



The University of
Nottingham

Discussion Papers in Economics

Discussion Paper
No. 11/03

Identifiability and estimation
of the sign of a covariate effect
in the competing risks model

S.M.S. Lo and R.A. Wilke

April 2011

Identifiability and estimation of the sign of a covariate effect in the competing risks model.*

Simon M.S. Lo[†]

Ralf A. Wilke[‡]

April 2011

*We thank Lutz Dümbgen for helpful discussions and the participants at numerous seminars for their comments. Wilke is supported by the Economic and Social Research Council through the *Bounds for Competing Risks Duration Models using Administrative Unemployment Duration Data* (RES-061-25-0059) grant.

[†]Lingnan University, Rm 218, Ho Sin Hang Building, Lingnan University, Hong Kong, E-mail: simonlo@ln.edu.hk .

[‡]University of Nottingham, School of Economics, University Park, Nottingham NG7 2RD, United Kingdom, Phone: +441159515248, E-mail: ralf.wilke@nottingham.ac.uk

Abstract

It is well known that the competing risks model is identified if the dependence structure between risks (the copula function) is known or assumed. Special cases include independence of risks or independent censoring. If the copula function is not specified, parameters of interest are only set identified. As these sets are often wide in applications, it is difficult to obtain informative results. In this paper we strike a balance between imposing too much and too little structure. By establishing a general link between observable changes in sub-distributions (cumulative incidence curves) and the sign of changes in marginal distributions (the causal treatment effect) we are able to show the identifiability of the latter if the copula function is independent of the varying covariate. This has two important implications: First, it is possible to obtain informative results even if the copula function is mainly unspecified or unknown. Second, the sign of the covariate effect tends to be invariant with respect to the chosen dependence structure. Our method is computationally very simple and our simulations suggest that it identifies and consistently estimates the sign of the treatment effect for large sets of duration times. An application to unemployment duration data illustrates the usefulness of our method for empirical research.

Keywords: dependent censoring, nonparametric estimation, bootstrap

JEL: C14, C24, C41

1 Introduction

The non-identifiability of the competing risks model (Cox, 1962; Tsiatis, 1975) implies that data alone is only partly informative for the identification of the parameters of interest. If we are, for example, interested in the marginal distribution of a latent competing variable, this functional can only be bounded (Peterson, 1976). See also Manski (2003) for partially identified probability distributions. Honoré and Lleras-Muney (2006) consider the accelerated failure time model and obtain tighter bounds on the marginal effect from discrete covariates. Point identification can be achieved by making assumptions on the marginal distributions and the dependence structure between the competing risks (Abbring and van den Berg, 2003) or by fully specifying a the dependence structure (a copula function) between the competing risks (Zheng and Klein, 1995). By performing a sensitivity analysis, Lo and Wilke (2010) observe that the sign of a covariate effect -

the *causal treatment effect* - is often the same for any assumed copula function while the magnitude of the effect varies considerably. Basing on this observation we take a different route in this paper by focusing on the identifiability of the sign of a covariate effect rather than its magnitude. Indeed, we can show that it is identifiable by exploiting variation in the cumulative incidence curves (CIC) and the survival function of the observed failure time under a mild condition. Our simulations and illustrations with data suggest that our identification approach provides substantially more informative results than the Peterson bounds. Our method is very simple and general as it is fast to compute and it works with existing non-, semi- and parametric estimators for the CICs (e.g. Jeong and Fine, 2007, and Peng and Fine, 2009). Although our approach is less informative than point estimates, we claim that it still is very informative for research in various disciplines such as biometrics, econometrics and social sciences. We therefore conceive our novel approach as a useful tool for a wide research community.

The structure of this paper is as follows: the next section introduces the model and presents our main identification results. Section 3 suggests nonparametric estimation and inference procedures. Section 4 presents simulation results. Section 5 illustrates our method by estimating the effect of various covariates on the job finding probability for unemployed individuals in Germany.

2 Identifiability

We first consider a model with two latent competing random variables T_1 and $T_2 \in \mathcal{T} \subset R_+$. A model with more than two competing risks is considered in Section 2.2. T_1 and T_2 are times to failure or times to the events 1 and 2 respectively. While T_1 and T_2 are not observable, $T = \min(T_1, T_2)$ and $\delta = \operatorname{argmin}_j T_j$ are observed. There is one observable binary covariate x which takes values $x = x_0$ (control group) and $x = x_1$ (treatment group). Data on (T, δ, x) enable the identifiability of the unknown cumulative incidence curves $Q_j(t; x) = \Pr(T \leq t, \delta = j | x)$, the unknown cause-specific crude hazard functions $\lambda_j(t; x) = \lim_{\Delta \rightarrow 0} \Pr(t \leq T \leq t + \Delta, \delta = j | T \geq t, x) / \Delta$ for risk $j = 1, 2$ and the unknown survival function $S(t; x) = \Pr(T > t | x)$ of $T = \min(T_1, T_2)$ for all t . The marginal distribution functions $F_j(t; x) = \Pr(T_j \leq t | x)$ and the marginal survival functions $S_j(t; x) = 1 - F_j(t; x)$ are also unknown for all j and t but they are not identifiable from data alone (Cox, 1962; Tsiatis, 1975).

Assumption 1 $S_j(t; x)$, $1 - Q_j(t; x)$ and $S(t; x)$ are continuous and strictly decreasing in t for

all j .

Let be S^{-1} the inverse of the functional S . We denote $\Delta_x F_j(t) = F_j(t; x_1) - F_j(t; x_0)$ as the covariate or treatment effect on $F_j(t; x_0)$ and its direction by $\text{sign} |\Delta_x F_j(t)|$. The operator $\text{sign} |w|$ equals to the sign of the variable w . It is +1, 0 or -1 if w is positive, zero or negative respectively. In our model $\Delta_x F_j(t)$ is unknown and not identified. However, in the following we show that $\text{sign} |\Delta_x F_j(t)|$ can be identified from the CICs and the survival function for some duration time under a very mild condition. Note that in contrast to popular semiparametric models such as the proportional hazard model there is no restriction on the nature of the treatment effect acting on the duration time as the sign of $\Delta_x F_j(t)$ can vary with t .

An integral part of the competing risks model is the dependence structure between the risks which is determined by the survival copula $C(s_1, s_2; x) = \Pr(S_1 \leq s_1, S_2 \leq s_2; x)$. The competing risks model is fully characterised by

$$\begin{aligned} Q_1(t; x) &= \int_0^{\zeta_{21}(s_1; x)} \int_{S_1(t; x)}^1 C''(s_1, s_2; x) ds_1 ds_2; \\ S(t; x) &= C\{S_1(t; x), S_2(t; x); x\}, \end{aligned} \quad (1)$$

where $\zeta_{21}(s_1; x) = S_2\{S_1^{-1}(s_1; x); x\}$ is a unique link function that defines the relationship between $S_1(t; x)$ and $S_2(t; x)$ for all t and x (Lemma A1 of Zheng and Klein, 1995). For more details see also Lo and Wilke (2010). The probability density function of the copula is denoted by $C''(s_1, s_2; x) = \partial^2 C(s_1, s_2; x) / \partial s_1 \partial s_2$. Model (1) implies that $S_1(t; x)$ and $S_2(t; x)$ are determined jointly by the copula function $S(t; x) = C\{S_1(t; x), S_2(t; x); x\}$ and the link function $S_2(t; x) = \zeta_{21}\{S_1(t; x); x\}$ as $Q_j(t; x)$ and $S(t; x)$ are directly identified from the data. When the copula function is known, the link function is fixed by $Q_1(t; x)$. Then the two unknowns $S_1(t; x)$ and $S_2(t; x)$ can be determined by solving the two equations for all t .

In this paper we consider a model with an unknown copula function $C(s_1, s_2; x)$. For this reason ζ_{21} and F_j for all j are not identified. Without imposing additional assumptions, the treatment effect can only be bounded by using the Peterson bounds for the marginal distributions:

$$\Delta_x Q_j(t) - Q_i(t; x_0) \leq \Delta_x F_j(t) \leq \Delta_x Q_j(t) + Q_i(t; x_1) \quad (2)$$

with $\Delta_x Q_j(t) = Q_j(t; x_1) - Q_j(t; x_0)$ for $j = 1, 2$ and $j \neq i$. We denote \mathcal{I}_j as the nonparametric identification set for $\text{sign} |\Delta_x F_j(t)|$. \mathcal{I}_j consists of all t for which $\Delta_x Q_j(t) - Q_i(t; x_0) > 0$ or $\Delta_x Q_j(t) + Q_i(t; x_1) < 0$ or the two former being equal to zero. As the Peterson bounds are often

wide, P_j is likely small and no informative result can be obtained in an application. In the following we show that it is possible to obtain a considerably larger, although different, identification set by imposing a restriction on the unknown copula function.

Assumption 2 $C(s_1, s_2; x_1) = C(s_1, s_2; x_0) = C(s_1, s_2)$.

Despite being difficult to test in applications, most popular parametric, semi-, and non-parametric duration models make stronger assumptions on the dependence structure which imply Assumption 2. This includes the Kaplan-Meier estimator, the accelerated failure time model and the (mixed) proportional hazard model. For more details see Bond and Shaw (2006). In their paper they derive bounds for covariate-time transformations under a condition which implies Assumption 2. Although Assumption 2 limits the covariate effect to $F_j(t; x)$ for all j , $\zeta_{21}(s_1; x)$ and $S(t; x)$ still change with x . Assumption 2 is a crucial condition for our approach to identification, as it builds up observable linkages between the observable changes in the CICs and the unobserved changes in the marginal distributions. It can be seen from (1) that, when the integrand C'' is unaffected by the treatment, the direction of the covariate effect on $Q_1(t; x)$ and $F_1(t; x)$ can be analyzed by the changes in the size of the two domains in the integration function of $Q_1(t; x)$. For the purpose of illustration let us consider a special case where the treatment acts positively on $F_1(t)$ while keeping $F_2(t)$ constant for all t . In this case, $S_1(t; x_1) < S_1(t; x_0)$ for all t , and $\zeta_{21}(s_1; x_1) > \zeta_{21}(s_1; x_0)$ for all s_1 . Thus,

$$\begin{aligned} Q_1(t; x_1) &= \int_0^{\zeta_{21}(s_1; x_1)} \int_{S_1(t; x_1)}^1 C'' ds_1 ds_2 > \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(t; x_1)}^1 C'' ds_1 ds_2 \\ &> \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(t; x_0)}^1 C'' ds_1 ds_2 = Q_1(t; x_0). \end{aligned}$$

Therefore, $Q_1(t; x)$ increases with $F_1(t; x)$ for all t . This special case illustrates that there is a link between the direction of the covariate effect on F_j and the sign of the effect on Q_j .

To fully develop this observation into a general result, we exploit the fact that $S_1(t; x)$ and $S_2(t; x)$ are jointly determined by the copula function and the link function. We consider an analytical 2-part decomposition of the change in the marginal distributions, where we isolate two effects. The first part is the counterfactual movement of the marginal distributions due to the change in the link function from $\zeta_{21}(s_1; x_0)$ to $\zeta_{21}(s_1; x_1)$ while keeping the copula function constant at $S(t; x_0)$. S_1 and S_2 at t move from $[S_1(t; x_0), S_2(t; x_0)]$ to $[S_1(t^c; x_1), S_2(t^c; x_1)]$ at some t^c , such

that S_1 and S_2 become the solutions of (i) $C\{S_1(t^c; x_1), S_2(t^c; x_1)\} = C\{S_1(t; x_0), S_2(t; x_0)\}$ or

$$S(t^c; x_1) = S(t; x_0) \quad (3)$$

and (ii) $S_2(t^c; x_1) = \zeta_{21}\{S_1(t^c; x_1); x_1\}$. This is denoted as the copula effect.

Definition 1 *The copula effect of a covariate change from x_0 to x_1 on the marginal distribution in model (1) is defined by $\Delta_x^c F_j(t) = F_j(t^c; x_1) - F_j(t; x_0)$, with t^c satisfying condition (3).*

The difference between t^c and t has an interpretation of a quantile treatment effect at the $S(t; x_0)$ -quantile. t^c and t are generally different if the link function $\zeta_{21}(s_1; x)$ changes in response to a covariate change, except when $t = t^c = 0$ at $S(t; x_0) = S(t^c; x_1) = 1$.

The second part of the decomposition characterises the changes in the marginal distributions to $[S_1(t, x_1), S_2(t, x_1)]$ which are due to the change from $S(t; x_0)$ to $S(t; x_1)$ while keeping the link function unchanged. S_1 and S_2 become solutions to (i) $C\{S_1(t; x_1), S_2(t; x_1)\} = S(t; x_1)$ and (ii) $S_2(t; x_1) = \zeta_{21}\{S_1(t; x_1); x_1\}$. We denote this as the marginal distribution effect.

Definition 2 *The marginal distribution effect of a covariate change from x_0 to x_1 on the marginal distribution $F_j(t, x_0)$ in model (1) is defined by $\Delta_x^m F_j(t) = F_j(t; x_1) - F_j(t^c; x_1)$, with t^c satisfying condition (3).*

A graphical presentation of $\Delta_x^m F_j(t)$ and $\Delta_x^c F_j(t)$ is given in the supplemental material. We obtain the following useful results.

Lemma 1 *Under Assumptions 1 and 2 we have for all $t \in \mathbb{T}$ in Model (1):*

1. $\Delta_x^c F_j(t)$ and $\Delta_x^m F_j(t)$ are unique for all j .

2.

$$\Delta_x F_j(t) = \Delta_x^c F_j(t) + \Delta_x^m F_j(t) \quad \text{for all } j. \quad (4)$$

3. $\text{sign} |\Delta_x^c F_j(t)|$ and $\text{sign} |\Delta_x^m F_j(t)|$ determine the sign of $\Delta_x F_j(t)$ for at least one j .

The uniqueness follows from the uniqueness of t^c . The decomposition in (4) follows directly from Definitions 1 and 2. For more details on the proof see the Appendix. As a next step we define observable analogues of the marginal distribution effect and the copula effect.

Definition 3 We define $\Delta_x^c Q_j(t) = Q_j(t^c; x_1) - Q_j(t; x_0)$ as the copula effect on the CIC for risk $j = 1, 2$ for all t in model (1), with t^c satisfying the condition (3).

Definition 4 We define $\Delta_x^m Q_j(t) = Q_j(t; x_1) - Q_j(t^c; x_1)$ as the marginal distribution effect on the CIC for risk $j = 1, 2$ for all t in model (1), with t^c satisfying condition (3).

A graphical illustration of $\Delta_x^m Q_j(t)$ and $\Delta_x^c Q_j(t)$ is also given in the supplemental material. The following two lemmas establish systematic relationships between the unobserved $\text{sign} |\Delta_x^c F_j(t)|$ and $\text{sign} |\Delta_x^m F_j(t)|$ and the observable analogues based on the CICs.

Lemma 2 In model (1) under Assumptions 1 and 2, we have for $j = 1, 2$ and for all t

$$\text{sign} |\Delta_x^m F_j(t)| = \text{sign} |\Delta_x^m Q_j(t)|. \quad (5)$$

Lemma 2 can be proved by using Definitions 2 and 4. $F_j(t; x)$ and $Q_j(t; x)$ are both strictly increasing functions in t , and thus $\text{sign} |\Delta_x^m F_j(t; x)|$ equals $\text{sign} |\Delta_x^m Q_j(t; x)|$ for any t^c and t . As the equivalent result for the copula effect does not hold for all t , we consider a restricted set only.

Definition 5 Let $\{\acute{t}_k\}$ for $k = 0, 1, 2, \dots$ be a finite sequence of $t \in \mathcal{T}$ such that $\Delta_x^c Q_j(\acute{t}_k) = 0$ with $\acute{t}_0 = 0$. Let $\{\grave{t}_k\}$ for $k = 1, 2, \dots$ be a finite sequence of $t \in \mathcal{T}$ such that $|\Delta_x^c Q_j(t)|$ has its first local maximum in the interval $(\acute{t}_{k-1}, \acute{t}_k]$. Moreover, $\mathbb{I}_j = \bigcup_{k \geq 0} [\acute{t}_k, \grave{t}_{k+1}] \subset \mathcal{T}$.

\mathbb{I}_j is observable since \acute{t}_k and \grave{t}_k are observed for all k . The following result suggests that \mathbb{I}_j is the identification set for the sign of the copula effect on the marginal distributions for risk j .

Lemma 3 In model (1) under Assumptions 1 and 2, we have for $j = 1, 2$ and for all $t \in \mathbb{I}_j$

$$\text{sign} |\Delta_x^c F_j(t)| = \text{sign} |\Delta_x^c Q_j(t)|. \quad (6)$$

The proof of Lemma 3 can be found in the Appendix. The intuition behind Lemma 3 can be illustrated by a special case where the link function $\zeta_{21}(s_1, x)$ is a monotone function in x . We have for all s_1 either (i) $\zeta_{21}(s_1, x_1) \leq \zeta_{21}(s_1, x_0)$ or (ii) $\zeta_{21}(s_1, x_1) \geq \zeta_{21}(s_1, x_0)$. Under Assumption 2, case (i) implies that the copula effect on $F_1(t)$ is negative and it is positive on $F_2(t)$ for all t .

See also Lemma A3 in the Appendix for a formal derivation. This implies $S_1(t^c; x_1) < S_1(t; x_0)$ and $S_2(t^c; x_1) > S_2(t; x_0)$ with $C\{S_1(t^c; x_1), S_2(t^c; x_1)\} = C\{S_1(t; x_1), S_2(t; x_1)\}$. It follows that

$$\begin{aligned} Q_1(t^c; x_1) &= \int_0^{\zeta_{21}(s_1; x_1)} \int_{S_1(t^c; x_1)}^1 C'' ds_1 ds_2 < \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(t^c; x_1)}^1 C'' ds_1 ds_2 \\ &< \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(t; x_0)}^1 C'' ds_1 ds_2 = Q_1(t; x_0). \end{aligned}$$

Analogous reasoning can be applied to case (ii). And thus the sign of the copula effect on $Q_j(t)$ and $F_j(t)$ are the same for all t . In the general case where $\zeta_{21}(s_1, x)$ is not a monotone function, say, $\zeta_{21}(s_1, x_1)$ and $\zeta_{21}(s_1, x_0)$ cross each other at a sequence of duration time $\{\dot{t}_k\} \in \mathcal{T}$ for $k = 0, 1, 2, \dots$, the relationship between the sign of $|\Delta_x^c Q_j(t)|$ and $|\Delta_x^c F_j(t)|$ can only be established when $\{\dot{t}_k\}$ is known. But as $\zeta_{21}(s_1, x)$ is unidentified, $\{\dot{t}_k\}$ can only be bounded. We show in the Appendix that \dot{t} is bounded by the observed interval $[\dot{t}_k, \dot{t}_{k+1}]$ for all $k \geq 0$. The sign of $|\Delta_x^c Q_j(t)|$ can therefore be identified for the observable sequence of sets $[\dot{t}_k, \dot{t}_{k+1}]$ for all $k \geq 0$ which is \mathcal{I}_j .

According to Lemmas 1, 2 and 3, $sign |\Delta_x F_j(t)|$ is determined by $sign |\Delta_x^m Q_j(t)|$ and $sign |\Delta_x^c Q_j(t)|$ for all $t \in \mathcal{I}_j$. Let us denote $\Delta_x \mathbf{Q}_j(\mathbf{t}) = [\Delta_x^c Q_j(t), \Delta_x^m Q_j(t)]'$. $\Delta_x \mathbf{Q}_j(\mathbf{t}) \succeq \mathbf{0}$ means that both of $\Delta_x^c Q_j(t)$ and $\Delta_x^m Q_j(t)$ are nonnegative but that at least one is non-zero. $\Delta_x \mathbf{Q}_j(\mathbf{t}) \preceq \mathbf{0}$ is defined analogously. Moreover, let $\mathcal{D}_j \subset \mathcal{T}$ consist of all t such that $\Delta_x^c Q_j(t) \times \Delta_x^m Q_j(t) \geq 0$. Then \mathcal{I}_j consists of all t such that $\Delta_x^m Q_j(t)$ and $\Delta_x^c Q_j(t)$ do not have an opposite sign. We are now able to state our main identification result.

Proposition 1 *Under Assumptions 1 and 2, the sign of the treatment effect in model (1) is identified for all $t \in \mathcal{I}_j \cap \mathcal{D}_j$. We have*

$$sign |\Delta_x F_j(t)| = \begin{cases} +1 & \text{if } \Delta_x \mathbf{Q}_j(\mathbf{t}) \succeq \mathbf{0}; \\ 0 & \text{if } \Delta_x \mathbf{Q}_j(\mathbf{t}) = \mathbf{0}; \\ -1 & \text{if } \Delta_x \mathbf{Q}_j(\mathbf{t}) \preceq \mathbf{0}. \end{cases} \quad (7)$$

Moreover, it is identified for at least one risk $j = 1, 2$ for all $t \in \mathcal{I}_j$.

Proposition 1 follows directly from all previous results. It suggest that $sign |\Delta_x F_j(t)|$ is identified for $t \in \mathcal{I}_j$ if $\Delta_x^m Q_j(t)$ and $\Delta_x^c Q_j(t)$ do not the have opposite sign.

2.1 Increasing the identification set

Proposition 1 suggests that the direction of the treatment effect is unidentified if $\Delta_x^c Q_j(t)$ and $\Delta_x^m Q_j(t)$ have an opposite sign or if $t \notin \mathcal{I}_j$. We now suggest an approach which enlarges the

identification set. As it does not require any additional assumption it should always be performed.

Rather than the treatment effect, we consider the reversed treatment effect

$$\Delta_{-x}F_j(t) = F_j(t; x_0) - F_j(t; x_1) = -\Delta_x F_j(t). \quad (8)$$

The reversed treatment effect is the change in $F_j(t; x)$ from x_1 to x_0 and, thus, has the opposite sign than $\Delta_x F_j(t)$. It is obvious that Proposition 1 also holds for the reversed treatment effect by exchanging the notation x_1 and x_0 . We denote this property as independence of the decomposition route. Let $\mathbb{I}_j(x) \cap \mathbb{D}_j(x)$ and $\mathbb{I}_j(-x) \cap \mathbb{D}_j(-x)$ be the identification sets for $\text{sign} |\Delta_x F_j(t)|$ and $\text{sign} |\Delta_{-x} F_j(t)|$ respectively. We obtain the following useful result.

Corollary 1 $\mathbb{I}_j(x) \cap \mathbb{D}_j(x) \neq \mathbb{I}_j(-x) \cap \mathbb{D}_j(-x)$.

We prove Corollary 1 by showing that for some t the sign of the treatment effect is unidentified, while the sign of the reversed treatment effect is identified. The proof is given in the Appendix.

Corollary 1 suggests that it is always better to compute both decomposition routes and take the union of the two identification sets. Since the underlying identification approach of the Peterson bounds in (2) and our method are also different, it is advisable to determine the identification sets for all approaches and take the union of all sets for each risk

$$\mathbb{Q}_j = \mathbb{P}_j \cup \left\{ \mathbb{I}_j(x) \cap \mathbb{D}_j(x) \right\} \cup \left\{ \mathbb{I}_j(-x) \cap \mathbb{D}_j(-x) \right\} \subset \mathbb{T}.$$

For all $t \in \mathbb{Q}_j \setminus \mathbb{P}_j$ we can also tighten the bounds for the magnitude. If the sign of the effect is negative (positive) the lower (upper) bound for the magnitude is the lower (upper) Peterson bound and the upper (lower) bound is 0.

2.2 Identifiability in a multi-risks model

In this section we extend the identification result of the previous section to a model with a finite number of risks $J > 2$. The observed failure time becomes $T = \min(T_1, \dots, T_J)$ and the indicator function is $\delta = \text{argmin}_j T_j$. Model (1) becomes

$$\begin{aligned} Q_j(t; x) &= \int_0^{\zeta_{Jj}(s_j; x)} \dots \int_0^{\zeta_{j+1j}(s_j; x)} \int_{S_j(t; x)}^1 \int_0^{\zeta_{j-1j}(s_j; x)} \dots \int_0^{\zeta_{1j}(s_j; x)} \frac{\partial^J C(s_1, \dots, s_J; x)}{\partial s_1 \dots \partial s_J} ds_1 \dots ds_J; \\ S(t; x) &= C^J(s_1, \dots, s_j, \dots, \dots, s_J; x). \end{aligned} \quad (9)$$

The J-survival copula is

$$C^J(s_1, \dots, s_J; x) = \Pr\{S_1(T_1; x) \leq s_1, \dots, S_J(T_J; x) \leq s_J; x\}. \quad (10)$$

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

$$C^2(s_j, s_{-j}; x) = \Pr\{S_j(T_j; x) \leq s_j, S_{-j}(T_{-j}; x) \leq s_{-j}; x\}. \quad (11)$$

The unknown marginal survival function for the pooled variable T_{-j} is $S_{-j}(t; x) = \Pr(T_{-j} > t; x)$. The observed failure time is unaffected as $T = \min(T_j, T_{-j})$, and the indicator function is modified as $\delta_j = j$ if the original $\delta = j$ and $\delta_j = -j$ if $\delta \neq j$. For any J-copula in (10), the existence of a 2-copula in (11) is guaranteed under the following assumption (Nelsen, 2006).

Assumption 3 *In the competing risks model defined by (9) the survival copula belongs to the Archimedean class.*

In this case the multi-risk model can be reduced into a two risks model as (1):

$$\begin{aligned} Q_j(t; x) &= \int_0^{\zeta_{-j,j}(s_j; x)} \int_{S_j(t; x)}^1 C'' ds_j \dots ds_{-j}; \\ S(t; x) &= C^2\{S_j(t; x); S_{-j}(t; x); x\}, \end{aligned} \quad (12)$$

with $s_{-j} = \zeta_{-j,j}(s_j; x)$ denotes the link function between $S_j(t; x)$ and $S_{-j}(t; x)$. For more details see Lo and Wilke (2010). Our identification approach for the two risks model can therefore be subsequently applied to (12) for $j = 1, \dots, J$, where the order of application does not matter. Note, however, that only one risk is of interest in the pooled risks model as the pooled risks are generally uninformative.

3 Estimation

For simplicity we outline the estimation procedure for a two risks model. Suppose we have for x_0 and x_1 two finite samples of observations of $T = \min\{T_1, T_2\}$ with latent failure time identically and independently distributed as $T_1 \sim F_1(T_1; x)$ and $T_2 \sim F_2(T_2; x)$, $\delta = \operatorname{argmin}_j T_j$ and a dependence structure $C\{S_1(t; x), S_2(t; x); x\}$. F_1 , F_2 and C are unknown in an application

and only (T, δ, x) is observed. We suggest an estimation procedure which involves several stages. First, we estimate $Q_j(t; x)$ and $S(t; x)$ with common nonparametric estimators (Kalbfleish and Prentice, 2002), although any existing consistent (non-)parametric estimator could be used. These estimators are then plugged into the equations of the previous section to obtain their sample analogues. In what follows we list all steps of the estimation procedure:

1. Define an equally spaced grid $\{t_0, t_1, \dots, t_M\}$.
2. Estimate $\hat{Q}_j(t; x_k)$ and $\hat{S}(t; x_k)$ for $j, k = 0, 1$ nonparametrically at all t in $\{t_0, t_1, \dots, t_M\}$.
3. Determine t^c by solving the sample analogue of equation (3) for all t in $\{t_0, t_1, \dots, t_M\}$. Since the estimated survival curves are step functions, we assume left continuity of $\hat{S}(t; x_k)$. Thus for any value of t_i, t_j, t_{j+1} on the time grid such that $\hat{S}(t_j, x_1) > \hat{S}(t_i, x_0) > \hat{S}(t_{j+1}, x_1)$. The solution to equation (3) is $t_i^c = t_{j+1}$.
4. Compute $\Delta_x^m \hat{Q}_j(t)$ by plugging t^c into $\hat{Q}_j(t; x)$ according to Definition 4, for all t in $\{t_0, t_1, \dots, t_M\}$.
5. Compute $\Delta_x^c \hat{Q}_j(t)$ by plugging t^c into $\hat{Q}_j(t; x)$ according to Definition 3, for all t in $\{t_0, t_1, \dots, t_M\}$.
6. For all j estimate \mathbb{I}_j from the sequences $\{\hat{t}_k\}$ and $\{\hat{t}_k\}$ according to Definition 5 by using $\hat{Q}_j^c(t; x)$. Estimate \mathbb{D}_j by using $\text{sign} |\Delta_x^m \hat{Q}_j(t)|$ and $\text{sign} |\Delta_x^c \hat{Q}_j(t)|$ for all j .
7. The sign of the time-specific treatment effect is then determined by the sample analogues of Proposition 1.

This procedure is applicable to both directions of the decomposition Δ_x and Δ_{-x} , which results in different identification sets for each risk. Before we consider large sample properties and inference, we briefly outline two modifications to improve the finite sample performance:

- Sampling variation in $\hat{Q}_j(t)$ also implies some random variation in $\Delta_x^c \hat{Q}_j(t)$. For this reason, the estimated sequence $\{\hat{t}_k\}$ has also some random variation. In particular since $\Delta_x^c \hat{Q}_j(t)$ is not smooth and has some peaks created by the noise in the data, the estimated first local extreme value between $\{\hat{t}_k\}_x$ and $\{\hat{t}_{k+1}\}_x$ is likely to occur before the actual value of $\{\hat{t}_{k+1}\}$. This implies that the estimated $\{\hat{t}_{k+1}\}$ as well as the size of the identification region are often downward biased in small samples. We suggest two alternative procedures to overcome this issue:

- Employ a smoothing technique for $\hat{Q}_j(t)$ in step 5 to eliminate small peaks in $\Delta_x^c \hat{Q}_j(t)$. 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 $\Delta_x^c \hat{Q}_j(t)$ between $\{\hat{t}_k\}$ and $\{\hat{t}_{k+1}\}$. In this case, we recommend in step 6 the use of the t corresponding to the estimated global extreme value between $\{\hat{t}_k\}$ and $\{\hat{t}_{k+1}\}$ as an estimator for the sequence $\{\hat{t}_k\}$. This method produces good results if the true copula effect does not have multiple local extreme values. Otherwise, the estimated $\{\hat{t}_k\}$ is upward biased.

3.1 Consistency

Proposition 2 *Assume that $\hat{S} \xrightarrow{p} S$ (convergence in probability) and $\hat{Q}_j \xrightarrow{p} Q$ for all j or assume almost sure convergence. Then under Assumptions 1, 2 and 3 we have for model (12)*

$$\text{sign} |\Delta_x \hat{F}_j(t)| = \begin{cases} +1 & \text{if } \Delta_x \hat{Q}_j(\mathbf{t}) \geq \mathbf{0}; \\ 0 & \text{if } \Delta_x \hat{Q}_j(\mathbf{t}) = \mathbf{0}; \\ -1 & \text{if } \Delta_x \hat{Q}_j(\mathbf{t}) \leq \mathbf{0}. \end{cases} \quad (13)$$

converges in probability (or almost surely) to $\text{sign} |\Delta_x F_j(t)|$ for all j and all $t \in \mathbb{I}_j \cap \mathbb{D}_j$.

Whether we have weak or strong consistency depends on the choice of the estimators for $\hat{S}(t, x)$ and $\hat{Q}_j(t, x)$. The proof is a straightforward application of the continuous mapping theorem (Van der Vaart, 1998, Theorem, 18.11) provided that $S(t, x)$ and $Q_j(t, x)$ are continuous for all j .

Proof: For simplicity we only show weak consistency, i.e. $\text{sign} |\Delta_x \hat{F}_j(t)| \xrightarrow{p} \text{sign} |\Delta_x F_j(t)|$. First note that provided $\hat{S}(t, x) \xrightarrow{p} S(t, x)$ and $S(t, x)$ is continuous, we have $\hat{S}^{-1}(\hat{S}(t, x_0), x_1) = \hat{t}^c \xrightarrow{p} t^c$. Then, since $\hat{Q}_j(t, x) \xrightarrow{p} Q_j(t, x)$, $Q_j(t, x)$ is continuous, and since $\hat{t}^c \xrightarrow{p} t^c$, we have $\Delta_x^c \hat{Q}_j(t) \xrightarrow{p} \Delta_x^c Q_j(t)$ and $\Delta_x^m \hat{Q}_j(t) \xrightarrow{p} \Delta_x^m Q_j(t)$ for all j . Then all stochastic components in the right hand side of equation (13) converge to their true values in probability, which completes the proof.

3.2 Inference

Under the assumption that observations are independent, we can perform a statistical test for the sign of the treatment effect $\Delta_x F_j(t)$ for all j . Let us consider the case of the estimated treatment effect being positive at t . In this case we test the null hypothesis (H_0) that the treatment effect $\Delta_x F_j(t)$ is non-positive against the alternative hypothesis (H_1) that it is positive:

$$H_0 : \Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) \notin \Omega_1 \quad \text{and} \quad H_1 : \Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) \in \Omega_1, \quad (14)$$

where $\Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) \in \Omega_1$ if $\Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) \succeq \mathbf{0}$. This set up is similar to the multiple end points problem, in which the treatment and control groups are compared with more than one response variable. For a review see Silvapulle and Sen (2005, Ch.9). There is one difficulty involved in implementing this test as the parameter set in the null hypothesis does not include the boundary points which are $\Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) = (0, R_+)$ or $\Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) = (R_+, 0)$. Thus the null hypothesis is not suitable for many statistical tests. One way to overcome this problem is to define some non-negative number ϵ such that the two response variables of a treatment are said to be practically positive if $\Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) > \epsilon$ and practically noninferior if $\Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) > -\epsilon$. In this spirit we define another parameter set of the alternative hypothesis Ω_2 ,

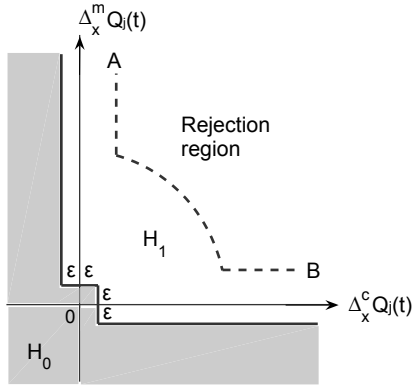
$$\Omega_2 = \{\Delta_{\mathbf{x}} \mathbf{Q}_j(\mathbf{t}) \in R^2 : [\Delta_x^c Q_j(t) > -\epsilon, \Delta_x^m Q_j(t) > -\epsilon] \cap [\Delta_x^c Q_j(t) > \epsilon \cup \Delta_x^m Q_j(t) > \epsilon]\}. \quad (15)$$

The parameter set of the null is represented by the shaded region in Figure 1. It is easy to see that the set Ω_2 converges to Ω_1 if $\epsilon \rightarrow 0$. The analytic joint distribution of $\Delta_x^c Q_j(t)$ and $\Delta_x^m Q_j(t)$ is complicated and difficult to derive. Moreover the null and alternative parameter space in (14) and (15) are some complicated composite of different one-tailed tests. We therefore suggest the convenient bootstrap test by Bloch et al. (2001) which is outlined in the supplemental material.

4 Simulations

In this section we present Monte Carlo results to demonstrate the applicability of the methods outlined in the previous sections. We consider a two risks model with a known closed form representation of the entire competing risks model, i.e. with known CICs, survival function of the minimum, marginal survival functions and copula function. We use the closed form expression given in Rivest and Wells (2001) which bases on the known copula generator of the

Figure 1: Parameter space of the null hypothesis, the alternative region and the rejection region



Archimedean copula, $\phi(s)$, and the known cause-specific crude hazard functions, $\lambda_j(t; x)$, $j = 1, 2$. We have $S(t; x) = \exp \left[- \int_0^t \lambda_1(u; x) + \lambda_2(u; x) du \right]$ and $Q_j(t; x) = \int_0^t \lambda_j(u; x) S(u; x) du$. Then the marginal survival function is given by (Rivest and Wells, 2001)

$$S_j(t; x) = \phi^{-1} \left[- \int_0^t \phi' [S(u; x)] S(u; x) \lambda_j(u; x) du \right]. \quad (16)$$

We choose a Frank's survival copula with parameter α . Note that the Frank's copula has the same function as its survival copula and it is the only Archimedean copula that possesses this property (Georges et al., 2001). In our simulations, the survival copula and the copula generator ϕ are specified as

$$\begin{aligned} C(s_1, s_2) &= \frac{1}{\alpha} \ln \left[1 + \frac{(e^{\alpha(1-s_1)} - 1)(e^{\alpha(1-s_2)} - 1)}{e^\alpha - 1} \right]; \\ \phi(s) &= \ln \left[\frac{e^{\alpha s} - 1}{e^\alpha - 1} \right]; \\ \phi^{-1}(t) &= \frac{1}{\alpha} \ln[1 + e^t(e^\alpha - 1)]. \end{aligned}$$

The cause-specific crude hazard functions follow log-logistic distributions with different parameters for each j and x . The model parameters are given in Table 1. Since we know the true S , S_j and Q_j for all j , we can easily assess the performance of our identification strategy. We first apply our identification method to the true values of S , Q_j and S_j for all j . Then we assess finite sample performance when we use nonparametric estimators for S and Q_j for all j . Although, the identification and estimation procedures outlined in the previous section can be applied for any $t \in R_+$, we restrict our analysis to grid points $\{0, 0.01, 0.02, \dots, 4.99, 5.00\}$.

Table 1: Parameters of the simulated competing risks model.

Function		Risks $j = 1$		Risks $j = 2$	
		$x = x_0$	$x = x_1$	$x = x_0$	$x = x_1$
Crude hazard function: $\lambda_j(t; x) = \frac{\theta_j \gamma_j (\theta_j t)^{\gamma_j - 1}}{1 + (\theta_j t)^{\gamma_j}}$	(γ_j, θ_j)	(8, 1)	(2, 0.5)	(1, 2)	(4, 1)
Copula generator: $\phi(s) = \ln \left[\frac{e^{\alpha s} - 1}{e^\alpha - 1} \right]$	α	0.3	0.3	0.3	0.3

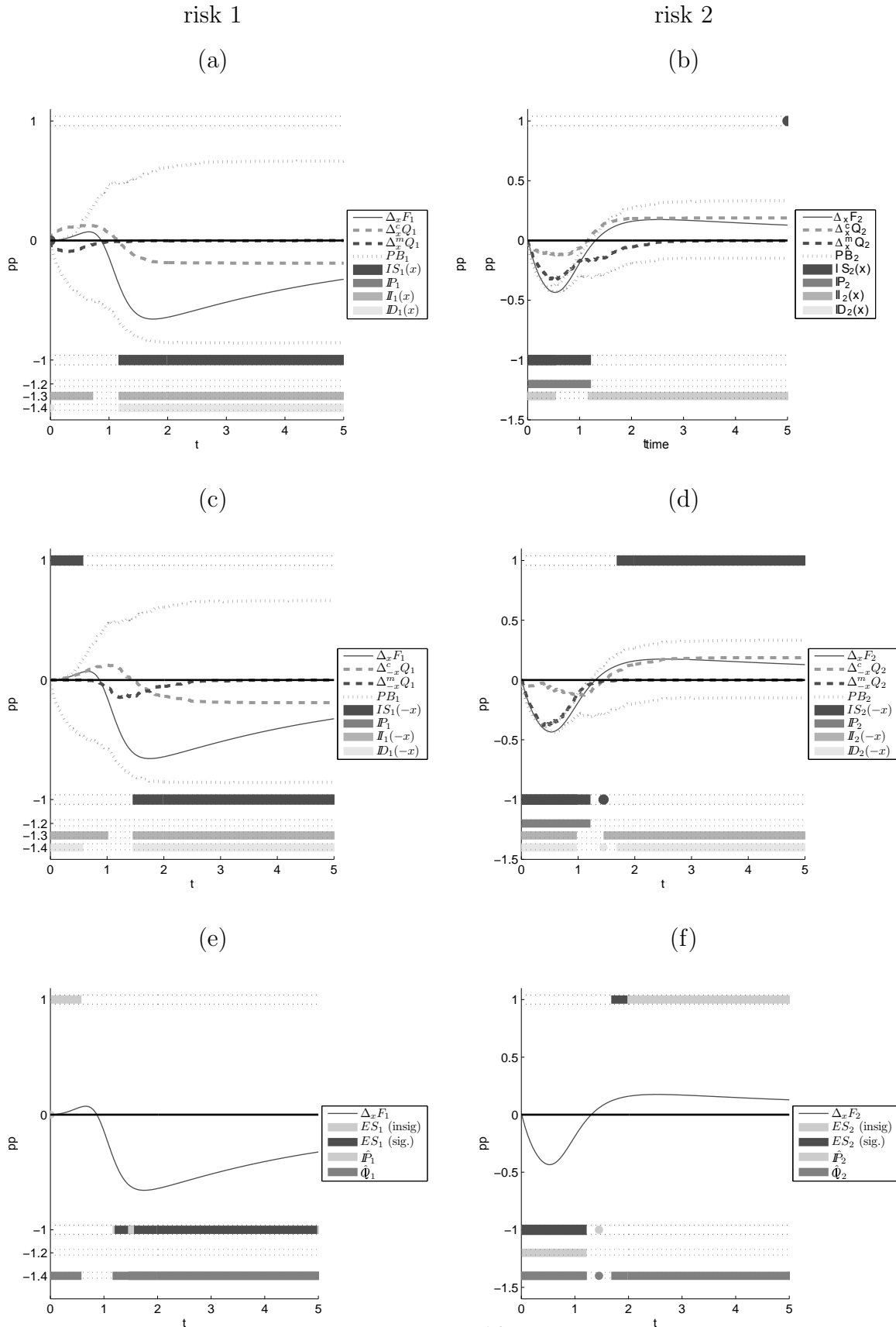
Identifiability Since we know Q_j and S_j for all j , it is straightforward to compute (t^c, \dot{t}, \hat{t}) , $\Delta_x^c Q_j(t)$, $\Delta_x^m Q_j(t)$, $\Delta_x^c F_j(t)$, $\Delta_x^m F_j(t)$ and $\Delta_x F_j(t)$ for any t on our grid.

Panels (a) and (b) of Figure 2 show the treatment effect ($\Delta_x F_j(t)$), the Peterson bounds (PB_j) and the corresponding nonparametric identification sets $\mathcal{I}P_j$. It also shows $\mathcal{I}I_j(x)$ and $\mathcal{I}D_j(x)$ using the copula effect and the marginal distribution effect. The sets $\mathcal{I}P_j$ are plotted as horizontal lines at -1.2, and $\mathcal{I}I_j(x)$ and $\mathcal{I}D_j(x)$ by horizontal lines at -1.3 and -1.4 respectively. The Figure also contains the identified sign ($IS_j(x)$) as horizontal lines at (1,0,-1) based on the union set of $\mathcal{I}I_j(x)$, $\mathcal{I}D_j(x)$ and $\mathcal{I}P_j$. Panels (c) and (d) contain the same sets for the reversed treatment effect Δ_{-x} to illustrate that the method of Section 2.1 indeed increases the union of all identification sets.

It can be seen that the sign of the treatment effect for risk 1 cannot be identified from the Peterson bounds for any t , while for risk 2 a negative treatment effect is identified from duration time zero to one. It is also evident that the resulting identification sets of our procedure are larger than the sets for the Peterson bounds. When we apply the procedure of Section 2.1 to increase the identification set it becomes apparent that it results in rather different sets IS_j , which indicates the usefulness of this procedure. As the sets $\mathcal{I}P_1$ and $\mathcal{I}P_2$ are only small subsets of \mathcal{Q}_1 and \mathcal{Q}_2 respectively, we also gain more insights about the magnitude of the effect as this is bounded by 0 from above (below) for risk 1 (2) for $t > 1.5$.

Estimation Next we simulate data with sample size 500 and estimate Q_j and S as outlined in Section 3. For inference we perform the bootstrap test outlined in Section 3.2 with 500 bootstrap samples and $\epsilon = 0.01$. The estimation results are given in panels (e) and (f) of Figure 2. The estimated sign (ES_j) is plotted as a black horizontal line at (-1,0,1) if the estimate is significant at the 5% level and it is plotted in grey otherwise. The sets $\hat{\mathcal{I}}P_j$ and $\hat{\mathcal{Q}}_j$ are plotted as horizontal lines at -1.2 and -1.5 respectively. When we compare the estimation results with the true values in panels (a)-(d), it is evident that there is no bias for the sign estimation and the identification

Figure 2: Identified sign (IS) and estimated sign (ES) of the treatment effect in a known two risks model.



sets are very similar to the true values.

5 Application

In this section we present an illustrative application to unemployment duration data from Germany. For the estimations we use a sample of the scientific use file version of the IAB employment sample (IABS) 2001 of the Institute for Employment Research (IAB). These administrative data are a 2% random sample of the German workforce subject to social security contributions in the period 1975-2001. The IABS 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 IABS see Hamann et al. (2004). From these data we extract all unemployment periods starting in 1998 or 1999 and having a foregoing employment period. We define unemployment as receipt of unemployment compensation from the German Federal Employment Agency. This leaves us with a sample of 107,522 observations. We consider a model with two risks: employment and other exits. We analyse the effect of various binary covariate changes such as gender, calendar time and employment history of individuals on the job finding probability. Results for the second risk are not presented as it does not have a direct interpretation.

Figures 3 and 4 present estimated nonparametric cumulative incidence curves, the resulting Peterson bounds and the estimated sign (ES) of the covariate effects on the marginal distribution of job finding. ES is plotted as horizontal lines at -1, 0 and 1 in dark grey if the estimate is significant and in light grey if it is insignificant. It is apparent from the left panels of the figures that estimated CICs change considerably in female, recall, previous unemployment, low wage and winter while the effect of calendar year 1999 (instead of 1998) is very small. The estimated Peterson bounds in the right panels are often wide and only partly informative. In the case of female, recall, low wage and winter the set \hat{P}_1 (horizontal line at -1.2) partly covers the interval up to 150-250 days while it is (almost) empty for other variables. In the case of winter the Peterson bounds suggest a change in the sign of the covariate effect at about 50 days. When applying our identification and estimation approaches of Section 3 the identification set \hat{Q}_1 (horizontal line at -1.5) is much larger than \hat{P}_1 for all variables. In particular it covers almost the entire set of

durations in the case of female, recall, previous unemployment and winter. This suggests that the identification set is enlarged substantially by imposing Assumptions 1-3 while still allowing for many different dependence structures. Due to the large sample size the estimated sign is mainly significant. For the calendar time variables winter and year 1999 we find evidence for a change in the sign of the covariate effect as unemployment duration progresses. Even though the estimated CICs for the years 1998 and 1999 are extremely similar, the set \mathbb{Q}_1 contains almost one third of all t . For all variables the set \mathbb{P}_j is much smaller than the set \mathbb{Q}_j . For this reason the identification of the sign of the effect can be also used to derive considerably tighter bounds for the magnitude of the effects for many durations.

Figure 3: Cumulative incidence for job finding (left) and estimated sign of covariate effects on the marginal distribution (right).

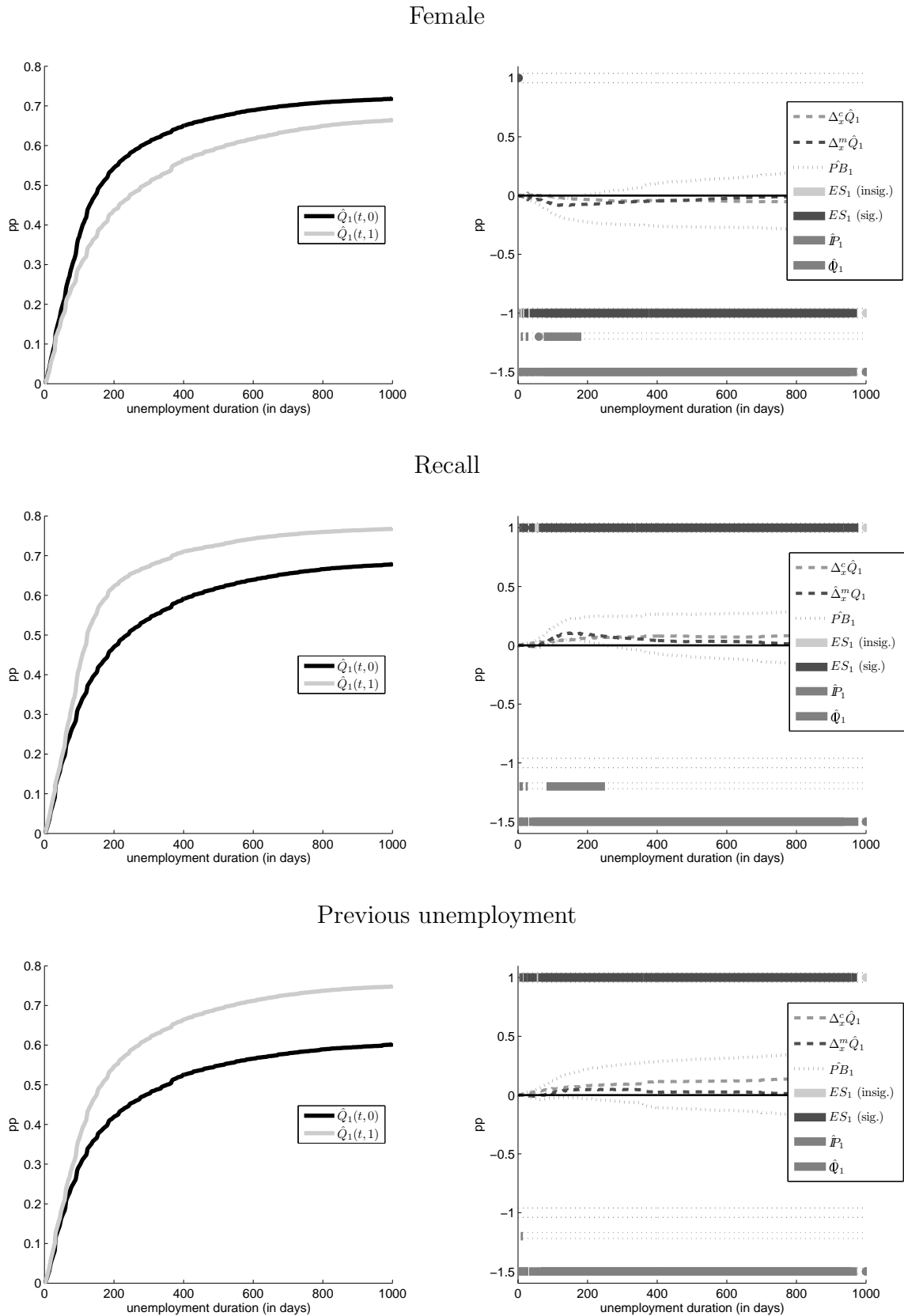
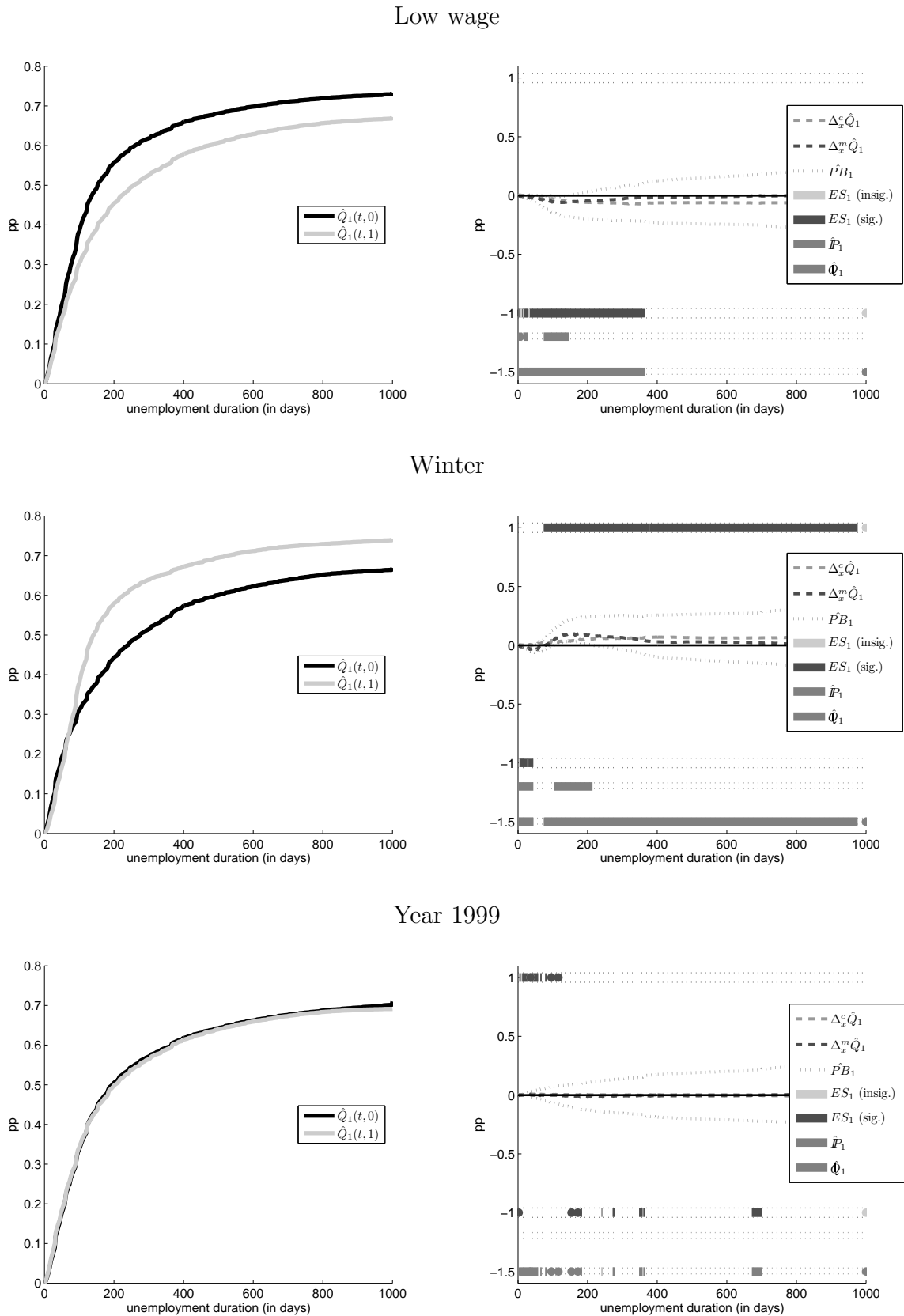


Figure 4: Cumulative incidence for job finding (left) and estimated sign of covariate effects on the marginal distribution (right).



Appendix: Proofs

Proof of Lemma 1: 1.(a) The copula effect and the marginal distribution effect in Definitions 1 and 2 are unique for all t if t^c is unique. The latter is guaranteed by the invertibility of $S(t)$.

1.(b) Result (4) follows from Definitions 1 and 2:

$$\begin{aligned}\Delta_x F_j(t) &= \Delta_x^c F_j(t) + \Delta_x^m F_j(t) \\ &= F_j(t^c; x_1) - F_j(t; x_0) + F_j(t; x_1) - F_j(t^c; x_1) \\ &= F_j(t; x_1) - F_j(t; x_0).\end{aligned}$$

In order to prove 1.(c) we first state and prove two additional Lemmas:

Lemma A1 *Under Assumptions 1 and 2 in model (1), we have $\text{sign} |\Delta_x^c F_1(t)| = -\text{sign} |\Delta_x^c F_2(t)|$ for all t .*

We prove Lemma A1 by stating the condition for t^c and t to satisfy Definitions 1 and 2:

$$S(t; x_0) = C\{S_1(t; x_0), S_2(t; x_0)\} = C\{S_1(t^c; x_1), S_2(t^c; x_1)\} = S(t^c; x_1). \quad (17)$$

A survival copula function is increasing in both of its arguments (see Theorems 2.2.7 and 2.4.4 in Nelsen, 2006). Then the copula effect of risk 1 being positive, i.e. $S_1(t^c; x_1) < S_1(t; x_0)$, implies the copula effect of risk 2 being negative as $S_2(t^c; x_1) < S_2(t; x_0)$ is required for condition (17) to hold. Similar reasoning applies if the copula effect of risk 1 is negative and the copula effect of risk 2 is positive. This proves that the copula effects for risk 1 and risk 2 have different signs for all t unless both are zero.

Lemma A2 *Under Assumptions 1 and 2 in model (1), we have $\text{sign} |\Delta_x^m F_1(t)| = \text{sign} |\Delta_x^m F_2(t)|$ for all t .*

Lemma A2 bases on Lemma A1 in Zheng and Klein (1995) which suggests that $s_2 = \zeta_{21}(s_1; x)$ is unique. This guarantees uniqueness of $s_2 = \zeta_{21}(s_1; x)$ for any given value s_1 and vice versa. $\zeta_{21}(s_1; x)$ is therefore a strictly increasing function in s_1 for all x . Suppose we have $F_1(t_a; x) < F_1(t_b; x)$ for any $t_a, t_b \in \mathbb{T}$ and thus $S_1(t_a; x) > S_1(t_b; x)$. Then by $s_2 = \zeta_{21}(s_1; x)$ being strictly increasing, we also have $S_2(t_a; x) > S_2(t_b; x)$, and thus $F_2(t_a; x) < F_2(t_b; x)$. The same reasoning applies to the cases of $F_1(t_a; x) > F_1(t_b; x)$ or $F_1(t_a; x) = F_1(t_b; x)$, which completes the proof.

These results are now used to prove 1.3). According to (4), the sign of $\Delta_x F_1(t)$ is determined by the sign of the copula and the marginal distribution effect if $\{\text{sign} |\Delta_x^c F_1(t)|, \text{sign} |\Delta_x^m F_1(t)|\}$

belongs to one of the following combinations: $\{1,1\}$, $\{1,0\}$, $\{0,1\}$, $\{0,0\}$, $\{-1,-1\}$, $\{-1,0\}$ or $\{0,-1\}$. Otherwise $\Delta_x^c F_1(t)$ and $\Delta_x^m F_1(t)$ have opposite signs, and in these cases the sign of $\Delta_x F_1(t)$ is undetermined (or ambiguous). In case $\Delta_x^c F_j(t) = 0$ for $j = 1, 2$ or $\Delta_x^m F_j(t) = 0$ for $j = 1, 2$, the sign of $\Delta_x F_1(t)$ and $\Delta_x F_2(t)$ are unambiguously determined. If both the copula and the marginal distribution effects are nonzero, we first consider the case that both $\Delta_x^c F_1(t)$ and $\Delta_x^m F_1(t)$ have the same sign. In this case, the sign of $\Delta_x F_1(t)$ is determined but not the sign of $\Delta_x F_2(t)$. It is because $- \text{sign} |\Delta_x^c F_2(t)| = \text{sign} |\Delta_x^c F_1(t)| = \text{sign} |\Delta_x^m F_1(t)| = \text{sign} |\Delta_x^m F_2(t)|$ according to Lemmas A1 and A2. Analogous reasoning applies if sign of $\Delta_x^c F_1(t)$ and $\Delta_x^m F_1(t)$ are different. As in this case, the sign of $\Delta_x F_2(t)$ is determined but not the sign of $\Delta_x F_1(t)$. It follows that $\text{sign} |\Delta_x F_j(t)|$ can be always determined for at least one j from $\text{sign} |\Delta_x^m F_j(t)|$ and $\text{sign} |\Delta_x^c F_j(t)|$ for all $t \in \mathcal{T}$. \square

Lemma A3 *Under Assumptions 1 and 2 in model (1), if there is some subset of duration time $M \subseteq \mathcal{T}$ such that the link function is monotone in all $t \in M$, we have $Z(t) = \zeta\{S_1(t); x_1 - S_1(t); x_0\} > 0$ or $Z(t) < 0$ for all $t \in M$. Then $Z(t) > 0$ implies $\Delta_x^c F_1(t) > 0$ and $Z(t) < 0$ implies $\Delta_x^c F_1(t) < 0$ for all $t \in M$.*

The proof of Lemma A3 uses Lemma A1 and Lemma A2. As the copula effect for risks 1 and 2 has an opposite sign for all t , we show that $\Delta_x^c F_1(t; x_1)$ cannot be negative if $Z(t) > 0$. This is because $\Delta_x^c F_1(t; x_1) < 0$ implies $\Delta_x^c F_2(t; x_1) > 0$, which means that $S_1(t^c; x_1) > S_1(t; x_0)$ and $S_2(t^c; x_1) < S_2(t; x_0)$ at all t . For some $\bar{t} > t^c$ we have $S_2(\bar{t}; x_1) < S_2(t^c; x_1)$. Moreover we know from Lemma A2 that there exists a \bar{t} such that $S_1(\bar{t}; x_1) < S_1(t^c; x_1)$ and fulfills the condition $S_1(\bar{t}; x_1) = S_1(t; x_0)$. Putting this together,

$$\zeta_{21}(S_1(\bar{t}; x_1); x_1) = S_2(\bar{t}; x_1) < S_2(t; x_0) = \zeta_{21}(S_1(t; x_0); x_0) = \zeta_{21}(S_1(\bar{t}; x_1); x_0).$$

This contradicts the initial condition $Z(t) > 0$ which requires $\zeta_{21}(s_1; x_1) > \zeta_{21}(s_1; x_0)$ for all s_1 . \square

Proof of Lemma 3: The proof comprises several steps. First, we define a sequence $\{\dot{t}_k\}$ and state several useful properties of it in Lemma A4. Second, we show that sign of the copula effect can be identified, if $\{\dot{t}_k\}$ was known (Lemma A5). Third, the unobserved \dot{t} can be bounded by intervals with observed endpoints $[\dot{t}_k, \dot{t}'_k]$ (Lemma A6). Fourth, we use these results to prove Lemma 3.

Definition A1 *We denote $\{\dot{t}_k\} \in \mathcal{T}$ for $k = 0, 1, 2, \dots$ as a finite sequence of durations t at which the link functions for the control group (x_0) and the treatment group (x_1) cross each other, i.e.*

$Z(\dot{t}_k) = \zeta_{21}\{S_1(\dot{t}_k); x_1\} - \zeta_{21}\{S_1(\dot{t}_k); x_0\} = 0$ and $\text{sign} |\lim_{t \rightarrow +\dot{t}_k} Z(t)| \neq \text{sign} |\lim_{t \rightarrow -\dot{t}_k} Z(t)|$ for all $k > 0$. We define $\dot{t}_0 = 0$ and the corresponding sequence $\{\dot{t}_k^c\}$ for the duration time t^c defined in Definition 1 such that $S(\dot{t}_k^c; x_1) = S(\dot{t}_k; x_0)$ for all k .

Lemma A4 Under Assumptions 1 and 2, $\{\dot{t}_k\} \in \mathcal{I}$ has the following properties:

1. The link function $S_2(t) = z_{21}\{S_1(t); x\}$ at any $t \in (\dot{t}_k, \dot{t}_{k+1}) \equiv \dot{\mathcal{I}}_k$, is a strictly monotone function in x , i.e. either (i) $Z(t) < 0$ or (ii) $Z(t) > 0$ for $t \in \dot{\mathcal{I}}_k$ and for all k . The link function is, however, generally non-monotone at any $t \in \{\dot{t}_m, \dot{t}_n\}$, for any $1 < m + 1 < n$.
2. The copula effect at $\{\dot{t}_k\}$ is zero for all k and all risks j , i.e.

$$\begin{aligned} \Delta_x^c F_j(\dot{t}_k) &= F_j(\dot{t}_k^c; x_1) - F_j(\dot{t}_k; x_0) = 0; \\ \text{thus} \quad S_j(\dot{t}_k^c; x_1) &= S_j(\dot{t}_k; x_0). \end{aligned} \tag{18}$$

3. The copula effect on the CIC at $\{\dot{t}_k\}$ is nonzero for any $k > 0$ and all risks j ,

$$\Delta_x^c Q_j(\dot{t}_k) = Q_j(\dot{t}_k; x_1) - Q_j(\dot{t}_k; x_0) \neq 0. \tag{19}$$

4. For any $t \in (\dot{t}_k, \dot{t}_{k+1})$ and $k > 0$,

$$\begin{aligned} A_1(t) &= \int_0^{\zeta_{21}(s_1; x_1)} \int_{S_1(t^c; x_1)}^{S_1(\dot{t}_k^c; x_1)} C'' ds_1 ds_2 - \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(t; x_0)}^{S_1(\dot{t}_k; x_0)} C'' ds_1 ds_2. \\ &\propto \Delta_x^c F_1(t; x). \end{aligned} \tag{20}$$

Lemma A4.1 holds for any link function $s_2 = \zeta_{21}(s_1)$ that is strictly increasing in s_1 . According to Definition A1 $Z(t)$ changes its sign at all $\{\dot{t}_k\}$. Lemma A4.2 can be proved by Definitions 1 and A1. $[S_1(\dot{t}_k^c; x_1), S_2(\dot{t}_k^c; x_1)]$ is the solution to (i) $C\{S_1(\dot{t}_k^c; x_1), S_2(\dot{t}_k^c; x_1)\} = C\{S_1(\dot{t}_k; x_1), S_2(\dot{t}_k; x_1)\}$, (ii) $S_2(\dot{t}_k^c; x_1) = \zeta_{21}\{S_1(\dot{t}_k^c; x_1); x_1\}$, and (iii) $\zeta_{21}\{S_1(\dot{t}_k^c; x_1); x_1\} = \zeta_{21}\{S_1(\dot{t}_k; x_0); x_0\}$. (ii) and (iii) imply $S_2(\dot{t}_k^c; x_1) = S_2(\dot{t}_k; x_0)$. As the copula function is strictly increasing in both of its arguments,

equation (i) holds only if $S_1(\dot{t}_k^c; x_1) = S_1(\dot{t}_k; x_1)$. This leads to equation (18) and completes the proof of Lemma A4.2. Lemma A4.3 can be proved by rewriting $\Delta_x^c Q_1(\dot{t}_k)$ as

$$\begin{aligned}\Delta_x^c Q_1(\dot{t}_k) &= Q_1(\dot{t}_k^c; x_1) - Q_1(\dot{t}_k; x_0) \\ &= \int_0^{\zeta_{21}(s_1; x_1)} \int_{S_1(\dot{t}_k^c; x_1)}^1 C'' ds_1 ds_2 - \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(\dot{t}_k; x_0)}^1 C'' ds_1 ds_2 \\ &= \int_{\zeta_{21}(s_1; x_0)}^{\zeta_{21}(s_1; x_1)} \int_{S_1(\dot{t}_k; x_0)}^1 C'' ds_1 ds_2.\end{aligned}\tag{21}$$

The last equality follows from (18). As $\zeta_{21}(s_1; x_0)$ equals to $\zeta_{21}(s_1; x_1)$ only at \dot{t}_l and \dot{t}_l^c , for all $0 \leq l \leq k$, respectively, $Z(t)$ is nonzero at all $t \in [0, \dot{t}_k]$. This completes the proof of Lemma A4.3. Lemma A4.4 can be shown by using equation (18) and the monotonicity property in Lemma A4.1, which implies that (i) $Z(t) < 0$ or (ii) $Z(t) > 0$ for all $t \in (\dot{t}_k, \dot{t}_{k+1})$. Case (i) implies a negative $\Delta_x^c F_1(t)$ as a result from Lemma A3 and $A_1(t)$ in (20) becomes

$$\begin{aligned}A_1(t) &= \int_0^{\zeta_{21}(s_1; x_1)} \int_{S_1(\dot{t}_k^c; x_1)}^{S_1(\dot{t}_k; x_0)} C'' ds_1 ds_2 - \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(\dot{t}_k; x_0)}^{S_1(\dot{t}_k; x_0)} C'' ds_1 ds_2 \\ &< \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(\dot{t}_k^c; x_1)}^{S_1(\dot{t}_k; x_0)} C'' ds_1 ds_2 - \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(\dot{t}_k; x_0)}^{S_1(\dot{t}_k; x_0)} C'' ds_1 ds_2 \\ &< \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(\dot{t}_k; x_0)}^{S_1(\dot{t}_k; x_0)} C'' ds_1 ds_2 - \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(\dot{t}_k; x_0)}^{S_1(\dot{t}_k; x_0)} C'' ds_1 ds_2 = 0.\end{aligned}$$

This completes the proof of Lemma A4.4.

Lemma A5 *Under Assumptions 1 and 2, we have for any t with $\dot{t}_k = \sup\{\dot{t}_l | \dot{t}_l \leq t\}$,*

$$\begin{aligned}\Delta_x^c Q_j(t) &= \Delta_x^c Q_j(\dot{t}_k) + A_j(t); \text{ and} \\ \text{sign} |\Delta_x^c F_j(t)| &= \text{sign} |\Delta_x^c Q_j(t) - \Delta_x^c Q_j(\dot{t}_k)|.\end{aligned}\tag{22}$$

for $j = 1, 2$, where $\pi_j^c(t) > 0$ and \dot{t}_l as in Definition A1.

In order to obtain (22), we rewrite the copula effect on the CIC for any $t \in [\dot{t}_k, \dot{t}_{k+1}]$ by distin-

guishing between the following two intervals for $t \in (0, t] = (0, \dot{t}_k] \cup (\dot{t}_k, t]$. Then

$$\begin{aligned}
\Delta_x^c Q_1(t) &= Q_1(t^c; x_1) - Q_1(t; x_0) \\
&= \int_0^{\zeta_{21}(s_1; x_1)} \int_{S_1(t^c; x_1)}^1 C'' ds_1 ds_2 - \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(t; x_0)}^1 C'' ds_1 ds_2; \\
&= \int_0^{\zeta_{21}(s_1; x_1)} \int_{S_1(t^c; x_1)}^{S_1(\dot{t}_k; x_1)} C'' ds_1 ds_2 + \int_0^{\zeta_{21}(s_1; x_1)} \int_{S_1(\dot{t}_k; x_1)}^1 C'' ds_1 ds_2 \\
&\quad - \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(t; x_0)}^{S_1(\dot{t}_k; x_0)} C'' ds_1 ds_2 - \int_0^{\zeta_{21}(s_1; x_0)} \int_{S_1(\dot{t}_k; x_0)}^1 C'' ds_1 ds_2; \\
&= A_1(t) + \Delta_x^c Q_1(\dot{t}_k).
\end{aligned}$$

The last line follows from (19) and the definition of $A_1(t)$ in Lemma A4.4. According to Lemma A4.4, the sign of $A_1(t)$ is equal to the sign of $\Delta_x^c F_1(t)$. And thus $\Delta_x^c F_1(t)$ has the same sign as $\Delta_x^c Q_1(t) - \Delta_x^c Q_1(\dot{t}_k)$ for any t such that $\dot{t}_k = \sup\{\dot{t}_l | \dot{t}_l \leq t\}$. This completes the proof of Lemma A5.

Lemma A6 $\dot{t}_l \in [\dot{t}_k, \dot{t}_k]$ for $l \geq k$ and for all k .

In order to prove Lemma A6, we show that $\Delta_x^c Q_j(t)$ has at least one local maximum (or minimum) for all $t \in \mathcal{I}$. This is when the link function $\zeta_{21}(S_1(t); x)$ is not monotone in x for all $t \in \mathcal{I}$ according to Lemma A4.1.

First, we consider the interval $t \in (0, \dot{t}_1) \equiv \dot{\mathcal{I}}_0$. From (22), we have $\Delta_x^c Q_1(t) = A_1(t)$. We know from Lemma A4.4 that $\Delta_x^c Q_1(t)$ has the same sign as $\Delta_x^c F_1(t)$. Without lost of generality, we consider the case $Z(t) > 0$ for all $t \in \dot{\mathcal{I}}_0$. According to Lemma A3 we have $\Delta_x^c F_1(t) < 0$, and thus $\Delta_x^c Q_1(t) < 0$ for all t in the first interval $\dot{\mathcal{I}}_0$. The same is true for $t = \dot{t}_1$ and $\Delta_x^c Q_1(\dot{t}_1) < 0$. Next, we consider the interval $t \in (\dot{t}_1, \dot{t}_2) \equiv \dot{\mathcal{I}}_1$. Using Definition A1, $Z(t)$ changes its sign in $\dot{\mathcal{I}}_1$, and thus $Z(t) < 0$. It follows that $\Delta_x^c F_1(t)$ changes its sign and becomes positive. From (22) and Lemma A4.4, at $t \in \mathcal{I}_1$,

$$\Delta_x^c Q_1(t) = \Delta_x^c Q_1(\dot{t}_1) + A_1(t) > \Delta_x^c Q_1(\dot{t}_1).$$

While we don't know whether $\Delta_x^c Q_1(t) < 0$ reaches its local minimum and starts to move up at some \dot{t}_1 (see Definition 5) prior to \dot{t}_1 , \dot{t}_1 should be no later than \dot{t}_1 . We therefore have $\dot{t}_1 \in (0, \dot{t}_1]$ and thus $\dot{t}_1 \leq \dot{t}_1$. At \dot{t}_2 , $\Delta_x^c Q_1(\dot{t}_2) = \Delta_x^c Q_1(\dot{t}_1) + A_1(\dot{t}_2)$. It is unclear whether $|A_1(\dot{t}_2)|$ is large enough to change the sign of $\Delta_x^c Q_1(t)$ from negative to zero or positive. We consider the simple

case where $|A_1(\dot{t}_2)| \geq |\Delta_x^c Q_1(\dot{t}_1)|$ and $\Delta_x^c Q_1(\dot{t}_2)$ is positive. Since $\Delta_x^c Q_1(\dot{t}_1) < 0$ and $\Delta_x^c Q_1(\dot{t}_2) \geq 0$, there exists a \dot{t}_1 such that $\Delta_x^c Q_1(\dot{t}_1) = 0$ with $\dot{t}_1 \in (\dot{t}_1, \dot{t}_2]$. It implies that $\dot{t}_1 < \dot{t}_1$.

Combining the above results \dot{t}_1 can be bounded by an interval with observable endpoints $[\dot{t}_1, \dot{t}_1)$. In the case where $|A_1(\dot{t}_2)| < |\Delta_x^c Q_1(\dot{t}_1)|$, \dot{t}_1 is only known to be in the interval $\dot{\mathcal{I}}_1, \dot{\mathcal{I}}_3$ or $\dot{\mathcal{I}}_5, \dots$. This is because $A_1(t)$ is positive only in these intervals and thus only when \dot{t}_1 is inside these intervals that could turn $\Delta_x^c Q_1(t)$ to be positive. If, for instance, the actual \dot{t}_1 is inside the interval $\dot{\mathcal{I}}_3$, the three unknowns \dot{t}_1, \dot{t}_2 and \dot{t}_3 and not just \dot{t}_1 are bounded by $[\dot{t}_1, \dot{t}_1)$. This, however, does not affect the identifiability of \dot{t}_k and \dot{t}_k for all k and the identification results of Lemma 3, except that the identification region will be reduced. The above reasoning can be carried over to all \dot{t}_k which completes the proof of Lemma A6.

We now prove Lemma 3 for different intervals of t by using Lemma A6. For simplicity we focus on the case $\dot{t}_k \in [\dot{t}_k, \dot{t}_k]$ for all k . For $t \in [0, \dot{t}_1]$, $\text{sign} |\Delta_x^c F_j(t)| = \text{sign} |\Delta_x^c Q_j(t)|$ as already shown. The lower bound of \dot{t}_1 is \dot{t}_1 and $\dot{t}_0 = 0$. Thus, Lemma 3 holds for $t \in [\dot{t}_0, \dot{t}_1]$. For $t \in [\dot{t}_1, \dot{t}_2]$, we observe only the upper bound of \dot{t}_1 and the lower bound of \dot{t}_2 which is the interval $[\dot{t}_1, \dot{t}_2]$. According to Definition 5, we have the following bounds for $\Delta_x^c Q_1(t)$ for $t \in [\dot{t}_1, \dot{t}_2]$:

$$\begin{aligned} \Delta_x^c Q_1(\dot{t}_1) < \Delta_x^c Q_1(t) < \Delta_x^c Q_1(\dot{t}_2) & \text{ for } \Delta_x^c Q_1(\dot{t}_2) > 0; \\ \Delta_x^c Q_1(\dot{t}_1) > \Delta_x^c Q_1(t) > \Delta_x^c Q_1(\dot{t}_2) & \text{ for } \Delta_x^c Q_1(\dot{t}_2) < 0. \end{aligned}$$

And with $\Delta_x^c Q_1(\dot{t}_1) = 0$ and Lemma A5, we have

$$\begin{aligned} 0 < \Delta_x^c Q_1(t) < \Delta_x^c Q_1(\dot{t}_1) + A_1(\dot{t}_2) & \text{ for } \Delta_x^c Q_1(\dot{t}_2) > 0; \\ 0 > \Delta_x^c Q_1(t) > \Delta_x^c Q_1(\dot{t}_1) + A_1(\dot{t}_2) & \text{ for } \Delta_x^c Q_1(\dot{t}_2) < 0. \end{aligned}$$

Note that $\Delta_x^c Q_1(t)$ is zero at \dot{t}_1 with $\dot{t}_1 \in (\dot{t}_1, \dot{t}_2)$ and thus $\Delta_x^c Q_1(\dot{t}_1)$ and $\Delta_x^c Q_1(\dot{t}_2)$, which are nonzero, must have different directions. It follows that

$$\begin{aligned} 0 < \Delta_x^c Q_1(t) < A_1(\dot{t}_2) & \text{ for } \Delta_x^c Q_1(\dot{t}_2) > 0 \\ 0 > \Delta_x^c Q_1(t) > A_1(\dot{t}_2) & \text{ for } \Delta_x^c Q_1(\dot{t}_2) < 0. \end{aligned}$$

Using Lemma A4.4, completes the proof of Lemma 3 for $t \in [\dot{t}_1, \dot{t}_2]$. The proof is analogous for all other values of k . \square

Proof of Corollary 1: We first state a Lemma that is required for the proof. As $F_j(t; x)$ is strictly increasing in t , we have the following useful property.

Lemma A7 $\Delta_x^m F_j(t)$ and $\Delta_{-x}^m F_j(t)$ have opposite signs for all t and j .

Similar to Lemma 1, the decomposition path of the reversed treatment effect is $\Delta_{-x} F_j(t) = \Delta_{-x}^c F_j(t) + \Delta_{-x}^m F_j(t)$ with $\Delta_{-x}^c F_j(t) = F_j(t^b; x_0) - F_j(t; x_1)$, $\Delta_{-x}^m F_j(t) = F_j(t; x_0) - F_j(t^b; x_0)$ and t^b such that $S(t^b, x_0) = S(t, x_1)$. We have either $t^c < t < t^b$ or $t^c > t > t^b$ because $S(t, x)$ is a strictly decreasing function in t . This implies

$$\begin{aligned} S(t; x_0) &= S(t^c; x_1) > S(t; x_1) = S(t^b; x_0) & \text{if } t^c < t; \\ S(t; x_0) &= S(t^c; x_1) < S(t; x_1) = S(t^b; x_0) & \text{if } t^c > t. \end{aligned}$$

Suppose that $\text{sign}|\Delta_x^c F_j(t)| = \text{sign}|\Delta_x^m F_j(t)| = +1$ for some t in $\mathbb{I}_j(x)$. $\text{sign}|\Delta_x^m F_j(t)|$ is then identified and positive. Suppose also that the same t belongs to $\mathbb{I}_j(-x)$. We now show that the sign of the reversed treatment effect cannot be positive then. According to Lemma A7, the sign of $\Delta_{-x}^m F_j(t)$ at t_k is then negative. This implies that the sign of the reversed treatment effect is either negative (when the sign of $\Delta_{-x}^c F_j(t)$ is negative) or undetermined (when the identified sign of $\Delta_{-x}^c F_j(t)$ is positive). This implies that the two decomposition routes produce different identification sets. While this suggests different $\mathbb{D}_j(x)$ and $\mathbb{D}_j(-x)$, it is also possible to show that $\mathbb{I}_j(x)$ and $\mathbb{I}_j(-x)$ differ. \square

References

- [1] Abbring, J.H. and G.J. van den Berg (2003) The identifiability of the mixed proportional hazards competing risks model, *Journal of the Royal Statistical Society B*, 65, 701–710.
- [2] 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.
- [3] Cox, D.R. (1962), *Renewal Theory*, London.
- [4] Georges, P., Lamy, A.G., Nicolas, E., Quibel, G. and Roncalli, T. (2001), Multivariate survival modelling: A unified approach with copulas. Working paper, Groupe de Recherche Opérationnelle Crédit Lyonnais,
- [5] Hamann, S., G. Krug, M. Köhler, W. Ludwig-Mayerhofer, and A. Hackett (2004), Die IAB-Regionalstichprobe 1975-2001: IABS-R01, *ZA-Information* Vol. 55, 36–42.

- [6] Honoré, B. and Lleras-Muney, A. (2006), Bounds in Competing Risks Models and the War on Cancer, *Econometrica*, 74, 1675–1698.
- [7] Jeong, J.H. and Fine, J.P. (2007), Parametric regression on cumulative incidence function, *Biostatistics*, 8, 184–196.
- [8] Kalbfleish, J.D. and Prentice, R.L. (2002), *The Statistical Analysis of Failure Time Data*, Wiley, Hoboken, N.J.
- [9] Lo, S.M.S. and Wilke, R.A. (2010), A copula model for dependent competing risks, *Journal of the Royal Statistical Society Series C*, 59, 359–376.
- [10] Manski, C.F. (2003) *Partial Identification of Probability Distributions*, New York: Springer-Verlag.
- [11] Nelsen, R.B. (2006), *An Introduction to Copulas*, 2nd Edition, Springer, New York.
- [12] Peng, L. and Fine, J.P. (2009), Competing Risks Quantile Regression, *Journal of the American Statistical Association*, 104, 1440–1453.
- [13] 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.
- [14] 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.
- [15] Silvapulle, M.J. and Sen, P.K. (2005), *Constrained Statistical Inference: Order, Inequality, and Shape Constraints*, New Jersey: Wiley.
- [16] Tsiatis, A. (1975), A Nonidentifiability Aspect of the Problem of Competing Risks, *Proceedings of the National Academy of Sciences*, 72, 20–22.
- [17] Van der Vaart, A.W. (1998), *Asymptotic Statistics*, Cambridge University Press, Cambridge.
- [18] Zheng, M. and Klein, J.P. (1995), Estimates of marginal survival for dependent competing risks based on assumed copula. *Biometrika*, 82, 127–138.