flexible parametric proportional competing risks model which permits unrestricted correlation among the risks. Unemployment duration data are then analysed using the flexible parametric duration and competing risks specifications. We find an important effect arising from the exhaustion of unemployment insurance and significantly different hazards for the two types of risks, new jobs and recalls

1. INTRODUCTION

Since Lancaster's (1979) paper on unemployment, duration models have become commonly used in econometrics. Heckman and Singer (1986) and Kiefer (1988) give recent surveys. 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.[1] Three possible reasons exist:

1. It is a continuous time specification while most duration data in econometrics are discrete where the discreteness may well be important, e.g. unemployment data.
2. 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 have very many simultaneous failures, e.g. unemployment data at 26 weeks.
3. 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.

[1] This remark does not mean that Cox's model is never used in econometrics; however, parametric specification of the baseline hazard is by far the most widely used technique. See Kiefer (1988) for a discussion of the Cox model and for reference to the use of duration models in econometrics.

0883–7252/90/010001–28$14.00

Received June 1988
Revised June 1989

In this paper we specify and estimate a flexible parametric proportional hazards (duration) model. The model specification is flexibly parametric 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 true parameters of the covariates are invariant to the length of time intervals which are chosen.[2] 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 convenient 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 utilize an identification theorem of Han and Hausman (1988), 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 two continuous variables among the covariates. This identification result should diminish the considerable confusion in the literature over whether competing risk models can always be specified as independent hazards models. We then specify a flexible parametric proportional competing risks 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 4 unemployment duration data are analysed using the flexible parametric 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 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.

2. SPECIFICATION AND ESTIMATION OF A FLEXIBLE DURATION MODEL

We assume observations of failure times over the discrete periods $t = 0, 1, 2, \ldots, T$ for individuals $i = 1, \ldots, n$. We assume that the predetermined variables of each individual X_i do

[2] The efficiency of the parameter estimates is, of course, affected by the length of time intervals used. However, the underlying parameters which we attempt to estimate are not affected.

not change with time.[3] Our specification begins with the proportional hazards specification of Prentice (1976); see also Kalbfleisch and Prentice (1980, Ch. 4) where the hazard function, which is the failure rate at time τ conditional upon survival to time τ,

$$\lambda_i(\tau) = \lim_{\Delta \to 0^+} \frac{P(\tau < t_i < \tau + \Delta \mid t_i > \tau)}{\Delta} = \lambda_0(\tau)\exp(-X_i\beta), \tag{1}$$

is specified in the log form of the integrated hazard as

$$\log \int_0^{t_i} \lambda_0(\tau)\, d\tau = X_i\beta + \varepsilon_i, \tag{2}$$

where ε_i takes an extreme value form. Now let

$$\log \int_0^{t} \lambda_0(\tau)\, d\tau = \delta_t \qquad t = 1, \ldots, T \tag{3}$$

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

$$\int_{\delta_{t-1} - X_i\beta}^{\delta_t - X_i\beta} f(\varepsilon)\, d\varepsilon. \tag{4}$$

We treat the logarithm of the integrated baseline hazards, δ_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 estimated the baseline hazard parameters jointly with the regression parameters, 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.[4]

Let the indicator variable y_{it} take on the values $y_{it} = 1$ if failure in period t occurs for person i, and $y_{it} = 0$ otherwise. The log-likelihood function takes the form[5]

$$\log L = \sum_i \sum_t y_{it} \log \int_{\delta_{t-1} - X_i\beta}^{\delta_t - X_i\beta} f(\varepsilon)\, d\varepsilon, \tag{5}$$

For with an extreme value distribution the likelihood function is of an 'ordered' logit form. For ε with a standard normal distribution, the likelihood function takes the familiar ordered probit form. The ordered probit model does not follow strictly from the porportional hazards specification of equations (1) and (2), but it will provide a very close approximation to the ordered logit specification. This result occurs because the extreme value distribution and the normal distribution are very similar except in the extreme tails of the distribution.

[3] Our method can be extended to the case of non-constant X in a straightforward manner. The value of X_t in each period is used or its mean during the period is used if X_t changes during the period. In the context of equation (2) below, the X's can be interacted with time period dummy variables in this situation. See Meyer (1987) or Sueyoshi (1987).

[4] Moffit'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.

[5] Right censoring of the observations is easily added to the log-likelihood function in the usual way by including a term which specifies the cumulative probability of not failing at the time the observation is censored.

Both ordered logit and ordered probit models are extremely easy to estimate (even on a PC).[6] 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 δ_t, together with their covariance matrix, are estimated, then the applied investigator 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 δ_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 4.

An additional advantage of this specification is that the linear form of the proportional hazards model in equation (2) allows handling jointly endogenous variables or errors in variables via instrumental variable techniques for the ordered probit specification. These potential problems commonly occur in empirical applications of duration models, e.g. Diamond and Hausman (1984). Many parametric hazard approaches used in econometrics so far do not permit 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. The evaluation of these multiple integrals would lead to an inordinate requirement of computer time for estimation given a sample size on the order of 1000 or more, which often occurs in econometrics. The multiple integral occurs because the 'risk set' at time t, corresponding to those individuals which have not yet failed at time t in the Cox model, corresponds to the choice set in the multinomial logit model where each person in the risk set corresponds to a single choice in the logit model. The Cox model chooses one person from the risk set corresponding to the next observed failure. Introduction of individual unobserved heterogeneity then requires introduction of a random variable for each person in the risk set, just as it does for each choice set in the logit model. At the beginning of the sample when the risk set is large, evaluation of integrals of order 1000 or more would make estimation impractical.

However, in the specification of equation (2), heterogeneity is straightforward to include, and involves only a single additional order of integration in equation (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).[7] Assume a gamma distribution with mean one and variance $\sigma^2 = 1/\theta$. Then rewrite equation (2) in exponential form as

$$\exp\left\{\log \int_0^t \lambda_0(\tau)\, \mathrm{d}\tau - X_i\beta\right\} = \exp(\varepsilon_i + w_i), \tag{6}$$

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

$$I_i(t) = \left\{\int_0^t \lambda_0(\tau)\, \mathrm{d}\tau\right\}\exp(-X_i\beta), \tag{7}$$

[6] The reason for this ease of estimation is that for the ordered logit model the integral in the likelihood function of equation (5) is known in closed form. For the ordered probit model, although the cumulative normal distribution does not exist in closed form, partial fraction expansions of four to six terms give an extremely high degree of accuracy. These partial fraction expansions are calculated very quickly on a computer. Many subroutine libraries contain the univariate normal integral calculation function.

[7] It is interesting to note that Lancaster (1979) mentioned the possibility of replacing his parametric baseline hazard specification with a nonparametric specification. However, he lacked sufficient data to estimate such a model.

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. Since $I_i(t)$ is monotonically increasing in t, the probability that the failure time is greater than some value t is the integral of the density of q_i over the portion of the support of q_i which is greater than $I_i(t)$. Thus, upon integrating we find

$$\int_{I_i(t)}^{\infty} g(q)\ \mathrm{d}q = [1 + (1/\theta)I_i(t)]^{-\theta} \tag{8}$$

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

$$I_i(t) = \exp(-X_i\beta)\exp(\delta_t). \tag{9}$$

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:[8]

$$\log L = \sum_i \log\left[1 + (1/\theta) \sum_t^{m_i} I_i(t)\right]^{-\theta}, \tag{10}$$

where m_i is the failure time for the ith individual.

Preliminary results indicate the maximum-likelihood estimation of with up to 50 δ_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, 1984b) 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.[9] Our empirical results indicate that a nonparametric specification of the hazard function reduces the sensitivity of the estimates to a parametric heterogeneity assumption. If these empirical results are found to hold in other applications, they 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 refer to theorems on the asymptotic statistical properties of the maximum-likelihood estimators for our specifications. The regularity conditions for the theorems and the proofs are contained in an earlier version of this paper, Han and Hausman (1986). The data setup is:

$$t = 1, 2, \ldots, T \text{ discrete periods.}$$

(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 $y = 0$ otherwise.

The likelihood function, from equation (5) then becomes

$$L = \prod_{i=1}^{N} \prod_{t=1}^{T} \left[\int_{\delta_{t-1} - X_i\beta}^{\delta_t - X_i\beta} f(\varepsilon)\ \mathrm{d}\varepsilon\right] y_{it}, \tag{11}$$

[8] The likelihood function can easily be extended to include right censoring as we discussed before.
[9] 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 (1987) has proven identification of a Heckman–Singer-type estimator in a model similar to the one developed in this paper.

We do all asymptotics for fixed T, e.g. $T = 50$ weeks, and as N becomes large. Under suitable regularity conditions we are able to derive the usual asymptotic properties of the maximum likelihood estimator (MLE). Let $\hat{\theta} = (\hat{\beta}, \hat{\delta}, \hat{\gamma})$ be the MLE where γ is the (finite) parameter vector which characterizes the unobserved heterogeneity. Then

$$\sqrt{N}(\hat{\theta} - \theta_0) \xrightarrow{A} N\left(0, \left[\lim_{N \to \infty} \frac{1}{N} \frac{\partial^2 \log L}{\partial \theta \partial \theta'}\right]^{-1}\right). \tag{12}$$

Thus, large sample inference follows in the usual way based on an estimate of the inverse of the Hessian of the log-likelihood function.

In this section of the paper we have specified a flexible parametric 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 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.

3. COMPETING RISKS 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 unemployment can end with either a new job or withdrawal from the labour 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 behaviour 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). See Kalbfleisch and Prentice (1980), pp. 173–175, who make a claim of non-identification and discuss the use of independent competing risk specifications. We have proven that identification does exist under quite weak regularity conditions (Han and Hausman, 1988). Here we specify and estimate bivariate ordered probit models 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 $k = 1, \ldots, K$ denote K competing risks. Introduce the latent random variable $Y_k^* \geqslant 0$ for $k = 1, \ldots, K$, 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 $F_k(x) = pr(Y_k^* \leqslant x)$. However, Y_k^* is a latent variable in general since it cannot necessarily be observed. Instead, only the minimum Y_k^* of the theoretical lifetime is observed:

$$Q = \min(Y_1^*, \ldots, Y_k^*) = \min_k \ Y_k^*. \tag{13}$$

The available information includes the actual outcomes and the fact that if Q is greater than c, then the survivor distribution function, $\bar{F}_q(c) = 1 - F_q(c)$ together with the probability density function gives the conditional hazard rate functions $\tilde{f}_q(c) = f_q(c)/\bar{F}_q(c)$ for the observable variable Q. 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, where δ_t^1 and δ_t^2 are usually functions of the original latent variables Y_1^* and Y_2^*, can then be written as

$$\begin{aligned} \delta_t^1 &= X_1\beta_1 + \varepsilon_1 \\ \delta_t^2 &= X_2\beta_2 + \varepsilon_2 \end{aligned} \tag{14}$$

where the δ_t^i are functions of the failure times.

Katz (1986) assumes that the stochastic disturbances, ε_1 and ε_2, are independent and thus treats equation (14) as two standard duration models which were discussed above. Diamond and Hausman (1984) allowed for dependence by assuming the δ_t^1 and δ_t^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 flexible parametric 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 (14). 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 then the model is identified in the sense that an observationally equivalent independent competing risks model does not exist. However, in many econometric appications such as the unemployment problem $X_1 = X_2$ so that this convenient identifying assumption does not exist.

However, in a separate paper (Han and Hausman, 1988), we prove identification of the bivariate competing risks model so long as at least two covariates are continuous and certain other regularity conditions are maintained, even if X_1 and X_2 are identical.[10] Thus, we have solved the long-standing identification problem for competing risks modes, at least for many econometric and statistical applications with covariates present. We apply this identification result to our discrete model of the preceding section.

We first consider the flexible parametric 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

$$\begin{aligned} \delta_{t_1}^1 &= -\log \int_0^{t_1} \lambda_0^1(s)\ \mathrm{d}s = X\beta_1 + \varepsilon_1 \\ \delta_{t_2}^2 &= -\log \int_0^{t_2} \lambda_0^2(s)\ \mathrm{d}s = X\beta_2 + \varepsilon_2. \end{aligned} \tag{15}$$

[10] The result generalizes to additional competing risks where the number of continuous variables must be at least as great as the number of competing risks.

Suppose that the failure type is of type 1 so that $t_1 = \min(t_1, t_2)$ the probability of this outcome, given the grouping of the underlying data, is

$$\int_{\delta_{t-1}^1 - X\beta_1}^{\delta_t^1 - X\beta_1} \int_{m(\varepsilon_1)}^{\infty} f(\varepsilon_1, \varepsilon_2)\, d\varepsilon_2\, d\varepsilon_1, \tag{16}$$

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,

$$X\beta_1 + \varepsilon_1^* = -\log \int_0^{t^*} \lambda_0^1(s)\, ds, \tag{17}$$

where $t^* \in (t-1, t)$,[11]

Thus, we find

$$X\beta_1 + \varepsilon_1^* = \delta_{t-1}^1 + k_1(\delta_t^1 - \delta_{t-1}^1), \tag{18}$$

where k_1 is defined by

$$\delta_{t^*}^1 = \delta_{t-1}^1 + k_1(\delta_t^1 - \delta_{t-1}^1), \tag{19}$$

We then use equation (2.6) to solve for k_1.

$$k_1 = (X\beta_1 + \varepsilon_t^* - \delta_{t-1}^1)/(\delta_t^1 - \delta_{t-1}^1), \tag{20}$$

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

$$\varepsilon_2^* \geqslant \delta_{t-1}^2 - X\beta_2 + \left[\frac{\delta_t^0 - \delta_{t-1}^2}{\delta_t^1 - \delta_{t-1}^1}\right] [\varepsilon_1^* - (\delta_{t-1}^1 - X\beta_1)] = m(\varepsilon_1^*). \tag{21}$$

We have thus solved for $m(\varepsilon_1)$ in equation (16), so that the probability of failure type 1 in period t is

$$\int_{\delta_{t-1}^l - X\beta_1}^{\delta_t^1 - X\beta_1} \int_{\delta_{t-1}^2 - X\beta_2 + h}^{\infty} f(\varepsilon_1, \varepsilon_2)\, d\varepsilon_2\, d\varepsilon_1 \tag{22}$$

$$h = \left[\frac{\delta_t^2 - \delta_{t-1}^2}{\delta_t^1 - \delta_{t-1}^1}\right] [\varepsilon_1 - (\delta_{t-1}^1 - X\beta_1)].$$

Thus, the specification estimates two sets of δ_t which gives a flexible 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 baseline cause specific hazards as it did in the previous Diamond and Hausman (1984) specification.

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

$$t = 1, 2, \ldots, T \text{ discrete periods}$$

$(t_1, d_i, X_{1i}, X_{2i})$ 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

[11] The linearity assumption is tested later to assure that it is accurate enough for estimation purposes. Note that it is equivalent to the assumption of a constant hazard within time intervals so that for a particular application it should be checked by varying the length of the time interval.

individual fails at time period t, and $y_{it}=0$ otherwise. The log likelihood function is

$$\log L=\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}\Bigg[(1-d_i)\log\int_{\delta_{t-1}^1-X_{1i}\beta_1}^{\delta_t^1-X_{1i}\beta_1}\int_{[\delta_t^2-X_{2i}\beta_2]+h_1}^{\infty} f(\varepsilon_1,\varepsilon_2)\,d\varepsilon_1\,d\varepsilon_2$$

$$+\,d_i\log\int_{\delta_{t-1}^2-X_{2i}\beta_2}^{\delta_t^2-X_{2i}\beta_2}\int_{[\delta_t^1-X_{1i}\beta_1]+h_2}^{\infty} f(\varepsilon_1,\varepsilon_2)\,d\varepsilon_1\,d\varepsilon_2\Bigg] \tag{23}$$

where

$-\infty<\delta_1^1<\delta_2^1<\ldots<\delta_T^1<\infty,$
$-\infty<\delta_1^2<\delta_2^2<\ldots<\delta_T^2<\infty,$
$h_1=[(\varepsilon_1-(\delta_t^1-X_{1i}\beta_1))\lambda_t]$
$h_2=[(\varepsilon_2-(\delta_t^2-X_{2i}\beta_2))/\lambda_t]$, and
$\lambda_t=(\delta_t^2-\delta_{t-1}^2)/(\delta_t^1-\delta_{t-1}^1)$ for $t=2,\ldots,T-1$ with $\lambda_1=\lambda_T=1$.

Again, under a 'standard' set of regularity conditions given in Han and Hausman (1986) we are able to derive the usual asymptotic properties of the maximum-likelihood estimator. The most stringent regularity condition basically requires that both X_1 and X_2 contain at least one continuous variable, although both vectors may be the same.[12] Note that requirement of a continuous regressor variable need not be satisfied in all econometric applications. Given the regularity conditions we again find that inference follows from the inverse of the Hessian matrix of the log-likelihood function where γ is the (finite) vector which parametrizes the stochastic terms in the specification, e.g. the standard bivariate normal distribution with unknown correlation coefficient:

Let $\hat\theta=(\hat\beta_1,\hat\beta_2,\hat\delta^1,\hat\delta^2,\hat\gamma)$ be the MLE with the error distribution

$$F(\varepsilon_1,\varepsilon_2\,|\,\gamma_0).\text{ The }\sqrt{N}\,(\hat\theta-\theta_0)\xrightarrow{A}N\left(0,\left[\lim_{N\to\infty}\frac{1}{N}\frac{\partial^2\log L}{\partial\theta\partial\theta'}\right]^{-1}\right). \tag{24}$$

As with the ordered probit specification of section 2, the bivariate normal specification does not follow strictly from the proportional hazards specification. Nevertheless, given the close approximation of the ordered probit to the ordered logit specification for empirical work, the approximation should be suitable.

We now consider a limited range of Monte-Carlo experiments to ascertain whether the linearity approximation used in equations (16)–(23) is sufficiently accurate for estimation purposes. The Monte-Carlo experiments also permit us to investigate whether the use of interval data to estimate an underlying continuous baseline hazard produces accurate results. We consider a two-hazard competing risks model with two covariates, which are identical, in the model specification of equation (15). The covariates are randomly generated from a bivariate normal distribution with means equal to zero, variances equal to 5, and a correlation coefficient of $0\cdot3$. The parameters of the covariates are assumed to be $\beta_1=(0\cdot5,0\cdot5)$ for the first hazard and $\beta_2=(0\cdot8,-0\cdot3)$ for the second hazard. The baseline hazards are assumed to be two-parameter Weibull distributions with location and shape parameter equal to $(0\cdot05,1\cdot1)$ for the first baseline hazard and $(0\cdot08,0\cdot9)$ for the second baseline hazard. Lastly, the stochastic disturbances in equation (15) are bivariate normal with zero means, unit variances, and a correlation coefficient which varies between $-0\cdot9$ up to $+0\cdot9$.

[12] Somewhat more stringent regularity conditions are required when X_1 and X_2 are the same. However, the presence of continuous regressor variables will continue to satisfy the more stringent conditions in most econometric applications.

The Monte-Carlo experiments were conducted with 1500 observations per replication with new covariates and new stochastic disturbances for each replication. We used 50 replications for each value of the correlation coefficient chosen. Lastly, we took each continuous failure time and rounded it up to the next integer. For example, if the minimum failure time is as 2·45, we record the data as 3, which is the way those data are typically recorded in many econometric applications.

Table I. Results of 50 Monte-Carlo experiments for bivariate competing risks

True ρ		Parameter β_{11} (0·5)	β_{12} (0·5)	β_{21} (0·8)	β_{22} (−0·3)	ρ
0·9	Mean	0·5086	0·5040	0·7996	−0·3123	0·8967
	Variance × 10,000	4·9669	8·2092	13·5569	9·6012	5·0225
	MSE × 10,000	5·7134	8·3710	13·5589	11·1072	5·1293
	Minimum	0·4572	0·4461	0·7183	−0·3726	0·8278
	Maximum	0·5569	0·5685	0·8955	−0·2396	0·9361
0·6	Mean	0·5062	0·4933	0·7925	−0·3004	0·6079
	Variance × 10,000	4·8056	7·0535	11·0861	12·5145	82·6522
	MSE × 10,000	5·1875	7·4965	11·6438	12·5160	83·2808
	Minimum	0·4699	0·4285	0·7235	−0·3736	0·3772
	Maximum	0·5546	0·5604	0·8983	−0·2242	0·8117
0·3	Mean	0·5019	0·4995	0·7973	−0·3129	0·2644
	Variance × 10,000	4·4120	5·7120	8·1531	9·2993	160·6331
	MSE × 10,000	4·4462	5·7142	8·2282	10·9623	173·3239
	Minimum	0·4562	0·4554	0·7320	−0·3672	0·0474
	Maximum	0·5401	0·5446	0·8879	−0·2234	0·5606
0·0	Mean	0·5050	0·4942	0·7935	−0·3075	−0·0152
	Variance × 10,000	5·8636	4·9330	15·8900	12·5523	297·4839
	MSE × 10,000	6·1158	5·2701	16·3078	13·1121	299·8075
	Minimum	0·4643	0·4310	0·7273	−0·3604	−0·3163
	Maximum	0·5562	0·5463	0·8887	−0·2316	0·4330
−0·3	Mean	0·5119	0·4954	0·8068	−0·3092	−0·3031
	Variance × 10,000	9·9785	5·8514	18·0158	15·1311	119·4919
	MSE × 10,000	11·3894	6·0651	18·4763	15·9830	119·5855
	Minimum	0·4689	0·3984	0·7297	−0·5234	−0·5000
	Maximum	0·6491	0·5467	0·9792	−0·2552	−0·0817
−0·6	Mean	0·5155	0·4925	0·7985	−0·3006	−0·5930
	Variance × 10,000	5·8712	3·6086	13·9034	4·4897	56·2736
	MSE × 10,000	8·2799	4·1708	13·9253	4·4931	56·7585
	Minimum	0·4739	0·4565	0·7283	−0·3369	−0·7691
	Maximum	0·5667	0·5265	0·8894	−0·2598	−0·4677
−0·9	Mean	0·5147	0·5037	0·8092	−0·3026	−0·8955
	Variance × 10,000	6·2495	3·0545	11·3757	3·3926	11·0308
	MSE × 10,000	8·4098	3·1926	12·2180	3·4592	11·2302
	Minimum	0·4772	0·4702	0·7311	−0·3420	−0·9430
	Maximum	0·5638	0·5439	0·8681	−0·2567	−0·8161

Overall, the results are extremely good. The estimates of the piecewise linear baseline hazards accord extremely well with the true underlying continuous hazards. Furthermore, the parameter estimates of the baseline hazards, the covariate parameters, and the correlation coefficient from the underlying bivariate normal distribution all have empirical distributions which are well approximated by the asymptotic distribution theory. Investigation of stem-and-leaf plots and normal probability plots demonstrated that the asymptotic distribution theory of normal distributions conformed well to our empirical findings.

The results of the Monte Carlo experiments are given in Table I. Note that we hold the covariate parameters at the same values and that we vary ρ, the correlation coefficient, from $+0\cdot 9$ to $-0\cdot 9$. The parameters estimates are extremely good with the mean estimated values very close to the true values. The maximum bias for the covariate parameters is $4\cdot 2$ per cent for β_{22} when $\rho = 0\cdot 3$. The maximum bias for the estimates of ρ is $12\cdot 6$ per cent, which occurs for the same experiment. Thus, the estimated mean square errors are close to the estimated variances. Lastly, an inspection of the extreme values of the estimates provides a demonstration that the intervals in which the estimates fell were not overly wide.

Thus, we conclude that the use of piecewise linear approximations to the hazard functions and the linear approximation used to estimate the competing risk models do not affect our ability to achieve quite accurate estimate of the underlying model parameters. Despite estimates up to 40–50 parameters for each hazard, we are still able to attain quite accurate estimates of the covariate parameters and the correlation coefficient of the stochastic disturbances. The potential advantage of the approach, which does not require *a priori* specification of the underlying hazard functions, is thus fulfilled with sample sizes of the magnitude typically found in econometric applications which utilize hazard and competing risk specifications.

4. EMPIRICAL 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 has assumed. By separating the overall re-employment 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 in 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 nonparametric estimation of the baseline hazards removing these restrictions, while also allowing stochastic dependence between the two hazards.

The sample is taken from waves 14 and 15 of the PSID. These two waves of the PSID include detailed questions on the unemployment experience of heads of households in the previous year. Information included allows determination of the last unemployment spell in the previous year, the duration of the unemployment spell in weeks, the reason for the unemployment spell, and the reason for the end of the spell which is either a new job, recall, or censorship of the spell at the interview date. Thus, the data are right censored, but no left censorship is present, which eliminates potentially difficult econometric problems. 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 occurred in either 1980 or 1981.

A potential problem of the PSID sampling frame is that long-duration spells will be oversampled because the individual's last unemployment spell must overlap the year previous to an interview to be included in the sample. 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. Another data problem occurs in the PSID. Information is available only on whether an individual received unemployment insurance (UI) during the unemployment spell. Information is not available which allows computation of weekly UI

Table II. 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 non-white	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 agriculture	0·034
Service	= 1 if in services	0·172
Construction	= 1 if in construction	0·180
Occupation dummy variables (at onset of spell)		
Labourer	= 1 if labourer or operative	0·508
Craft	= 1 if craftsman 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).

benefit levels or replacement rates which have been used in much previous research. Thus, estimates of the effect of replacement rates in leading to longer duration of unemployment spells cannot be made from our 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 II. Note that our sample is identical to the Katz (1986) sample with non-whites being oversampled because of the sample frame of the PSID. Recalls are the most important way in which unemployment ends in the sample: 57 per cent of the spells end in recall. Of the remaining spells, 23 per cent end in a new job while the remaining 20 per cent of unemployment spells are censored by the interview date. The basic outcomes are given in Table III. Note the increases in exit from unemployment at 26 and 39 weeks, which are the exhaustion points of UI benefits.

Table III. 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

Continued

Table III. (Continued)

Weeks	New job	Recall	Censored	Total
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
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

The low coverage of UI benefit, 63.6 per cent, is notable, although previous research by Burtless (1983) has also pointed out this fact. Individuals who received UI in our sample are much likely to be white males who were laid off from a manufacturing firm. The non-UI recipients are more likely to be low-income workers with unstable labour market histories. This result is expected given the qualification requirements for UI coverage; nevertheless, the rather significant differences have not been widely recognized previously.

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 1. Here both recalls and new jobs are grouped

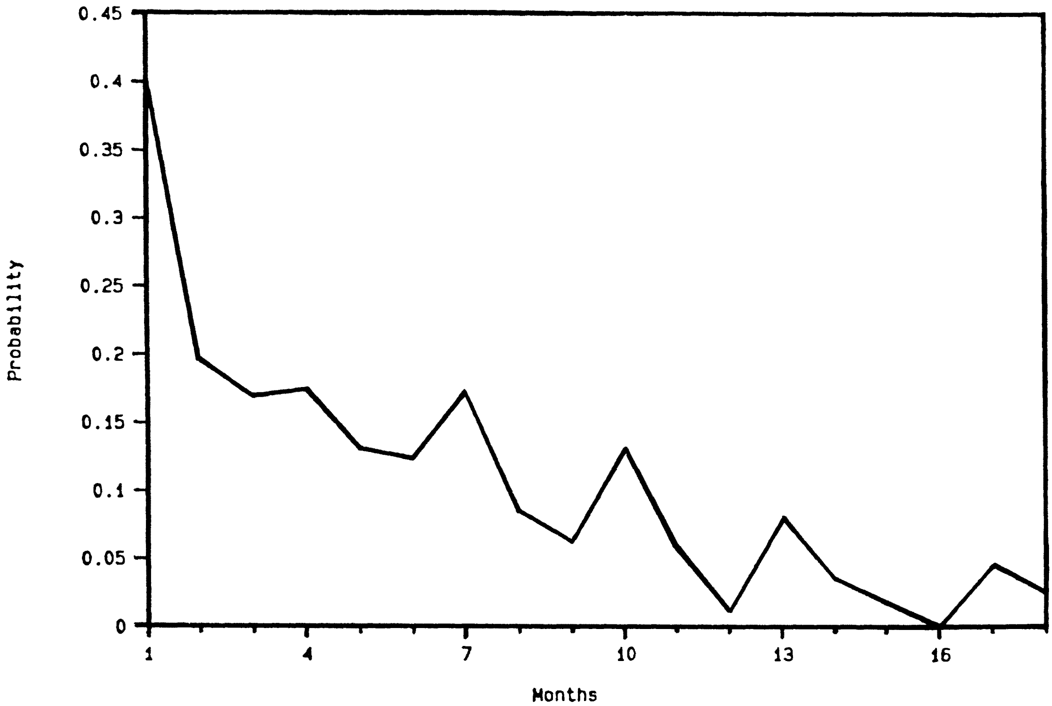

Figure 1. Sample hazard rates for re-employment, single risk model

together as exits from unemployment. Note the prominent spikes at the UI exhaustion points of 26 weeks and 39 weeks (near 7 months and 10 months in Figure 1).

While the rules for UI benefits differ greatly across states, during the sample period 10 states had maximum benefit periods of 26 weeks. In most other states the maximum duration of benefits was also 26 weeks, although some workers could qualify for additional weeks of benefits. Federally subsidized benefits extended UI benefits up to 39 weeks if the state unemployment rate exceeded a given threshold value over the previous 13-week period. Thus, the UI benefit exhaustions points of 26 and 39 weeks arise from the underlying state and federal rules which govern benefit duration.[13] The spikes in the estimated hazard correspond to expected individual behaviour when UI benefits are exhausted, but their large size may well lead to a reconsideration of models of individual search behaviour for jobs while unemployed.

In Figure 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. The presence of spikes in the estimated hazards again occurs at the UI exhaustion limits of 26 and 39 weeks, with mùch greater prominence in the new job hazard than in the recall hazard. Given the degree of employer control over the recall decision, the presence of a greater effect of UI on new jobs than on recalls would be expected.

In Figures 3 and 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

[13] Katz (1986) found that almost all UI recipients whose unemployment spells ended at 39 weeks lived in states during time periods when the extended benefits of 39 weeks were in place.

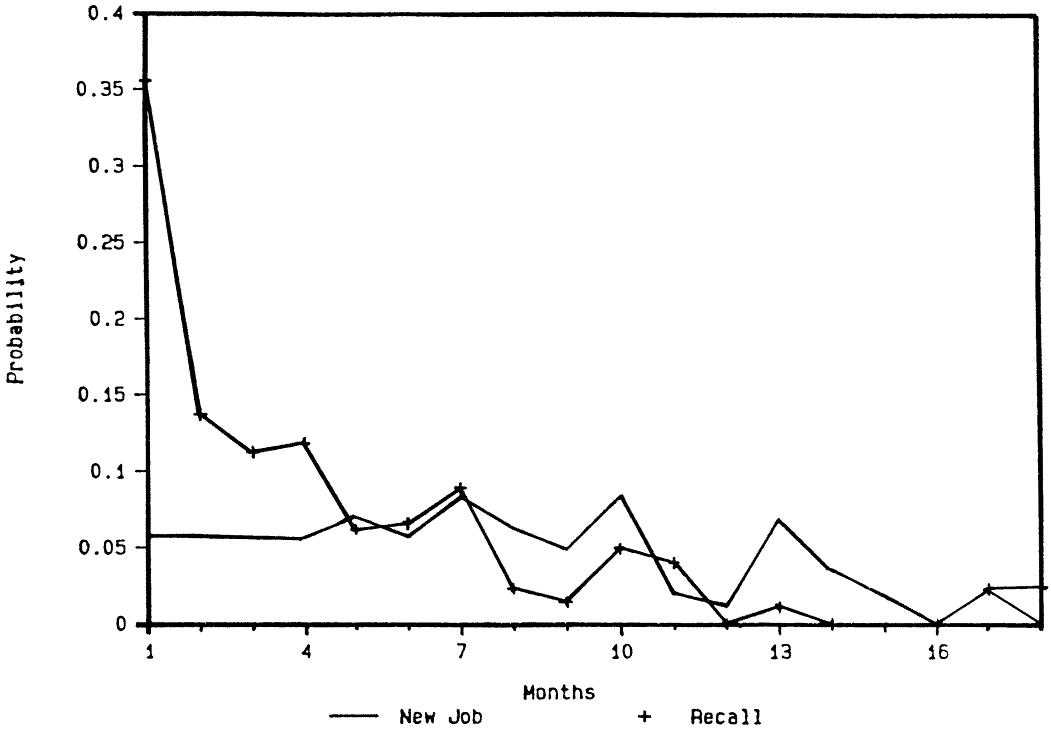

Figure 2. Sample hazard rates for re-employment, dual-risk model

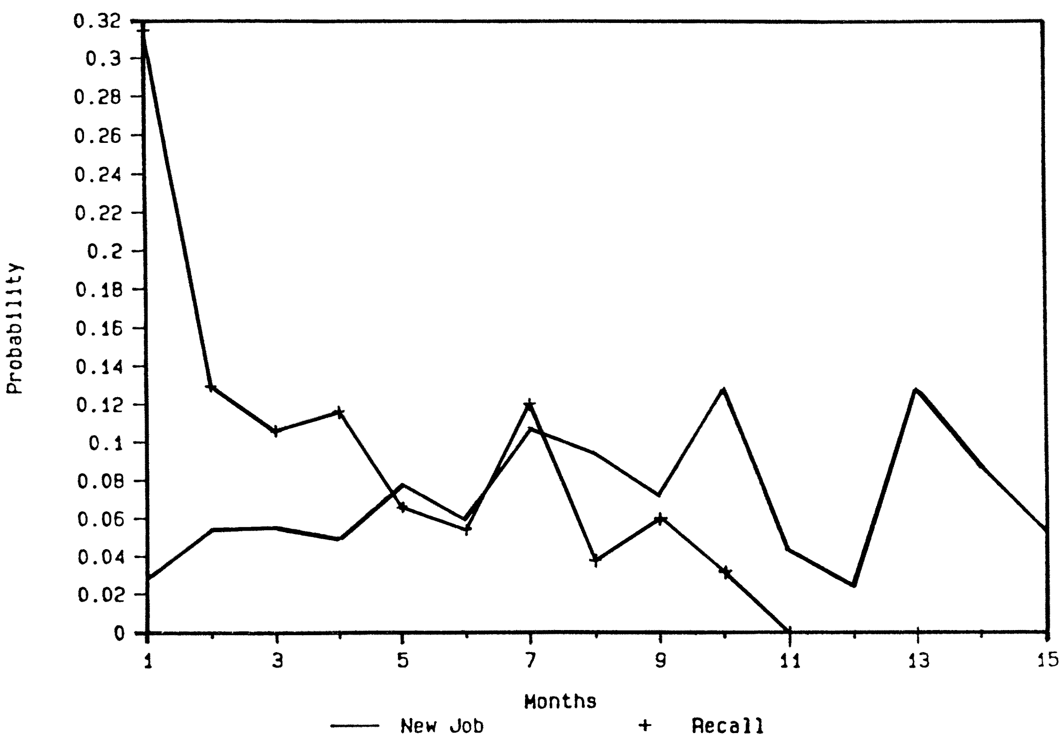

Figure 3. Sample hazard rates for re-employment, dual-risk model—individuals receive UI

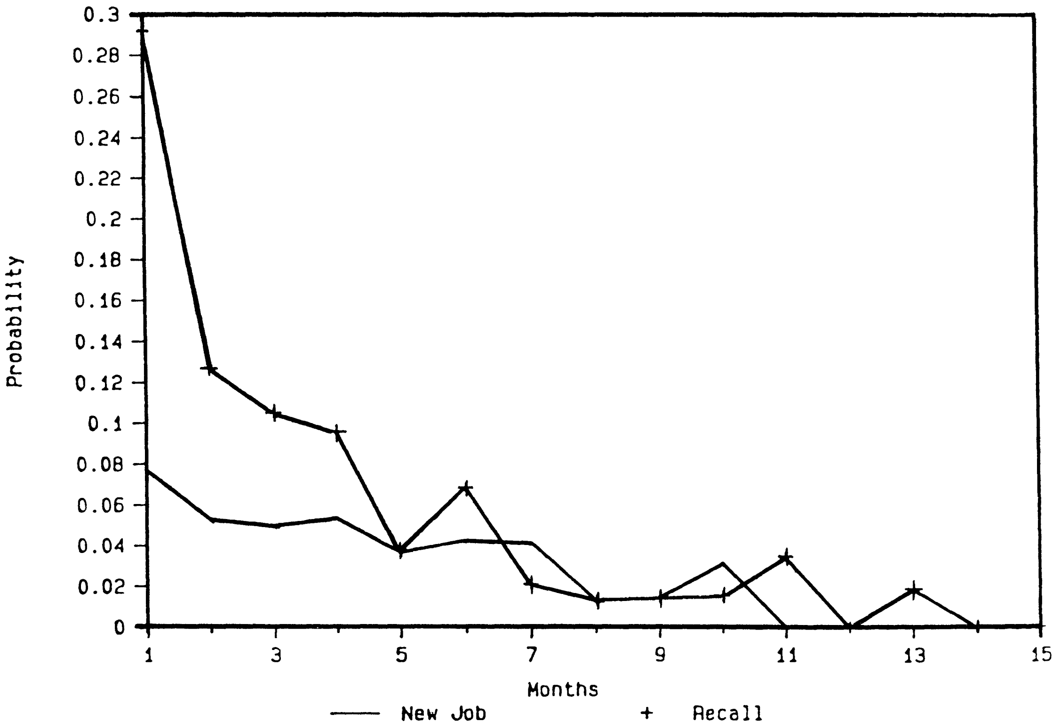

Figure 4. Sample hazard rates for re-employment, dual-risk model—individuals do not receive UI

weeks which appeared in the total sample in Figure 4. 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 exits from unemployment. A possible reason for the presence of the increased hazard near 26 weeks in the recall hazard for UI recipients in Figure 3 is that unemployed individuals who receive UI are able to await a potential recall offer until their UI benefits are exhausted. Employers may also use the UI system during business downturns to attempt to retain their skilled work force, especially if UI is imperfectly experience-rated as much previous economic literature has indicated.

We now turn to estimation of the flexible parametric single risk model from section 2. We first estimate the single-risk model without heterogeneity from equation (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 re-estimate the model allowing for parametric heterogeneity using the one-parameter gamma function as in equation (8). In addition to 17 predetermined variables, we also estimate 40 weekly values of δ_t for the baseline hazard. Tests using different (right) censoring points for the baseline hazard demonstrated that the results are insensitive to the endpoint of the baseline hazard, as would be expected given its nonparametric specification. We find these results to be an important feature of our model specification.

The estimates are given in Table IV.[14] First consider the two middle columns of Table IV

[14] Note that the original hazard specification of equation (1) includes the covariates as $\exp(-X_i\beta)$ so the signs of all coefficients are reversed, e.g. a positive UI coefficient *lowers* the failure rate and therefore increases the duration of unemployment.

Table IV. Parameter estimates—single risk model

Variable (standard error)	Normal (probit)	Extreme value	Extreme w/ heterogeneity
Age	−0·012	−0·012	−0·022
	(0·003)	(0·003)	(0·007)
Sex	0·133	0·208	0·172
	(0·102)	(0·117)	(0·187)
Education	0·008	0·004	0·022
	(0·018)	(0·019)	(0·032)
Race	0·330	0·348	0·580
	(0·072)	(0·080)	(0·150)
Married	−0·113	−0·130	−0·204
	(0·086)	(0·098)	(0·157)
UI	0·029	0·026	0·055
	(0·013)	(0·014)	(0·025)
Craft	−0·046	−0·031	−0·128
	(0·093)	(0·096)	(0·170)
Clerical	−0·074	−0·110	−0·100
	(0·118)	(0·139)	(0·207)
Professional	−0·256	−0·373	−0·414
	(0·222)	(0·241)	(0.375)
Manager	−0·168	−0·236	−0·263
	(0·189)	(0·209)	(0·331)
Equipment	0·140	0·100	0·200
	(0·119)	(0·125)	(0·219)
Durables	0·168	0·109	0·299
	(0·114)	(0·120)	(0·212)
Trade	0·489	0·403	0·887
	(0·138)	(0·162)	(0·266)
Transportation	0·464	0·395	0·847
	(0·140)	(0·152)	(0·269)
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
Observations	1055		

which are the specifications which do not include heterogeneity. UI has the expected effect of leading to longer spells of unemployment. Females have longer unemployment spells as do non-whites. Married persons and older workers both have shorter spells of unemployment. Note that the effect of being married is about 4 times the size of the UI effect, while the effect of being non-white is about 10 times the UI effect. Thus, while UI is certainly an important determinant of unemployment duration, the sociodemographic variables have a considerably larger effect. The industry variables are mostly estimated precisely. Unemployment durations are longer in service, construction, retail, and manufacturing jobs to white-collar jobs. 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.

We now allow for heterogeneity in the last column of Table IV. 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 based on the log-likelihood values. The effect of UI benefits is again important and of about the same size as in the previous models when compared to sociodemographic variables such as age, race, and marriage (after rescaling the coefficients to take account of the increased variance). The industry variables have approximately the same estimated effect as in the models without heterogeneity.

We now consider whether the inclusion of heterogeneity has affected our previous results in an important manner. We infer that allowing for parametric heterogeneity has only a minor effect on the estimated hazard functions. These results are demonstrated graphically in Figure 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 V, where the cumulative distribution functions and asymptotic standard errors are presented from an average across the 1055 individuals using the parameter estimates from Table IV. 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

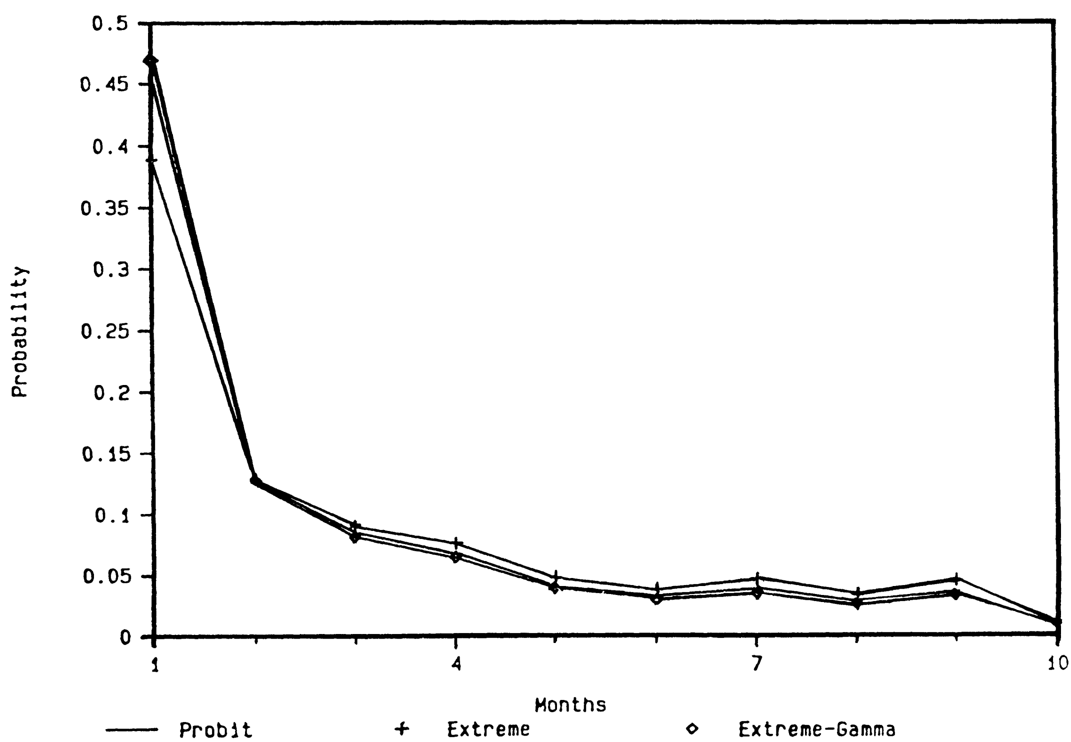

Figure 5. Estimated hazard rates for re-employment, single-risk, various models

the baseline hazard function. Both the estimated effects of the covariates and the estimated baseline hazard functions are very similar as Table IV and Figure 5 demonstrate. Table V demonstrates that when the covariate effects and baseline hazards are combined, the estimated models with and without heterogeneity yield almost identical results. Thus, when we consider

Table V. Cumulative distribution functions (CDF) for the extreme value and heterogeneity distributions average across complete sample: 1055 individuals

	Extreme value		Heterogeneity	
Weeks	CDF	Standard error	CDF	Standard 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.

Table VI. Competing risks: parameter estimates—UI interaction model

	New job		Recall	
Age	0·0099	(0·0037)	−0·0145	(0·0087)
Education	−0·0798	(0·0309)	0·0364	(0·0195)
Race	0·3588	(0·1243)	0·2661	(0·0801)
Dependents	0·0546	(0·0364)	−0·0023	(0·0265)
UI	0·2451	(0·1051)	0·0871	(0·0779)
Married	−0·1534	(0·1525)	−0·2646	(0·0938)
UI Interaction				
26 weeks	−1·924	(0·478)	−2·530	(0·479)
39 weeks	−3·555	(1·930)	−3·526	(1·131)
ρ	0·057	(1·18)		
Log LF	−3286·364			
Observations	1055			

Note: Asymptotic standard errors in parentheses.

competing risks specifications we do not attempt to include separate distributions for unobserved heterogeneity in our models.[15]

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 (15) and (16) where the distribution function is assumed to be joint standard normal. The normal assumption leads to a bivariate ordered probit model which can be viewed as a generalization of the univariate ordered probit model which provides a close approximation to the original ordered logit model arising from the proportional hazard specification of equation (1). Our model specification only provides a (close) approximation to equation (2), but it permits unrestricted correlation among the stochastic disturbances in equation (15). Straightforward generalizations of equation (2) to the bivariate case typically allow for only quite limited correlation in equation (15). Thus, the bivariate ordered probit model allows for flexible correlation among the risks, but does so at the potential cost of using an approximation to the proportional hazard specification. However, given the accumulated experience of the very similar results that ordered logit and ordered probit models yield in the binomial outcome case, the cost of the approximation should be quite small.[16]

Given the bivariate normal assumption, the only unknown parameter of the distribution function is ρ, the correlation coefficient. The log-likelihood function is given in equation (23) from which we do maximum-likelihood estimation. We specify the model in two different ways: the first specification again allows for 40 parameters for each of the two competing risks 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 Kaplan–Meier estimates demonstrated that the UI interaction variables are quite important.

[15] Inclusion of heterogeneity in the competing risks specification of equation (23) is straightforward, at least in principle. However, estimation would require numerical evaluation of integrals which would be quite time-consuming in practice.

[16] Of course, the similarity here between the logit and probit models arises because of the binomial nature of the outcomes. In the discrete choice model where numerous choices (outcomes) are present, probit and logit models can yield very different results, as previous empirical research has demonstrated.

The parameter estimates for the semiparametric competing risks model are given in Table VI. Note that the presence of UI has an important effect for both new jobs and for recalls. The effect of UI on new jobs durations is much more precisely estimated than the UI effect on recall durations. 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.

The effect of sociodemographic variables is again quite important in the competing risks model. Race and marital status also have important effects with the estimated directions as expected. Being married now has a much larger effect for recalls than it does for new jobs, especially relative to the effect of UI. Similarly, the effect of being non-white has an effect about 1·5 times larger than UI for new jobs exits, but an effect over 3 times larger than UI for recalls. Thus, the estimated effect of the covariates is quite different across new jobs and recalls. Lastly note that increasing age reduces unemployment duration for recalls, but increases unemployment duration for new jobs. Both effects are estimated relatively precisely. Again, note that the UI interaction effects at both 26 and 39 weeks are estimated to be quite large and significant. Thus, we once again conclude that the effect of UI exhaustion has important effects on exit probabilities from unemployment.[17]

We next consider the question of the importance of Katz's (1986) assumption of independence of the hazards and the assumption of a one-parameter Weibull specification for the baseline hazard. The estimated ρ is 0·057 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. Table VII presents 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. We graph the baseline hazard estimates from the estimates of the semiparametric hazard model in Figure 6 for individuals who receive UI, and Figure 7 for individuals who do not receive UI. 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. But even for individuals who do not receive UI, the new job hazard which is graphed in Figure 7 is essentially flat with a decrease around 7 months which is then followed by an increasing hazard. Similarly, the recall hazard has periods of both increases and decreases in Figure 7.

In contrast to the flexible shape of the hazard function permitted by our specification, 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 flexible parametric 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 behaviour.

We now turn to a formal test of the Weibull specification since the graphical evidence, while suggestive, is not definitive. 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 flexible parametric hazard estimates from the bivariate

[17] We have also estimated the competing risks model with the industry dummy variables that we used in Table IV for the single-hazard model. The industry dummy variables are jointly significant, but the results on the other variables are very similar. Thus, we present the results from the simpler specification.

Table VII. Estimated cumulative distribution functions

	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

ordered probit specification as $\hat{\delta}^1$ and $\hat{\delta}^2$. These parameter estimates are the basis for the hazard functions which are graphed in Figures 6 and 7, and from which the cumulated functions are calculated in Table VII. From our previous theorems we know that

$$\begin{bmatrix} \hat{\delta}^1 \\ \hat{\delta}^2 \end{bmatrix}^A \sim N(\delta, \Omega), \tag{25}$$

where Ω corresponds to the lower right-hand block of the inverse of the Fisher information

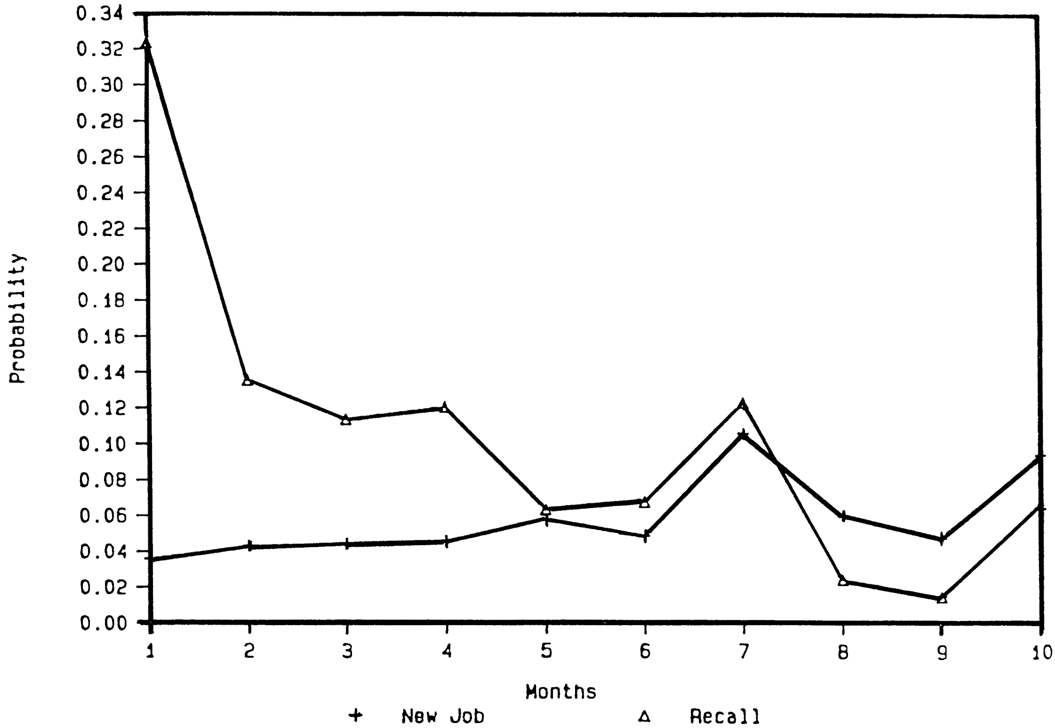

Figure 6. Estimated hazard rates for re-employment, dual-risk model—individuals receive UI

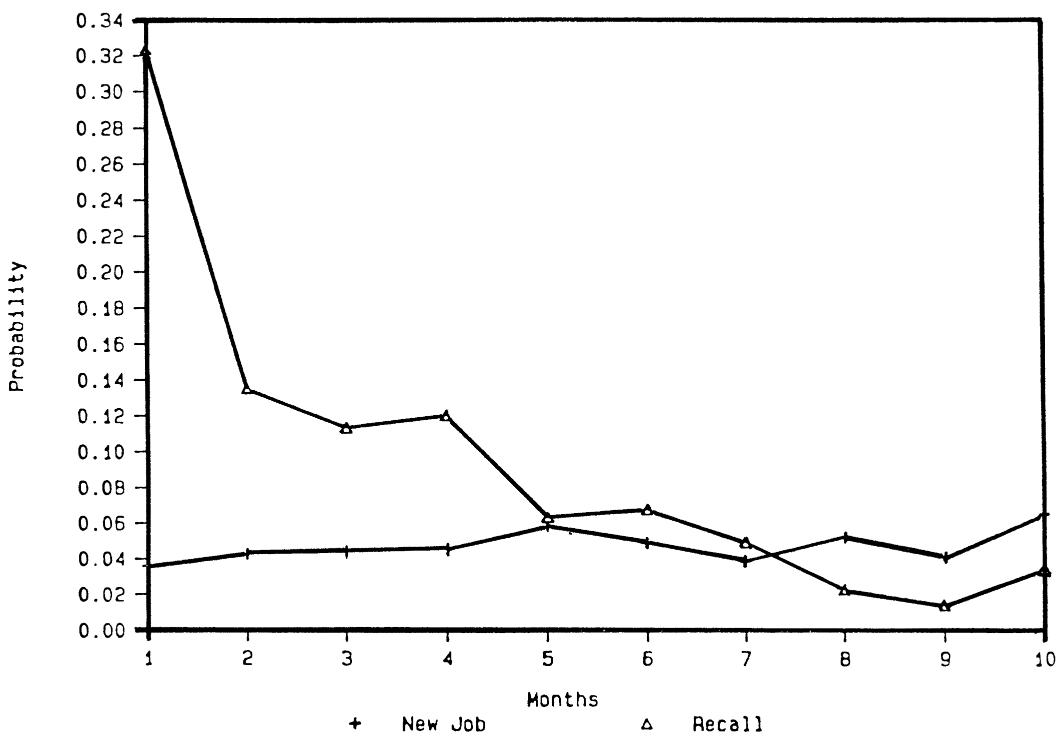

Figure 7. Estimated hazard rates for re-employment, dual-risk model—individuals do not receive UI

matrix. The Weibull specification can be written

$$\delta^j = \alpha_{j1} t^{\alpha_{j2}-1} = g_j(\alpha_j) \text{ for } \alpha_j > 0 \text{ for } j = 1, 2. \tag{26}$$

To estimate the unknown α_j we use minimum chi square estimation

$$\min_{\alpha_1, \alpha_2} W = \begin{bmatrix} \hat{\delta}^1 - g_1(\alpha_1) \\ \hat{\delta}^2 - g_2(\alpha_2) \end{bmatrix}' \hat{\Omega} - 1 \begin{bmatrix} \hat{\delta}^1 - g_1(\alpha_1) \\ \hat{\delta}_2 - g_2(\alpha_2) \end{bmatrix} \tag{27}$$

Table VIII. Weibull Specification Test

	New job		Recall	
Non-UI recipients				
α_1	0·007	(0·005)	0·077	(0·007)
α_2	1·253	(0·251)	0·596	(0·051)
		$\chi^2 = 98\cdot954$		
UI recipients				
α_1	0·005	(0·005)	0·076	(0·010)
α_2	1·446	(0·345)	0·645	(0·068)
		$\chi^2 = 117\cdot620$		

Note: Asymptotic standard errors are in parentheses.

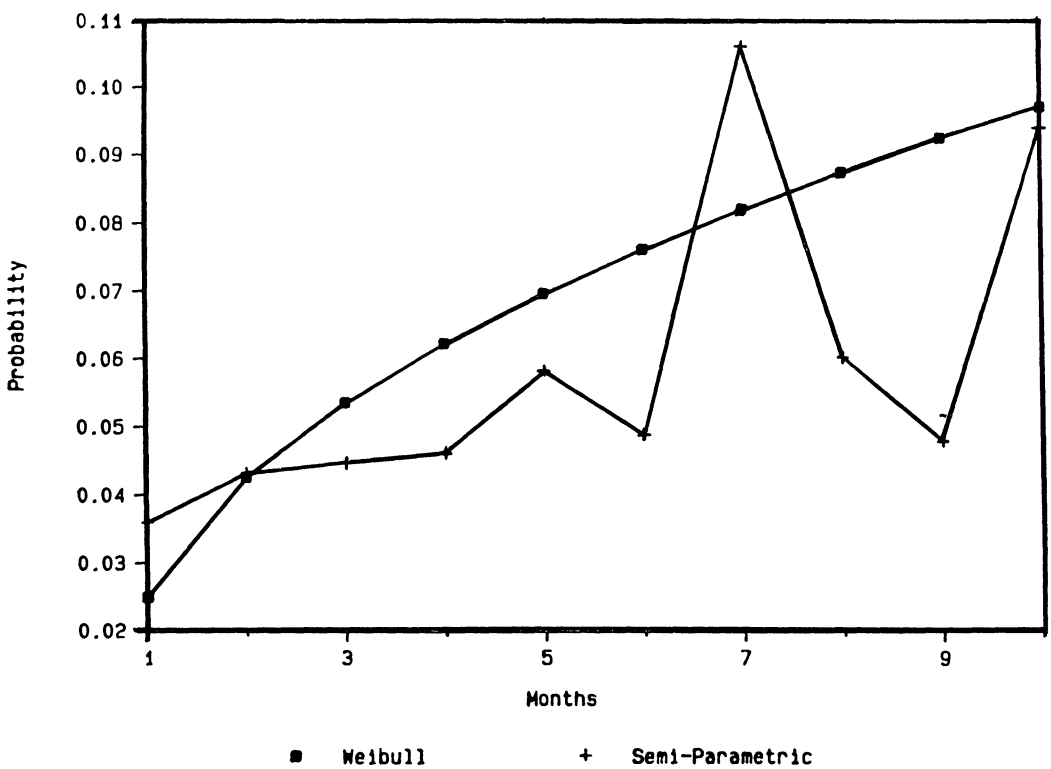

Figure 8. Comparison of semiparametric baseline hazard and fitted two-parameter Weibull hazard for new job duration—individuals receive UI

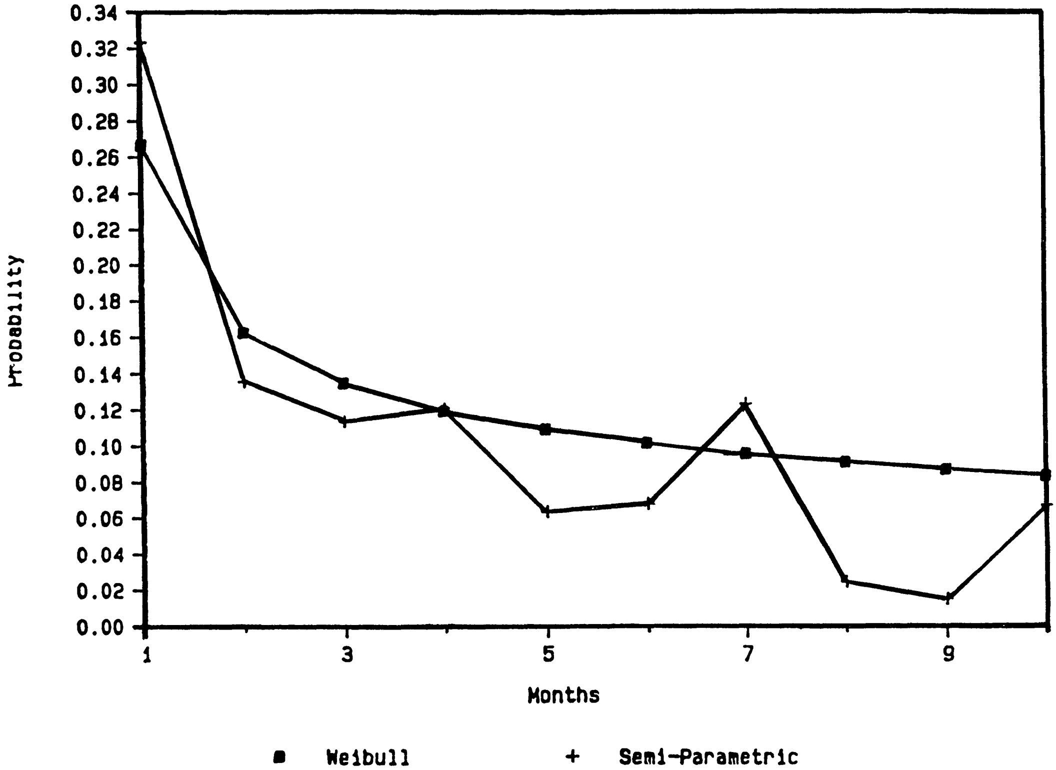

Figure 9. Comparison of semiparametric baseline hazard and fitted two-parameter Weibull hazard for recall duration—individuals receive UI

Note that the resulting estimates are asymptotically efficient. That is, the minimum χ^2 estimates have the same asymptotic distribution as maximum-likelihood estimates of the restricted models; see Ferguson (1958) and Rothenberg (1973, Ch. 2). The estimated α_j for the UI and non UI samples are given in Table VIII. The value of W is then distributed under the null hypothesis of the Weibull specification as asymptotic 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. Figures 8 and 9 show graphs of the estimated semiparametric baseline hazards together with the Weibull estimates of the baseline hazards.[18]

We recommend this general approach to estimation and testing of baseline hazard functions. Flexible parametric estimation puts no restrictions on the shape of the underlying hazard function in discrete data. Given the nonparametric hazard estimates, the econometrician can then test any particular functional hazard specification given our approach. Difficult analysis of residuals is eliminated since the flexible parametric 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

[18] 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.

different baseline hazard specification. Tests of duration dependence are thus easy to carry out without undue parametric restrictions.

6. CONCLUSIONS

Our approach begins with the Cox proportional hazards model which has not 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 are 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 once a flexible parameter specification of the baseline hazard is used. Attention to unobserved heterogeneity may be less important when nonparametric hazard specifications are used.

We then extend the flexible parametric 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 specify a flexible parametric 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 labour market behaviour 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 effect. These findings point to the need for more realistic models of labour market behaviour for the unemployed.

ACKNOWLEDGEMENTS

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

REFERENCES

Bailey, K. (1984), 'Asymptotic equivalence between the Cox estimator and the general maximum likelihood estimator of regression and survival parameters', *Annals of Statistics*, **12**, 730–736.

Burtless, G. (1983), 'Why is insured unemployment so low', *Brookings Papers on Economic Activity*, **1**, 225–253.

Clayton, D., and J. Cuzick (1985), 'Multivariate generalizations of the proportional hazards model', *Journal of Royal Statistical Society*, *A*, **148**, 82–117.

Cox, D. R. (1959), 'The analysis of exponentially distributed life-times with two types of failures', *Journal of the Royal Statistical Society*, *B*, **21**, 411–421.

Cox, D. R. (1972), 'Regression models and life-tables', *Journal of Royal Statistical Society*, *B*, **26**, 186–220.

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

Diamond, P. A., and J. A. Hausman (1984), 'The retirement and unemployment behavior of older men', in H. Aaron and G. Burtless (eds), *Retirement and Economic Behavior*, Brookings Institute, Washington, DC.

Ferguson, T. S. (1958), 'A method of generating best asymptotically normal estimates with applications to the estimation of bacterial densities', *Annals of Mathematical Statistics*, **29**, 1046–1062.

Han, A., and J. A. Hausman (1986), 'Semiparametric estimator of duration and competing risk models', MIT Working Paper.

Han A., and J. A. Hausman (1988), 'Identification of continuous and discrete competing risk models', mimeo.

Heckman, J., and B. Singer (1984a), 'Econometric duration analysis', *Journal of Econometrics*, **24**, 63–132.

Heckman, J., and B. Singer (1984b), 'A method for minimizing the impact of distributional assumptions in econometric models for duration data', *Econometrica*, **52**, 271–320.

Heckman, J., and B. Singer (1986), 'Econometric analysis of longitudinal data', in Z. Grilichis and M. Intriligator (eds), *The Handbook of Econometrics*, Vol. 3, North Holland, Amsterdam.

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

Katz, L. (1986), 'Layoffs, recall and the duration of unemployment', Working Paper No. 1825, National Bureau of Economic Research, Cambridge, Massachusetts.

Kiefer, N. (1988), 'Economic duration data and hazard functions', *Journal of Economic Literature*, **26**, 646–679.

Lancaster, T. (1979), 'Econometric methods for the duration of unemployment', *Econometrica*, **47**, 939–956.

Manton, K. G., E. Stallard, and J. W. Vaupel (1986), 'Alternative models for the heterogeneity of mortality risks among the aged', *Journal of American Statistical Society*, **81**, 635–644.

Meyer, B. (1987), 'Semiparametric estimation of duration models', MIT Ph.D thesis.

Moffit, R. (1985), 'Unemployment insurance and the distribution of unemployment spells', *Journal of Econometrics*, **28**, 85–101.

Prentice, R. L. (1976), 'A generalization of the probit and logit models for dose response curves', *Biometrics*, **32**, 761–768.

Rothenberg, T. J. (1973), *Efficient Estimation with A Priori Information*, Yale University Press, New Haven.

Sueyoshi, G. (1987), 'Competing risks model of economic behavior', MIT Ph.D thesis.

Tsiatis, A. A. (1975), 'A nonidentifiability aspect of the problem of competing risks', *Proceedings of the National Academy of Sciences, USA*, **72**, 20–22.