Identifiability of the Sign of Covariate Effects in the Competing Risks Model

Lo, Simon M.S.; Wilke, Ralf

Link to publication in CBS Research Portal

Identifiability of the Sign of Covariate Effects in the Competing Risks Model

Simon M.S. Lo and Ralf Wilke

Journal article (Accepted version)

Identifiability of the sign of covariate effects in the competing risks model*

Simon M.S. Lo[†]

Ralf A. Wilke[‡]

August 2016

Abstract

We present a new framework for the identification of competing risks models, which also include Roy models. We show that by establishing a Hicksian-type decomposition, the direction of covariate effects on the marginal distributions of the competing risks model can be identified under weak restrictions. Our approach leaves the marginal distributions and their joint distribution completely unspecified, except that the latter is invariant in the covariates. Results from simulations and two data examples suggest that our method often outperforms existing comparable approaches in terms of the range of durations for which the direction of the covariate effect is identified, particularly for long duration.

Keywords: dependent censoring, copula, identifiability

[†]Lingnan University, E–mail: simonlo@ln.edu.hk

[‡]Copenhagen Business School, Department of Economics, E–mail: rw.eco@cbs.dk

1   Introduction

A feature of the competing risks model is that only the transition to one risk (or failure because of one cause of death) is observed. This is the risk with shortest realised duration. The latent duration for the other risks are therefore not observed. The non-identifiability of the competing risks model means that observed data alone do not contain sufficient information to identify the marginal distributions of the latent durations (Cox, 1962; Tsiatis, 1975). This identification problem is closely related to the identification problem of the Roy model (Roy, 1951), where an individual faces different potential wage distributions in different economic sectors but only the wage in the chosen sector (maximum potential wage) is being observed.

The joint distribution of the latent durations can be viewed as a copula function of the marginal distributions (Schweizer and Sklar, 1983). Most previous studies focus on the identifiability of the marginal distributions. Peterson (1976) in his seminal paper derives bounds for the marginal distributions in absence of any knowledge about them and the copula function. These bounds are typically too wide for informative results, particularly for longer durations as their width (difference between the upper and lower bound) increases with duration. When the copula is known, Zheng and Klein (1995) show that the marginal distributions are nonparametrically identified. Given that full knowledge about the copula is a strong requirement, many existing studies consider an intermediate approach. In particular, the copula is unknown but independent of the covariates (the copula invariance assumption). In this scenario identification results are obtained by exploiting linkages between variations in covariates and variations in the observed durations. The copula invariance assumption ensures that by changing the covariates, changes in observed durations stem solely from changes in the marginal distributions but not from changes in the copula function. By relying on exclusion restrictions or considering certain classes of regression models that restrict the effect of covariates on the marginal distributions, a number of studies have derived widely regarded identification results. For instance, Heckman and Honoré (1989) show for proportional hazard models and accelerated failure time models that marginal distributions are identified semiparametrically, provided that the variations induced by the covariates are sufficiently large. Heckman and Honoré (1990) establish this result for a corresponding Roy model. Abbring and van den Berg (2003) derive similar results for the semiparametric mixed proportional hazard model. In their model the copula function be-

longs to a Laplace transform of an unknown mixture distribution. Using the accelerated failure time model, Honoré and Lleras-Muney (2006) obtain bounds for the marginal effect of discrete covariates on latent durations. Lee and Lewbel (2013) show that the accelerated failure time model is identified provided that a certain rank condition is satisfied. Relying on exclusion restrictions, Henry and Mourifie (2014) derive bounds for the marginal distributions in the Roy model. Park (2015) identifies the joint distribution of the latent outcome variables in the Roy model when an instrumental variable is available. Apart from theoretical studies, the copula invariance assumption is also commonly made in empirical economic analysis. The most popular example is the mixed proportional hazard model using finite mass point specification for the unknown mixture distribution (Heckman and Singer, 1984), see e.g. Butler et al. (1989), Carling et al. (1996), Meghir and Whitehouse (1997), Dolton and van der Klaauw (1999), Steiner (2001), D'Addio and Rosholm (2005), Alba-Ramirez et al. (2007). Other empirical studies using the copula invariance assumption include the independent risks model and parametric copula models. See for example Carling et al. (1996), Mealli and Pudney (1996), and Burda et al. (2015).

We consider a more general model in this paper than the above mentioned studies, although we maintain the copula invariance assumption. First, the marginal distributions in our model are nonparametric and therefore it is not limited to specific classes of duration models such as the proportional hazard models or the accelerated failure time model. This is a practical advantage as these models impose parametric restrictions on the marginal distributions, which may be violated in applications. Second, our model does not rely on exclusion restriction nor requires instrumental variables which could be either difficult to justify or might not be available in an application.

In this paper we establish a Hicksian-type decomposition of covariate effects on marginal distributions. We develop a general link between the observable sign of covariate effects on subdistributions (cumulative incidence functions, CIF) and the unobservable sign of covariate effects on the marginal distributions. We show that under rather weak restrictions the sign of covariate effects on the marginal distributions is identifiable for some set of durations.

Definition 1 *The identification set is defined as the set of durations for which the sign of a covariate effect on the marginal distributions is identified. Identification set A is larger (smaller) than identification set B if $B \subset A$ ($A \subset B$).*

At a glance our approach shares some similarities with the approach proposed by Bond and Shaw (2006). Under the copula invariance assumption they derive bounds for the covariate-time transformation (CTT). These bounds can be used to identify the sign of the covariate effect. However, there are three major differences between our decomposition approach and the CTT. First, these two methods produce different identification sets. In order to make the difference apparent, we restate Bond and Shaw's approach using our analytical framework. Second, the width of the bounds for the CTT increases with duration. This implies that, similar to the Peterson bounds, the bounds for the CTT tend to be less informative at longer durations, but this is not the case for our approach. Third, the bounds for the CTT require an additional non-testable order assumption which restricts the role of covariates on the marginal distributions in a non-trival way. This order assumption implies that the propensity of one risk will either increase or decrease for all durations when a covariate changes. In the context of the Roy model, this implies that the utility for one state increases more or decreases less than that for the other state irrespective of the level of outcome variables when a covariate changes. As a by-product of rewriting Bond and Shaw's (2006) approach, we accommodate a feature of our approach into their model which obviates their order assumption.

In our simulation studies and two real-data illustrations, our proposed decomposition approach tends to produce the largest identification set among the considered methods. We illustrate that a proposed combination of the various methods is even more appealing for empirical research if the direction rather than the magnitude of the covariate effect is of main interest. Our real-data illustration also provides evidence for changes in the sign of covariate effects at different durations, highlighting the importance of using a more flexible model rather than the proportional hazard and accelerated failure time model for the marginal distributions. These findings are useful for empirical research that utilises competing risks models as well as the Roy model.

The structure of this paper is as follows: Section 2 introduces the model and presents the identification results. Section 3 explores the performance of the considered approach by means of simulations. Section 4 investigates the empirical performance with two data examples.

2 Identifiability

We consider a model with two latent competing random variables T_1 and $T_2 \in \mathbb{R}_+$. T_1 and T_2 are latent durations to events 1 and 2 respectively. A competing risks model with more than two risks is considered in Section 2.2. $\boldsymbol{X} \in \mathbb{R}^K$ is a vector of continuous covariates x_k, $k = 1, ..., K$. A model with discrete $\boldsymbol{X}$ is considered in Section 2.1. The marginal survival function (latent survival) of T_j is $S_j(t; \boldsymbol{x}) = \Pr(T_j > t | \boldsymbol{x})$, with $\boldsymbol{X} = \boldsymbol{x}$. The joint survival distribution of the latent durations is $S(t_1, t_2; \boldsymbol{x}) = \Pr(T_1 > t_1, T_2 > t_2 | \boldsymbol{x})$. Let $T = \min(T_1, T_2)$ be the observed minimum and $\delta = \arg\min_j\{T_j\}$ be the risk indicator. When $\delta = 1$, latent duration T_2 is censored by T_1, and vice versa. Define the cumulative incidence function (CIF) as $Q_j(t; \boldsymbol{x}) = \Pr(T \leq t, \delta = j | \boldsymbol{x})$, the cause-specific crude hazard function as $\lambda_j(t; \boldsymbol{x}) = \lim_{\Delta \to 0} \Pr(t \leq T \leq t + \Delta, \delta = j | T \geq t, \boldsymbol{x})/\Delta$ for risk $j = 1, 2$ and the survival function of T (overall survival) as $S(t; \boldsymbol{x}) = \Pr(T > t | \boldsymbol{x}) = 1 - Q_1(t; \boldsymbol{x}) - Q_2(t; \boldsymbol{x})$.

Assumption 1 $S_j(t; \boldsymbol{x}) : [0, \infty] \to [0, 1]$ and $S(t; \boldsymbol{x}) : [0, \infty] \to [0, 1]$ are continuous and strictly decreasing in t for all j with inverses denoted by S_j^{-1} and S^{-1} respectively. $Q_j(t; \boldsymbol{x})$ is continuous and strictly increasing in t for all j with inverse denoted by Q_j^{-1}. S_j, S_j^{-1}, S, S^{-1}, Q_j, and Q_j^{-1} are differentiable with respect to $\boldsymbol{x}$.

Definition 2 The copula function $C(u_1, u_2) = \Pr(U_1 \leq u_1, U_2 \leq u_2) : [0, 1]^2 \to [0, 1]$ is a joint distribution of two uniform random variables (U_1, U_2) with density function $\kappa(u_1, u_2)$.

See Nelsen (2006) for more details on copulas.

According to Sklar's theorem (Schweizer and Sklar, 1983), the joint distribution of the latent durations T_1 and T_2 can be represented by a copula function of the latent survivals , i.e.

$$
\begin{aligned}
S(t_1, t_2; \boldsymbol{x}) &= \Pr(T_1 > t_1, T_2 > t_2 | \boldsymbol{x}) \\
&= \Pr(S_1(T_1; \boldsymbol{x}) \leq S_1(t_1; \boldsymbol{x}), S_2(T_2; \boldsymbol{x}) \leq S_2(t_2; x) | \boldsymbol{x}) \\
&= C(S_1(t_1; \boldsymbol{x}), S_2(t_2; \boldsymbol{x}); \boldsymbol{x}).
\end{aligned} \tag{1}
$$

The copula function characterises the dependence structure between the latent survivals.

Definition 3 Let $u_2 = \zeta_1(u_1; \boldsymbol{x}) = S_2(S_1^{-1}(u_1; x); \boldsymbol{x}) : [0, 1] \to [0, 1]$ be a continuous and strictly increasing link function, which uniquely defines the relationship between $u_1 = S_1(t; \boldsymbol{x})$

5

and $u_2 = S_2(t; \boldsymbol{x})$ for all t and $\boldsymbol{x}$. The link function is differentiable with respect to $\boldsymbol{x}$ with its inverse defined as $u_1 = \zeta_2(u_2; \boldsymbol{x}) = \zeta_1^{-1}(u_2; \boldsymbol{x})$.

Continuity, monotonicity, uniqueness, and differentiability of the link function are guaranteed by Assumption 1. The link function plays the role of determining the propensity of risk 1 such that $\Pr(T_1 \leq T_2; \boldsymbol{x}) = \Pr(S_1^{-1}(U_1; \boldsymbol{x}) \leq S_2^{-1}(U_2; \boldsymbol{x}); \boldsymbol{x}) = \Pr(U_2 \leq \zeta_1(U_1; \boldsymbol{x}); \boldsymbol{x}) = \int_0^1 \int_0^{\zeta_1(u_1; \boldsymbol{x})} \kappa(u_1, u_2; \boldsymbol{x}) du_2 du_1$. In the context of the Roy model, the link function can be viewed as a nonlinear and nonseparable selection equation (Henry and Mourifie, 2014) in which $u_j = S_j(t; \boldsymbol{x})$ is the utility function of the outcome variable T_j.

The copula is unknown but assumed to satisfy the following condition.

Assumption 2 $C(u_1, u_2; \boldsymbol{x}) = C(u_1, u_2)$ *for all* $\boldsymbol{x}$, u_1, *and* u_2,

Given (1) and Assumption 2 the competing risks model is fully characterised by the following system of equations:

$$
\begin{aligned}
S(t; \boldsymbol{x}) &= \Pr(T > t; \boldsymbol{x}) = \Pr(U_1 \leq S_1(t; \boldsymbol{x}), U_2 \leq S_2(t; \boldsymbol{x})) \\
&= \int_0^{S_1(t; \boldsymbol{x})} \int_0^{S_2(t; \boldsymbol{x})} \kappa(u_1, u_2) \, du_2 \, du_1 = C(S_1(t; \boldsymbol{x}), S_2(t; \boldsymbol{x}));
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
Q_1(t; \boldsymbol{x}) &= \Pr(T \leq t, \delta = 1; \boldsymbol{x}) = \Pr(U_1 > S_1(t; \boldsymbol{x}), U_2 \leq \zeta_1(U_1; \boldsymbol{x})) \\
&= \int_{S_1(t; \boldsymbol{x})}^1 \int_0^{\zeta_1(u_1; \boldsymbol{x})} \kappa(u_1, u_2) \, du_2 \, du_1.
\end{aligned}
\tag{3}
$$

A graphical presentation of the problem using the unit square is given in Figure 1. A similar graphical presentation of the competing risk model can be found in Zheng and Klein (1995) and for the Roy model in Henry and Mourifie (2014).

In our model $(T, \delta, \boldsymbol{x})$ are observed and $S(t; \boldsymbol{x})$, $Q_j(t; \boldsymbol{x})$ and $\lambda_j(t; \boldsymbol{x})$ are identified nonparametrically. $S_j(t; \boldsymbol{x})$, $S(t_1, t_2; \boldsymbol{x})$, $\zeta_1(\cdot; \boldsymbol{x})$, $C(u_1, u_2)$, and $\kappa(u_1, u_2)$ are unknown and not identified but somehow restricted due to Assumptions 1 and 2. Instead of considering the identifiability of these functionals we focus on the identifiability of the sign of a covariate effect on $S_j(t; \boldsymbol{x})$. The idea of our approach is to use the observable direction of the covariate effect on $Q_j(t; \boldsymbol{x})$ to identify the sign of the covariate effect on $S_j(t; \boldsymbol{x})$. One can see from $Q_1(t; \boldsymbol{x}) = \Pr(T \leq t, T_1 \leq T_2; \boldsymbol{x}) = \Pr(U_1 > S_1(t; \boldsymbol{x}), U_2 \leq \zeta_1(U_1; \boldsymbol{x}))$ that a covariate effect on

Figure 1: Graphical presentation of a competing risks model.
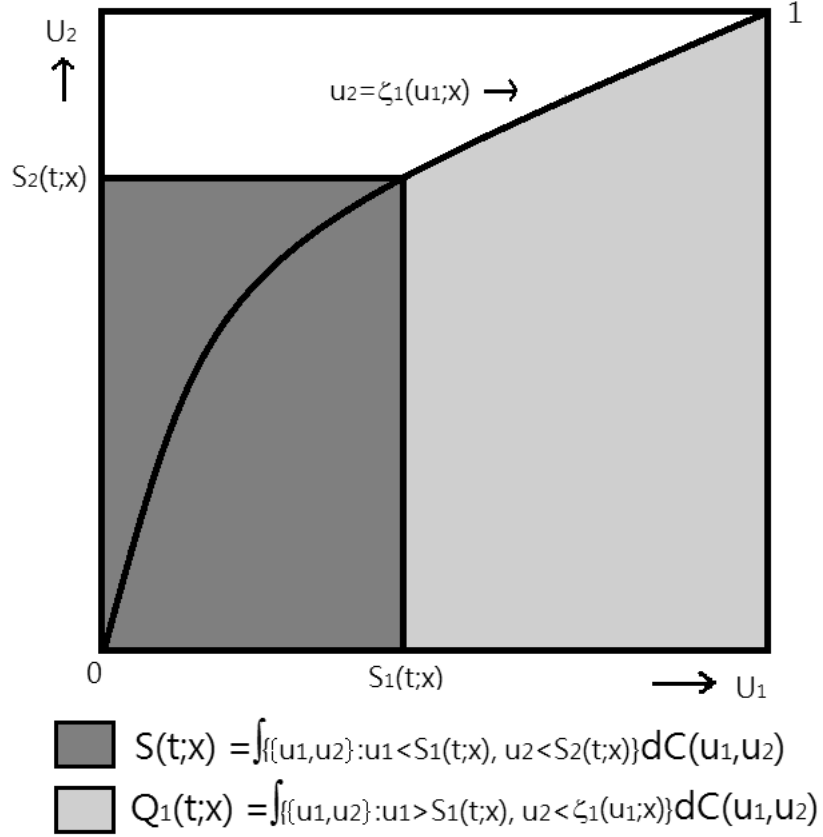
the cumulative incidence function is driven by both the changes in $S_1(t; \boldsymbol{x})$ and $\zeta_1(u_1; \boldsymbol{x})$ for all $u_1 \in [S_1(t; \boldsymbol{x}), 1]$. It is possible that a negative covariate effect on S_j results in a negative effect on Q_j, when the negative effect driven by the link function overrides the positive effect driven by the negative covariate effect on S_j. Therefore, an identified sign of the covariate effect of Q_j is not immediately informative about the sign of the covariate effect on S_j. We propose a Hicksian-type decomposition approach of the covariate effect which makes the relationship between the sign of the covariate effect of S_j and Q_j tractable. For this purpose, we reformulate the competing risks models characterised by (2) and (3) as the following constrained maximisation problem (compare also Figure 2):

$$(S_1(t; \boldsymbol{x}), S_2(t; \boldsymbol{x})) \quad = \quad \underset{(u_1, u_2)}{\arg \max} \, u_1 + u_2 \tag{4}$$

$$\text{subject to: (i) } C(u_1, u_2) \leq S(t; \boldsymbol{x}) \text{ and (ii) } u_2 = \zeta_1(u_1; \boldsymbol{x}).$$

7

Intuitively, risks 1 and 2 are two groups that compete for survival by time t. The competition

Figure 2: Latent survivals under a given copula function and link function.
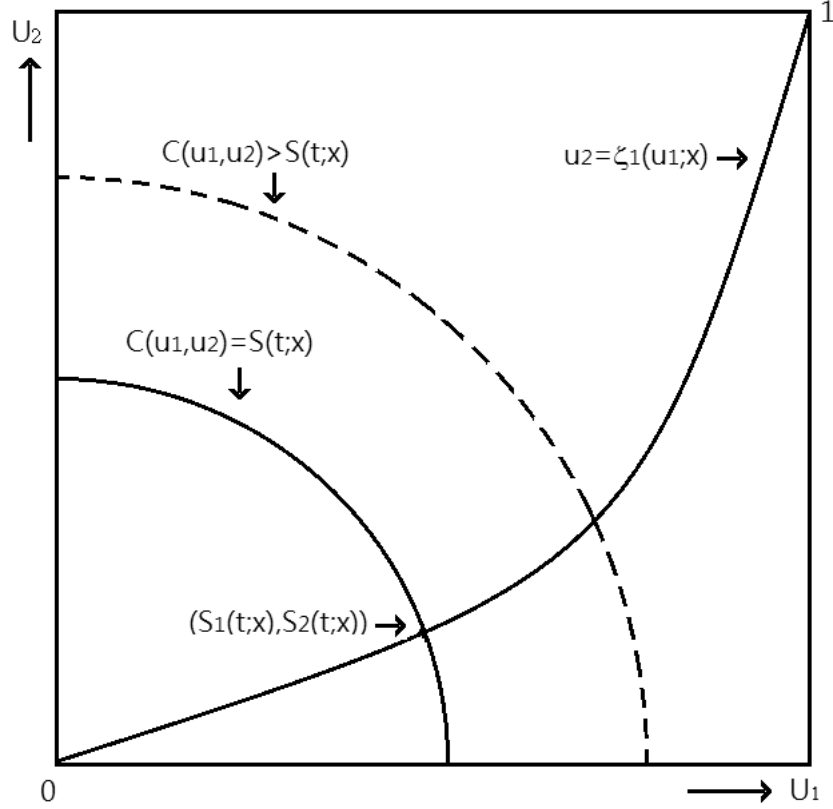
is subject to two constraints. First, there is a maximum level of overall survivor $c = S(t; \boldsymbol{x})$ such that $C(S_1(t; \boldsymbol{x}), S_2(t; \boldsymbol{x}))$ is no greater than c. Second, for a given copula, the composition of non-survivor for risk 1 and 2 at each t is fixed by the link function, and thus $S_2(t; \boldsymbol{x}) = \zeta_1(S_1(t; \boldsymbol{x}); \boldsymbol{x})$. These two constraints together determine the value of survival for each risk at each t. When $\boldsymbol{x}$ changes, it changes the value of the maximum level of overall survivor $c = S(t; \boldsymbol{x})$ for each given t and the link function $\zeta_1(\cdot; \boldsymbol{x})$ simultaneously. Thus, $S_1(t; \boldsymbol{x})$ and $S_2(t; \boldsymbol{x})$ also attain new values. We consider the partial effect of a covariate x_k in the following and suppress the index k for convenience.

Definition 4 *Let $\Delta_x S_j(t; \boldsymbol{x}) = \partial S_j(t; \boldsymbol{x})/\partial x$ be the covariate effect of x on the latent survival S_j at t given $\boldsymbol{x}$.*

Definition 5 *Let $\Delta_x \zeta_1(t; \boldsymbol{x}) = \partial \zeta_1(u_1; \boldsymbol{x})/\partial x$ be the covariate effect of x on the link function at $u_1 = S_1(t; \boldsymbol{x})$ given $\boldsymbol{x}$.*

Definition 6 *The duration function $D(c; \boldsymbol{x})$ is defined by the minimum duration time to keep the value of the overall survival no greater than $c \in [0, 1]$ given $\boldsymbol{x}$, i.e.*

$$D(c; \boldsymbol{x}) = \inf\{v \in \mathbb{R}_+ : S(v; \boldsymbol{x}) \le c\}. \tag{5}$$

Due to Assumption 1 the inverse of $S(t; \boldsymbol{x})$ exists. This implies that $D(c; \boldsymbol{x})$ exists and is unique. Since $S(t; \boldsymbol{x})$ is differentiable w.r.t. x, $D(c; \boldsymbol{x})$ is also differentiable w.r.t. x.

Definition 7 *The Hicksian latent survival function, $S_j^*(c(t); \boldsymbol{x})$, is defined by the value of the latent survival function when S is held constant at $c(t) = S(t; \boldsymbol{x})$, i.e.*

$$S_j^*(c(t); \boldsymbol{x}) = S_j(D(c(t); \boldsymbol{x}); \boldsymbol{x}). \tag{6}$$

Differentiating both sides of (6) with respect to x and rearranging, we obtain

$$\frac{\partial S_j(t; \boldsymbol{x})}{\partial x} = \frac{\partial S_j^*(c(t); \boldsymbol{x})}{\partial x} - \frac{\partial S_j(t; \boldsymbol{x})}{\partial t} \frac{\partial D(c(t); \boldsymbol{x})}{\partial x}. \tag{7}$$

The covariate effect on the latent survival S_j can therefore be decomposed into two parts by isolating the effect on the link function and the overall survival. We call the first part the link function effect. This is the change in the Hicksian latent survival S_j^* due to the change in the link function while holding $c(t) = S(t; \boldsymbol{x})$ constant.

Definition 8 *Let $\Delta_x^l S_j(t; \boldsymbol{x}) = \partial S_j^*(c(t); \boldsymbol{x})/\partial x$ be the link function effect of a change in x on latent survival of risk j at t given $\boldsymbol{x}$.*

The link function effect can be thought of as a compensated substitution effect between risks 1 and 2. A movement from $\boldsymbol{x}_0$ to $\boldsymbol{x}_1$ will change the value of overall survival from $c = S(t; \boldsymbol{x}_0)$ to $S(t; \boldsymbol{x}_1)$. An adjustment of the duration time from t to $D(c; \boldsymbol{x}_1)$ is necessary to 'compensate' the induced change in the overall survival in order to hold the value of the overall survival constant.

The second part is called the duration effect, which is the change in the latent survival S_j due to moving the duration time that is required to push the level of overall survival to the new level while holding the link function constant.

Definition 9 *Let $\Delta^d_x S_j(t; \boldsymbol{x}) = -\partial S_j(t; \boldsymbol{x})/\partial t \times \partial D(c(t); \boldsymbol{x})/\partial x$ be the duration effect of a change in the covariate x on the latent survival of risk j at t given $\boldsymbol{x}$.*

A graphical illustration of the decomposition for a move from $\boldsymbol{x}_0$ to $\boldsymbol{x}_1$ is given in Figure 3.

Figure 3: Decomposition of Covariate Effect

A: Link Function Effect
B: Duration Effect
C: (Total) Covariate Effect
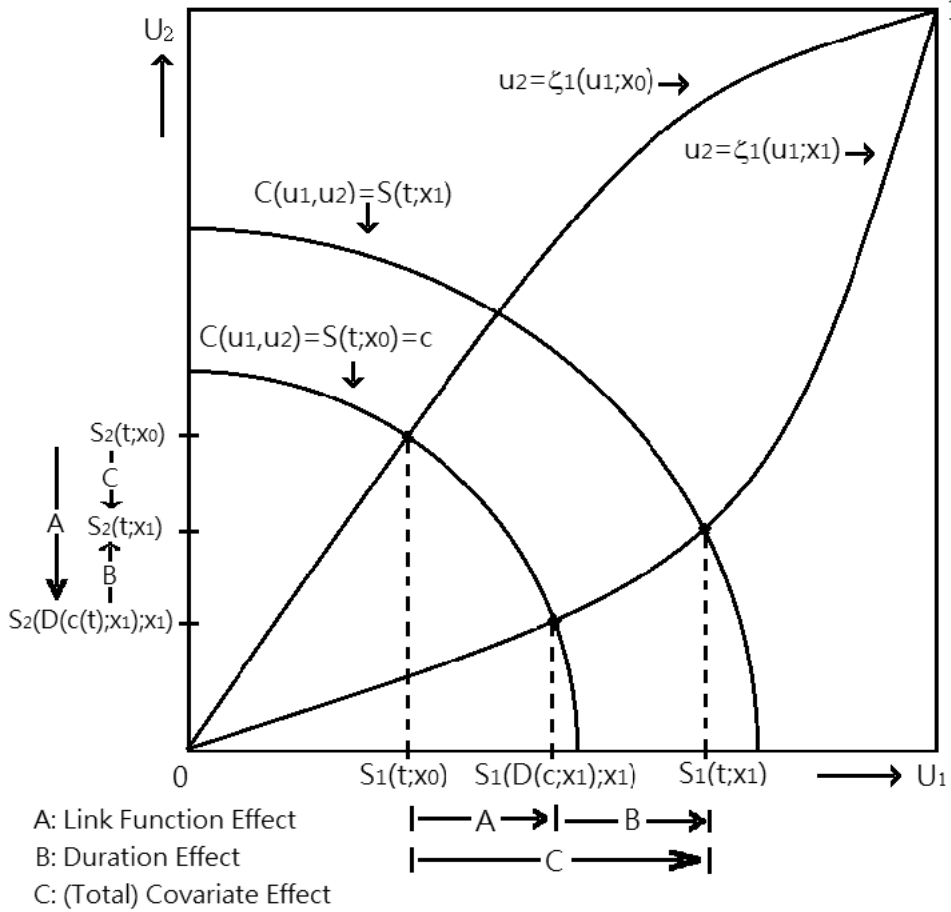
Let *sign* $|z|$ be the sign operator of z. This means that it is $+1$, 0, or -1 if z is positive, zero, or negative respectively.

Lemma 1 *Under Assumptions 1 and 2 the following holds for the competing risks model characterised by equations (2)-(3):*

1. *There is a unique decomposition of the covariate effect on $S_j(t; \boldsymbol{x})$ for all t and $\boldsymbol{x}$ and $j = 1, 2$:*

$$\Delta_x S_j(t; \boldsymbol{x}) = \Delta^l_x S_j(t; \boldsymbol{x}) + \Delta^d_x S_j(t; \boldsymbol{x}). \tag{8}$$

2. *$sign\, |\Delta^l_x S_1(t; \boldsymbol{x})| = -\, sign\, |\Delta^l_x S_2(t; \boldsymbol{x})| = -\, sign\, |\Delta_x \zeta_1(t; \boldsymbol{x})|$ for all t and $\boldsymbol{x}$.*

10

3. $sign\,|\Delta_x^d S_1(t; \boldsymbol{x})| = sign\,|\Delta_x^d S_2(t; \boldsymbol{x})|$ *for all t and $\boldsymbol{x}$.*

4. $sign\,|\Delta_x S_j(t; \boldsymbol{x})|$ *for at least one risk j can be determined by $sign\,|\Delta_x^l S_1(t; \boldsymbol{x})|$ and $sign\,|\Delta_x^d S_1(t; \boldsymbol{x})|$ for all t and $\boldsymbol{x}$.*

We provide a sketch of the proof in Appendix A.I.

Lemma 1.4 suggests that the direction of the covariate effect can be identified when it is known that the link function effect and the duration effect do not have opposite signs. Although the sign of the link function effect and the duration effect are unknown, we show next that the link function (duration) effect of the latent survivals can be identified by the link function (duration) effect of the CIF for some subsets of t. For this reason we define a similar decomposition for Q_j.

Definition 10 *The Hicksian cumulative incidence function, $Q_j^*(c(t); \boldsymbol{x})$, is defined by the value of the cumulative incidence function when the value of the overall survival is fixed at $c(t) = S(t; \boldsymbol{x})$, i.e.*

$$Q_j^*(c(t); \boldsymbol{x}) = Q_j(D(c(t); \boldsymbol{x}); \boldsymbol{x}). \tag{9}$$

Analogously to equation (7), the covariate effect on Q_j can be decomposed into two parts.

Definition 11 *Let $\Delta_x^l Q_j(t; \boldsymbol{x}) = \partial Q_j^*(c(t); \boldsymbol{x})/\partial x$ be the link function effect of a change in the covariate x on the cumulative incidence function for risk j at t given $\boldsymbol{x}$.*

Definition 12 *Let $\Delta_x^d Q_j(t; \boldsymbol{x}) = -\partial Q_j(t; \boldsymbol{x})/\partial t \times \partial D(c(t); \boldsymbol{x})/\partial x$ be the duration effect of a change in the covariate x on the cumulative incidence function for risk j at t given $\boldsymbol{x}$.*

Since the duration effect comes solely from the change in the duration while holding the link function constant, and, under Assumption 1, $S_j(t; \boldsymbol{x})$ is a decreasing function in t and $Q_j(t; \boldsymbol{x})$ is an increasing function in t, it is immediately clear that the sign of the duration effect of the latent survival is always opposite to the sign of the duration effect of the cumulative incidence function.

Lemma 2 *Under Assumptions 1 and 2 and for $j = 1, 2$ we have for the competing risks model characterised by equations (2)-(3): $sign\,|\Delta_x^d S_j(t; \boldsymbol{x})| = -\,sign\,|\Delta_x^d Q_j(t; \boldsymbol{x})|$ for $j = 1, 2$ and for all t and $\boldsymbol{x}$.*

11

The equivalent result can be established for the link function effect but it requires an additional monotonicity assumption.

Assumption 3 *The link function is a monotonic function in x, i.e. $\Delta_x \zeta_1(t; \boldsymbol{x}) < 0$ or > 0 for all t and $\boldsymbol{x}$.*

Given Lemma 1.2, Assumption 3 implies that the link function effect on S_j has the same direction for all t. Specifically, $\Delta_x^l S_1(t; \boldsymbol{x}) < 0$ for all t implies $\Delta_x \zeta_1(t; \boldsymbol{x}) > 0$ for all t and vice versa. As the effect of the covariate on the two arguments of $Q_1^*(c; \boldsymbol{x}) = \Pr(U_1 > S_1^*(c; \boldsymbol{x}), U_2 \leq \zeta_1(U_1; \boldsymbol{x}))$ leads to the same direction of the change in $Q_1^*(c; \boldsymbol{x})$, the sign of $\Delta_x^l Q_1(t; \boldsymbol{x})$ can be unambiguously determined. This can be summarised as follows:

Lemma 3 *Under Assumptions 1, 2 and 3 and for $j = 1, 2$, we have for the competing risks model characterised by equations (2)-(3):*

1. *$sign |\Delta_x^l S_j(t; \boldsymbol{x})| = - sign |\Delta_x^l Q_j(t; \boldsymbol{x})|$ for all t and $\boldsymbol{x}$.*

2. *$sign |\Delta_x S_j(t; \boldsymbol{x})|$ can be determined by $sign |\Delta_x^l Q_j(t; \boldsymbol{x})|$ and $sign |\Delta_x^d Q_j(t; \boldsymbol{x})|$ for at least one of the risks for all t and $\boldsymbol{x}$.*

The proof can be found in Appendix A.I.

Intuitively speaking, Assumption 3 is similar to the monotonicity assumption in the Roy model as discussed by Park (2015). It implies that the propensity of risk 1 will either increase or decrease for all $t \in \mathbb{R}_+$ when a covariate changes. It can be shown that some popular duration models, e.g. the accelerated failure time model, are compatible with the restrictions of Assumption 3. While convenient, Assumption 3 is rather restrictive in applications. For instance, unemployment research has found that the hazard rate of being recalled to the previous employer and the hazard for taking up a new job have very different patterns of duration dependence (see e.g. Alba-Ramirez, Arranz, and Munoz-Bullon, 2007). Specifically, as unemployment duration increases, the hazard of finding a new job remains relatively high, but the recall hazard rate drops quickly and becomes very low. Different individual and job characteristics will affect the relative propensity of recall and new job in different directions depending on the length of unemployment duration. This then violates Assumption 3. While Assumption 3

is the key identification assumption in Park (2015), a relaxation of Assumption 3 in our model will only restrict the validity of Lemma 3 from the entire support of t to some subsets of t.

Relaxing Assumption 3 implies that the link function can change direction at some t. We define the sequence of zero-cutting point(s) of the link function as follows.

Definition 13 *Let $\{\dot{t}_k\}$ for $k = 1, 2, 3, \ldots$ be a sequence of t such that $\dot{t}_1 = 0$ and, for all $k > 1$, $\Delta_x \zeta_j(\dot{t}_k; \boldsymbol{x}) = 0$ and there exists some $\epsilon > 0$ such that $\Delta_x \zeta_j(s; \boldsymbol{x}) \neq 0$ for all $s \in [\dot{t}_k - \epsilon, \dot{t}_k)$. These conditions imply that if $\Delta_x \zeta_j(t; \boldsymbol{x}) = 0$ for all $t \in [t_a, t_b]$ for some $t_b > t_a \geq 0$, only the left end point of this interval t_a enters the sequence $\{\dot{t}_k\}$.*

This sequence is the same for all j because the covariate effect on the link function is zero for both j at $\{\dot{t}_k\}$ due to Lemma 1.2. Since the link function is unidentified, $\{\dot{t}_k\}$ is unidentified. But for $\Delta_x^l Q_j(t; \boldsymbol{x})$, its zero cut-off point(s) and first local turning (maximum or minimum) point(s) after the zero cut-off point(s) are all identified. These observable quantities can be used to identify the direction of the link function effect on S_j for some subset of t.

Definition 14 *(i) Let $\{\acute{t}_k\}$ for $k = 1, 2, 3, \ldots$ be a sequence of t such that $\Delta_x^l Q_j(\acute{t}_k; \boldsymbol{x}) = 0$ and there exists some $\epsilon > 0$ such that $\Delta_x^l Q_j(s; \boldsymbol{x}) \neq 0$ for all $s \in (\acute{t}_k, \acute{t}_k + \epsilon]$. These conditions imply that if $\Delta_x^l Q_j(t; \boldsymbol{x}) = 0$ for all $t \in [t_a, t_b]$ for some $t_b > t_a \geq 0$, only the right end point of this interval t_b enters the sequence $\{\acute{t}_k\}$. (ii) Let $\{\grave{t}_{j,k}\}$ for $k = 1, 2, 3, \ldots$ be a sequence of t such that $\grave{t}_{j,k} = \inf\{t \in (\acute{t}_k, \acute{t}_{k+1}) : \Delta_x^l Q_j(t; \boldsymbol{x}) \geq \Delta_x^l Q_j(s; \boldsymbol{x}) \text{ or } \Delta_x^l Q_j(t; \boldsymbol{x}) \leq \Delta_x^l Q_j(s; \boldsymbol{x}) \text{ for all } s \in [t - \epsilon, t + \varepsilon] \text{ for some } \epsilon, \varepsilon > 0\}$. (iii) $\mathbb{I}_{j,k} = [\acute{t}_k, \grave{t}_{j,k}]$ and $\mathbb{I}_j = \bigcup_{k \geq 1} \mathbb{I}_{j,k}$ for $j = 1, 2$ and $k = 1, 2, 3, \ldots$.*

The sequence $\{\acute{t}_k\}$ is the same for all j because the link function effect on Q_j is zero for both j at $\{\acute{t}_k\}$ due to Lemma 1.2 and Lemma 3.1. Since $\Delta_x^l Q_j(t; \boldsymbol{x})$ is identified, $\{\acute{t}_k\}$, $\{\grave{t}_{j,k}\}$, $\mathbb{I}_{j,k}$, and $\mathbb{I}_j$ are identifiable. The following lemma establishes that in absence of Assumption 3 the validity of Lemma 3 is restricted to $t \in \mathbb{I}_j$.

Lemma 4 *Under Assumptions 1 and 2 and for $j = 1, 2$, we have for the competing risks model characterised by equations (2)-(3):*

   1. For $k = 1, 2, 3, \ldots$, $\mathbb{I}_{j,k}$ is a subset in the interval $[\dot{t}_l, \dot{t}_{l+1}]$ for some l.

   2. $sign |\Delta_x^l S_j(t; \boldsymbol{x})| = - sign |\Delta_x^l Q_j(t; \boldsymbol{x})|$, for all $t \in \mathbb{I}_j$.

13

Each set $I\!I_{j,k}$ is contained in one of the interval $[\dot{t}_l, \dot{t}_{l+1}]$ for $l = 1, 2, 3, \ldots$. However, not every interval in the sequence $[\dot{t}_l, \dot{t}_{l+1}]$ will contain an element in the sequence $I\!I_{j,k}$. For more details see the remark on Lemma 4.1 in Appendix A.I. The proof of Lemma 4 can be found in Appendix A.I.

Let us denote $\boldsymbol{\Delta_x Q_j}(\mathbf{t}; \mathbf{x}) = [\Delta_x^l Q_j(t; \boldsymbol{x}), \Delta_x^d Q_j(t; \boldsymbol{x})]'$ for $j = 1, 2$. $\boldsymbol{\Delta_x Q_j}(\mathbf{t}; \mathbf{x}) \gneq \mathbf{0}$ means that both of $\Delta_x^l Q_j(t; \boldsymbol{x})$ and $\Delta_x^d Q_j(t; \boldsymbol{x})$ are non-negative but that at least one is non-zero. $\boldsymbol{\Delta_x Q_j}(\mathbf{t}; \mathbf{x}) \lneq \mathbf{0}$ is defined analogously.

Definition 15 $I\!D_j$ consists of all t such that $\Delta_x^l Q_j(t; \boldsymbol{x}) \times \Delta_x^d Q_j(t; \boldsymbol{x}) \geq 0$.

Then $I\!D_j$ consists of all t such that $\Delta_x^l Q_j(t; \boldsymbol{x})$ and $\Delta_x^d Q_j(t; \boldsymbol{x})$ do not have the opposite sign.

Definition 16 $G_j = I\!I_j \bigcap I\!D_j$.

We now state our main identification result for the identification set $\mathbb{G}_j$.

Proposition 1 *Under Assumptions 1 and 2 and for $j = 1, 2$, the sign of the covariate effect on S_j is identified in the competing risks model characterised by equations(2)-(3) for all $t \in \mathbf{G}_j$:*

$$
sign\,|\Delta_x S_j(t; \boldsymbol{x})| \;=\; \begin{cases} -1 & if \;\; \boldsymbol{\Delta_x Q_j}(\mathbf{t}; \mathbf{x}) \gneq \mathbf{0}; \\ 0 & if \;\; \boldsymbol{\Delta_x Q_j}(\mathbf{t}; \mathbf{x}) = \mathbf{0}; \\ +1 & if \;\; \boldsymbol{\Delta_x Q_j}(\mathbf{t}; \mathbf{x}) \lneq \mathbf{0}. \end{cases} \tag{10}
$$

Proposition 1 follows directly from Lemmas 1, 2, and 4.

2.1 Increasing the identification set

In this section we consider approaches to increase the set of durations for which the direction of the covariate effect is identified:

1. Reversed application of our decomposition approach.

2. Bounding unknown functionals without copula invariance (Peterson, 1976).

3. Bounding unknown functionals with copula invariance (Bond and Shaw, 2006).

14

The considered approaches are appealing because they do not require additional restrictions on the model. Since these are only applicable for discrete covariates, we focus in the following on the partial effect of a discrete x_k moving from $\boldsymbol{X} = \boldsymbol{x}_0$ to $\boldsymbol{X} = \boldsymbol{x}_1$. Again, we suppress the index k for the ease of notation. We restate a number of definitions in analogy to Section 2.

Definition 17 *For a change in x inducing a movement from $\boldsymbol{x}_0$ to $\boldsymbol{x}_1$*

- *the covariate effect on S_j (compare Definition 4) is $\Delta_x S_j(t; \boldsymbol{x}_0) = S_j(t; \boldsymbol{x}_1) - S_j(t; \boldsymbol{x}_0)$,*

- *the covariate effect on the link function (compare Definition 5) is $\Delta_x \zeta_1(t; \boldsymbol{x}_0) = \zeta_1(S_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1)) - \zeta_1(S_1(t; \boldsymbol{x}_0); \boldsymbol{x}_0)$,*

- *the duration function (compare Definition 6) to keep the overall survival at $c = S(t; \boldsymbol{x}_0)$ is $D(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$,*

- *the link function effect (compare Definition 8) is $\Delta_x^l S_j(t; \boldsymbol{x}_0) = S_j(D(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1); \boldsymbol{x}_1) - S_j(t; \boldsymbol{x}_0)$,*

- *the duration effect (compare Definition 9) is $\Delta_x^d S_j(t; \boldsymbol{x}_0) = S_j(t; \boldsymbol{x}_1) - S_j(D(S(t; \boldsymbol{x}_0); x_1); \boldsymbol{x}_1)$,*

- *the link function effect of Q_j (compare Definition 11) is $\Delta_x^l Q_j(t; \boldsymbol{x}_0) = Q_j(D(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1); \boldsymbol{x}_1) - Q_j(t; \boldsymbol{x}_0)$,*

- *the duration effect of Q_j (compare Definition 12) is $\Delta_x^d Q_j(t; \boldsymbol{x}_0) = Q_j(t; \boldsymbol{x}_1) - Q_j(D(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1); \boldsymbol{x}_1)$.*

It is straightforward to restate Proposition 1 for the case of discrete covariates and a presentation is therefore omitted. Instead, we focus on how the identification set can be increased.

(1) Decomposition of the reversed covariate effect   A simple expansion of the identification set can be achieved by applying our proposed decomposition to the reversed covariate effect.

Definition 18 *The reversed covariate effect on S_j for a discrete movement from $\boldsymbol{x}_0$ to $\boldsymbol{x}_1$ is*

$$\Delta_{-x} S_j(t; \boldsymbol{x}_0) = S_j(t; \boldsymbol{x}_0) - S_j(t; \boldsymbol{x}_1) = -\Delta_x S_j(t; \boldsymbol{x}_0). \tag{11}$$

Clearly, $\Delta_{-x}S_j(t; \boldsymbol{x}_0)$ has the opposite sign than $\Delta_x S_j(t; \boldsymbol{x}_0)$. It is also obvious that Proposition 1 can be carried over to the reversed covariate effect by exchanging the notation $\boldsymbol{x}_1$ and $\boldsymbol{x}_0$. We denote this property as independence of the decomposition route. Let $\mathbb{G}_j(x)$ and $\mathbb{G}_j(-x)$ be the identification sets for $sign\,|\Delta_x S_j(t; \boldsymbol{x}_0)|$ and $sign\,|\Delta_{-x}S_j(t; \boldsymbol{x}_0)|$ respectively. We obtain the following useful result:

Corollary 1 $\mathbb{G}_j(x) \neq \mathbb{G}_j(-x)$.

The proof is given in Appendix A.I. We show in the proof that there exists some set of t for which the sign of the covariate effect is unidentified, while the sign of the reversed covariate effect is identified. Corollary 1 suggests that it is always better to compute both decomposition routes and take the union of the two identification sets.

The identification set can be further enlarged by applying an approach that relies on bounds for unknown functionals.

(2) Peterson Bounds. Peterson bounds can be constructed by applying the Fréchet-Hoeffding bounds for the joint survival distribution in (1), i.e.

$$W(S_1(t_1; \boldsymbol{x}), S_2(t_2; \boldsymbol{x})) \quad \leq C(S_1(t_1; \boldsymbol{x}), S_2(t_2; \boldsymbol{x}); \boldsymbol{x}) \leq \quad M(S_1(t_1; \boldsymbol{x}), S_2(t_2; \boldsymbol{x})), \qquad (12)$$

with $W(s_1, s_2) = \max\{s_1 + s_2 - 1, 0\}$ is the lower Fréchet-Hoeffding bound for the copula function and $M(s_1, s_2) = \min\{s_1, s_2\}$ is the upper Fréchet-Hoeffding bound. The lower (upper) Fréchet-Hoeffding bound corresponds to the case where S_1 and S_2 are perfectly negatively (positively) correlated. The corresponding bounds for the latent survivals in (2)-(3) are the Peterson bounds, i.e.

$$S(t; \boldsymbol{x}) \leq S_j(t; \boldsymbol{x}) \leq 1 - Q_j(t; \boldsymbol{x}). \qquad (13)$$

The upper (lower) bound in (13) is attained when the copula attains its upper (lower) Fréchet-Hoeffding bound and the copula attains its lower (upper) Fréchet-Hoeffding bound and they do not require Assumption 2. C is therefore allowed to vary freely in $\boldsymbol{x}$. Thus, competing risks may be perfectly positively correlated with one value of $\boldsymbol{x}$ while they are perfectly negatively correlated at another value of $\boldsymbol{x}$.

16

For a discrete movement from $\boldsymbol{x}_0$ to $\boldsymbol{x}_1$, the Peterson bounds for $\Delta_x S_j(t; \boldsymbol{x}_0)$ are given by

$$-\Delta_x Q_j(t; \boldsymbol{x}_0) - Q_i(t; \boldsymbol{x}_0) \leq \ \Delta_x S_j(t; \boldsymbol{x}_0) \ \leq -\Delta_x Q_j(t; \boldsymbol{x}_0) + Q_i(t; \boldsymbol{x}_1), \qquad (14)$$

for $i \neq j$. Equivalent bounds for the covariate effect of continuous x cannot be derived.

Definition 19 *Let $\mathbb{P}_j$ be the identification set for $sign\,|\Delta_x S_j(t; \boldsymbol{x}_0)|$ obtained by the Peterson bounds. $\mathbb{P}_j$ consists of all t for which $-\Delta_x Q_j(t; \boldsymbol{x}_0) - Q_i(t; \boldsymbol{x}_0) > 0$ or $-\Delta_x Q_j(t; \boldsymbol{x}_0) + Q_i(t; \boldsymbol{x}_0) < 0$ for $j \neq i$ or the two former being equal to zero.*

One characteristic of the Peterson bounds in (14) is that the difference between the lower and upper bound, i.e. $-\Delta_x Q_j(t; \boldsymbol{x}_0) + Q_i(t; \boldsymbol{x}_1) - (-\Delta_x Q_j(t; \boldsymbol{x}_0) - Q_i(t; \boldsymbol{x}_0)) = Q_i(t; \boldsymbol{x}_1) + Q_i(t; \boldsymbol{x}_0)$, is an increasing function of t. This implies that the bounds in (14) tend to be less informative for greater values of t. In contrast, the identification set in Proposition 1 is a function of $\Delta_x^l Q_j(t; \boldsymbol{x})$ and $\Delta_x^d Q_j(t; \boldsymbol{x})$, which are generally not monotonic in t. Thus there are no mechanics which make our decomposition approach less informative for greater t. However, it is possible that some subsets of $\mathbb{P}_j$ are not included in $\mathbb{G}_j(x)$. For instance, consider some $t \in \mathbb{P}_1 \cap \mathbb{P}_2$ such that $sign\,|\Delta_x S_j(t; \boldsymbol{x}_0)|$ is identified as positive for both $j = 1, 2$; but from Lemma 1, the sign of only one risk can be identified with our decomposition approach. This appears to be a limitation of our decomposition approach. However, by applying the reversed decomposition there may be some durations for which the sign of the covariate effect is identified for either risk. An example is illustrated with simulations in Section 3.

(3) Bounds for Covariate Time Transformation. Bond and Shaw (2006) consider the so-called covariate-time transformation for a discrete movement from $\boldsymbol{x}_0$ to $\boldsymbol{x}_1$ under Assumptions 1 and 2.

Definition 20 *The covariate-time transformation (CTT) is $\phi_j(t; \boldsymbol{x}_0) = S_j^{-1}(S_j(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$.*

The difference between $\phi_j(t; \boldsymbol{x}_0)$ and t can be interpreted as the $S_j(t; \boldsymbol{x}_0)$-quantile treatment effect on the latent duration. The sign of this difference also corresponds to the direction of the covariate effect, i.e. $sign\,|\Delta_x S_j(t; \boldsymbol{x}_0)| = sign\,|\phi_j(t; \boldsymbol{x}_0) - t|$. But as the S_j's are not identified, the CTT are also unidentified. Bond and Shaw (2006) show that the CTT can be bounded provided that the following order assumption holds:

17

Assumption 4 $\phi_2(t; \boldsymbol{x}_0) < \phi_1(t; \boldsymbol{x}_0)$ *or* $\phi_2(t; \boldsymbol{x}_0) > \phi_1(t; \boldsymbol{x}_0)$ *for all* $t \geq 0$.

The following result establishes the equivalence of this order assumption and the monotonicity of the link function (Assumption 3).

Lemma 5 $\zeta_1(u_1; \boldsymbol{x}_0) < \zeta_1(u_1; \boldsymbol{x}_1)$ *iff* $\phi_2(t; \boldsymbol{x}_0) > \phi_1(t; \boldsymbol{x}_0)$ *at any* $u_1 = S_1(t; \boldsymbol{x}_0)$, *and vice versa.*

According to Lemma 5, the order assumption has the same implication for the competing risks model and the Roy model as Assumption 3 (see above for the discussion of Assumption 3). Bounds for the CTT can then be derived by using Assumption 3 and by exploiting the link between the observable changes in the Q_j's and the changes in the unobservable S_j's. Suppose that $\phi_2(t; \boldsymbol{x}_0) \leq \phi_1(t; \boldsymbol{x}_0)$ for all t and thus $\zeta_1(u_1; \boldsymbol{x}_1) \leq \zeta_1(u_1; \boldsymbol{x}_0)$ for all u_1, we have

$$
\begin{aligned}
Q_1(\phi_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1) &= \int_{S_1(\phi_1(t;\boldsymbol{x}_0);\boldsymbol{x}_1)}^{1} \int_{0}^{\zeta_1(u_1;\boldsymbol{x}_1)} \kappa(u_1, u_2) \, du_2 \, du_1 \\
&\leq \int_{S_1(t;\boldsymbol{x}_0)}^{1} \int_{0}^{\zeta_1(u_1;\boldsymbol{x}_0)} \kappa(u_1, u_2) \, du_2 \, du_1 = Q_1(t; \boldsymbol{x}_0); \text{ and} \quad (15) \\
S(\phi_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1) &= \int_{0}^{S_1(\phi_1(t;\boldsymbol{x}_0);\boldsymbol{x}_1)} \int_{0}^{\zeta_1(S_1(\phi_1(t;\boldsymbol{x}_0);\boldsymbol{x}_1);\boldsymbol{x}_1)} \kappa(u_1, u_2) \, du_2 \, du_1 \\
&\leq \int_{0}^{S_1(t;\boldsymbol{x}_0)} \int_{0}^{\zeta_1(S_1(t;\boldsymbol{x}_0);\boldsymbol{x}_0)} \kappa(u_1, u_2) \, du_2 \, du_1 = S(t; \boldsymbol{x}_0). \quad (16)
\end{aligned}
$$

The bounds for the CTT are therefore

$$
S^{-1}(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1) \leq \phi_1(t; \boldsymbol{x}_0) \leq Q_1^{-1}(Q_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1) \quad (17)
$$

$$
Q_2^{-1}(Q_2(t; \boldsymbol{x}_0); \boldsymbol{x}_1) \leq \phi_2(t; \boldsymbol{x}_0) \leq S^{-1}(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1). \quad (18)
$$

For a given t the sign of the covariate effect on S_j may be obtained as follows: If the minimum of $Q_j^{-1}(Q_j(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ and $S^{-1}(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ is greater than t, one can conclude that $\phi_j(t; \boldsymbol{x}_0) > t$ and thus $S_j^{-1}(S_j(t; \boldsymbol{x}_0); \boldsymbol{x}_1) > t$. It follows that $S_j(t; \boldsymbol{x}_0) < S_j(t; \boldsymbol{x}_1)$ which implies a positive covariate effect on S_j. Similarly, the covariate effect on S_j is negative when the maximum of $Q_j^{-1}(Q_j(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ and $S^{-1}(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ is smaller than t. Otherwise the direction of the covariate effect on S_j is not identified. Compared with our decomposition approach this approach has two disadvantages: First, similar to the Peterson bounds, the difference between

the bounds for the CTT increase when t becomes greater, making it less likely that the sign of the covariate effect can be identified. Specifically, when t approaches infinity, the lower (upper) bound of $\phi_1(t; \boldsymbol{x}_0)$ ($\phi_2(t; \boldsymbol{x}_0)$) approaches infinity. Second, (15) and (16) are only valid under the restrictions of Assumption 3 which cannot be verified in an application. While it is possible to detect some rejections of the order assumption, an observed rejection does not constitute a sufficient condition. In particular, whenever the values of $Q_1^{-1}(Q_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ and $Q_2^{-1}(Q_2(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ in (17) and (18) change their order at some observed t^*, it can be certain that the order assumption is violated at some $t < t^*$. But, the mere fact that $Q_1^{-1}(Q_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ and $Q_2^{-1}(Q_2(t; \boldsymbol{x}_0); x_1)$ do not change their order before t^* does not imply that there is no violation at some $t < t^*$.

In what follows we propose a modification of the approach by Bond and Shaw which does not require Assumption 3. Instead it uses the observation that the order assumption is not violated for $t \in [0, \grave{t}_{j,1}]$. This is a consequence of Lemma 5, keeping Definitions 13 and 14 in mind. The set of t for which the sign of the covariate effect is identified is then:

Definition 21 *Let $\mathbb{B}_j(\boldsymbol{x})$ be the identification set for which sign $|\Delta_x S_j(t; \boldsymbol{x}_0)|$ is identified by the modified Bond and Shaw's approach under Assumptions 1 and 2. $\mathbb{B}_j(x)$ consists of all $t \in [0, \grave{t}_{j,1}]$ s.t. $\min\{Q_j^{-1}(Q_j(t; \boldsymbol{x}_0); \boldsymbol{x}_1), S^{-1}(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1)\} > t$ or $\max\{Q_j^{-1}(Q_j(t; \boldsymbol{x}_0); \boldsymbol{x}_1), S^{-1}(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1)\} < t$ or the former two being equal to zero. Similarly, let $\mathbb{B}_j(-x)$ be the identification set for sign $|\Delta_{-x} S_j(t; \boldsymbol{x}_0)|$ derived by the modified Bond and Shaw's approach.*

To sum up we have now defined five identification sets for risk j: $\mathbb{P}_j$, $\mathbb{B}_j(x)$, $\mathbb{B}_j(-x)$, $\mathbb{G}_j(x)$ and $\mathbb{G}_j(-x)$. The overall identification set is obtained by taking their union:

Definition 22 *For $j = 1, 2$ let $\mathbb{U}_j$ be the set of t for which the sign of the covariate on S_j is identified by at least one of the approaches:*

$$
\begin{aligned}
\mathbb{U}_j &= \mathbb{U}_j(x) \bigcup \mathbb{U}_j(-x) \ with \\
\mathbb{U}_j(x) &= \mathbb{P}_j \bigcup \mathbb{B}_j(x) \bigcup \mathbb{G}_j(x) \\
\mathbb{U}_j(-x) &= \mathbb{P}_j \bigcup \mathbb{B}_j(-x) \bigcup \mathbb{G}_j(-x).
\end{aligned}
$$

In Sections 3 and 4 we explore with simulations and data examples how the size of the sets $\mathbb{P}_j$, $\mathbb{B}_j$, and $\mathbb{G}_j$ compare in practice and whether the size of $\mathbb{U}_j$ is large enough to obtain practically

19

informative results.

2.2 Identifiability in a multi-risks model

In this subsection we extend the model of Section 2 to a model with a finite number of risks $J(> 2)$. The observed failure time becomes $T = \min(T_1, \ldots, T_J)$ and the indicator function is $\delta = \arg\min_j\{T_j\}$. The link function is defined as $u_i = \zeta_{i,j}(u_j; \boldsymbol{x}) = S_i(S_j^{-1}(u_j; \boldsymbol{x}); \boldsymbol{x})$. Equations (2)-(3) becomes

$$
\begin{aligned}
S(t; \boldsymbol{x}) &= C^J(S_1(t; \boldsymbol{x}), \ldots, S_J(t; \boldsymbol{x})); \\
Q_j(t; \boldsymbol{x}) &= \int_{S_j(t; \boldsymbol{x})}^1 \int_0^{\zeta_{1,j}(u_j; \boldsymbol{x})} \ldots, \int_0^{\zeta_{j-1,j}(u_j; \boldsymbol{x})} \int_0^{\zeta_{j+1,j}(u_j; \boldsymbol{x})} \\
&\qquad \ldots \int_0^{\zeta_{J,j}(u_j; \boldsymbol{x})} \kappa(u_1, \ldots, u_J) \, du_J \ldots \, du_{j+1} \, du_{j-1} \ldots \, du_1 \, du_j.
\end{aligned}
\tag{19}
$$

The J-copula is

$$
C^J(u_1, \ldots, u_J) = \Pr(S_1(T_1; \boldsymbol{x}) \le u_1, \ldots, S_J(T_J; \boldsymbol{x}) \le u_J).
\tag{20}
$$

To carry over the identification results for the model with $J = 2$, we follow the risk pooling approach by Lo and Wilke (2010). Suppose that we want to identify the sign of the covariate effect on the j'th risk. By conceptually pooling all other risks into a single risk, we generate an unobserved new variable $T_{-j} = \min(T_1, \ldots, T_{j-1}, T_{j+1}, \ldots, T_J)$. This is then a two risks model with a 2-copula

$$
C^2(u_j, u_{-j}) = \Pr(S_j(T_j; \boldsymbol{x}) \le u_j, S_{-j}(T_{-j}; \boldsymbol{x}) \le u_{-j}).
\tag{21}
$$

The unknown marginal survival function for the pooled variable T_{-j} is $S_{-j}(t; \boldsymbol{x}) = \Pr(T_{-j} > t; \boldsymbol{x})$. The observed failure time is unaffected as $T = \min(T_j, T_{-j})$, and the indicator function is modified as $\delta_j = j$ if $\delta = j$ and $\delta_j = -j$ if $\delta \ne j$. For any J-copula in (20), the existence of a 2-copula in (21) is guaranteed under the following assumption (Nelsen, 2006).

Assumption 5 *In the competing risks model defined by (19), the copula belongs to the Archimedean class.*

In this case the multi-risk model can be reduced into a two risks model as (2)-(3):

$$
\begin{aligned}
S(t; \boldsymbol{x}) &= C^2(S_j(t, \boldsymbol{x}), S_{-j}(t; \boldsymbol{x})); \\
Q_j(t; \boldsymbol{x}) &= \int_{S_j(t;\boldsymbol{x})}^{1} \int_{0}^{\zeta_{-j,j}(u_j;\boldsymbol{x})} \kappa(u_{-j}, u_j) du_{-j} \ du_j,
\end{aligned}
\tag{22}
$$

with $u_{-j} = \zeta_{-j,j}(u_j; \boldsymbol{x})$ denotes the link function between $S_j(t; \boldsymbol{x})$ and $S_{-j}(t; \boldsymbol{x})$. For more details see Lo and Wilke (2010). The identification approaches for the two risks model in Section 2 can therefore be subsequently applied to (22) for $j = 1, \ldots, J$, where the order of application does not matter. Note, however, that only the non-pooled risk is of interest in the pooled risks model as a pooled risk is generally uninformative.

3 Simulation Study

In this section we explore the practical performance of the methods outlined in Section 2 with the help of a simulation study. We consider a two risks model with a known closed form representation of the entire competing risks model. This means Q_j, S, S_j and C for $j = 1, 2$ are fully known. We consider the closed form expression given in Rivest and Wells (2001) for an Archimedean copula generator with parameters θ, $\phi_\theta(s)$, and the known cause-specific crude hazard functions, $\lambda_j(t; x)$, $j = 1, 2$. For simplicity, we consider a model with one binary x. We have $S(t; x) = \exp\left[-\int_0^t \lambda_1(u; x) + \lambda_2(u; x) \ du\right]$ and $Q_j(t; x) = \int_0^t \lambda_j(u; x) S(u; x) \ du$, and

$$
S_j(t; x) = \phi_\theta^{-1}\left[-\int_0^t \phi_\theta'[S(u; x)] S(u; x) \lambda_j(u; x) \ du\right].
\tag{23}
$$

We consider two one-parameter copulas in our simulations: Frank and Clayton (see e.g. Nelsen, 2006, for details). These copulas are characterised by different tail dependencies: the Frank copula has no upper and lower tail dependence, while the Clayton copula has lower tail dependence. The copula generators ϕ_θ are given in Table 1. We consider four simulation designs in which θ of these copulas is chosen such that Kendall's τ equals to the following four values: -0.8, -0.4, 0.4 and 0.8.

We consider three specifications of the cause-specific crude hazard functions:

(i) Odd-rate transformation model (Dabrowska and Doksum, 1988) with Weibull baseline

Table 1: Copula generators of Frank and Clayton copula.

Copula	Copula generator	Support of parameter
Frank	$\phi_\theta(s) = -\ln((\exp(-\theta s) - 1)/(\exp(-\theta) - 1))$	$\theta \in (-\infty, \infty) \setminus \{0\}$
Clayton	$\phi_\theta(s) = (s^{-\alpha} - 1)\theta$	$\theta > 0$

Table 2: Parameters of the simulated competing risks model.

Models	Parameters	Risk $j = 1$		Risk $j = 2$	
		$x = x_0$	$x = x_1$	$x = x_0$	$x = x_1$
Odd-rate transformation model	$(\nu_j, \rho_j, \gamma_j)$	(1, 10, 1)	(1,2,2)	(0.5, 1,1)	(1.5, 2, 2)
Log-logistic proportional odds model	(ν_j, ρ_j)	(2, 2)	(5,1.8)	(2, 2.5)	(3, 1.5)
Log-normal accelerated failure model	(ν_j, ρ_j)	(1.2, -0.5)	(1,-0.5)	(1.2, 1)	(1.3, 1.5)

such that $\lambda_j(t; x) = \nu_j \rho_j t^{(\rho_j - 1)}(1 + \gamma_j \nu_j t^{\rho_j})^{-1}$ with $\nu_j, \rho_j \in \mathbb{R}_+$ and $\gamma_j \in \mathbb{R}$,

(ii) Log-logistic proportional odds model such that $\lambda_j(t; x) = \nu_j \rho_j (\nu_j t)^{(\rho_j - 1)}(1 + (\nu_j t)^{\rho_j})^{-1}$ with $\nu_j, \rho_j \in \mathbb{R}_+$, and

(iii) Log-normal accelerated failure time model such that $\lambda_j(t; x) = (f((\log t - \rho_j)/\nu_j)(\nu_j t(1 - \mathcal{F}((\log t - \rho_j)/\nu_j)))^{-1})$ with $\rho_j \in \mathbb{R}$, and $\nu_j \in \mathbb{R}_+$ where f and $\mathcal{F}$ are the probability and the cumulative density function of the standardized normal distribution respectively.

Table 2 gives the parameters for the models that we use in our simulations. Since we know the true S, S_j and Q_j for all j, we can easily assess the performance of the different considered identification approaches by comparing their identification sets. We compute $\Delta_x^l Q_j(t; x)$, $\Delta_x^d Q_j(t; x)$, $\Delta_x^l S_j(t; x)$, $\Delta_x^d S_j(t; x)$ and $\Delta_x S_j(t; x)$ for $t \in \{0, 0.005, 0.01, \ldots, 1.995, 2.000\}$ and the sequences $\{\acute{t}_k\}$, $\{\grave{t}_{j,k}\}$.

Figure 4 presents the results for the Frank copula using the odd-rate transformation model with $\tau = 0.4$. Panels (a) and (b) show the covariate effect $(\Delta_x S_j(t; x))$, the link function effect of the CIF $(\Delta_x^l Q_j(t; x))$, the duration effect of the CIF $(\Delta_x^d Q_j(t; x))$, and the Peterson bounds (PB_j). They also report the identification set derived from the Peterson bounds $(\mathbb{P}_j)$, the modified Bond and Shaw approach $(\mathbb{B}_j(x))$, the decomposition approach $(\mathbb{G}_j(x))$, and their union $(\mathbb{U}_j(x))$. The set $\mathbb{P}_j$ is marked as horizontal lines in grey color at the value of the vertical axis of -1.2, while unmarked intervals indicate the range of duration in which the sign is unidentified. Similarly, the identification sets $\mathbb{B}_j(x)$, $\mathbb{G}_j(x)$, and $\mathbb{U}_j(x)$ are marked as horizontal lines in different grey colors at -1.3, -1.4, and -1.6 respectively. The Figure also

reports the direction of the identified sign for t in $\mathbb{U}_j(x)$. When the sign is identified as positive, zero, or negative at t, $IS_j(x)$ (identified sign) is marked as the horizontal lines at 1, 0, -1, respectively at that t. Panels (c) and (d) report equivalent results for the reversed covariate effect $\Delta_{-x}S_j(t; x)$. This includes the corresponding link function effect $(\Delta^l_{-x}Q_j(t; x))$, the duration effect $(\Delta^d_{-x}Q_j(t; x))$, and the identification sets using the Peterson bounds $(\mathbb{P}_j)$, the modified Bond and Shaw approach $(\mathbb{B}_j(-x))$, the decomposition approach $(\mathbb{G}_j(-x))$ and their union $(\mathbb{U}_j(-x))$. Panels (e) and (f) present the union of the various identification approaches in different directions. This means $\mathbb{B}_j = \mathbb{B}_j(x) \bigcup B_j(-x)$, $\mathbb{G}_j = \mathbb{G}_j(x) \bigcup \mathbb{G}_j(-x)$, $\mathbb{U}_j = \mathbb{U}_j(x) \bigcup \mathbb{U}_j(-x)$, and $IS_j = IS_j(x) \bigcup IS_j(-x)$.

Figure 4(a) shows that the upper and lower Peterson bounds (PB_1) contain the value zero at all t and thus $\mathbb{P}_1$ is an empty set (unmarked). The identification set for the modified Bond and Shaw approach, $\mathbb{B}_1(x)$ (compare Definition 21), is restricted to $t \in [0, \grave{t}_{1,1}]$. This is confirmed in Panel (a). $\mathbb{B}_1(x)$ does not contain values of t greater than $\grave{t}_{1,1}$, this is the time at which the maximum of the link function effect $\Delta^l_x Q_1$ occurs in the interval $[\grave{t}_1, \acute{t}_2]$. In contrast, the sign of the covariate effect is also identified for $t \in [\acute{t}_2, \grave{t}_{1,2}]$ when the decomposition approach is used. For this reason, $\mathbb{G}_j(x)$ includes values of t in $[0, \grave{t}_{1,1}] \bigcup [\acute{t}_2, 2]$ for which $\Delta^l_x Q_1(t)$ and $\Delta^d_x Q_1(t)$ have the same direction. Panel (a) shows that the decomposition approach provides the largest identification set and coincides with the union $\mathbb{U}_j(x)$.

Panel (c) shows the same upper and lower Peterson bounds (PB_1) as in Panel (a), as the Peterson bounds are identical in the reversed direction. In contrast $\Delta^l_{-x}Q_1$ and $\Delta^d_{-x}Q_1$ differ compared to Panel (a). The computed identification sets for the modified Bond and Shaw approach $(\mathbb{B}_1(-x))$ and the decomposition approach $(\mathbb{G}_1(-x))$ are therefore different from those in Panel (a). This illustrates the usefulness of Corollary 1. Similar to Panel (a), the identification set of the decomposition approach in Panel (c) is the largest. Panel (e) shows that combining $\mathbb{U}_1(x)$ and $\mathbb{U}_1(-x)$ produces larger identification sets $\mathbb{U}_1$ and IS_1. This illustrates the usefulness of combining the three approaches.

Panel (b) shows the results for risk 2. It can be seen that $\mathbb{P}_2(x)$ and $\mathbb{B}_2(x)$ consist of some t which are not contained in $\mathbb{G}_2(x)$. This is the set of t for which the link function and the duration effect have different directions. At the same time, $\mathbb{G}_2(x)$ includes some set of t which is not included in $\mathbb{P}_2(x)$ and $\mathbb{B}_2(x)$. Similarly, Panels (d) and (f) show that the three approaches partly complement each other. Notably, the decomposition approach is particularly

23

Figure 4: Identified sign (IS) of the covariate effect in a known two risks model using Frank copula and odd-rate transformation model with $\tau = .4$: Risk 1 (left) and Risk 2 (right).
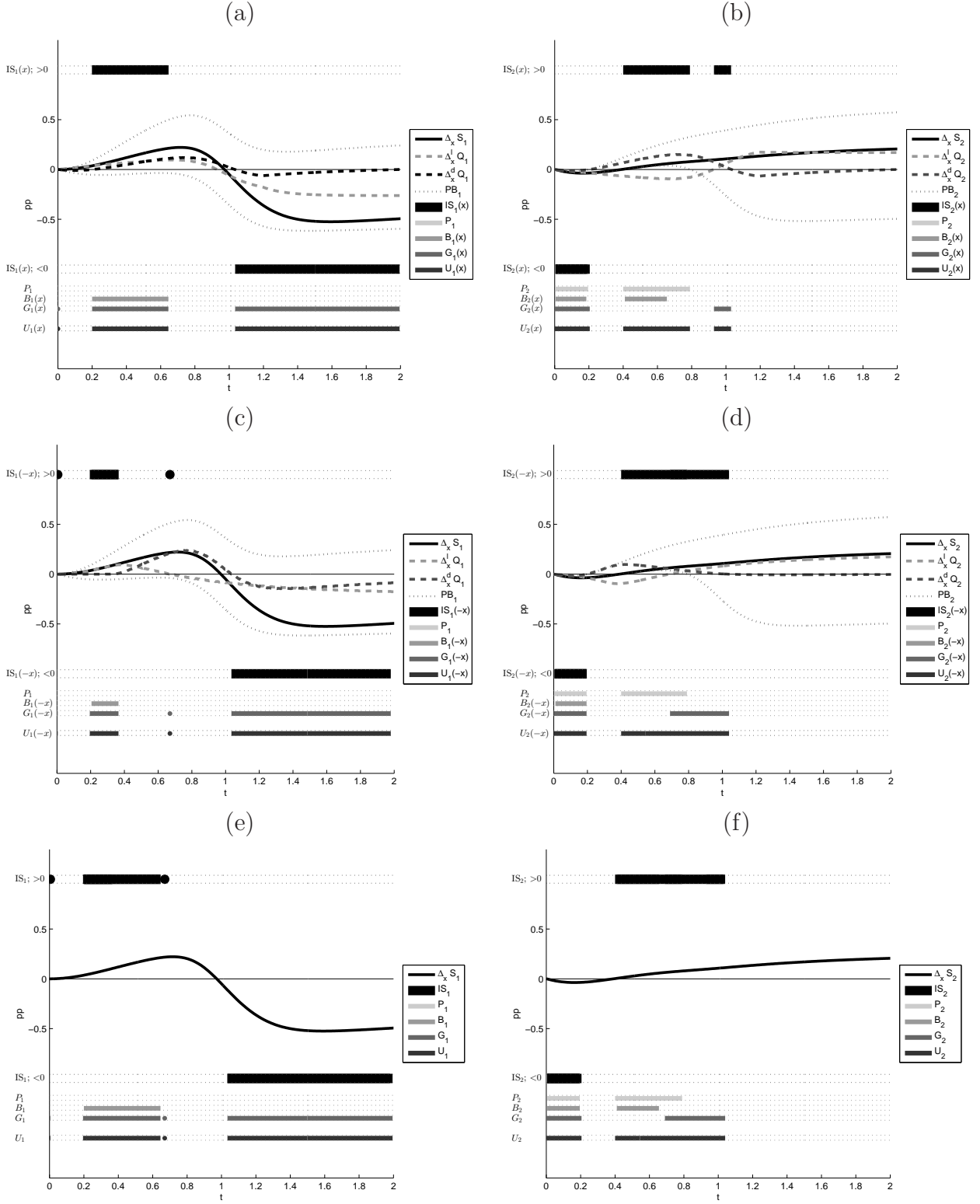
useful for identifying the sign of the covariate effect at longer durations. The application of the decomposition and the reversed decomposition in Figure 4 did not generate sets of durations at which the sign of the effect is identified for both risks at the same time. This means the sets $\mathbb{G}_1$ and $\mathbb{G}_2$ are disjoint.

We also show the results for a Clayton copula and log-logistic proportional odds model in Figure S1 in the supplementary material. The result patterns are similar to those in Figure 4 but there is an overlap of the sets $\mathbb{G}_1$ and $\mathbb{G}_2$ for $t \in [.18, .22]$, meaning that the direction of the covariate effect is identified for both risks in this interval.

Table 3 summarises the results using different combinations of copulas and marginal distributions with $\tau = .4$. Tables S1-S3 in the supplementary material show more results for $\tau = -0.8$, $\tau = -0.4$ and $\tau = 0.8$. In these tables, the relative size of the identification set using each approach is measured as the size of the respective identification set divided by the size of $\mathbb{T}$ ($\mathbb{T}$ is the support on which the model has been evaluated, see below for details). This is the share of $\mathbb{T}$ identified by the respective approach. The relative size of the identification set that is uniquely identified by each approach is measured by the size of the subset of the respective identification set that is not contained in the identification sets of the other methods (unique of $\mathbb{T}$). The latter shows the additional contribution of each of the approaches to the union of identification sets. If the latter is zero, it means that there is no additional gain from using this method when the other methods have already been applied. Since the support of T has no upper limit, we define set $\mathbb{T}$ as the set of duration from 0 to the value at which one of the S_j becomes very flat, particularly when S_j decreases by less than 1×10^{-6} when t increases by 0.05. The tables show that the order of the relative size of the identification sets using different approaches is quite robust in the choice of τ and C. More specifically, in almost all cases our proposed decomposition approach produces the largest identification set. In the case it does not generate the largest set, it is only marginally smaller than the largest set. Moreover, in all of the cases our proposed decomposition approach makes the largest additional contribution among all methods. It is therefore the method that provides the most distinctive information.

Table 3: Relative size of identification sets in simulated competing risks models ($\tau = 0.4$).

	Frank Copula		Clayton Copula	
	Share of $\mathbb{T}^*$	Unique of $\mathbb{T}^{**}$	Share of $\mathbb{T}^*$	Unique of $\mathbb{T}^{**}$
Odd-rate				
transformation model				
$\mathbb{T}$	[0,3.59]		[0,3.03]	
Risk $j = 1$				
$\mathbb{P}_1$	0%	0%	0%	0%
$\mathbb{B}_1$	12.5%	0%	14.8%	0%
$\mathbb{G}_1$	84.1%	71.6%	81.2%	66.4%
$\mathbb{U}_1$	84.1%		81.2%	
Risk $j = 2$				
$\mathbb{P}_2$	16.4%	2%	14%	4.2%
$\mathbb{B}_2$	12.4%	0%	14.7%	0%
$\mathbb{G}_2$	15.7%	7.4%	18.6%	8.7%
$\mathbb{U}_2$	23.8%		28.2%	
Log-logistic				
proportional odds model				
$\mathbb{T}$	[0,5.805]		[0,2.71]	
Risk $j = 1$				
$\mathbb{P}_1$	6.5%	2%	14%	4.2%
$\mathbb{B}_1$	4.6%	0%	9.7%	0%
$\mathbb{G}_1$	91.2%	87.7%	81.2%	71.5%
$\mathbb{U}_1$	93.2%		85.5%	
Risk $j = 2$				
$\mathbb{P}_2$	7.8%	1.2%	16.8%	2.6%
$\mathbb{B}_2$	1.5%	0%	3.1%	0%
$\mathbb{G}_2$	10.9%	5.8%	23.4%	12.3%
$\mathbb{U}_2$	13.6%		29.1%	
Log-normal				
accelerated failure model				
$\mathbb{T}$	[0,24.59]		[0,19.945]	
Risk $j = 1$				
$\mathbb{P}_1$	1.9%	0%	2.6%	0%
$\mathbb{B}_1$	94%	0.1%	91.8%	0.2%
$\mathbb{G}_1$	92.6%	0.2%	89.8%	0.4%
$\mathbb{U}_1$	94.2%		92.1%	
Risk $j = 2$				
$\mathbb{P}_2$	0%	0%	0%	0%
$\mathbb{B}_2$	8%	0%	11%	0%
$\mathbb{G}_2$	8.2%	0.2%	11.3%	0.3%
$\mathbb{U}_2$	8.2%		11.3%	

*: size of respective identification set divided by the size of $\mathbb{T}$.

**: size of the subset of the respective identification set that is not contained in the identification sets of the other methods divided by the size of $\mathbb{T}$. Measure of *additional* contribution.

4 Data Examples

We illustrate the usefulness of the proposed decomposition approach with the help of two data examples. The sets $\mathbb{P}_j$ and $\mathbb{B}_j$ (and the associated signs of the covariate effect) can be easily determined using the estimated Q_j. The set $\mathbb{G}_j$ is estimated by a multiple-step procedure. First, $Q_j(t; \boldsymbol{x})$ and $S(t; \boldsymbol{x})$ are estimated with a consistent estimator. See for example Kalbfleisch and Prentice (2002) for nonparametric estimation in stratified samples for discrete $\boldsymbol{x}$. For continuous $\boldsymbol{x}$, nonparametric kernel estimator can be applied. See for example Fermanian (2003). These estimates are then plugged into the population equations of Section 2 to obtain their sample analogues. The estimate for $\mathbb{G}_j$ and the direction of the covariate are obtained from the sample analogues of the relevant equations. The estimation procedure is outlined step by step in Appendix A.II. A discussion of asymptotic properties and a bootstrap based inference procedure can be found in the discussion paper version of this paper (Lo and Wilke, 2011).

We use two samples of unemployment durations and consider the effect of various discrete covariates on the probability of taking up employment. All other exit states are pooled into a second risk using the method discussed in Section 2.2. Results for the second risk are not presented because it does not have a direct interpretation. The first sample we use is the illustrating data for competing risks models in the textbook of Cameron and Trivedi (2005). These data have been originally used in the study by McCall (1996) and come from the U.S. Current Population Survey's Displaced Worker Supplement. They contain monthly information on three destination states for the displaced jobless individual. We only consider transitions into the risk "full-time employment". In addition to observed transitions there is independent censoring because not all joblessness periods were terminated by the time of the interview. There are 3,343 observations with 1,073 transitions into full-time employment and 913 transitions into the pooled risk. We report results only for the first 24 months as there are hardly any transitions for longer durations. We consider the effect of three dummy variables: receipt of unemployment insurance (ui), being married ($married$) and being female ($female$). The second sample is extracted from German administrative labour market data. In particular, we use a sample of the scientific use file version of the sample of the integrated labour market biographies (SIAB) 1975-2010 of the Institute for Employment Research (IAB), Germany. These

administrative data are a 2% random sample of the workforce in Germany that contributed to the social insurance in the period 1975-2010. The SIAB contains daily information about periods of dependent employment and claim periods for unemployment compensation along with basic information about the individual (such as gender, wage, age and employment history) and the employing firm (such as business sector and location). For more information on the SIAB see Dorner et al. (2010). From these data we extract all unemployment benefit claim periods starting in 2007 and 2008. We define unemployment benefit duration as receipt of unemployment benefits (Arbeitslosengeld) from the German Federal Employment Agency. This leaves us with a sample of 95,271 observations. We only consider transitions into employment (with contributions to the social insurance). We observe 56,843 transitions into employment and 38,428 transitions into the pooled risk. There is no censoring. We report results only for the first 730 days as there are hardly any transitions to employment beyond that point. For these data we consider the effect of five dummy covariates: unemployment starts during winter period, i.e. November-February (*winter*), unemployment starts in calendar year 2008 (*year2008*), a previous unemployment period of the unemployed had been terminated by a recall to the former employer (*previous recall*), the person has been unemployed in the past (*previous unemployment*) and gender (*female*).

We present estimation results in Table 4 and Figure 5. The table summarises the results for all variables by providing information about the relative size of the identification sets and their uniqueness. It can be seen that the decomposition approach produces the largest identification set in almost all cases. It is also apparent that the decomposition approach in most cases produces the largest additional contribution. These observations confirm the findings of the simulations.

Figure 5 presents more detailed insights for two variables: *ui* (receipt of unemployment insurance, McCall data) and *winter* (SIAB). In particular, it shows plots for the estimated nonparametric cumulative incidence functions ($\hat{Q}_1$) in the left panel. The right panel contains the estimated sign of the covariate effects on the marginal distribution of job finding (ES_1), and the estimated identification sets $\hat{\mathbb{P}}_1$, $\hat{\mathbb{B}}_1$, $\hat{\mathbb{G}}_1$, and $\hat{\mathbb{U}}_1$. ES_1 is plotted as horizontal lines at 0.6, 0 and -0.6 when the effect is estimated as positive, zero, and negative, respectively. $\hat{\mathbb{P}}_1$, $\hat{\mathbb{B}}_1$, $\hat{\mathbb{G}}_1$, and $\hat{\mathbb{U}}_1$ are plotted at -0.7, -0.8, -0.9, and -1 respectively. For comparison these plots also report the estimated covariate effects under the assumption of independent competing

Figure 5: Cumulative incidence for job finding (left) and estimated sign of covariate effects on the marginal distribution (right).

McCall data: receipt of unemployment insurance (ui)
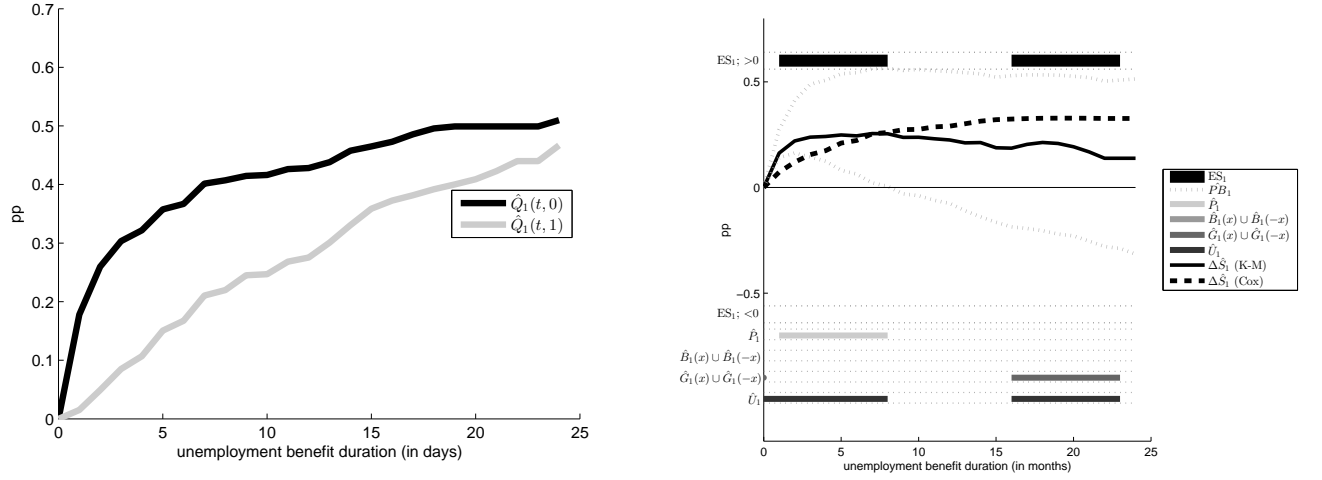
SIAB: *winter*
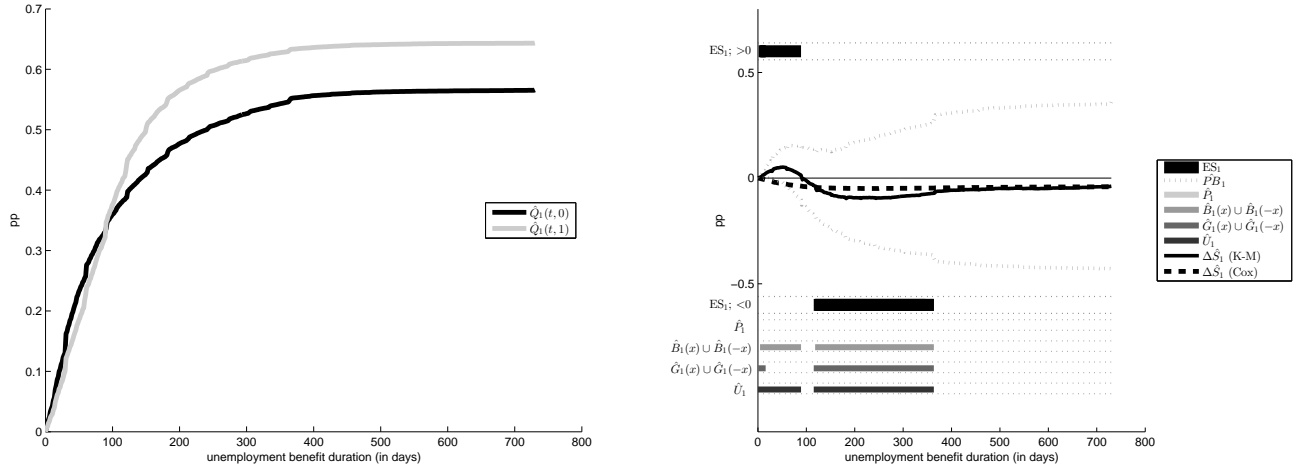
risks (Kaplan-Meier estimator, K-M) and under the assumption of independent competing risks and parametric restrictions on the marginal survivals (Cox-proportional hazards model, Cox). The K-M estimator and Cox model are very frequently applied in empirical research and it is therefore of interest to see how the imposition of stronger restrictions leads to changes in the results.

It is apparent from the left panel of Figure 5 that the estimated Q_j change considerably in the two variables. The cumulative incidence for $ui=1$ is lower for all durations. This suggests that unemployment insurance claimants have lower incidences of taking up a job irrespective of the length of unemployment. The cumulative incidence for $winter=1$ is lower for shorter

Table 4: Relative size of estimated identification sets in the data examples.

	McCall data Risk: Full-time employment $\mathbb{T} = [0, 24]$ (months)		SIAB Risk: Employment $\mathbb{T} = [0, 730]$ (days)	
	Share of $\mathbb{T}^*$	Unique of $\mathbb{T}^{**}$	Share of $\mathbb{T}^*$	Unique of $\mathbb{T}^{**}$
Covariate	*ui*		*winter*	
$\mathbb{P}_1$	32%	32%	0%	0%
$\mathbb{B}_1$	0%	0%	45.6%	10%
$\mathbb{G}_1$	36%	36%	36.5%	0.1%
$\mathbb{U}_1$	68%		46.6%	
Covariate	*married*		*year2008*	
$\mathbb{P}_1$	0%	0%	0%	0%
$\mathbb{B}_1$	44%	0%	41.7%	0%
$\mathbb{G}_1$	96%	52%	57.7%	16%
$\mathbb{U}_1$	96%		57.7%	
Covariate	*female*		*previous recall*	
$\mathbb{P}_1$	0%	0%	28.7%	0%
$\mathbb{B}_1$	0%	0%	99.4%	0%
$\mathbb{G}_1$	32%	32%	99.8%	0.2%
$\mathbb{U}_1$	32%		99.8%	
Covariate			*previous unemployment*	
$\mathbb{P}_1$			0%	0%
$\mathbb{B}_1$			99.0%	0%
$\mathbb{G}_1$			99.8%	0.8%
$\mathbb{U}_1$			99.8%	
Covariate			*female*	
$\mathbb{P}_1$			0%	0%
$\mathbb{B}_1$			1.4%	1.4%
$\mathbb{G}_1$			94.6%	94.6%
$\mathbb{U}_1$			96%	

*: size of respective identification set divided by the size of $\mathbb{T}$.

**: size of the subset of the respective identification set that is not contained in the identification sets of the other methods divided by the size of $\mathbb{T}$. Measure of *additional* contribution.

durations but is higher for longer durations. This implies that *winter* affects the incidence of reemployment differently at different lengths of unemployment duration. But it remains to be seen whether the estimated directions of the covariate effects on the latent survival distribution of reemployment follow the same patterns.

Regarding the effect of *ui*, the Peterson bounds in the right panel Figure 5 reveal a positive covariate effect at shorter durations. This means that the receipt of unemployment insurance decreases the reemployment rate at unemployment duration between 2 to 8 months. The decomposition approach identifies a positive effect at long durations, More specifically, the receipt of unemployment insurance decreases the reemployment rate between 16 and 24 months of unemployment duration. The modified Bond and Shaw procedure does not provide informative results in this case. This example illustrates clearly that the decomposition approach produces the most distinctive information compared to the other approaches. Namely, the decomposition approach is particularly useful for identification at longer duration. When the decomposition approach is not employed, one cannot identify the effect of *ui* on long-term unemployment (16 to 23 months) without making assumptions on the copula function (*K-M*) or the latent survival function (*Cox*). This distinctive feature of the decomposition approach could be particularly relevant for empirical applications in which the covariate effect at longer duration is of major interest, for instance, when the impact of a policy on curbing long-term unemployment is examined.

Another interesting observation can be made for the variable *winter* in Figure 5. The results for the sets $\mathbb{B}_1$ and $\mathbb{G}_1$ show that the estimated direction of the covariate effect changes in duration. The re-employment rate for a short-term unemployed individual (< 90 days) is lower for those who started unemployment in the winter. But their re-employment rate during medium-term unemployment (110 - 370 days) is higher. In contrast, the Cox model, due to its restriction implied by the proportionality of hazards, does not reveal this pattern of duration dependence but suggests a negative effect for all durations. Indeed, the estimated partial effect of the Cox model is not even located within the Peterson bounds for short durations. This highlights the practical relevance of using a model with less parametric restrictions on the role of the covariates for marginal survival distributions.

The Kaplan-Meier estimator, although it does not require parametric restrictions on the latent survival distribution, does require that the competing risks are independent. It is therefore

a special case and included in the Peterson bounds. Indeed, in Figure 5 the Kaplan-Meier based estimates locate near the center of the Peterson bounds. For the variable *winter*, the Kaplan-Meier based estimate coincides with the results of the decomposition approach for short-term (< 90 days) and medium-term (90 - 370 days) unemployment. While it also suggests a positive effect for long-term unemployment (> 370 days), the decomposition approach does not provide evidence of such an effect. This shows that the K-M estimator based results for long-term unemployment are not robust with respect to the degree of dependence (τ) in this example.

These results have shown that our proposed decomposition method produces useful insights into the direction of the covariate effect. The estimated sign is robust to the functional form of the latent survival distribution as well as the degree of dependence between the competing risks.

References

Abbring, J.H. and van den Berg, G.J. (2003) The identifiability of the mixed proportional hazards competing risks model, *Journal of the Royal Statistical Society, B*, 65, 701–710.

Alba-Ramirez, A., Arranz, J. M., and Munoz-Bullon, F. (2007) Exits from unemployment: Recall or new job. *Labour Economics*, 14, 788–810.

Bond, S.J. and Shaw, J.E.H. (2006) Bounds on the covariate-time transformation for competing-risks survival analysis, *Lifetime Data Analysis*, 12, 285–303.

Berger, R.L. (1997), Likelihood ratio tests and intersection union tests. In S. Panchapakesan and N. Balakrishnan (eds), *Advances in Statistical Decision Theory and Applications*, 225–237, Boston: Birkhäuser.

Bloch, D.A., Lai, T.L., Tubert-Bitter, P. (2001) One-sided tests in clinical trials with multiple endpoints. *Biometrics*, 57, 1039–1047.

Burda, M., Harding, M., and Hausman, J. (2015) A Bayesian Semiparametric Competing Risk Model with Unobserved Heterogeneity? *Journal of Applied Econometrics*, 30, 353–376.

Butler, J. S., Anderson, K. H., and Burkhauser, R. V. (1989) Work and Health after Retirement: A Competing Risks Model with Semiparametric Unobserved Heterogeneity. *The Review of Economics and Statistics*, 71, 46–53.

Carling, K., Edin, P.A., Harkman, A., and Holmlund, B. (1996) Unemployment duration, unemployment benefits, and labor market programs in Sweden. *Journal of Public Economics*, 59, 313–334.

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

D'Addio, A.C., Rosholm, M. (2005) Exits from Temporary Jobs in Europe: A competing risks Analysis. *Labour Economics*, 12, 449–468.

Dabrowska, D.M. and Doksum, K.A. (1988) Estimation and testing in a two-sample generalized odds-rate model, *Journal of the American Statistical Association*, 83, 744–749.

Dolton, P. and van der Klaauw, W. (1999) The turnover of teachers: A competing risks explanation. *The Review of Economics and Statistics*, 81, 543–552.

Dorner, M, Heining, J., Jacobebbinghaus, P. and Seth, S. (2010), The sample of integrated labour market biographies, *Schmollers Jahrbuch*, 130, 599—608.

Fermanian, J. (2003) Nonparametric estimation of competing risks models with covariates, *Journal of Multivariate Anaysis*, 85, 156–191.

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

Heckman, J. and Honoré, B. (1990) The empirical content of the Roy Model. *Econometrica*, 58, 1121–1149.

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

Henry, M. and Mourifie, I. (2014) Sharp bounds in the binary Roy Model. Working Paper.

Honoré, B. and Lleras-Muney, A. (2006) Bounds in competing risks models and the war on cancer. *Econometrica*, 74, 1675–1698.

Kalbfleisch, J.D. and Prentice, R.L. (2002) *The Statistical Analysis of Failure Time Data*, Wiley, Hoboken, N.J..

Lee, S. and Lewbel, A. (2013) Nonparametric Identification of Accelerated Failure Time Competing Risks Models. *Econometric Theory*, 29, 905–919.

Lo, S.M.S. and Wilke, R.A. (2010) A copula model for dependent competing risks, *Journal of the Royal Statistical Society, C*, 59, 359–376.

Lo, S.M.S. and Wilke, R.A. (2011) Identifiability and estimation of the sign of a covariate effect in the competing risks model, *Discussion Papers in Economics*, No. 11/03, University of Nottingham, UK.

Mealli, F. and Pudney, S. (1996) Occupational pensions and job mobility in Britain: Estimation of a random-effects competing risks model. *Journal of Applied Econometrics*, 11, 293–320.

Meghir, C. and Whitehouse, E. (1997) Labour market transitions and retirement of men in the UK. *Journal of Econometrics*, 79, 327–354.

Nelsen, R.B. (2006) *An Introduction to Copulas*, 2nd Edition, Springer, New York.

Park, B. G. (2015) Nonparametric Identification and Estimation of the Extended Roy Model, Working Paper.

Peterson, A.V. (1976) Bounds for a Joint Distribution With Fixed Sub-Distribution Functions: Application to Competing Risks, *Proceedings of the National Academy of Science*, 73, 11–13.

Rivest, L. and Wells, M.T. (2001) A Martingale Approach to the Copula-Graphic Estimator for the Survival Function under Dependent Censoring, *Journal of Multivariate Analysis*, 79, 138–155.

Roy, A. D. (1951) Some Thoughts on the Distribution of Earnings, *Oxord Economic Papers (New Series)*, 3, 135–146.

Schweizer, B. and Sklar, A. (1983) *Probabilistic Metric Spaces*, New York: North-Holland.

Silvapulle, M.J. and Sen, P.K. (2005) *Constrained Statistical Inference: Order, Inequality, and Shape Constraints*, New Jersy: Wiley.

Steiner, V. (2001) Unemployment Persistence in the West German Labor Market: Negative Duration Dependence or Sorting? *Oxford Bulletin of Economics and Statistics*, 63, 91–113.

Tsiatis, A. (1975) A Nonidentifiability Aspect of the Problem of Competing Risks, *Proceedings of the National Academy of Sciences*, 72, 20–22.

Zheng, M. and Klein, J.P. (1995) Estimates of marginal survival for dependent competing risks based on assumed copula. *Biometrika*, 82, 127–138.

Appendix

A.I: Proofs

Proof of Lemma 1:

Lemma 1.1 is a direct result of (7) and Definitions 8 and 9. The uniqueness is guaranteed by the uniqueness of $D(c; \boldsymbol{x})$, which is ensured by Assumption 1. The first equality of Lemma 1.2 can be proved by differentiating the copula function $C(S_1^*(c; \boldsymbol{x}), S_2^*(c; \boldsymbol{x})) = c$ with respect to x while keeping c constant. This results in $\frac{\partial C}{\partial S_1^*}\frac{\partial S_1^*}{\partial x} + \frac{\partial C}{\partial S_2^*}\frac{\partial S_2^*}{\partial x} = 0$ and thus $\frac{\partial S_2^*}{\partial x} / \frac{\partial S_1^*}{\partial x} = -\frac{\partial C}{\partial S_1^*} / \frac{\partial C}{\partial S_2^*} < 0$. The second equality of Lemma 1.2 can be proved by differentiating $S_2^*(c; \boldsymbol{x}) = \zeta_1(S_1^*(c; \boldsymbol{x}); \boldsymbol{x})$ with respect to x, such that $\frac{\partial S_2^*}{\partial x} = \frac{\partial \zeta_1}{\partial S_1^*}\frac{\partial S_1^*}{\partial x} + \frac{\partial \zeta_1}{\partial x}|_{u_1=S_1^*}$. After rearranging, we have $\frac{\partial \zeta_1}{\partial x}|_{u_1=S_1^*} = \frac{\partial S_2^*}{\partial x} - \frac{\partial \zeta_1}{\partial S_1^*}\frac{\partial S_1^*}{\partial x}$. Since $\frac{\partial \zeta_1(u)}{\partial u} > 0$ and $sign|\frac{\partial S_2^*}{\partial x}| = -sign|\frac{\partial S_1^*}{\partial x}|$, we have the result. Lemma 1.3 follows from Assumption 1 that S_j is decreasing in t. Lemma 1.4 follows directly from Lemmas 1.1, 1.2 and 1.3. For example, if both $\Delta_x^d S_1(t; \boldsymbol{x})$ and $\Delta_x^l S_1(t; \boldsymbol{x})$ are positive (negative), this implies from 1.2 and 1.3 that $\Delta_x^d S_2(t; \boldsymbol{x}) > (<)0$ and $\Delta_x^l S_2(t; \boldsymbol{x}) < (>)0$. In such case $\Delta_x S_1(t; \boldsymbol{x})$ is positive (negative) and the sign of $\Delta_x S_2(t; \boldsymbol{x})$ cannot be determined. Similarly, if $\Delta_x^d S_1(t; \boldsymbol{x}) < (>)0$ and $\Delta_x^l S_1(t; \boldsymbol{x}) > (<)0$, the sign of $\Delta_x S_1(t; \boldsymbol{x})$ is undetermined while $\Delta_x S_2(t; \boldsymbol{x})$ is negative (positive). Second, if $\Delta_x^d S_1(t; \boldsymbol{x}) > (<)0$ and $\Delta_x^l S_1(t; \boldsymbol{x}) = 0$, this implies from 1.2 and 1.3 that $\Delta_x^d S_2(t; \boldsymbol{x}) > (<)0$ and $\Delta_x^l S_2(t; \boldsymbol{x}) = 0$. In this case, $\Delta_x S_1(t; \boldsymbol{x})$ and $\Delta_x S_2(t; \boldsymbol{x})$ are positive (negative). Similarly, if $\Delta_x^d S_1(t; \boldsymbol{x}) = 0$ and $\Delta_x^l S_1(t; \boldsymbol{x}) > (<)0$, $\Delta_x S_1(t; \boldsymbol{x})$ is positive (negative) while $\Delta_x S_2(t; \boldsymbol{x})$ is negative (positive). Third, if both $\Delta_x^d S_1(t; \boldsymbol{x})$ and $\Delta_x^l S_1(t; \boldsymbol{x})$ are zero, $\Delta_x^d S_2(t; \boldsymbol{x})$ and $\Delta_x^l S_2(t; \boldsymbol{x})$ are also zero, the sign $\Delta_x S_1(t; \boldsymbol{x})$ and $\Delta_x S_2(t; \boldsymbol{x})$ are then known to be zero. To conclude, the sign of one risk can be determined in the first case, while in cases 2 and 3 the sign of both risks can be determined. $\qquad\square$

Proof of Lemma 3:

We prove Lemma 3.1 by differentiating $Q_1^*(c(t); \boldsymbol{x}) = \int_{S_1^*(c(t), \boldsymbol{x})}^{1} \int_0^{\zeta_1(u_1; \boldsymbol{x})} \kappa(u_1, u_2) du_2 du_1$ w.r.t. x:

$$
\begin{aligned}
\frac{\partial Q_1^*(c(t); \boldsymbol{x})}{\partial x} &= \int_{S_1^*(c(t); \boldsymbol{x})}^{1} \frac{\partial}{\partial x} \zeta_1(u_1; \boldsymbol{x}) \kappa(u_1, \zeta_1(u_1; \boldsymbol{x})) du_1 \\
&\quad - \frac{\partial}{\partial x} S_1^*(c(t); \boldsymbol{x}) \int_0^{\zeta_1(S_1^*(c(t), \boldsymbol{x}); \boldsymbol{x})} \kappa(S_1^*(c(t); \boldsymbol{x}), u_2) du_2.
\end{aligned}
\tag{24}
$$

According to Assumption 3 and Lemma 1.2, $sign\, |\frac{\partial}{\partial x} \zeta_1(u_1; \boldsymbol{x})| = -\, sign\, |\frac{\partial}{\partial x} S_1^*(c(t); \boldsymbol{x})|$ for all $u_1 \in [S_1^*(c(t); \boldsymbol{x}), 1]$, or, equivalently, for all $t_1 \in [0, t]$. We have therefore

$$
\begin{aligned}
sign\, |\Delta_x^l Q_1(t; \boldsymbol{x})| &= -\, sign\, |\Delta_x^l S_1(t; \boldsymbol{x})| \times sign\, \Bigg| \int_{S_1^*(c; \boldsymbol{x})}^{1} \kappa(u_1, \zeta_1(u_1; \boldsymbol{x})) du_1 \\
&\quad + \int_0^{\zeta_1(S_1^*(c, \boldsymbol{x}); \boldsymbol{x})} \kappa(S_1^*(c; \boldsymbol{x}), u_2) du_2 \Bigg|
\end{aligned}
\tag{25}
$$

for all $t \in \mathbb{R}_+$. Since the integrals in (25) are positive, the proof for risk 1 is complete. For risk 2 it suffices to prove the following: Since $S(t; \boldsymbol{x}) = 1 - Q_1(t; \boldsymbol{x}) - Q_2(t; \boldsymbol{x})$, we have $\frac{\partial Q_1^*(c(t); \boldsymbol{x})}{\partial x} = -\frac{\partial Q_2^*(c(t); \boldsymbol{x})}{\partial x}$. Substituting this and Lemma 1.2 in (25), we have $sign\, |\Delta_x^l Q_2(t; \boldsymbol{x})| = -\, sign\, |\Delta_x^l S_2(t; \boldsymbol{x})|$. Lemma 3.2 is a direct result from Lemmas 1.4, 2 and 3.1. $\qquad\square$

Lemma 6 *Under Assumptions 1 and 2 and for $k = 1, 2, ...$, we have for the competing risks model characterised by equations (2)-(3):*

1. *$\Delta_x^l S_j(t; \boldsymbol{x}) = 0$ for all $t \in \{\dot{t}_k\}$.*

2. *Either $sign\, |\Delta_x^l S_j(t; \boldsymbol{x})| \geq 0$ or $sign\, |\Delta_x^l S_j(t; \boldsymbol{x})| \leq 0$ for all $t \in (\dot{t}_k, \dot{t}_{k+1})$.*

3. *$sign\, |\Delta_x^l S_j(m; \boldsymbol{x})| = -\, sign\, |\Delta_x^l S_j(n; \boldsymbol{x})|$ for all $m \in (\dot{t}_k, \dot{t}_{k+1})$ and $n \in (\dot{t}_{k+1}, \dot{t}_{k+2})$.*

4. *$sign\, |\Delta_x^l A_j(\dot{t}_k, t; \boldsymbol{x})| = -\, sign\, |\Delta_x^l S_j(t; \boldsymbol{x})|$ for all $t \in (\dot{t}_k, \dot{t}_{k+1})$ with*

$$
A_j(t_1, t_2; \boldsymbol{x}) = \int_{S_j(t_1; \boldsymbol{x})}^{S_j(t_2; \boldsymbol{x})} \int_0^{\zeta_j(u_j)} \kappa(u_1, u_2) du_{-j} du_j
\tag{26}
$$

for any $t_2 > t_1 \geq 0$ and $\Delta_x^l A_j(t_1, t_2; \boldsymbol{x})$ as the link function effect of $A_j(t_1, t_2; \boldsymbol{x})$.

Proof of Lemma 6:

Lemmas 6.1, 6.2 and 6.3 follow directly from Definition 13 and Lemma 1.2. We prove Lemma 6.4 as follows: For any $t \in (\acute{t}_k, \acute{t}_{k+1})$, rewrite $Q_j(t; \boldsymbol{x})$ in (3) as

$$Q_j(t; \boldsymbol{x}) = Q_j(\acute{t}_k; \boldsymbol{x}) + A_j(\acute{t}_k, t; \boldsymbol{x}). \tag{27}$$

The link function effect of (27) becomes

$$\Delta_x^l Q_1(t; \boldsymbol{x}) = \Delta_x^l Q_1(\acute{t}_k; \boldsymbol{x}) + \Delta_x^l A_1(\acute{t}_k, t; \boldsymbol{x}), \text{ with} \tag{28}$$

$$\Delta_x^l A_1(\acute{t}_k, t; \boldsymbol{x}) = \int_{S_1^*(c(t); \boldsymbol{x})}^{S_1^*(c(\acute{t}_k); \boldsymbol{x})} \frac{\partial}{\partial x} \zeta_1(u_1; \boldsymbol{x}) \kappa(u_1, \zeta_1(u_1; \boldsymbol{x})) du_1$$

$$- \frac{\partial}{\partial x} S_1^*(c(t); \boldsymbol{x}) \int_0^{\zeta_1(S_1^*(c(t), \boldsymbol{x}); \boldsymbol{x})} \kappa(S_1^*(c(t); \boldsymbol{x}), u_2) du_2. \tag{29}$$

Equation (29) holds because of Lemma 1.1. From Lemma 1.2, $sign \, |\Delta_x \zeta_1(t; \boldsymbol{x})| = - sign \, |\Delta_x^l S_1(t; \boldsymbol{x})|$ for all $t \in (\acute{t}_k, t)$. The sign of (29) equals to

$$sign \, |\Delta_x^l A_1(\acute{t}_k, t; \boldsymbol{x})| = - sign \, |\Delta_x^l S_1(t; \boldsymbol{x})| \times sign \left| \int_{S_1^*(c(t); \boldsymbol{x})}^{S_1^*(c(\acute{t}_k); \boldsymbol{x})} \kappa(u_1, \zeta_1(u_1; \boldsymbol{x})) du_1 \right.$$

$$\left. + \int_0^{\zeta_1(S_1^*(c(t), \boldsymbol{x}); \boldsymbol{x})} \kappa(S_1^*(c(t); \boldsymbol{x}), u_2) du_2 \right|$$

$$= - sign \, |\Delta_x^l S_1(t; \boldsymbol{x})|. \tag{30}$$

This completes the proof. □

Proof of Lemma 4:

We first prove Lemma 4.1. For any $k > 0$, assume without loss of generality that $\acute{t}_k \in [\acute{t}_l, \acute{t}_{l+1}]$ for some $l > 0$. We have from (28) and Definition 14

$$\Delta_x^l Q_j(\acute{t}_k; \boldsymbol{x}) = \Delta_x^l Q_j(\acute{t}_l; \boldsymbol{x}) + \Delta_x^l A_j(\acute{t}_l, \acute{t}_k; \boldsymbol{x}) = 0. \tag{31}$$

For all $t \in (\acute{t}_k, \grave{t}_{l+1}]$, we have therefore

$$\Delta_x^l Q_j(t; \boldsymbol{x}) \quad = \quad \Delta_x^l Q_j(\acute{t}_k; \boldsymbol{x}) + \Delta_x^l A_j(\acute{t}_k, t; \boldsymbol{x}) = \Delta_x^l A_j(\acute{t}_k, t; \boldsymbol{x}). \tag{32}$$

The first turning point $\grave{t}_{j,k}$ is located in $(\acute{t}_k, \grave{t}_{l+1})$ if $\Delta_x^l A_j(\acute{t}_k, t; x)$ is a non-monotonic (non-increasing or non-decreaseing) function in $t \in (\acute{t}_k, \grave{t}_{l+1})$. Otherwise $\grave{t}_{j,k} = \grave{t}_{l+1}$ if $\Delta_x^l A_j(\acute{t}_k, t; \boldsymbol{x})$ is monotonic (increasing or decreasing) in $t \in (\acute{t}_k, \grave{t}_{l+1})$. This is because $sign\, |\Delta_x^l A_j(\grave{t}_{l+1}, t; \boldsymbol{x})| = -sign\, |\Delta_x^l A_j(\acute{t}_k, \grave{t}_{l+1}; \boldsymbol{x})|$ in $\Delta_x^l Q_j(t; \boldsymbol{x}) = \Delta_x^l A_j(\acute{t}_k, \grave{t}_{l+1}; \boldsymbol{x}) + \Delta_x^l A_j(\grave{t}_{l+1}, t; \boldsymbol{x})$ for all $t \in (\grave{t}_{l+1}, \grave{t}_{l+2})$, and $\Delta_x^l Q_j(t; \boldsymbol{x})$ changes direction at $\grave{t}_{l+1}$. To conclude, we have $[\acute{t}_k, \grave{t}_{j,k}] \in [\grave{t}_l, \grave{t}_{l+1}]$ in all cases. This completes the proof of Lemma 4.1.

Next, we prove Lemma 4.2. According to Definition 14, we have for $t \in (\acute{t}_k, \grave{t}_{j,k})$,

$$\Delta_x^l Q_j(\acute{t}_k; \boldsymbol{x}) < \Delta_x^l Q_j(t; \boldsymbol{x}) < \Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) \quad \text{if} \quad \Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) > 0$$
$$\Delta_x^l Q_j(\acute{t}_k; \boldsymbol{x}) > \Delta_x^l Q_j(t; \boldsymbol{x}) > \Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) \quad \text{if} \quad \Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) < 0.$$

Since $\Delta_x^l Q_j(\acute{t}_k; \boldsymbol{x}) = 0$ and by writing $\Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) = \Delta_x^l Q_1(\grave{t}_k; \boldsymbol{x}) + \Delta_x^l A_1(\grave{t}_k, \grave{t}_{j,k}; \boldsymbol{x})$, we have

$$0 < \Delta_x^l Q_j(t; \boldsymbol{x}) < \Delta_x^l Q_1(\grave{t}_k; \boldsymbol{x}) + \Delta_x^l A_1(\grave{t}_k, \grave{t}_{j,k}; \boldsymbol{x}) \quad \text{if} \quad \Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) > 0$$
$$0 > \Delta_x^l Q_j(t; \boldsymbol{x}) > \Delta_x^l Q_1(\grave{t}_k; \boldsymbol{x}) + \Delta_x^l A_1(\grave{t}_k, \grave{t}_{j,k}; \boldsymbol{x}) \quad \text{if} \quad \Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) < 0.$$

$\Delta_x^l Q_j(t; \boldsymbol{x})$ is zero at $\acute{t}_k$ with $\acute{t}_k \in (\grave{t}_k, \grave{t}_{j,k})$. Thus $\Delta_x^l Q_j(\grave{t}_k; \boldsymbol{x})$ and $\Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x})$, which are non-zero, must have different directions. It follows that

$$0 < \Delta_x^l Q_j(t; \boldsymbol{x}) < \Delta_x^l A_1(\grave{t}_k, \grave{t}_{j,k}; \boldsymbol{x}) \quad \text{if} \quad \Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) > 0$$
$$0 > \Delta_x^l Q_j(t; \boldsymbol{x}) > \Delta_x^l A_1(\grave{t}_k, \grave{t}_{j,k}; \boldsymbol{x}) \quad \text{if} \quad \Delta_x^l Q_j(\grave{t}_{j,k}; \boldsymbol{x}) < 0.$$

Using Lemma 6.4 we have for all $t \in [\acute{t}_k, \grave{t}_{j,k}]$ and for all $k = 1, 2...$

$$sign\, |\Delta_x^l Q_j(t; \boldsymbol{x})| \quad = \quad sign\, |\Delta_x^l A_j(\acute{t}_k, t; \boldsymbol{x})| = sign\, |\Delta_x^l S_j(t; \boldsymbol{x})|. \tag{33}$$

This completes the proof of Lemma 4.2. □

Remark on Lemma 4.1:

It is possible that there is more than one zero cut-off point in $[\acute{t}_l, \acute{t}_{l+1}]$ for some l such that all $I\!\!I_{j,k}, I\!\!I_{j,k+1}, ..., I\!\!I_{j,k+m}$ are included in $[\acute{t}_l, \acute{t}_{l+1}]$ for some $k, l = 1, 2, ...$ and $m = 1, 2, ...$ It is also possible that there exists no zero cut-off point in $[\acute{t}_l, \acute{t}_{l+1}]$ for some l. For example, we assume without loss of generality that $I\!\!I_{j,k} \in [\acute{t}_l, \acute{t}_{l+1}]$ for some k and l and consider the location of $I\!\!I_{j,k+1}$. If $|\Delta_x^l A_j(\acute{t}_{l+1}, t; \boldsymbol{x})| < |\Delta_x^l A_j(\acute{t}_k, \acute{t}_{l+1}; \boldsymbol{x})|$ in $\Delta_x^l Q_1(t; \boldsymbol{x}) = \Delta_x^l A_j(\acute{t}_k, \acute{t}_{l+1}; x) + \Delta_x^l A_j(\acute{t}_{l+1}, t; \boldsymbol{x})$ for all $t \in (\acute{t}_{l+1}, \acute{t}_{l+2})$, the negative (positive) value of $\Delta_x^l A_j(\acute{t}_{l+1}, t; \boldsymbol{x})$ is not large enough to counterbalance the positive (negative) value of $\Delta_x^l A_j(\acute{t}_k, \acute{t}_{l+1}; \boldsymbol{x})$. In this case there is no zero cut-off point in $[\acute{t}_{l+1}, \acute{t}_{l+2}]$, and thus $I_{j,k+1}$ can only be located in $[\acute{t}_{l+m}, \acute{t}_{l+m+1}]$ for some odd integer $m > 1$. To conclude, there is no particular restriction on the values of k and l in $I\!\!I_{j,k} \in [\acute{t}_l, \acute{t}_{l+1}]$.

Proof of Corollary 1:

It suffices to show an example in which the sign of the covariate effect is identified while the reversed covariate effect is not. Consider the case of crossing link functions at $\acute{t}_1 = 0$, some $\acute{t}_2$, and $\acute{t}_3$ such that $\zeta_1(S_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1) > \zeta_1(S_1(t; \boldsymbol{x}_0); \boldsymbol{x}_0)$ for all $t \in (0, \acute{t}_2)$ and $\zeta_1(S_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1) < \zeta_1(S_1(t; \boldsymbol{x}_0); \boldsymbol{x}_0)$ for all $t \in (\acute{t}_2, \acute{t}_3)$. Suppose also there exists a t^* such that $S_1(t^*; \boldsymbol{x}_0) > S_1(\acute{t}_2; \boldsymbol{x}_0)$ and $S_1(D(S(\acute{t}_2; \boldsymbol{x}_0); \boldsymbol{x}_1); \boldsymbol{x}_1) > S_1(t^*; \boldsymbol{x}_1)$, and thus

$$D(S(\acute{t}_2; \boldsymbol{x}_0); \boldsymbol{x}_1) < t^* < \acute{t}_2. \tag{34}$$

Given Definition 6 and the continuity of $S(t; \boldsymbol{x})$, $S_1(D(S(\acute{t}_2; \boldsymbol{x}_0); \boldsymbol{x}_1); \boldsymbol{x}_1) = S_1(\acute{t}_2; \boldsymbol{x}_0)$, we have therefore from (34) $S_1(t^*; \boldsymbol{x}_1) < S_1(D(S(\acute{t}_2; \boldsymbol{x}_0); \boldsymbol{x}_1); \boldsymbol{x}_1) = S_1(\acute{t}_2; \boldsymbol{x}_0) < S_1(t^*; \boldsymbol{x}_0)$ which implies a negative covariate effect on $S_1(t^*; \boldsymbol{x}_0)$. We will show the case that the decomposition approach identifies the sign of the covariate effect for risk 1 at t^*, but the reversed covariate effect is not identified at t^*.

From (34), we have $S(D(S(t^*; \boldsymbol{x}_0); \boldsymbol{x}_1), \boldsymbol{x}_1) = S(t^*; \boldsymbol{x}_0) > S(\acute{t}_2; \boldsymbol{x}_0) = S(D(S(\acute{t}_2; \boldsymbol{x}_0); \boldsymbol{x}_1); \boldsymbol{x}_1)$, and thus $D(S(t^*; \boldsymbol{x}_0); \boldsymbol{x}_1) < D(S(\acute{t}_2; \boldsymbol{x}_0); \boldsymbol{x}_1)$. Together with (34), we have $t^* > D(S(t^*; \boldsymbol{x}_0); \boldsymbol{x}_1)$. This implies that the duration effect is negative, i.e. $\Delta_x^d S_1(t^*; \boldsymbol{x}_0) < 0$. Under the assumption that $\zeta_1(S_1(t, \boldsymbol{x}_0), \boldsymbol{x}_1) > \zeta_1(S_1(t, \boldsymbol{x}_0), \boldsymbol{x}_0)$ for all $t \in (0, \acute{t}_2)$ and from (34) $t^* < \acute{t}_2$, we have $\Delta_x \zeta_1(t^*; \boldsymbol{x}) < 0$. Given Lemma 1.2, the link function effect is also negative, i.e.

$\Delta_x^l S_1(t^*; \boldsymbol{x}_0) < 0$. The sign of the covariate effect is therefore identified. Next, we consider the sign of the reversed covariate effect from a movement from $\boldsymbol{x}_1$ to $\boldsymbol{x}_0$.

From (34), we have $S(D(S(t^*; \boldsymbol{x}_1); \boldsymbol{x}_0), \boldsymbol{x}_0) = S(t^*; \boldsymbol{x}_1) < S(D(S(\dot{t}_2; \boldsymbol{x}_0); \boldsymbol{x}_1); \boldsymbol{x}_1) = S(\dot{t}_2; \boldsymbol{x}_0)$, we have $D(S(t^*; \boldsymbol{x}_1); \boldsymbol{x}_0) > \dot{t}_2$. Together with (34), we have $D(S(t^*; \boldsymbol{x}_1); \boldsymbol{x}_0) > t^*$. This implies that the duration effect is positive, i.e. $\Delta_{-x}^d S_1(t^*; \boldsymbol{x}_0) > 0$. However, under the assumption that $\zeta_1(S_1(t, \boldsymbol{x}_0), \boldsymbol{x}_1) < \zeta_1(S_1(t, \boldsymbol{x}_0), \boldsymbol{x}_0)$ for all $t \in (\dot{t}_2, \dot{t}_3)$, we have from Lemma 1.2 that the link function effect is negative, i.e. $\Delta_{-x}^l S_1(t^*; \boldsymbol{x}_0) < 0$. And thus the sign for the reversed covariate effect is unidentified. $\qquad \square$

Proof of Lemma 5:

From Definitions 3 and 20:

$$
\begin{aligned}
\phi_2(t; \boldsymbol{x}_0) &< \phi_1(t; \boldsymbol{x}_0) \iff \\
S_2^{-1}(S_2(t; \boldsymbol{x}_0); \boldsymbol{x}_1) &< S_1^{-1}(S_1(t; \boldsymbol{x}_0); x_1) \iff \\
S_2(t; \boldsymbol{x}_0) &> S_2(S_1^{-1}(S_1(t; \boldsymbol{x}_0); x_1); x_1) \iff \\
\zeta_1(S_1(t; \boldsymbol{x}_0); \boldsymbol{x}_0) &> \zeta_1(S_1(t; \boldsymbol{x}_0); \boldsymbol{x}_1).
\end{aligned}
$$

This holds vice versa and completes the proof. $\qquad \square$

A.II: Estimation of $\mathbb{G}_j$ and the sign of the covariate effect.

We apply the following multiple-step estimation procedure in the data examples in Section 4 to estimate the effect of a discrete covariate.

1. Define an equally spaced time grid $\{t_1, t_2, \ldots, t_M\}$.

2. Estimate $\hat{Q}_j(t; \boldsymbol{x}_k)$ and $\hat{S}(t; \boldsymbol{x}_k)$ for $j = 1, 2$ and $k = 0, 1$ nonparametrically at all t in $\{t_1, t_2, \ldots, t_M\}$.

3. Compute $\hat{D}(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ in Definition 17 by solving the sample analogue of equation (5) for all t in $\{t_1, t_2, \ldots, t_M\}$.

4. Compute $\hat{\Delta}_x^l Q_j(t; \boldsymbol{x}_0)$ and $\hat{\Delta}_x^d Q_j(t; \boldsymbol{x}_0)$ by plugging $\hat{D}(S(t; \boldsymbol{x}_0); \boldsymbol{x}_1)$ into $\hat{Q}_j(t; \boldsymbol{x})$ according to Definition 17 for all t in $\{t_1, t_2, \ldots, t_M\}$.

5. Compute $\hat{\mathbb{I}}_j$ from the estimated sequences $\{\acute{t}_k\}$ and $\{\grave{t}_{j,k}\}$ according to Definition 14 by using $\hat{\Delta}_x^l Q_j(t; \boldsymbol{x}_0)$ for all j.

6. Compute $\hat{\mathbb{D}}_j$ by using $\hat{\Delta}_x^l Q_j(t; \boldsymbol{x}_0)$ and $\hat{\Delta}_x^d Q_j(t; \boldsymbol{x}_0)$ for all j.

7. The sign of the covariate effect at $t \in \hat{\mathbb{G}}_j$ is then determined by the sample analogue of Proposition 1.

This procedure is applicable to both directions of the decomposition Δ and Δ_{-x}, which provides $\hat{\mathbb{G}}_j(x)$ and $\hat{\mathbb{G}}_j(-x)$ for each risk. There are two modifications to improve the finite sample performance: Sampling variation in $\hat{Q}_j(t; \boldsymbol{x})$ also imply some random variation in $\hat{\Delta}_x^l Q_j(t; \boldsymbol{x})$. For this reason, the estimated sequence $\{\grave{t}_{j,k}\}$ has also some random variation. In particular since $\hat{\Delta}_x^l Q_j(t; \boldsymbol{x})$ is not smooth and has some peaks created by sampling errors, the estimated first local extreme value between $\{\acute{t}_k\}$ and $\{\acute{t}_{k+1}\}$ is likely to occur before the actual value of $\{\grave{t}_{j,k+1}\}$. This implies that the estimated $\{\grave{t}_{j,k+1}\}$ as well as the size of the identification region are likely downward biased in small samples. We suggest two alternative procedures to overcome this issue:

- Employ a smoothing technique for $\hat{Q}_j(t; \boldsymbol{x})$ in step 5 to eliminate small peaks in $\hat{\Delta}_x^l Q_j(t; \boldsymbol{x})$. Although it can eliminate peaks due to random sampling, it can also eliminate the true

extreme values if the chosen degree of smoothing is too large. As with any smoothing technique there is some arbitrariness involved and it is difficult to determine the optimal degree of smoothing.

- Impose an additional assumption that there are no multiple extreme values of $\hat{\Delta}_x^l Q_j(t; \boldsymbol{x})$ between $\{\acute{t}_k\}$ and $\{\acute{t}_{k+1}\}$. In this case, we recommend in step 5 using the estimated global extreme value between $\{\acute{t}_k\}$ and $\{\acute{t}_{k+1}\}$ as an estimator for the sequence of $\{\grave{t}_{j,k}\}$. This method produces good results if the true link function effect does not have multiple local extreme values. Otherwise, the estimated $\{\grave{t}_{j,k}\}$ is upward biased.