IDENTIFICATION OF A COMPETING RISKS MODEL WITH UNKNOWN TRANSFORMATIONS OF LATENT FAILURE TIMES

Sokbae Lee

THE INSTITUTE FOR FISCAL STUDIES

DEPARTMENT OF ECONOMICS, UCL

c**e**mmap working paper CWP17/05

Identification of a Competing Risks Model with Unknown Transformations of Latent Failure Times

Sokbae Lee[*]

Centre for Microdata Methods and Practice

Institute for Fiscal Studies

and

Department of Economics

University College London

London, WC1E 6BT, UK.

Email: l.sokbae@ucl.ac.uk

July 2005

Abstract

This paper is concerned with identification of a competing risks model with unknown transformations of latent failure times. The model in this paper includes, as special cases, competing risks versions of proportional hazards, mixed proportional hazards, and accelerated failure time models. It is shown that covariate effects on latent failure times, cause-specific link functions, and the joint survivor function of the disturbance terms can be identified without relying on modelling the dependence between latent failure times parametrically nor using an exclusion restriction among covariates. As a result, the paper provides an identification result on the joint survivor function of the latent failure times conditional on covariates.

Key words: Competing risks model; Identification; Transformation model.

1 Introduction

This paper is concerned with identification of a competing risks model with unknown transformations of latent failure times. Suppose that there are J competing causes of failure indexed by the integers 1 to J with corresponding latent failure times $(T_1, \ldots T_J)$. One observes the duration to the first failure and the corresponding cause of failure, denoted by $Y = \min_j T_j$ and $\Delta = \arg\min_j T_j$, along with explanatory variables. It is well known (see, for example, Cox (1962) and Tsiatis (1975)) that the distribution of latent failure times is nonparametrically unidentified. Heckman and Honoré (1989) and Abbring and Van den Berg (2003), among others, demonstrate that one can break this nonidentification result by considering a certain class of models for latent failure times and by exploiting sufficiently independent variations of latent failure times with explanatory variables.

The main purpose of this paper is to provide weak restrictions that are sufficient to identify important features of a competing risks model. The model and restrictions imposed in this paper are quite different from those of Heckman and Honoré (1989) and Abbring and Van den Berg (2003) and can be viewed as an alternative modelling framework. Specifically, we consider a transformation model for each latent failure time and also get around a difficult problem of identifying the scale factor of covariate effects (or equivalently, the scale factor of a link function). Our identification result is sufficiently weak in a sense that it is not needed to have a parametric form of dependence nor an exclusion restriction on covariates.

It is assumed in this paper that each latent failure time T_j is generated by a linear transformation regression model:

$$H_j(T_j) = X'\beta_j + U_j, \quad j = 1, \ldots, J, \tag{1}$$

where $H_j(\cdot)$ is an unknown, differentiable, strictly increasing function with a derivative $h_j(\cdot)$, X a d-dimensional vector of continuous explanatory variables (not including a constant term), β_j a d-dimensional vector of unknown parameters, and U_j is an unobserved random variable that is independent of X. It is also assumed that the distribution of U_j is unknown and U_j may depend on each other.

The model (1) includes, as special cases, competing risks versions of proportional hazards, mixed proportional hazards, and accelerated failure time models and may be called a *competing risks transformation model*. For example, a mixed proportional hazards competing risks model can be expressed as a special case of (1) with $U_j = \alpha_j + \varepsilon_j$, where α_j is a cause-specific frailty term, ε_j is an unobserved random variable that has CDF

$F_j(\varepsilon) = 1 - \exp(-e^\varepsilon)$ and $\exp[H_j(t)]$ is the integrated baseline hazard function (Clayton and Cuzick (1985) and Abbring and Van den Berg (2003)). An accelerated failure time competing risks model may be presented as (1) in which $H_j(t) = \log t$ (Heckman and Honoré (1989)). In survival analysis, transformation models have been studied intensively for single risks data (see, e.g., Cheng, Wei, and Ying (1995, 1997) and Horowitz (1996, 1999)). Fine (1999) uses a transformation model to analyze the cumulative incidence function for competing risks data.

This paper provides identification results on the three objects in (1): β_j, H_j, and the joint survivor function of U_j's. Given the knowledge of these objects, we can identify the joint survivor function of $(T_1, \dots T_J)$ conditional on $X = x$. Therefore, we can carry out a counterfactual analysis of latent failure times, which would be impossible using only observed data.

The paper is organized as follows. Section 2 provides an identification result for β_j. Section 3 presents additional identification results on H_j, the joint survivor function of U_j's, and the joint survivor function of $(T_1, \dots T_J)$ conditional on $X = x$. Concluding remarks are given in Section 4.

2   Identification of β_j

This section provides conditions under which β_j in (1) is identified. β_j is a vector of parameters that measure the effects of X on latent failure time T_j. Since H_j and the distribution of U_j are unknown, β_j is identified only up to some location and scale normalization. Also, because H_j can be different from H_k, β_j is not directly comparable to β_k for $j \neq k$. Therefore, only the direction of β_j is identified and ratios between components of β_j give relative importance of components of X.

In this paper, we assume that $d \geq 2$. Otherwise, there is nothing left to identify since the scale factor has to be normalized. To identify β_j, define $S(t|x) = \mathrm{Prob}(Y > t | X = x)$ and $Q_j(t|x) = \mathrm{Prob}(T_j > t, \Delta = j | X = x)$ for $j = 1, \dots, J$. Also, let $p(x)$ denote the probability density function of X, S_X the support of X, and x_k and $\beta_j^{(k)}$ the k-th components of x and β_j, respectively. For $k = 1, \dots, d$, and $j = 1, \dots, J$, define

$$A_k(t, x) = \frac{\partial S(t|x)}{\partial x_k} p(x),$$

$$B_j(t, x) = -\frac{\partial Q_j(t|x)}{\partial t} p(x),$$

$B(t, x) = [B_1(t, x), \ldots, B_J(t, x)]'$, and $\beta^{(k)} = (\beta_1^{(k)}, \beta_2^{(k)}, \ldots, \beta_J^{(k)})'$ for $k = 1, \ldots, d$. Thus, $\beta^{(k)}$ is a $(J \times 1)$ vector of unknown cause-specific coefficients of the k-th component of X.

Observe that $A_k(t, x)$ and $B(t, x)$ are identified directly from the data. In the following, we will achieve identification of β_j by showing that β_j is a functional of $A_k(t, x)$ and $B(t, x)$. To do so, we make the following assumptions:

Assumption 1 (Identification of β_j). *(a) $(U_1, \ldots, U_J)$ are continuously distributed and independent of X but may be arbitrarily correlated with one another.*
(b) X is a $d(\geq 2)$-dimensional vector of continuous explanatory variables and has a joint probability density function $p(x)$ that is positive on S_X except on the boundary.
(c) $H_j(\cdot)$ is an unknown, differentiable, strictly increasing function with a derivative $h_j(\cdot)$.
(d) For each $j = 1, \ldots, J$,

$$\int [w_T(t)/h_j(t)]dt = 1, \tag{2}$$

where $w_T(t)$ is a weight function with compact support S_T.
(e) Assume that as functions of $x \in S_X$, components of $B(t, x)$ are linearly independent for every $t \in S_T$.

Condition (a) allows for arbitrary correlations among U_j. Location normalization is achieved by excluding an intercept term in X (see condition (b)). Condition (c) is convenient in analyzing a transformation model. Scale normalization is accomplished by condition (d). This assumption is useful to create averaging effects, so that a sample analog estimator based on our identification result converges in probability at a rate of $n^{-1/2}$, where n is the sample size. The same type of scale normalization is used for similar reasons in Horowitz (2001) and Horowitz and Lee (2004). Condition (e) amounts to assuming that cause-specific sub-densities of latent failure times conditional on $X = x \in S_X$ are linearly independent for every $t \in S_T$. There are no possible values of t and x satisfying this condition if β_j, H_j, α_j, and F_j are identical over $j = 1, \ldots, J$; however, this is not an interesting case to use a competing risks model.

The following theorem gives a constructive identification result for β_j.

Theorem 1. *Let Assumption 1 hold. Then for each $k = 1, \ldots, d$, $\beta^{(k)}$ can be expressed as*

$$\beta^{(k)} = \int w_T(t) E[B(t, X)B(t, X)']^{-1} E[B(t, X)A_k(t, X)] \, dt. \tag{3}$$

Proof of Theorem 1. Let $f(u_1, \ldots, u_J)$ denote the joint probability density function of $(U_1, \ldots, U_J)$. Notice that

$$
\begin{aligned}
S(t|x) &= \Pr(H_j(T_j) > H_j(t) \text{ for all } j | X = x) \\
&= \Pr(U_j > H_j(t) - x'\beta_j \text{ for all } j) \\
&= \int_{H_1(t)-x'\beta_1}^{\infty} \cdots \int_{H_J(t)-x'\beta_J}^{\infty} f(u_1, \ldots, u_J) \, du_1 \cdots du_J
\end{aligned}
$$

and

$$
\begin{aligned}
Q_j(t|x) &= \Pr(H_j(T_j) > H_j(t) \text{ and } H_l(T_l) > H_l(T_j) \text{ for all } l \neq j | X = x) \\
&= \Pr(U_j > H_j(t) - x'\beta_j \text{ and } U_l > H_l(T_j) - x'\beta_l \text{ for all } l \neq j) \\
&= \int_{H_j(t)-x'\beta_j}^{\infty} \underbrace{\int_{H_1[H_j^{-1}(x'\beta_j+u_j)]-x'\beta_1}^{\infty} \cdots \int_{H_J[H_j^{-1}(x'\beta_j+u_j)]-x'\beta_J}^{\infty}}_{J-1 \text{ integrals excluding } j} f(u_1, \ldots, u_J) \underbrace{du_1 \cdots du_J}_{du_j \text{ is excluded}} \, du_j
\end{aligned}
$$

for $j = 1, \ldots, J$. By differentiation,

$$
\frac{\partial Q_j(t|x)}{\partial t} = -h_j(t) \underbrace{\int_{H_1(t)-x'\beta_1}^{\infty} \cdots \int_{H_J(t)-x'\beta_J}^{\infty}}_{J-1 \text{ integrals excluding } j} f(u_1, \ldots, u_{j-1}, H_j(t) - x'\beta_j, u_{j+1}, \ldots, u_J) \underbrace{du_1 \cdots du_J}_{du_j \text{ is excluded}}
$$

and

$$
\frac{\partial S(t|x)}{\partial x_k} = \sum_{j=1}^{J} \beta_j^{(k)} \underbrace{\int_{H_1(t)-x'\beta_1}^{\infty} \cdots \int_{H_J(t)-x'\beta_J}^{\infty}}_{J-1 \text{ integrals excluding } j} f(u_1, \ldots, u_{j-1}, H_j(t) - x'\beta_j, u_{j+1}, \ldots, u_J) \underbrace{du_1 \cdots du_J}_{du_j \text{ is excluded}},
$$

where x_k and $\beta_j^{(k)}$ are the k-th components of x and β_j, respectively. It follows that

$$
\frac{\partial S(t|x)}{\partial x_k} = \sum_{j=1}^{J} \frac{-\partial Q_j(t|x)}{\partial t} \frac{\beta_j^{(k)}}{h_j(t)}. \tag{4}
$$

Multiplying by $p(x)$ both right and left sides of (4) gives

$$
\frac{\partial S(t|x)}{\partial x_k} p(x) = \sum_{j=1}^{J} \frac{-\partial Q_j(t|x)}{\partial t} p(x) \frac{\beta_j^{(k)}}{h_j(t)}. \tag{5}
$$

To express identifying relationships compactly, define, for $k = 1, \ldots, d$, let

$$
b_k(t) = \left(\frac{\beta_1^{(k)}}{h_1(t)}, \frac{\beta_2^{(k)}}{h_2(t)}, \ldots, \frac{\beta_J^{(k)}}{h_J(t)} \right)'.
$$

5

Then it follows from (5) that

$$A_k(t, x) = B(t, x)' b_k(t).$$

To identify $\beta^{(k)}$, write

$$B(t, x) A_k(t, x) = B(t, x) B(t, x)' b_k(t). \tag{6}$$

To solve for $b_k(t)$, substitute the random vector X for x in (6) and take expectations to obtain

$$E[B(t, X) A_k(t, X)] = E[B(t, X) B(t, X)'] b_k(t). \tag{7}$$

By the assumption that components of $B(t, x)$ are linearly independent for every $t \in S_T$, we have that $E[B(t, X) B(t, X)']$ is nonsingular for every $t \in S_T$. Therefore, under the scale normalization (2), $\beta^{(k)}$ can be expressed as in the equation (3), which proves the theorem. □

The equation (3) could be used as the basis for a sample analog estimator of $\beta^{(k)}$. Semiparametric estimation of $\beta^{(k)}$ can be carried out by replacing unknown population quantities in (3) with suitable nonparametric estimators.

It can be seen from the expression of $S(t|x)$ in the proof of Theorem 1 that the expectation of Y conditional on $X = x$ belongs to the class of multiple-index models (see, for example, Ichimura and Lee (1991)). Typically, certain exclusion restrictions (for example, certain components of parameters are zero) are needed for multiple-index models to achieve identification of parameters. As shown by Heckman and Honoré (1989), Abbring and Van den Berg (2003), and equation (3), exclusion restrictions are not required for the identification of semiparametric competing risks models.

It is important to notice that there exists an important difference between identification results of Heckman and Honoré (1989) and Abbring and Van den Berg (2003) and one obtained in Theorem 1. Those of Heckman and Honoré (1989) and Abbring and Van den Berg (2003) are based on the arguments of letting $t \to 0$, thereby implying that corresponding estimation methods would be based on only observations with failure times close to zero. An estimator of Femanian (2003, Section 4) is such an example. This is mainly because the scale factor has to be identified in the setup of Heckman and Honoré (1989) and Abbring and Van den Berg (2003). The difficulty of identifying the scale factor is not specific to competing risks models. A similar problem arises in a single-risk mixed proportional hazard model (see, for example, Horowitz (1999)).

It is also important to note that the continuity of the distribution of X is not needed for identification of the covariate effects in the framework of Heckman and Honoré (1989) and Abbring and Van den Berg (2003). This is due to the fact that in their framework, the joint survivor function of latent failure times can be written as an exponential function, which does not necessarily hold in our model (1). Finally, we note that (3) is expressed in terms of a density-weighted form (see also (5)), which would be convenient to construct a resulting sample analog estimator.

When all components of X are discrete, it is unclear whether one can point-identify covariate effects. However, some recent studies show that it is possible to derive bounds for covariate effects in some competing risks models with discrete covariates. For example, Bond and Shaw (2003) obtain bounds for covariate effects under the assumption that the copula associated with the joint distribution of latent failure times is invariant to the value of covariates. Abbring and Van den Berg (2005) apply the result of Bond and Shaw (2003) to bound the treatment effects on duration outcomes. Honoré and Lleras-Muney (2004) derive bounds in an accelerated failure time competing risks model with discrete covariates.

3   Identification of H_j and the joint distribution of $(U_1, \ldots, U_J)$

This section presents conditions under which H_j and the joint distribution of U_j's are identified. These results combined along with that for β_j provide the identification of the joint survivor function of $(T_1, \ldots, T_J)$ conditional on $X = x$.

We first consider identification of H_j. Notice that $H_j(t)$ can be unbounded as $|t| \to \infty$, so that we only focus on identification of H_j on a compact interval $S_T = [\underline{t}, \overline{t}]$. It is necessary to impose location normalization on H_j. It is achieved here by assuming that there exists $t_0 \in [\underline{t}, \overline{t}]$ such that $H_j(t_0) = 0$ for each j.

To present an identification result for H_j, we need some additional notation. Define $H(t) = (H_1(t), \ldots, H_J(t))'$ and for two J-dimensional vectors $a = (a_1, \ldots, a_J)$ and $b = (b_1, \ldots, b_J)$, let $a./b = (a_1/b_1, \ldots, a_J/b_J)$ denote an element-by-element division operator. The following theorem gives a constructive identification result for H_j.

Theorem 2. *Let Assumption 1 hold with $S_T = [\underline{t}, \overline{t}]$. Suppose that all the components of $\beta^{(k)}$ are non-zero. In addition, assume that $H_j(t_0) = 0$ for some $t_0 \in [\underline{t}, \overline{t}]$ for each j. Then for any k, $H(t)$ can be expressed as*

$$H(t) = \int_{t_0}^{t} \beta^{(k)}./\{E[B(s,X)B(s,X)']^{-1} E[B(s,X)A_k(s,X)]\}\, ds. \tag{8}$$

for any $t \in [\underline{t}, \overline{t}]$.

Proof of Theorem 2. It follows from (7) and Assumption 1 (e) that

$$h(t) = \beta^{(k)}./\{E[B(t,X)B(t,X)']^{-1}E[B(t,X)A_k(t,X)]\}, \tag{9}$$

where $h(t) = [h_1(t), \ldots, h_J(t)]'$. Then the theorem follows by integrating both sides of the equation above from t_0 to t and invoking the location normalization $H_j(t_0) = 0$. The condition that all the components of $\beta^{(k)}$ are non-zero is necessary to determine $H(t)$ uniquely. $\qquad\square$

Notice that we achieve identification of $H_j(t)$ only on a compact interval rather than $\mathbf{R}$, where $\mathbf{R}$ is the entire real line. This is because $H_j(t)$ could be unbounded and $E[B(t,X)B(t,X)']$ might be singular for a very small or very large t. Also notice that components of $H(t)$ are over-identified since (8) holds for any $k = 1$ provided that corresponding components of $\beta^{(k)}$ are non-zero.

We now turn to identification of the joint distribution of $(U_1, \ldots, U_J)$. Define

$$S_U(u_1, \ldots, u_J) = \Pr(U_1 > u_1, \ldots, U_J > u_J)$$

to be the joint survivor function of $U = (U_1, \ldots, U_J)$. In addition, define $Z = (Z_1, \ldots, Z_J)$, where $Z_j \equiv X'\beta_j$ for each $j = 1, \ldots, J$. To identify $S_U(u_1, \ldots, u_J)$, we use Z_j to create exclusion-restriction-type variations. Specifically, if β_j's are identified and there are no functionally deterministic relationships among Z_j's, we can exploit independent variations of one of Z_j's given other Z_j's. Finally, let $S_{Y|Z}(t|z_1, \ldots, z_J) = \Pr(Y > t|Z_1 = z_1, \ldots, Z_J = z_J)$ denote the conditional survivor function of Y given $(Z_1, \ldots, Z_J) = (z_1, \ldots, z_J)$. Note that $S_{Y|Z}(t|z_1, \ldots, z_J)$ is well-defined because Z_j's are not functionally dependent. Also note that once β_j is identified for each j, $S_{Y|Z}(t|z_1, \ldots, z_J)$ is identified directly from the data.

Theorem 3. *Let Assumption 1 hold. Suppose that β_j and $\{H(t) : t \in [\underline{t}, \overline{t}]\}$ are identified. Assume that there are no functionally deterministic relationships among components of Z and the support of Z is $\mathbf{R}^J$. Then for any values of $(u_1, \ldots, u_J)$, $S_U(u_1, \ldots, u_J)$ can be expressed as*

$$S_U(u_1, \ldots, u_J) = \int w_U(t) S_{Y|Z}(t|H_1(t) - u_1, \ldots, H_J(t) - u_J) \, dt \tag{10}$$

with a predetermined weight function $w_U(t)$ with support in an interior of $[\underline{t}, \overline{t}]$.

Proof of Theorem 3. Recall that $f(u_1, \ldots, u_J)$ denotes the joint probability density function of $(U_1, \ldots, U_J)$. On one hand, notice that

$$S_U(u_1, \ldots, u_J) = \int_{u_1}^{\infty} \cdots \int_{u_J}^{\infty} f(u_1, \ldots, u_J) \, du_1 \cdots du_J.$$

On the other hand, observe that

$$\begin{aligned}
S_{Y|Z}(t|z_1, \ldots, z_J) &= \Pr(H_j(T_j) > H_j(t) \text{ for all } j | Z_1 = z_1, \ldots, Z_J = z_J) \\
&= \Pr(U_j > H_j(t) - z_j \text{ for all } j) \\
&= \int_{H_1(t)-z_1}^{\infty} \cdots \int_{H_J(t)-z_J}^{\infty} f(u_1, \ldots, u_J) \, du_1 \cdots du_J.
\end{aligned}$$

In view of the expressions of $S_U(u_1, \ldots, u_J)$ and $S_{Y|Z}(t|z_1, \ldots, z_J)$,

$$S_U(u_1, \ldots, u_J) = S_{Y|Z}(t|H_1(t) - u_1, \ldots, H_J(t) - u_J)$$

for any $t \in [\underline{t}, \overline{t}]$. Therefore, the theorem follows immediately since given t, $S_U(u_1, \ldots, u_J)$ is identified for any values of $(u_1, \ldots, u_J)$ thanks to the support condition on Z. $\qquad\square$

The infinite support condition on Z is useful to identify $S_U(u_1, \ldots, u_J)$ for any values of $(u_1, \ldots, u_J)$. Without this condition, the identification result would be limited to a subset of the support of $(U_1, \ldots, U_J)$.

We are now ready to state the main result of this paper.

Theorem 4. *Let Assumption 1 hold. Suppose that β_j, $\{H(t) : t \in [\underline{t}, \overline{t}]\}$, and $\{S_U(u_1, \ldots, u_J) : -\infty < u_j < \infty, \ j = 1, \ldots, J\}$ are identified. Then for any $(t_1, \ldots, t_J) \in [\underline{t}, \overline{t}]^J$,*

$$Pr(T_1 > t_1, \ldots, T_J > t_J | X = x) = S_U(H_1(t_1) - x'\beta_1, \ldots, H_J(t_J) - x'\beta_J) \qquad (11)$$

Proof of Theorem 4. This is straightforward to prove since

$$\begin{aligned}
\Pr(T_1 > t_1, \ldots, T_J > t_J | X = x) &= \Pr(U_j > H_j(t_j) - x'\beta_j \text{ for each } j) \\
&= \int_{H_1(t_1)-x'\beta_1}^{\infty} \cdots \int_{H_J(t_J)-x'\beta_J}^{\infty} f(u_1, \ldots, u_J) \, du_1 \cdots du_J \\
&= S_U(H_1(t_1) - x'\beta_1, \ldots, H_J(t_J) - x'\beta_J).
\end{aligned}$$

$$\square$$

This theorem shows that one can identify the underlying distribution of latent failure times conditional on explanatory variables. One weakness of this result is that the joint distribution is not identified for all the values of $(t_1, \ldots, t_J)$. Thus, we fail to achieve identification of the marginal distribution of T_j.

9

4   Conclusions

This paper has shown that a transformation model can be used to identify important features of a dependent competing risks model that includes, as special cases, competing risks versions of proportional hazards, mixed proportional hazards, and accelerated failure time models. In particular, all the identification results are constructive, thereby suggesting semiparametric methods for estimating the competing risks transformation model in (1). Sample analog estimation can be carried by replacing unknown population quantities in identification results with some suitable nonparametric estimators, e.g. kernel estimators. Implementation of this estimation is a topic for future research.

References

Abbring, J.H. and G.J. Van den Berg (2003) The identifiability of the mixed proportional hazards competing risks model. *J. R. Statist. Soc. B*, 65, 701-710.

Abbring, J.H. and G.J. Van den Berg (2005) Social experiments and IV with duration outcomes, Working paper, Department of Economics, Free University, Amsterdam.

Bond, S.J. and J.E.H. Shaw (2003) Bounds on the covariate-time transformation for competing risks survival analysis, Working paper, Department of Statistics, University of Warwick, Coventry.

Cheng, S.C., L.J. Wei, and Z. Ying (1995) Analysis of transformation models with censored data, *Biometrika*, 82, 835-845.

Cheng, S.C., L.J. Wei, and Z. Ying (1997) Predicting survival probabilities with semiparametric transformation models, *J. Amer. Statist. Assoc.*, 92, 227-235.

Clayton, D. and J. Cuzick (1985) Multivariate generalization of the proportional hazard model, *J. R. Statist. Soc. A*, 148, 82-117.

Cox. D.R. (1962) *Renewal Theory*, London: Methuen.

Femanian, J.-D. (2003) Nonparametric estimation of competing risks models with covariates, *J. Multivariate Anal.*, 85, 156-191.

Fine, J.P. (1999) Analyzing competing risks data with transformation models, *J. R. Statist. Soc. B*, 61, 817-830.

Heckman, J.J. and B.E. Honoré (1989) The identifiability of the competing risks model, *Biometrika*, 76, 325-330.

B.E. Honoré and A. Lleras-Muney (2004) Bounds in competing risks models and the war on cancer, Working paper, Department of Economics, Princeton University, Princeton.

Horowitz, J.L. (1996) Semiparametric estimation of a regression model with an unknown transformation of the dependent variable, *Econometrica*, 64, 103-137.

Horowitz, J.L. (1999) Semiparametric estimation of a proportional hazard model with unobserved heterogeneity, *Econometrica*, 67, 1001-1028.

Horowitz, J.L. (2001) Nonparametric estimation of a generalized additive model with an unknown link function, *Econometrica*, 69, 499-513.

Horowitz, J.L. and S. Lee (2004) Semiparametric estimation of a panel data proportional hazards model with fixed effects, *J. Econometrics*, 119, 155-198.

Ichimura, H. and L.-F.Lee (1991) Semiparametric least squares estimation of multiple index models: single equation estimation. In: Barnett W A, Powell J, and Tauchen G (eds) *Nonparametric and Semiparametric Methods in Econometrics and Statistics.* Cambridge University Press, Cambridge, pp. 3-49.

Tsiatis, A. (1975) A nonidentifiability aspect of the problem of competing risks, *Proc. Nat. Acad. Sci.*, 72, 20-22.