DISCUSSION PAPER SERIES

IZA

IZA DP No. 8839

Identification of the Timing-of-Events Model with Multiple Competing Exit Risks from Single-Spell Data

Bettina Drepper
Georgios Effraimidis

February 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Identification of the Timing-of-Events Model with Multiple Competing Exit Risks from Single-Spell Data

Bettina Drepper
Tilburg University
and IZA

Georgios Effraimidis
University of Southern Denmark

Discussion Paper No. 8839
February 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 8839
February 2015

ABSTRACT

Identification of the Timing-of-Events Model with Multiple Competing Exit Risks from Single-Spell Data*

This note describes how the (single-spell) identification result of the timing-of-events model by Abbring and Van den Berg (2003b) can be extended to a model that accommodates several competing exit risks. The extended model can be used for example to distinguish between the different effects of a benefit sanction on several competing exit risks out of unemployment such as ‘finding work’ vs. ‘exiting the labor force’. By allowing for a flexible dependence structure between competing exit risks and the duration until entry into treatment, the model can take account of selection into treatment and dependencies between competing exit risks by way of unobservables.

JEL Classification: C41, C31, J64

Keywords: competing risks, treatment effects, multivariate duration analysis, mixed proportional hazard, timing-of-events

Corresponding author:

Bettina Drepper
Tilburg University
Department of Econometrics & OR
Warandelaan 2
5037 AB Tilburg
The Netherlands
E-mail: b.drepper@uvt.nl

* We thank Jaap Abbring, Gerard J. van den Berg, Petyo Bonev, Andrew Chesher, Bo Honoré, Enno Mammen, Sebastian Klein and Geert Ridder for helpful comments. The research of Bettina Drepper is financially supported by the Netherlands Organisation for Scientific Research (NWO) through Vici grant 016.125.624.

1 Introduction

This note presents an identification result for a model that combines two popular multivariate duration models: the mixed proportional hazard (MPH) competing risks model and the timing-of-events approach.[1]

Competing risks models are used when the duration of interest can end in several mutually exclusive exit states (competing risks) and exit rates are dependent by way of unobservables. Popular fields of application include mortality studies where a lifespan can be ended by different competing causes of death (see, e.g. Escarela and Carriere, 2003; Honoré and Lleras-Muney, 2006), as well as labor market studies where spells of unemployment can end in different exit states, such as full-time vs. part-time employment (McCall, 1997), finding a new job vs. a recall to the same firm (Han and Hausman, 1990; Katz and Meyer, 1990) or employment vs. non-participation (Van den Berg et al., 2007). It is very common that such competing exit risks are dependent conditional on covariates; i.e. the decision to look for employment or exit the labor force is often driven by similar unobserved attributes of the job searcher, such as skills, preferences or motivation.

The timing-of-events approach is designed to identify the effect of an endogenous treatment on the subsequent rate to exit the state of interest. Multiple empirical studies have used this approach to evaluate the effect of active labor market programs or benefit sanctions on the rate of finding work (e.g. see Lalive et al., 2002; Van den Berg et al., 2004; Abbring et al., 2005; Rosholm and Svarer, 2008).[2] Endogeneity of the treatment is a common problem in many applications, since the speed at which the treatment occurs is often influenced by similar unobserved characteristics that also affect the outcome variable, e.g.

[1] The identification of competing risks models is addressed by Heckman and Honoré (1989) and Abbring and Van den Berg (2003a), while identification of the timing-of-events model is addressed in Abbring and Van den Berg (2003b).

[2] Other applications include the effect of patent grants on the timing of licensing by start-up technology entrepreneurs (Gans et al., 2008), the effect of cannabis use on cocaine use (Van Ours, 2003) and school dropout (Van Ours and Williams, 2009), the effect of child birth on relationship duration (Svarer and Verner, 2008), and the effect of bereavement on the spouse (Van den Berg et al., 2011) or co-twins's survival (Van den Berg and Drepper, 2012).

how quickly a job searcher enters into a labor market program depends on his skill level, preferences and motivation, which also affect his success in finding work. This endogeneity can be accounted for by adding the hazard of entering into treatment as a second equation to the model while allowing for dependent unobservables across both equations. Abbring and Van den Berg (2003b) show that when the hazard rates are multiplicative in all its components (MPH-like structure), the resulting bivariate duration model can be identified from single-spell (non-repeated spell) data.

By design, the timing-of-events model accounts for the effect of an endogenous treatment[3] on a *single* exit risk out of unemployment such as finding full-time employment. Exit states that occur other than the one exit of interest (e.g. exiting the labor force) are dealt with by right-censoring the unemployment spells at the point of exit (e.g., see Van den Berg et al., 2004; Abbring et al., 2005). This convenient solution is only valid under the strong assumption that the competing exit risks are independent conditional on covariates and the treatment. This assumption is easily violated in many applications, such as with competing exits out of unemployment, all of which are usually influenced by unobserved attributes of the job searcher. Ignoring this dependence leads to a misspecification of the likelihood function and thus to incorrect statistical inference.

In this note, we present a new identification result that extends the timing-of-events approach by Abbring and Van den Berg (2003b) to accommodate the different effects of an endogenous treatment on multiple[4] competing exit risks that can be dependent by way of unobserved characteristics.[5] Our result relies on similar assumptions as the timing-of-events model with a single exit risk and we allow for the same flexibility in each component of the hazard rates, including the treatment effect functions.

Identification of the timing-of-events model with J competing exit risk equations has to

[3]Identification of the effects of multiple treatments on a single exit risk can be directly achieved by making use of the result of Abbring and Van den Berg (2003b) and is not the focus of this note.

[4]The number of different competing exit risk equations that can be identified depends on the degree of covariate variation in the dataset (see Section 3).

[5]Identification is achieved from single-spell data where treatment effects are assumed to be homogeneous across units.

be achieved from a limited set of observable distribution functions. Note that with a single exit risk, the distribution of the outcome duration is fully observable, because other exit states are dealt with by assuming independent right-censoring. However, when one accounts for the dependence induced by unobservables between J mutually exclusive exit states, the joint distribution of the J corresponding outcome durations is not fully observable, i.e. they are latent outcomes. Only the joint distribution of their minimum and the type of exit is observed.[6] In this note, we show that all components of the $J + 1$ hazard equations including J exit risk specific treatment effect functions can be identified from the observable distributions in single-spell (non-repeated spell) data.

Some empirical studies estimate timing-of-events models with multiple competing exit risks or larger models that embed it as a sub-model. Such examples include evaluating the effect of benefit sanctions on the rate of finding work vs. leaving the labor force (e.g. see Arni et al., 2013) and estimating a home ownership effect on the rate of finding a local vs. a geographically distant job (Munch et al., 2006, 2008; Battu et al., 2008)[7]. These studies have to rely on multiple-spell data to identify the effects of the endogenous treatment, as well as making the restrictive assumption that unobserved characteristics such as the skill level, preferences and motivation of the job searcher remain constant across repeated spells of unemployment.[8] In some applications, multiple-spells are only available for a subset of the sample; indeed, such data is never available in some fields, as with life-span data or data on first-time substance use[9].

In this note, identification is derived from single-spell data. In view of this, the presented result is relevant for empirical work that evaluates endogenous treatments with single-spell

[6]It is a well-known property of competing risks models that the joint distribution of the minimum and the indicator of the type of exit is not sufficient to identify the joint distribution of the J outcome durations as long as no additional assumptions on their dependence structure is imposed (Cox, 1959; Tsiatis, 1975).

[7]See also Van Leuvensteijn and Koning (2004), who measure the home ownership effect on the rate to transition from employment to a new job vs. unemployment vs. non-participation.

[8]The identification result of Abbring and den Berg (2003) for the MPH competing risks model allows for endogenous covariates (such as a binary treatment variable) if multiple-spell data with constant unobservables across repeated spells is available.

[9]The identification result of this paper is used in Drepper and Effraimidis (2013) to identify the interdependencies between siblings' first-time use of marijuana.

datasets or multiple-spell datasets, where the assumption of constant unobservables across repeated spells is considered too restrictive.

We introduce the timing-of-events model with competing exit risks in the next section and present the corresponding identification result in Section 3, before Section 4 concludes.

2 Timing-of-events model with two competing exit risks

The first component of the model proposed in this note is the MPH competing risks model. For ease of exposition, the case with two competing exit risks ($J = 2$) is presented here. The generalization of our result to more than two competing exit states is straightforward and briefly addressed at the end of Section 3.

At time $t_0 = 0$, a unit enters into the state of interest (e.g. a worker enters into unemployment). In this state, for each point in time $t \in \mathbb{R}_+$ the unit is exposed to two competing exit risks (e.g. 'finding work' vs. 'exiting the labor force'). The non-negative random variables Y_1 and Y_2 denote the two corresponding latent durations until the unit exits to state 1 or 2. Once the first exit takes place, the other duration is right-censored at that point.[10] Thus, the joint distribution of Y_1 and Y_2 is not fully observed. Instead, one observes the joint distribution of the minimum of Y_1 and Y_2 and the indicator of the first realized exit: (Y, I), with $Y = \min_{j\in\{1,2\}}(Y_j)$ and $I = \arg\min_{j\in\{1,2\}}(Y_j)$.[11]

The two hazard rates that correspond to the latent durations Y_1 and Y_2 are each assumed to follow the well-known MPH structure:

$$\begin{aligned} \theta_1(t|x, V_1) &= \lambda_1(t)\ \phi_1(x)\ V_1 \\ \theta_2(t|x, V_2) &= \lambda_2(t)\ \phi_2(x)\ V_2 \end{aligned} \tag{1}$$

[10]Note that right-censoring is not independent here, due to the existence of dependent unobservables.

[11]The distribution of (Y, I) is not sufficient to identify the joint distribution of (Y_1, Y_2) as long as the dependence structure between Y_1 and Y_2 remains unrestricted (Cox, 1959; Tsiatis, 1975). This nonidentification result can be overturned when sufficient covariate variation is available (Heckman and Honoré, 1989). Abbring and Van den Berg (2003a) show that the necessary variation in covariates can be relaxed considerably in the mixed proportional hazard competing risks model.

The functions λ_1 and λ_2 capture risk-specific duration dependence. Conditional on the vector of covariates x that enters into both hazard rates, a dependence is introduced between θ_1 and θ_2 through a vector of non-negative random variables $V = (V_1, V_2)$, which is jointly drawn from a bivariate distribution G. In the job search example, V_1 reflects how the unobserved characteristics such as the skills, preferences or motivation of the job searcher influence her hazard rate to find a new job, while V_2 reflects the influence of the same unobserved characteristics on her hazard rate to exit the labor force. Naturally, V_1 and V_2 are dependent in many applications.

We now introduce an endogenous treatment to the MPH competing risk model in (1). Let S denote the spell duration from t_0 to the time when the unit enters into treatment (e.g. the time spent in unemployment before the case worker imposes a benefit sanction). Once the unit enters into treatment at time $S = s$, the two competing exit hazards of 'finding work' vs. 'exiting the labor force' are affected for all $t > s$. The dynamic effects of this treatment enter through two risk-specific treatment effect functions, $\delta_1(t|S,x)$ and $\delta_2(t|S,x)$.

Model I (Timing-of-events model with two competing exit risks):
The hazard rates of $Y_1|(S,x,V_1)$ *and* $Y_2|(S,x,V_2)$ *are given by*

$$\theta_1(t|S,x,V_1) = \lambda_1(t)\ \phi_1(x)\ \delta_1(t|S,x)^{\mathbb{I}(t>S)}\ V_1$$
$$\theta_2(t|S,x,V_2) = \lambda_2(t)\ \phi_2(x)\ \delta_2(t|S,x)^{\mathbb{I}(t>S)}\ V_2,$$

where $\mathbb{I}$ *is the indicator function. The hazard rate of* $S|(x,V_S)$ *is given by*

$$\theta_S(s|x,V_S) = \lambda_S(s)\ \phi_S(x)\ V_S.$$

The random vector $(V_1\ V_2\ V_S)'$ *is jointly drawn from the trivariate cumulative distribution function* G*.*

The hazard of entering into treatment θ_S is influenced by covariates x, as well as unobserved characteristics of the job searcher (V_S). The dependence between V_S, V_1 and V_2 is captured by the trivariate distribution G, which - aside from a finite means assumption - remains unrestricted in the identification result in Section 3.

A crucial feature of MODEL I is that it allows for a completely separate treatment effect function $\delta_j(t|S,x)$ for each of the two competing exit hazards $j = 1, 2$. For example, some labor market programs may be rather ineffective in increasing the rate to find work yet may have a strong impact on the decision of the job searcher to exit the labor force. It is not only the immediate impact of the treatment that may strongly differ across competing exit risks, but also their dynamics. A labor market program that succeeds in increasing the human capital of the job searcher will have long-lasting positive effects on the hazard to find work, while the potential effect on the risk to exit the labor force may be short-lived. Furthermore, the dependence of the treatment effect on the elapsed duration spent in unemployment at the time of treatment as well as on covariates x may differ across competing exit risks.[12]

In the special case that the state of interest has only one relevant exit state ($J = 1$), MODEL I reduces to the well-known (single-risk) timing-of-events model. In this case, the distribution of the single outcome duration Y_1 is fully observed. Abbring and Van den Berg (2003b) show that all functions of their (single-risk) model can be identified from the set of observed distributions: (Y_1, S) for $Y_1 > S$ and (Y_1) for $Y_1 < S$.[13]

Once an additional competing exit risk is introduced ($J = 2$), the joint distribution of (Y_1, Y_2) is not fully observed due to the nature of competing risks, which are mutually exclusive and may be dependent by way of unobservables. Instead, only the distribution (Y, I) is observed, with $Y = \min_{j\in\{1,2\}}(Y_j)$ and $I = \arg\min_{j\in\{1,2\}}(Y_j)$. In the following section, we show that all model components of MODEL I including the exit risk-specific

[12]The timing-of-events model with competing exit risks has some similarities to the event history approach to program evaluation proposed by Abbring (2008). The main difference is that the model proposed in this note does not rely on the semi-Markov assumption that the transition model with multiple states by Abbring (2008) is based on. This assumption rules out a dependence of the treatment effects on the time spend in unemployment which is a characteristic feature of the timing-of-events approach.

[13]Note that the distribution of the treatment S is not fully observed, because once a job searcher finds employment, her duration until treatment is right-censored.

treatment effect functions δ_1 and δ_2 are identified from the observed distributions: (Y, I, S) for $Y > S$ and (Y, I) for $Y < S$.

Estimation methods of the proposed timing-of-events model with competing exit risks are similar to those used for the well-established single-risk case. Functional form assumptions specific to the application are imposed on the different components of the hazard rate, whereby parametric maximum likelihood methods can be applied. In applications in economics, the multidimensional unobserved heterogeneity distribution is often approximated using the mass-point approach introduced by Heckman and Singer (1984). Naturally, the complexity of the unobserved heterogeneity distribution increases with each additional competing exit risk equation. Restrictions on the dependence structure between competing exit risks can help to estimate models with a larger number of equations.

3 Main result

Before stating the main identification result, we present the following standard technical conditions regarding the underlying model (for all $j = 1, 2$).

Assumption 1 *Each covariate effect function* $\phi_j : \mathbb{X} \rightarrow (0, \infty)$, $\phi_S(x) : \mathbb{X} \rightarrow (0, \infty)$ *is a continuous function with* $\phi_j(x^*) = \phi_S(x^*) = 1$ *for some* $x^* \in \mathbb{X}$.
Furthermore, $(\phi_1(x), \phi_2(x), \phi_S(x); x \in \mathbb{X})$ *contains a non-empty open subset of* $\mathbb{R}^3_+$.

Assumption 2 *The functions* $\lambda_j : \mathbb{R}_+ \rightarrow (0, \infty)$ *and* $\lambda_S : \mathbb{R}_+ \rightarrow (0, \infty)$ *are measurable. The integrated baseline hazard rates* $\Lambda_j(t) := \int_0^t \lambda_j(\omega) d\omega$ *and* $\Lambda_S(t) := \int_0^t \lambda_S(\omega) d\omega$ *exist and are finite for all* $t > 0$ *with* $\Lambda_j(t^*) = \Lambda_S(t^*) = 1$ *for some particular* $t^* > 0$.

Assumption 3 *The trivariate cumulative distribution function* G *does not depend on* x *and*

$$\mathbb{E}(V_j) < \infty, \quad \mathbb{E}(V_S) < \infty.$$

Assumption 4 *The treatment effect function* $\delta_j : \{(t, \tau) \in \mathbb{R}_+^2 : t > \tau\} \times \mathbb{X} \to (0, \infty)$ *is measurable. Moreover, the quantities*

$$\Upsilon_j(t|s, x) := \int_s^t \lambda_j(\omega)\delta_j(\omega|s, x)d\omega,$$
$$\Delta_j(t|s, x) := \int_0^t \delta_j(\omega|s, x)d\omega$$

exist, are finite, and are either cadlag or caglad in s.

Assumption 1 ensures that there is sufficient variation of the covariate effects across the two competing exit durations and the duration to treatment. This assumption is discussed further below. Assumption 2 addresses the functional form of the baseline hazard. The function space is restricted to integrable functions, which is in line with most applied work. Assumption 2 includes, for example, the case of piecewise constant, Weibull or Gombertz baseline hazard specifications, which are widely used in empirical studies. Assumption 3 is a common assumption in single-spell mixed proportional hazard type models (e.g. Elbers and Ridder, 1982). Assumption 4 deals with measurability and finiteness conditions of the treatment effect functions. These conditions are not restrictive in the sense that they allow for a wide range of parametric families.

Assumptions 1 - 4 are almost identical to the Assumptions of Abbring and Van den Berg (2013). The main difference results from the extension to multiple competing exit risks, which requires more covariate variation (Assumption 1) and increases the dimensions of the unobserved heterogeneity distribution G (Assumption 3).[14]

We now introduce the main result. Recall that if the realization of the treatment occurs before the first exit, i.e. $Y > S$, we observe (S, Y, I) and if $Y < S$, we only observe (Y, I). Let $-j = 2$ if $j = 1$ and $-j = 1$ if $j = 2$. In a large dataset, the following subsurvival

[14] Assumption 4 is slightly weaker than Assumption 4 of Abbring and Van den Berg (2003b) since instead of continuity, we only impose continuity either from the left or from the right.

functions are observable (for $j = 1, 2$)

$$Q_{Y_j,S}(y, s|x) := \mathbb{P}(Y_j > y, Y_{-j} > Y_j, S > s, Y > S|x), \tag{2}$$

$$Q_{Y_j}(y|x) := \mathbb{P}(Y_j > y, Y_{-j} > Y_j, S > Y|x) \tag{3}$$

for all $(y, s, x) \in \mathbb{R}^2_+ \times \mathbb{X}$. Define, $Q_S(y, s|x) := \mathbb{P}(Y > y, S > s, Y > S|x) = Q_{Y_1,S}(y, s|x) + Q_{Y_2,S}(y, s|x)$ and let $Q^0_S(s|x) = Q_S(0, s|x)$. Note that the distribution of (S, Y, I) for $Y > S$, and (Y, I) for $Y < S$ is fully characterized by (2) and (3). Next, we state the main result of the paper:

Proposition 1: *Let the Assumptions 1-4 hold. Then, the functions Λ_1, ϕ_1, Λ_2, ϕ_2, Λ_S, ϕ_S, G, Δ_1, and Δ_2 are identified from the observable functions $\{Q_{Y_1}, Q_{Y_2}, Q_{Y_1,S}, Q_{Y_2,S}\}$.*

For ease of exposition, Model I and correspondingly Proposition 1 is presented here for the special case of $J = 2$. It is straightforward to extend this result to J competing exit risks, where J is a positive finite integer. In this case, Assumption 1 has to be adapted such that $(\phi_1(x), ..., \phi_J(x), \phi_S(x); x \in \mathbb{X})$ has to contain a nonempty open subset in $\mathbb{R}^{J+1}_+$. Note that this assumption on available covariate variation across the $J+1$ model equations limits to the number of competing exit risks that can be included in the model. For example, for $J = 2$ and three covariates with $\phi_q(x) = exp(x'\beta_q)$ $\forall q \in \{1, 2, S\}$, it is sufficient for Assumption 1 that $(\beta_1\ \beta_2\ \beta_S)$ has full rank and $\mathbb{X}$ contains a non-empty open set in $\mathbb{R}^3_+$. Thus in most applications, where this rank condition is fulfilled, $J+1$ continuous covariates will generate sufficient variation for Assumption 1 to hold.[15]

Proof. [Proof of Proposition 1] The joint distribution of the identified minimum of (Y_1, Y_2, S) and the identity of this smallest duration is fully characterized by $\{Q_{Y_1}, Q_{Y_2}, Q^0_S\}$ (Tsiatis, 1975). From this, it follows that under Assumptions 1-3 the functions $\Lambda_1, \phi_1, \Lambda_2, \phi_2, \Lambda_S, \phi_S$,

[15] For a detailed discussion of this assumption for two equations, see Abbring and Van den Berg 2003a.

and G are identified from $\{Q_{Y_1}, Q_{Y_2}, Q_S^0\}$ (Abbring and Van den Berg, 2003a).

In the sequel, we focus on the identification of Δ_1 and Δ_2. Let $\mathcal{L}_G$ express the trivariate Laplace transform of the random vector $(V_1\ V_2\ V_S)'$. For almost all $y \in \mathbb{R}_+$ and all $x \in \mathbb{X}$, we have

$$\frac{\partial Q_{Y_j}(y|x)}{\partial y} = \mathcal{L}_G^j(\phi_1(x)\Lambda_1(y), \phi_2(x)\Lambda_2(y), \phi_S(x)\Lambda_S(y))\lambda_j(y)\phi_j(x), \tag{4}$$

where the notation $\mathcal{L}_G^j$ represents the corresponding partial derivative for $j \in \{1, 2\}$. Additionally, for almost all $(s, y) \in \mathbb{R}_+^2$ with $y > s$ and all $x \in \mathbb{X}$, we have

$$\begin{aligned}\frac{\partial^2 Q_{Y_j,S}(y, s|x)}{\partial s \partial y} &= \mathcal{L}_G^{j,3}\left(\phi_1(x)(\Lambda_1(y) + \Upsilon_1(y|s, x)), \phi_2(x)(\Lambda_2(y) + \Upsilon_2(y|s, x)), \phi_S(x)\Lambda_S(s)\right) \\ &\quad \times \lambda_S(s)\phi_S(x)\phi_j(x)\lambda_j(y)\delta_j(y|s, x), \end{aligned}\tag{5}$$

where $\mathcal{L}_G^{j,3}$ denotes the corresponding mixed partial derivative for $j \in \{1, 2\}$.

The above equations imply that for any $y \in \mathbb{R}_+$ and all $x \in \mathbb{X}$, we have

$$\lambda_j(y) = \left[\mathcal{L}_G^{(j)}(\phi_1(x)\Lambda_1(y), \phi_2(x)\Lambda_2(y), \phi_S(x)\Lambda_S(y))\phi_j(x)\right]^{-1} \frac{\partial Q_{Y_j}(y|x)}{\partial y}. \tag{6}$$

Similarly, we obtain for each $(s, y) \in \mathbb{R}_+^2$ with $y > s$ and all $x \in \mathbb{X}$ and $j \in \{1, 2\}$

$$\begin{aligned}\lambda_j(y)\delta_j(y|s, x) = \Big[&\mathcal{L}_G^{(1j)}(\phi_1(x)(\Lambda_1(y) + \Upsilon_1(y|s, x)), \phi_2(x)\left(\Lambda_2(y) + \Upsilon_2(y|s, x)\right), \\ &\phi_s(x)\Lambda_S(s))\lambda_S(s)\phi_S(x)\phi_j(x)\Big]^{-1} \frac{\partial^2 Q_{Y_j,S}(y, s|x)}{\partial s \partial y}. \end{aligned}\tag{7}$$

For the remainder of the proof, we fix s and x. Define $\Lambda_s := \Lambda_S(s)$. Moreover, let $\mathcal{H}_j(y) := \Lambda_j(y)$ and $\mathcal{Q}_j(y) := \frac{\partial Q_{Y_j}(y|x)}{\partial t}$ for $0 \leq y \leq s$, and $\mathcal{H}_j(y) := \Lambda_j(s) + \Upsilon_j(y|s, x)$ and $\mathcal{Q}_j(y) := \frac{\partial^2 Q_{Y_j,S}(y,s|x)}{\partial s \partial y}$ for $y > s$. Finally, we define $g_j{:=}\lambda_S(s)\phi_S(x)\phi_j(x)$ and supress the dependence of ϕ_j on x.

Hence, for almost all $y \in (0, \infty)$ we have a system of two differential equations in the sense of Carathéodory (1918) (see Walter, 1998), i.e.

$$\begin{aligned}&\frac{d}{dy}\mathcal{H}(y) = f\left(y, \mathcal{H}(y)\right), \\ &\quad \mathcal{H}(\tau) = \gamma_\tau \in \mathbb{R}_+^2, \text{ for some specific } \tau \in (0, s) \qquad \text{(initial conditions)}, \end{aligned}\tag{8}$$

where $\mathcal{H} := (\mathcal{H}_1 \ \mathcal{H}_2)'$ and $f := (f_1 \ f_2)'$, with

$$f_j(y, \mathcal{H}) = \begin{cases} \left[\mathcal{L}_G^{(j)}\left(\phi_1 \mathcal{H}_1, \phi_2 \mathcal{H}_2, \phi_3 \Lambda_S(y)\right) \phi_j\right]^{-1} \mathcal{Q}_j(t) & \text{if } 0 < y \leq s, \\ \left[\mathcal{L}_G^{(j3)}\left(\phi_1 \mathcal{H}_1, \phi_2 \mathcal{H}_2, \phi_3 \Lambda_S\right) g_j\right]^{-1} \mathcal{Q}_j(t) & \text{if } y > s. \end{cases}$$

Note that for given $(s, x) \in \mathbb{R}_+ \times \mathbb{X}$ we can choose a $\tau \in (0, s)$ which yields the initial conditions $\mathcal{H}(\tau) = (\mathcal{H}_1(\tau) \ \mathcal{H}_2(\tau))' = (\Lambda_1(\tau) \ \Lambda_2(\tau))' = \gamma_\tau$ as the functions Λ_1 and Λ_2 have been already identified (see the first paragraph). Moreover, the rest quantities on the right-hand side from the above equation are identified by the first step of the current proof (see the fist paragraph). Furthermore, the quantity $\mathcal{Q}_j$ is observed from the data. By making use of Lemma 1, the above system of differential equations has a unique solution for each $x \in \mathbb{X}$ and almost all $s \in \mathbb{R}_+$. Recall that $\Upsilon_1(y|s, x)$ and $\Upsilon_2(y|s, x)$ are either cadlag or caglad with respect to y. The above discussion implies that the quantities Υ_1 and Υ_2 are uniquely identified. By definition, the latter yields identification of Δ_1 and Δ_2. ■

4 Conclusion

This note presents identification conditions for a multivariate duration model that combines the MPH competing risks model with the timing of events model. The resulting model can be seen as a competing risks model that allows for an endogenous treatment time variable, or as an extension of the timing-of-events model to multiple competing exit risks. Identification is achieved from single (non-repeated) spell data using similar assumptions as in the single-risk version of Abbring and Van den Berg (2003b). A crucial feature of the extended model is that it allows identifying separate treatment effect functions for each competing exit risk equation.

References

Abbring, J. H. (2008). The event-history approach to program evaluation. In *Advances in Econometrics*, pp. 33 – 55.

Abbring, J. H. and G. J. V. den Berg (2003). The identifiability of the mixed proportional hazards competing risks model. *Journal of the Royal Statistical Society. Series B 65*(3), 701–710.

Abbring, J. H. and G. J. Van den Berg (2003). The nonparametric identification of treatment effects in duration models. *Econometrica 71*(5), 1491–1517.

Abbring, J. H., G. J. van den Berg, and J. C. van Ours (2005). The Effect of Unemployment Insurance Sanctions on the Transition Rate from Unemployment to Employment. *Economic Journal 115*(July), 602–630.

Arni, P., R. Lalive, and J. V. Ours (2013). How effective are unemployment benefit sanctions? Looking beyond unemployment exit. *Journal of Applied Econometrics 28*, 1153–1178.

Battu, H., A. Ma, and E. Phimister (2008, March). Housing Tenure, Job Mobility and Unemployment in the UK. *The Economic Journal 118*(527), 311–328.

Carathéodory, C. (1918). Vorlesungen über reelle Funktionen. *Leipzig: Teubner*.

Cox, D. R. (1959). The Analysis of Exponentially Distributed Life-Times with Two Types of Failure. *Journal of the Royal Statistical Society Series B 21*(2), 411–421.

Drepper, B. and G. Effraimidis (2013). Social Interaction Effects in Duration Models.

Escarela, G. and J. F. Carriere (2003). Fitting competing risks with an assumed copula. *Statistical Methods in Medical Research 12*, 333–349.

Gans, J. S., D. H. Hsu, and S. Stern (2008). The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. *Management Science*.

Han, A. and J. A. Hausman (1990). Flexible Parametric Estimation of Duration and Competing Risk Models. *Journal of Applied Econometrics 5*(1), 1–28.

Heckman, J. and B. Singer (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica 52*(2), 271–320.

Heckman, J. J. and B. E. Honoré (1989). The identifiability of the competing risks model. *Biometrika 76*(2), 325–330.

Honoré, B. E. and A. Lleras-Muney (2006). Bounds in Competing Risks Models and the War on Cancer. *Econometrica 74*(6), 1675–1698.

Katz, F. and B. D. Meyer (1990). The impact of the potential duration of unemployment benefits on the duration of unemployment. *Journal of Public Economics 41*, 45–72.

Lalive, R., J. C. van Ours, and J. Zweimuller (2002). The Effect of Benefit Sanctions on the Duration of Unemployment.

McCall, B. P. (1997). The Determinants of FullTime versus PartTime Reemployment following Job Displacement. *Journal of Labor Economics 15*(4), 714–734.

Munch, J. R., M. Rosholm, and M. Svarer (2006, October). Are Homeowners Really More Unemployed? *The Economic Journal 116*(514), 991–1013.

Munch, J. R., M. Rosholm, and M. Svarer (2008, January). Home ownership, job duration, and wages. *Journal of Urban Economics 63*(1), 130–145.

Rosholm, M. and M. Svarer (2008, June). The Threat Effect of Active Labour Market Programmes. *Scandinavian Journal of Economics 110*(2), 385–401.

Svarer, M. and M. Verner (2008). Do children stabilize relationships in Denmark? *Journal of Population Economics 21*, 395–417.

Tsiatis, A. (1975, January). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences 72*(1), 20–22.

Van den Berg, G. J. and B. Drepper (2012). A unique bond: Does losing your co-twin affect your remaining life-span?

Van den Berg, G. J., M. Lindeboom, and F. Portrait (2011, July). Conjugal bereavement effects on health and mortality at advanced ages. *Journal of Health Economics 30*(4), 774–94.

Van den Berg, G. J., B. Van der Klaauw, and J. C. Van Ours (2004). Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics 22*(1), 211–241.

Van den Berg, G. J., a. G. C. Van Lomwel, and J. C. Van Ours (2007, March). Nonparametric estimation of a dependent competing risks model for unemployment durations. *Empirical Economics 34*(3), 477–491.

Van Leuvensteijn, M. and P. Koning (2004, May). The effect of home-ownership on labor mobility in the Netherlands. *Journal of Urban Economics 55*(3), 580–596.

Van Ours, J. C. (2003, July). Is cannabis a stepping-stone for cocaine? *Journal of health economics 22*(4), 539–54.

Van Ours, J. C. and J. Williams (2009). Why parents worry: initiation into cannabis use by youth and their educational attainment. *Journal of Health Economics 28*, 132–142.

Walter, W. (1998). *Ordinary Differential Equations*. Springer.

Appendix

The appendix presents a technical result that is employed for the proof of the main result. Let $\mathcal{H}_\rho : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous as well as almost everywhere differentiable function for $\rho = A, B$, and define $\mathcal{H} := (\mathcal{H}_A \; \mathcal{H}_B)'$. Consider also the functions $\mathcal{Q}_\rho(t) : \mathbb{R}_+ \to \mathbb{R}_+$, $r_\rho : \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}_+$, and let $f_\rho(t, \mathcal{H}) := \mathcal{Q}_\rho(t) \; r_\rho(t, \mathcal{H})$ for $\rho = A, B$. We study the following system of first order differential equations

$$\frac{d}{dt}\mathcal{H}(t) = f\left(t, \mathcal{H}(t)\right),$$
$$\mathcal{H}(\tau) = \gamma_\tau, \text{ for some specific } \tau \in (0, \infty) \qquad \text{(initial conditions).} \tag{A.1}$$

Note that a similar problem has been studied by Abbring and Van den Berg (2003a); however, our problem is different as the function r_ρ also depends on the variable t. The next Lemma establishes existence and uniqueness of a solution for (A.1).

Lemma 1 *Consider the initial value problem (A.1). Suppose that i) $\mathcal{Q}_\rho(t)$ is measurable and integrable function for any $t \in \mathbb{R}_+$, ii) $r_\rho\left(t, \mathcal{H}\right)$ is continuously differentiable in $\mathcal{H}$ for any $t > 0$, and iii) $\partial r_j\left(t, \mathcal{H}\right)/\partial \mathcal{H}$ is continuous for all $\mathcal{H} \in \mathbb{R}_+^2$. Then, there exists a unique solution to the (A.1).*

Proof. Let $\mathbb{S} = \mathbb{T} \times \mathbb{H}$ with $\mathbf{T} = [\tau, \tau + a]$ for some $a > 0$ and $\mathbb{K} \subset (0, \infty)^2$ to be a closed ball. By the imposed conditions of the Lemma, we know that $\mathcal{Q}_A(t)$ and $\mathcal{Q}_B(t)$ are measurable and integrable functions for any $t \in \mathbb{R}_+$. Hence, we can claim that $f\left(t, \mathcal{H}\right)$ is continuous as a function of $\mathcal{H}$ in $\mathbb{H}$ for fixed t, and integrable as well as measurable as a function of t over $\mathbb{T}$ for fixed $\mathcal{H}$ (i.e. f satisfies the Carathéodory conditions). Our goal is to prove that f satisfies the following generalized Lipschitz condition for $(t, \mathcal{H}), (t, \mathcal{H}^*) \in \mathbb{S}$

$$\|f(t, \mathcal{H}) - f(t, \mathcal{H}^*)\| \leq l(t) \left\|\mathcal{H} - \mathcal{H}^*\right\|, \tag{A.2}$$

where the function $l(t)$ is measurable and integrable over $\mathbb{T}$. Here, we use $\|.\|$ to denote the Frobenius norm for a matrix.

Define $r := (r_A \; r_B)'$. By employing the Frobenius norm inequality and the fact that the sign of $\mathcal{Q}_A(t)$ and $\mathcal{Q}_B(t)$ is the same for each $t \in \mathbb{T}$, we obtain by simple algebra

$$\|f(t,\mathcal{H}) - f(t,\mathcal{H}^*)\| = \sqrt{|\mathcal{Q}_A^2(t) + \mathcal{Q}_B^2(t)|}\, \|r(t,\mathcal{H}) - r(t,\mathcal{H}^*)\|. \tag{A.3}$$

Given that $r_\rho(t,\mathcal{H})$ $(\rho = A, B)$ is continuously differentiable in $\mathcal{H}$ for any $t \in \mathbb{T}$ and $\frac{\partial r_\rho(t,\mathcal{H})}{\partial \mathcal{H}}$ is continuous in t for all $\mathcal{H} \in \mathbf{H}$, it will hold for some particular positive constant $\mathcal{C}_1 < \infty$ $\sup_{(t,\mathcal{H})\in\mathbb{S}} \left|\frac{\partial r_\rho(t,\mathcal{H})}{\partial \mathcal{H}_\rho}\right| < \mathcal{C}_1$. This implies for $t \in \mathbb{T}$ $\sup_{\mathcal{H}\in\mathbf{H}} \left|\frac{\partial r_\rho(t,\mathcal{H})}{\partial \mathcal{H}_\rho}\right| < \mathcal{C}_1$ Hence, by the mean value theorem, we get for $(t,\mathcal{H}),(t,\mathcal{H}^*) \in \mathbb{S}$ and some positive constant $\mathcal{C}_2 < \infty$

$$\|r(t,\mathcal{H}) - r(t,\mathcal{H}^*)\| \leq \mathcal{C}_2 \|\mathcal{H} - \mathcal{H}^*\|. \tag{A.4}$$

Therefore, combining the inequalities (A.3) and (A.4), we get (A.2) with $l(t) = \mathcal{C}_3\sqrt{|\mathcal{Q}_A^2(t) + \mathcal{Q}_B^2(t)|}$ for a positive constant $\mathcal{C}_3 < \infty$. Note that the measurability and integrability of the functions $\mathcal{Q}_1(t)$ and $\mathcal{Q}_2(t)$ over $\mathbb{T}$, also imply the measurability and integrability of $l(t)$ over $\mathbb{T}$.

The above discussion shows that the conditions of theorem §10.XX of Walter (1998) are satisfied and thus the (A.1) is uniquely solved with respect to $\mathcal{H}_j(t)$ for $t \in (0,\infty)$ and $j = A, B$. ■