SURVIVAL ANALYSIS FOR THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

I-li Lu and Lynn Weidman, Bureau of the Census
I-li Lu, Bureau of the Census, Washington, DC 20233

1. Introduction

Denote by T the spell duration of an individual participating in a government benefits program. T measures the length of time that an individual takes to exit the program. There are a variety of methods one can use to gain an understanding of the behavior of T. To investigate its behavior, we use data extracted from the 1987 panel of the Survey of Income and Program Participation (SIPP).

SIPP is a longitudinal panel survey conducted by the Census Bureau and designed to provide data on income distribution and poverty at the national level (see Nelson, et al, 1984, for an overview of the SIPP). The data collected in SIPP are often employed in the study of cost and effectiveness of Federal programs. Policy makers also entertain the data to evaluate the policies that motivate household independence form the welfare programs. Knowledge of the distribution of T permits us to make direct inference on the estimated cost of a relevant program. The dynamic behavior of T and the extent to which household characteristics affect the distribution of T, are crucial for our understanding of the effects of proposed changes in program regulations and benefit levels.

While duration data extracted form the SIPP have been used frequently to achieve the above purposes (Bane and Welsh, 1985; Fields and Jakubson, 1985; Short, 1985, 1992; Ross, 1988), practitioners often find that they suffer some inherent problems. Here are some of the complications normally being recognized in the analysis of spell durations when the spell data are extracted from the longitudinal panel survey: (1) multiple occurrences, (2) random left truncation, (3) random right censorship, (4) random number of repeated occurrences, (5) dependencies of truncation of censoring mechanisms of the spell durations.

There are a variety of methods proposed and developed to circumvent these problems in different contexts. Allison (1982, 1984) proposed solutions for (1) and (3) in the analysis of event histories; Flinn and Heckman (1982) developed an econometric model-specific solution for (2); Cox and Oakes (1984) developed an analysis for (3) by treating censoring variables as predetermined constants; Turnbull (1974), Tuma and Hannan (1982), and recently, Sun (1992) developed approaches to censored and interval-truncated data. All approaches generally lead to either biases or suboptimal uses of the information in the estimation process. Rarely has the analysis been conducted in the way that it is general enough to not only optimize the use of information, but also to encompass all of the problems above.

Therefore, as the major objective of this paper, we attempt to develop a procedure that will accomplish this task. We construct a general procedure that is applicable to any data as long as they have similar characteristics to those observed in a longitudinal panel survey. The procedure makes use of the EM algorithm (Dempster, Laird and Rubin, 1977) and iteratively maximizes the appropriate likelihood function if it cannot be determined explicitly.

Parametric models that capture the effects of possible time varying factors are well developed for lifetime data (see Fleming and Harrington, 1991; Cox and Oakes, 1984; and Lawless, 1982). We classify them in three categories: accelerated failure time models (Kalbfleisch and Prentice, 1980; Lawless, 1982), proportional hazards models (Cox, 1972), and Markov models (Tuma, 1976; Tuma and Hannan, 1979). We focus our attention on the accelerated failure time models and define a generalized accelerated failure time model for the multivariate spell duration. We then develop a parametric estimation procedure that treats multiple spells as spell vectors having joint multivariate distribution when dependencies among the spells cannot be ignored.

2. Parametric Models for Survival Analysis

To start our discussion, we assume that there are n independent sampling units and for each sampling unit i we observe m_i spells, with the last spell possibly censored. We also assume all spells begin at some observed starting point t=1 except for the first spell which might be truncated from the left.

Let the hazard function $h(t) = \Pr\{T = t | T \geq t\}$ be the probability that a unit exits the program at time t when in fact this unit is still at risk of an exit at time t. Set $\mathbf{t}_i = (t_{i1},...,t_{im})'$ to be a vector of times and define $\mathrm{T_i} = (T_{i1}, T_{i2},...,T_{im_i})'$ as the spell length vector associated with unit i and write its joint distribution function as

$$\mathrm{F_i(t_i)} = \Pr(T_{i1} \leq t_{i1}, T_{i2} \leq t_{i2},..., T_{im_i}). \tag{2.0}$$

Accelerated Failure Time

Let $\mathbf{T}_0 = (T_{01},...,T_{0m})'$ denote the vector of spell durations with baseline distribution corresponding to zero value covariates. Also let

$$\mathbf{X} = \begin{pmatrix} \mathbf{X_1} & 0 & \cdots & 0 \\ 0 & \mathbf{X_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \mathbf{X_m} \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \beta_m \end{pmatrix}$$

be the matrices of covariates and parameters respectively, where for $1 \leq j \leq m, \mathbf{X}_j = (X_{j1},...,X_{jp})'$, are covariate vectors associated with the jth spell and $\boldsymbol{ß}_j = (ß_{j1},...,ß_{jp})'$ are parameter vectors to be determined.

The following definition is a generalization of univariate accelerated failure time models (Cox and Oakes, 1984; and Lawless, 1982).

Definition 1. *A model is said to exhibit generalized accelerated failure times if the existence of a nonzero matrix of covariates* $\mathbf{X} = (\mathbf{X}_1,...,\mathbf{X}_m)$ *implies that the spell duration vector is* $\mathbf{T} = \Lambda\mathbf{T}_0$, *where* $\Lambda = \exp(\mathbf{B}'\mathbf{X})$ *is the* m x m *diagonal matrix with jth diagonal element*

λ_j = exp ($\mathbf{ß}'_j\mathbf{X}_j$).

Now define, for each sampling unit i = 1,2,...,n,

	δ_{il}	= 1	if the first spell is not left-truncated
		= 0	otherwise,
and	δ_{ir}	= 1	if the last spell is not right-censored
		= 0	otherwise,

To begin our discussion, we will assume that for each sampling unit, a repeated spell is completely described by the characteristics associated with that unit within the duration of the spell. We use the following assumption.

Assumption 1. *For each sampling unit i, conditionary on the explanatory variables,* $\mathbf{X}_{ij}$ *,* j = 1,...,m_i*, the spells are independent. In other words,* T_{ij} *is independent of* T_{ij} *, for* $j \neq j'$ *when* $\mathbf{X}_{ij}$ *, and* $\mathbf{X}_{ij}$ *, are known.*

The assumption is not unreasonable, as long as we have in our data, a large number of explanatory variables to capture most of the behavior differences among spells. When the assumption is true, we may drop the subscript for the β and write **B** as

$$\mathbf{B} = \begin{pmatrix} \beta & 0 & \cdots & 0 \\ 0 & \beta & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \beta \end{pmatrix}.$$

If a unit i already participated in a program before the survey starts, we will denote the truncation time T_{i0} as the length of time between the actual program entry time and the time the survey began. And, throughout, for i = 1, 2,...,n, we denote T_{i1}^* as the left-truncated spells and $T_{iM_i} = \min(T_{iM_i}, t_{im_i})$ as the right-censored spells, which is censored by a random amount of time t_{im_i}. We will also use the following notation for the spell vectors.

$$\mathbf{T_i^*} = (T_{i1}^*, T_{i2}, ..., T_{iM_i})'$$

$$\mathbf{T_i^0} = (T_{i0} + T_{i1}^*, T_{i2}, ..., T_{iM_i}^*)'$$

$$\mathbf{T_{0i}^*} = (T_{i0}, \mathbf{T_i^{*\prime}})'$$

$$\mathbf{T_n^0} = (T_{10}, T_{20}, ..., T_{n0})'$$

$$\delta_i = (\delta_{il}, \delta_{ir})'$$

3. Estimating the General Model

Following the notation in section 2, if a unit already participated in a program before the survey starts, and if T_{i1}^* is the spell observed to be truncated, we have

$$\Pr(T_{i1}^{*} = t_{i1} | T_{i0} = t_{i0}) = \frac{\Pr(T_{i1} = t_{i0} + t_{i1})}{S_{i1}(t_{i0})}, \tag{3.1}$$

where S_{i1} is the survival function for T_{i1}.

For ease of notation, we will replace the probability function by the joint density function and write $f^{*}(\mathbf{T}_i^{*}, \delta_i, M_i, | ß, \mathbf{X})$ as the joint density of the observed data. The following results are applicable to both discrete and continuous variables.

Closely following the development of Little and Rubin (1987) in the analysis of missing data, and assuming all the measure theoretical difficulties are not present, we decompose the conditional joint density of $(\mathbf{T}_{0i}^{*}, \delta_i, M_i)$ as

$$f(\mathbf{T}_{0i}^{*}, \delta_i, M_i | \beta, \mathbf{X}) =$$

$$f_l(\delta_{il} | \mathbf{T}_{0i}^{*}, \delta_{ir}, M_i, \theta_l) f_{\mathrm{T}}(\mathbf{T}_{0i}^{*}, \delta_{ir}, M_i | \beta, \mathbf{X}),$$

where we use θ_l as the parameter for δ_{il} , Therefore,

$$f^{*}(\mathbf{T}_{i}^{*}, M_i | \beta, \mathbf{X}) =$$

$$\int f_l(\delta_{il} | \mathbf{T}_{0i}^{*}, \delta_{ir}, , M_i, \theta_l) f_T(\mathbf{T}_{0i}^{*}, \delta_{ir}, M_i | \beta, \mathbf{X})\, dT_{i0}$$

Now, we use the following assumption.

Assumption 2. *The truncation mechanism does not depend on the initial spell truncation time.*

Clearly, when Assumption 2 is satisfied, above equation becomes

$$f^{*}(\mathbf{T}_{i}^{*}, \delta_i, M_i | \beta, \mathbf{X}) =$$

$$f_l(\delta_{il} | \mathbf{T}_{0i}^{*}, \delta_{ir}, , M_i, \theta_l) \int f_T(\mathbf{T}_{0i}^{*}, \delta_{ir}, M_i | \beta, \mathbf{X})\, dT_{i0}$$

We now use the notation $\theta \perp \lambda$ to mean that θ is independent of λ. Therefore, if $\theta_l \perp ß$, the estimation problem for ß is simply reduced to the maximization problem for the likelihood function deduced from the joint density function $\prod_{i=1}^{n} \int f_T(\mathbf{T}_{0i}^{*}, \delta_{ir}, M_i | \beta, \mathbf{X}) dT_{i0}$

To simplify the problem further, we apply Bayes theorem again, for a random sample of size n,

$$\prod_{i=1}^{n} \int f_T(\mathbf{T}_{0i}^{*}, \delta_{ir}, M_i | \beta, \mathbf{X}) dT_{i0} =$$

$$\prod_{i=1}^{n} \int f_M(M_i | \mathbf{T}_{0i}^{*}, \delta_{ir}, \theta_M) f_T(\mathbf{T}_{0i}^{*}, \delta_{ir} | \beta, \mathbf{X}) dT_{i0}$$

Again, we use the following mild assumption.

Assumption 3. *For i=1,...,n, the number of spells M_i are independent of the initial spell truncation times T_{i0}.*

Denote the M_i x 1 unit vector by $\mathbf{1}_i$. Then for i = 1,...,n, the M_i are inevitably dependent upon $\mathbf{1}_i^{\prime}\mathbf{T}_{\mathbf{i}}^{*}$, and are restricted by the length of the panel. Therefore, this assumption would be violated if the T_{i1}^{*} is not independent of t_{i0}. And this is certainly true by equation (3.1). For illustration, we will assume that the Assumption 3 is true. We have

$$\prod_{i=1}^{n}\int f_T(\mathbf{T}_{\mathbf{0i}}^{*},\delta_{ir},M_i\,|\,\beta,\mathbf{X})dT_{i0} =$$

$$\prod_{i=1}^{n} f_M(M_i\,|\,\mathbf{T}_{\mathbf{0i}}^{*},\delta_{ir},\theta_M)\int f_T(\mathbf{T}_{\mathbf{0i}}^{*},\delta_{ir}\,|\,\beta,\mathbf{X})dT_{i0}.$$

Further, suppose that $\theta_M \perp$ ß, then equivalently, it becomes sufficient for us to maximize the likelihood function deduced from

$$\prod_{i=1}^{n}\int f_T(\mathbf{T}_{\mathbf{0i}}^{*},\delta_{ir},\,|\,\beta,\mathbf{X})dT_{i0}.$$

Denote g($\cdot$ | ß,$\mathbf{X}$) as the joint density function for the vector $(T_{i1},...,T_{iM_i}^{*},\delta_{ir})^{\prime}$. Then applying Bayes theorem again, we obtain

$$f_T(\mathbf{T}_{\mathbf{0i}}^{*},\delta_{ir}\,|\,\beta,\mathbf{X})$$

$$= f_T(\mathbf{T}_{\mathbf{i}}^{*},\delta_{ir}\,|\,T_{i0},\beta,\mathbf{X})f_0(T_{i0}\,|\,\beta,\mathbf{X})$$

$$= \frac{g(\mathbf{T}_{\mathbf{i}}^{\mathbf{0}},\delta_{ir}\,|\,\beta,\mathbf{X})}{S_{i1}(T_{i0}|\beta,\mathrm{X})}\, f_0(T_{i0}\,|\,\beta,\mathbf{X})$$

Where the second equality follows from equation (3.1),

$$S_{i1}(T_{i0}\,|\,\beta,\mathbf{X}) = \int_{T_{i0}}^{\infty} g_1(t\,|\,\beta,\mathbf{X})dt,$$

with

$$g_1(t\,|\,\beta,\mathbf{X}) = \int g(\mathbf{t}_{\mathbf{i}}^{\mathbf{0}},\delta_{ir}\,|\,\beta,\mathbf{X})\,dt_{i2}...dt_{iM_i}d\delta_{ir},$$

the marginal density function of T_{i1}.

Finally, we define the likelihood function deduced from this reduced model to be

$$L(\beta|\mathbf{X},\mathbf{T}^*,\delta_{ir}) \propto \int \prod_{i=1}^{n} \frac{g(\mathbf{T_i^0},\delta_{ir}|\beta,\mathbf{X})}{S_{i1}(T_{i0}|\beta,\mathbf{X})} f_0(T_{i0}|\beta,\mathbf{X})\, d\mathbf{T_0}. \tag{3.2}$$

We want to maximize (3.2) with respect to ß. That is, we want to find a solution for $\max_{ß} L(ß|X,T^*)$. In general, this is very difficult to do directly when the expressions for g, f_0, and S_{i1} are complex. To present a general solution for this problem, we will use the EM algorithm (Dempster, Laird, and Rubin, 1977) indirectly at each step to maximize the conditional expected log likelihood function with the augmented data matrix that includes as its arguments the truncation times $\mathbf{T}_n^0$ and the observed spell matrix $\mathbf{T}^*$.

Let $\mathbf{T} = (\mathbf{T}_{0i}^{*\prime})$, i=1,...,n be the augmented data matrix that includes the truncation times. Define the augmented data log likelihood function to be

$$ll(\beta|\mathbf{X},\mathbf{T}) \propto \log \prod_{i=1}^{n} \frac{g(\mathbf{T_i^0},\delta_{ir}|\beta,\mathbf{X})}{S_{i1}(T_{i0}|\beta,\mathbf{X})} f_0(T_{i0}|\beta,\mathbf{X}). \tag{3.3}$$

Given the current estimate $\hat{\beta}_{(j)}$ for ß , we want to find $\hat{\beta}_{(j-1)}$ which is the solution of

$$\max_{\beta} E\left\{ ll(\beta|\mathbf{T},\mathbf{X})\,|\,\mathbf{T}^*,\hat{\beta}_{(j)},\mathbf{X}\right\},$$

that is, by (3.3), at each step, to solve

$$\max_{\beta} E\left\{ \sum_{i=1}^{n} \log\left[\frac{g(\mathbf{T_i^0},\delta_{ir}|\beta,\mathbf{X}) f_0(\mathrm{T}_{i0}|\beta,\mathbf{X})}{S_{i1}(T_{i0}|\beta,\mathbf{X})}\right] \Big| \mathrm{T}^*,\hat{\beta}_{(j)},\mathbf{X}\right\}. \tag{3.4}$$

Now equation (3.4) is the objective function when Assumption 3′ is satisfied. If this assumption is not true, then (3.4) becomes

$$\max_{\beta} E\left\{ \sum_{i=1}^{n} \log\left[\frac{g(\mathbf{T_i^0},\delta_{ir}|\beta,\mathbf{X}) f_0(\mathrm{T}_{i0}|\beta,\mathbf{X}) f_{\mathrm{M}}(M_i|\mathbf{T^*_{0i}},\delta_{ir},\theta_M)}{S_{i1}(T_{i0}|\beta,\mathbf{X})}\right] \Big| \mathbf{T}^*,\hat{\beta}_{(j)},\mathbf{X}\right\}.$$

Clearly, the task now becomes how to specify the density functions (probability functions, in the discrete case) for g, f_0, and possibly f_M. But by Assumption 1, we can write g as a product of marginal density functions. Therefore, since the objective function involves S_{i1}, i.e. the survival function of T_{i1}, some obvious candidates are those distributions that have explicit expressions for the survival function. Therefore, Weibull, Gompertz-Makeham, compound exponential, orthogonal polynomial, and log logistic distributions are certainly the classes of distributions that could be entertained as the baseline distributions for the marginal spells. We illustrate in examples with Weibull and log logistic distributions. In the last example we use a distribution with U-shaped hazard function.

Example 1. Suppose we want to fit the accelerated lifetime model with a baseline Weibull distribution, which has the following density, hazard rate, and survival functions respectively,

$$f(t) = kt^{k-1}\exp(-t^k), \quad h(t) = kt^{k-1}, \quad S(t) = \exp(-t^k).$$

Assuming T_{i0} has the same distribution as T_{i1} then

$$\log f(t_{i0}, t_{i1}^{*}) = \log\left[\frac{g(t_{i0}+t_{i1}^{*} \mid \beta' \mathbf{X}_{\mathbf{i1}}) f_0(t_{i0} \mid \beta' \mathbf{X}_{\mathbf{i1}})}{S_{i1}(t_{i0} \mid \beta' \mathbf{X}_{\mathbf{i1}})}\right] \quad = 2\log k - 2k\beta' \mathbf{X}_{\mathbf{i1}} + (k-1)[\log(t_{i0}+t_{i1}^{*}) + \log t_{i0}] - \exp(-k\beta' \mathbf{X}_{\mathbf{i1}})(t_{i0}+t_{i1}^{*})^{k}.$$

Therefore, the step of finding conditional expectation for the log likelihood function becomes to find $\mathrm{E}\left[(t_{i0}+t_{i1}^{*})^{k} \mid \mathbf{t}_{\mathbf{i1}}^{*}, \hat{\beta}_{(j)}' \mathbf{X}_{\mathbf{i1}}\right]$, and then substitute it for the value of $(t_{i0}+t_{i1}^{*})^{k}$.

Example 2. For a baseline log logistic distribution, which has the following density, hazard rate, and survival functions respectively,

$$f(t) = \frac{kt^{k-1}}{(1+t^{k})^2}, \quad h(t) = \frac{kt^{k-1}}{1+t^{k}}, \quad S(t) = \frac{1}{1+t^{k}}.$$

Assuming T_{i0} has the same distribution as T_{i1}, then

$$\log f(t_{i0}, t_{i1}^{*}) = \log\left[\frac{g(t_{i0}+t_{i1}^{*} \mid \beta' \mathbf{X}_{\mathbf{i1}}) f_0(t_{i0} \mid \beta' \mathbf{X}_{\mathbf{i1}})}{S_{i1}(t_{i0} \mid \beta' \mathbf{X}_{\mathbf{i1}})}\right]$$

$$= 2\log k + k\beta' \mathbf{X}_{\mathbf{i1}} + (k-1)[\log(t_{i0}+t_{i1}^{*}) + \log t_{i0}] \quad - 2\log\left\{\exp(k\beta' \mathbf{X}_{\mathbf{i1}}) + (t_{i0}+t_{i1}^{*})^{k}\right\} - \log\left\{\exp(k\beta' \mathbf{X}_{\mathbf{i1}}) + t_{i0}^{k}\right\}$$

And we need to find

$$\mathrm{E}\left[\log\left\{\exp(k\beta' \mathbf{X}_{\mathbf{i1}}) + (t_{i0}+t_{i1}^{*})^{k}\right\} \mid t_{i1}^{*}, \hat{\beta}_{(j)}' \mathbf{X}_{\mathbf{i1}}\right]$$

and

$$\mathrm{E}\left[\log\left\{\exp(k\beta' \mathbf{X}_{\mathbf{i1}}) + t_{i0}^{k}\right\} \mid t_{i1}^{*}, \hat{\beta}_{(j)}' \mathbf{X}_{\mathbf{i1}}\right]$$

to replace $\log\{\exp(k\boldsymbol{\beta}'\mathrm{X}_{i1})+(t_{i0}+t_{i1}^{*})^{k}\}$ and $\log\left\{\exp(k\beta' \mathbf{X}_{\mathbf{i1}}) + t_{i0}^{k}\right\}$ respectively.

Example 3. For a U-shaped hazard function, we use

$$f(t) = kt^{k-1}\exp(1+t^{k}-\exp(t^{k})), \quad h(t) = kt^{k-1}\exp(1-\exp(t^{k})), \quad S(t) = \exp(1-\exp(t^{k})).$$

Assuming T_{i0} has the same distribution as T_{i1}, then

$$\log f(t_{i0}, t_{i1}^{*}) = \log\left[\frac{g(t_{i0}+t_{i1}^{*} \mid \beta' \mathbf{X}_{\mathbf{i1}}) f_0(t_{i0} \mid \beta' \mathbf{X}_{\mathbf{i1}})}{S_{i1}(t_{i0} \mid \beta' \mathbf{X}_{\mathbf{i1}})}\right] \quad = 1 + 2\log k + 2k\beta' \mathbf{X}_{\mathbf{i1}} + \exp(k\beta' \mathbf{X}_{\mathbf{i1}}) t_{i0}^{k} + \exp(k\beta' \mathbf{X}_{\mathbf{i1}})(t_{i0}+t_{i1}^{*})^{k}$$

$$- \exp\left\{(t_{i0}+t_{i1}^{*})^{k} \exp(k\beta' \mathbf{X}_{\mathbf{i1}})\right\}$$

Here we need to find

$$\mathrm{E}\left[(t_{i0}+t_{i1}^{*})^{k} \mid t_{i1}^{*}, \hat{\beta}_{(j)}' \mathbf{X}_{\mathbf{i1}}\right],$$

$$\mathrm{E}\left[t_{i0}^{k} \mid t_{i1}^{*}, \hat{\beta}_{(j)}' \mathbf{X}_{\mathbf{i1}}\right],$$

and

$$E\left[\exp\left\{(t_{i0}+t_{i1}^{*})^{k}\exp(k\beta'\mathbf{X_{i1}})\right\}\mid t_{i1}^{*},\hat{\beta}_{(j)}'\mathbf{X_{i1}}\right]$$

then substitutes them for the values of $(t_{i0}+t_{i1}^{*})^{k}$, t_{i0}^{k}, and $\exp\left\{(t_{i0}+t_{i1}^{*})^{k}\exp(k\beta'\mathbf{X_{i1}})\right\}$ respectively.

4. Concluding Remark

In this paper, we developed and summarized a parametric estimation procedure for the distributions of spell durations extracted from the SIPP data set. We presented an explicit solution for handling the left-truncated spells that normally are treated as nonexistent of are discarded in the analysis of spell durations. Also, the number of spells for each sampling unit can also be assumed to be stochastic and depend on the length of spells for each unit.

In the sequel of this paper, we will complete the task of estimation for the models we developed. First, we will simulate the data with different distributional assumptions and check for the consistency of the models. Then, we will apply these models to the actual data to see if there is empirical support for our models. Finally, we will use topical module or recipient history data to evaluate the accuracy of the models. We also will like to extend our ideas into the proportional hazard and the markov models.

Allison, P.D. (1982), "Discrete-Time Methods for the Analysis of Event Histories", *Sociological Methodology*, San Francisco: Jossey-Bass, 61-98.

Allison, P.D. (1984), *Event History Analysis: Regression for Longitudinal Event Data*, Beverly Hills, CA: Sage University Press.

Bane, M.J., Welsh, J. (1985), "SIPP's potential Contributions to Policy Research on Children", *Journal of Economic and Social Measurement*, **v13, #3 & 4**, 273-279.

Cox, D.R. (1972), "Regression Models and Life-Tables", *J.R. Statist. Soc*., **B, 34**, 187-220.

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

Dempster, A.P., Laird, N.M., and Rubin D.B. (1977), "Maximum Likelihood for Incomplete Data Via the EM Algorithm", *J.R. Statist. Soc*., **B, 39**, 1-38.

Fields, G.S., and Jakubson, G.H. (1985), "Labor Market Analysis Using SIPP", *Journal of Economic and Social Measurement*, **v13, #3 & 4**, 281-286.

Fleming, T.R., and Harrington, D.P. (1991), *Counting Processes and Survival Analysis*, New York: John Wiley.

Flinn, C.J., Kleckman, J.J. (1982), "Models for the Analysis of Labor Force Dynamics", *Advances in Econometrics*. **v1**, 35-95.

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

Lawless, J.F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: John Wiley.

Little, R.J.A., and Rubin, D.B. (1987), *Statistical Analysis with Missing Data*, New York: John Wiley.

Nelson, D., McMillen, D.B., and Kaprzyck, D. (1985), "An Overview of the Survey of Income and Program Participation: Update 1", *SIPP Working Paper Series No. 8401,* Washington, D.C.: U.S. Bureau of the Census.

Ross, C.M. (1988), "The Role of Employment in Exits from the Food Stamp Program", *Proceedings of the Soc. Stat. Sec., ASA.*

Short, K., Woodrow, K. (1985), "Exploration of the Application of Hazard Model in Analyzing the SIPP: Labor Force Transition", *ASA Meeting*, Nov. 1985.

Short, K., Shea, M. (1992), "Spell Durations of Food Stamp and AFDC Participation", *Proceedings of the Soc. Stat. Sec., ASA*, to appear.

Sun, J. (1992), "Empirical Estimation of Distribution Function with Truncated and Doubly Interval-Censored Data and its Application to AIDS Studies", *Contributed papers given at the 55th IMS Annual Meeting*, Ontario, Canada: University of Waterloo.

Tuma, N.B. (1976), "Rewards, Resources and the Rate of Mobility: A Nonstationary Multivariate Stochastic Model", *American Sociological review,* **41**, 338-360.

Tuma, N.B., Hannan, M.T. (1979), "Approaches to the Censoring Problem in the Analysis of Event Histories", *Sociological Methodology,* San Francisco: Jossey-Bass.

Turnbull, B.W. (1974), "Nonparametric Estimation of a Survival Function with Doubly Censored Data", *Journal of the American Statistical Association*, **69**, 74-80.