Nonparametric Identification of Dynamic Treatment Effects in Competing Risks Models

Bettina Drepper*
Georgios Effraimidis†

August 2012

Abstract

We introduce a dynamic treatment to the mixed proportional hazard competing risks model and allow for selection on unobservables. Our model may for example be used to simultaneously evaluate the effect of a benefit sanction on different competing exit risks such as 'finding work' vs. 'exiting the labor force'. We account for the endogeneity of the timing at which the individual enters into treatment by adding the hazard rate of the duration to treatment as an additional equation to the competing risks model. We present a new identification result of this model for single-spell duration data.

*University of Mannheim, Center for Empirical Economics and Econometrics. Address for correspondence: L7, 3-5, 68161 Mannheim, Germany. Email: bdrepper@mail.uni-mannheim.de

†University of Southern Denmark, Institut for Virksomhedsledelse og konomi. Address for correspondence: Campusvej 55, 5230 Odense M, Denmark. Email: georgiosnl@gmail.com

Acknowledgements: We thank Gerard J. van den Berg, Petyo Bonev, Andrew Chesher, Bo Honoré, Enno Mammen, Sebastian Klein, Geert Ridder and Johannes Schoch for helpful comments.

1 Introduction

The mixed proportional hazard competing risks model is a widely applied model in numerous empirical studies with duration data. In this note we introduce a dynamic treatment effect to this model and allow for selection on unobservables. The model allows to distinguish between different effects of a treatment on mutually exclusive exits. One example is the effect of a benefit sanction on the transition rate out of unemployment. The main idea of imposing sanctions is to put additional pressure on job searchers to increase their search effort. Empirical work is usually focused on measuring this effect on the transition rate into employment. However, imposing a sanction can also have an adverse effect. If unemployed individuals face too high pressure on the labor market, they are more likely to exit the labor force. In order to evaluate the full effect of benefit sanctions, the effects on both transition rates to 'finding work' and 'exit the labor force' have to be taken into account. However, the latter effect of the sanction is often ignored in empirical work due to the methodological difficulties and thus the statistical inference is generally incorrect. Another application of the proposed model can be found in mortality studies. In particular, it is often useful to distinguish between different effects of a medical treatment or health-related activity(e.g. abortion, birth) at some point in life on different causes of death. Identification of treatment effects in these cases usually has to be derived from single-spell mortality data since multiple-spell data are rarely available, and therefore the existing multivariation models cannot be employed for addressing such important questions.

Most of the studies, which make use of the competing risks model, consider only the exit of interest and right-censor the corresponding duration in case the first exit occurs due to one of the other risks. This approach relies on the rather strong assumption that conditional on observable characteristics, all competing exits are independent of each other. In labor economcs, for instance, researchers right-censor the unemployment spell if an individual has a different destination state (see Van den Berg, 2004) other than employment. If however the two competing exit risks of 'finding work' and 'exit for a different reason' are dependent of each other due to unobservable characteristics, this will generaly lead to wrong statistical results[1]. Our model deals with this issue by allowing for a flexible dependence structure between the different competing exit risks and the duration to treatment. In the field of treatment evaluation, the decision when to be assigned to a treatment (e.g. the assignment of a benefit sanction by a caseworker) is often driven by unobservable characteristics of the job seeker such as ability, motivation and preferences. These characteristics can simultaneously affect the different exit hazards leading to an endogeneity problem. We include the hazard rate of the duration to treatment as an additional equation in our model and account for the influence of unobservable characteristics in each equation while allowing for a flexible dependence structure. This way we control for selection on unobservables.

We present a new identification result for this model. Intuitively, the result relies on the assumption that the unobservable influences which simultaneously affect the competing exit hazards of interest and the transition hazard into treatment are time-invariant. Conditional on the realization of the unobservable influences, the exogenous variation in the timing of the treatment can be exploited to identify the causal effect of the treatment. This allows to distinguish between the two effects of selection on unobservable characteristics and the causal treatment effect. Following a main strand in the literature, our result makes use of a mixed proportional hazard type structure.

The identification strategy builds on results which have been developed by Abbring and Van den Berg (2003a, 2003b) on the identification of the timing-of-events model. Our model incorporates this original model as a special case with only one risk (no competing risks). Our contribution is to extend this model to the case of several competing exit durations which are allowed to be dependent due to unobservable characteristics and we allow for different effects of the treatment on different competing exit risks.

The competing risks model with a dynamic endogenous treatment is used by Arni et al. (2009) with multiple-spell data. They evaluate the effect of unemployment benefit sanctions

[1] Similar assumptions are used in mortality studies focusing on a specific death cause.

on the hazard to find a new job and on the hazard to leave the labor force. Identification of their model relies on the result on multiple-spell mixed proportional hazard competing risks models (see Abbring and Van den Berg, 2003b). This approach however imposes strong assumptions on the available dataset. Multiple spells have to be observed for most of the units (strata) in the sample and the treatment has to vary sufficiently across spells. Furthermore, in many applications like in models of unemployment spells unobservable characteristics such as skill level, motivation and preferences are often not constant across repeated unemployment spells. In this note we show that the mixed proportional hazard competing risks model with an endogenous dynamic treatment effect can be identified from single-spell data.

In the next Section we introduce the model under consideration. Section 3 presents the identification result and Section 4 provides some short discussion. We conclude in Section 5. Finally, the Appendix contains a Lemma which is employed to derive the identification result.

2 The mixed proportional hazard competing risks model with an endogenous treatment

In the following we will describe the mixed proportional hazard competing risks model for the case of two competing exit risks (see Abbring and Van den Berg, 2003b) and introduce an endogenous treatment to this model. The extension to more than two exit risks and/or multiple endogenous treatment effects is straightforward and will not be discussed here.

At time $t_0 = 0$, the unit of interest enters into some state (e.g. a worker enters into unemployment). In this state, for all $t \in \mathbb{R}_+$ the unit faces 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 durations until the unit exits to destination 1 or 2. Once the first exit takes place, the other duration is censored at that point. Consequently, the full joint distribution of (Y_1, Y_2) is not identifiable in case the duration variables Y_1 and Y_2 are not independent of each other. Instead, one observes (Y, I), with $Y = \min_{j\in\{1,2\}}(Y_j)$ and $I = \arg\min_{j\in\{1,2\}}(Y_j)$. The dependence between the two durations Y_1 and Y_2 is modeled via a bivariate mixed proportional hazard model. This models specifies that Y_1 and Y_2 are independent conditional on x, which denotes the realization of the vector of the observed characteristics X, and also unobservable influences V. Here, $V = (V_1\ V_2)'$ is a vector of unobserved non-negative random variables jointly drawn from the bivariate distribution function G with $\mathbb{P}(V_1 > 0, V_2 > 0) > 0$. The hazard rates of $Y_1|(X = x, V_1)$ and $Y_2|(X = x, V_2)$ are given by

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

Here, the functions $\lambda_j(t)$ and $\phi_j(x)$ capture the dependence of the exit hazard j on duration t and on observable characteristics x, respectively, for $j \in \{1, 2\}$.

In the following we introduce an additional source of dependence between Y_1 and Y_2 to the competing risk model (1). Let S denote the duration until the unit enters into treatment (e.g. time spent in unemployment before the case worker imposes a benefit sanction). Once the unit enters into treatment, the two subsequent exit hazards are affected. A crucial feature of our model is that we allow for different effects of the treatment on the two competing exit hazards. Furthermore, the timing of the treatment can be influenced by unobservable characteristics (e.g. skills, motivation, preferences) which also can affect the two exit hazards. We account for this endogeneity by including the hazard rate of S as an additional equation in model (1) and allow for flexible dependence between the unobservable influences across the three equations. This leads to

Model: *The hazard rates of* $Y_1|(S, X = x, V_1)$ *and* $Y_2|(S, X = x, V_2)$ *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* $S|(X = x, V_S)$ *is given by*

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

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

Here, the effects of the treatment on the subsequent hazard rates of the two competing exit risks is captured by the functions $\delta_1(t|s, x)$ and $\delta_2(t|s, x)$. The two effects can vary over time t, depend on the time of the realization s of the treatment S and on covariates x.

3 Main result

Before stating the main identification result we present the following technical conditions regarding the underlying model.

Assumption 1 *We assume that* $\phi_1 : \mathbb{X} \to (0, \infty)$, $\phi_2 : \mathbb{X} \to (0, \infty)$, $\phi_S(x) : \mathbb{X} \to (0, \infty)$ *are continuous functions with* $\phi_1(x^*) = \phi_2(x^*) = \phi_S(x^*) = 1$ *for some* $x^* \in \mathbb{X}$*. Further,* $(\phi_1(x), \phi_2(x), \phi_S(x); x \in \mathbb{X})$ *contains a nonempty open subset of* $\mathbb{R}^3_+$*.*

Assumption 2 *The functions* $\lambda_1 : \mathbb{R}_+ \to (0, \infty)$, $\lambda_2 : \mathbb{R}_+ \to (0, \infty)$, *and* $\lambda_S : \mathbb{R}_+ \to (0, \infty)$ *are measurable. The integrated baseline hazard rates* $\Lambda_1(t) := \int_0^t \lambda_1(\omega) d\omega$, $\Lambda_2(t) := \int_0^t \lambda_2(\omega) d\omega$, *and* $\Lambda_S(t) := \int_0^t \lambda_S(\omega) d\omega$ *exist and are finite for all* $t > 0$ *with* $\Lambda_1(t^*) = \Lambda_2(t^*) = \Lambda_S(t^*) = 1$ *for some particular* $t^* > 0$,.

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

$$\mathbb{E}(V_1) < \infty, \quad \mathbb{E}(V_2) < \infty, \quad \mathbb{E}(V_S) < \infty.$$

Assumption 4 *The functions* $\delta_1 : \{(t, \tau) \in \mathbb{R}^2_+ : t > \tau\} \times \mathbb{X} \to (0, \infty)$ *and* $\delta_2 : \{(t, \tau) \in \mathbb{R}^2_+ : t > \tau\} \times \mathbb{X} \to (0, \infty)$ *are measurable. Moreover, the quantities*

$$\begin{aligned}
\Upsilon_1(t|s, x) &:= \int_s^t \lambda_1(\omega)\delta_1(\omega|s, x) d\omega, \\
\Upsilon_2(t|s, x) &:= \int_s^t \lambda_2(\omega)\delta_2(\omega|s, x) d\omega, \\
\Delta_1(t|s, x) &:= \int_0^t \delta_1(\omega|s, x) d\omega, \\
\Delta_2(t|s, x) &:= \int_0^t \delta_2(\omega|s, x) d\omega
\end{aligned}$$

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. For a detailed discussion of this assumption for two durations see Abbring and Van den Berg (2003b). Assumption 2 concerns the functional form of the baseline hazard. We restrict the function space to integrable functions which is not very restrictive in applied work. Assumption 2 includes, for example, the case of piecewise constant, Weibull, 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 several parametrc families.

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). In a large dataset the following subsurvival functions are observable

$$Q_{Y_1,S}(y, s|x) := \mathbb{P}(Y_1 > y, Y_2 > Y_1, S > s, Y > S|x), \tag{2}$$

$$Q_{Y_2,S}(y, s|x) := \mathbb{P}(Y_2 > y, Y_1 > Y_2, S > s, Y > S|x), \tag{3}$$

$$Q_{Y_1}(y|x) := \mathbb{P}(Y_1 > y, Y_2 > Y_1, S > Y|x), \tag{4}$$

$$Q_{Y_2}(y|x) := \mathbb{P}(Y_2 > y, Y_1 > Y_2, S > Y|x) \tag{5}$$

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)-(5).

The identifiability of the model under study requires that the mapping from the model determinants to the subsurvival functions (2) - (5) is invertible. Put differently, from the information given by the observable data, the model determinants have to be uniquely determined. This leads to the following

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* $\{Q_{Y_1}, Q_{Y_2}, Q_{Y_1,S}, Q_{Y_2,S}\}$.

Proof. The joint distribution of the identified minimum of (Y_1, Y_2, S) and the identity of this smallest duration is fully characterized by $\{Q_{Y_1}, Q_{Y_2}, Q_S^0\}$ (Tsiatis, 1975). From this it follows that under the Assumptions 1-3 the functions $\Lambda_1, \phi_1, \Lambda_2, \phi_2, \Lambda_S, \phi_S$, and G are identified from $\{Q_{Y_1}, Q_{Y_2}, Q_S^0\}$ (see Abbring and Van den Berg, 2003b).

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

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

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

$$\begin{aligned}\frac{\partial^2 Q_{Y_j,S}(y, s|x)}{\partial s \partial y} &= \mathcal{L}_G^{j,3}\left(\phi_1(x)(\Lambda_1(y) + \Upsilon_1(y|s, x)), \phi_2(x)(\Lambda_2(y) + \Upsilon_2(y|s, x)), \phi_S(x)\Lambda_S(s)\right) \\ &\quad \times \lambda_S(s)\phi_S(x)\phi_j(x)\lambda_j(y)\delta_j(y|s, x), \end{aligned}\tag{7}$$

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

$$\lambda_j(y) = \left[\mathcal{L}_G^{(j)}(\phi_1(x)\Lambda_1(y), \phi_2(x)\Lambda_2(y), \phi_S(x)\Lambda_S(y))\phi_j(x)\right]^{-1} \frac{\partial Q_{Y_j}(y|x)}{\partial y}. \tag{8}$$

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

$$\begin{aligned}\lambda_j(y)\delta_j(y|s, x) = \Big[&\mathcal{L}_G^{(1j)}(\phi_1(x)(\Lambda_1(y) + \Upsilon_1(y|s, x)), \phi_2(x)\left(\Lambda_2(y) + \Upsilon_2(y|s, x)\right), \\ &\phi_s(x)\Lambda_S(s))\lambda_S(s)\phi_S(x)\phi_j(x)\Big]^{-1} \frac{\partial^2 Q_{Y_j,S}(y, s|x)}{\partial s \partial y}. \end{aligned}\tag{9}$$

For the remainder of the proof we fix s, x. Define $\mathcal{H}_j(y) := \Lambda_j(y)$ 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 and Λ_S on x and s, respectively.

Hence, for almost all $y \in (0, \infty)$ we have a system of two differential equations in the sense of Carath$\acute{e}$odory (1918), i.e.,

$$\frac{d}{dy}\mathcal{H}(y) = f\left(y, \mathcal{H}(y)\right),$$

$$\mathcal{H}(\tau) = \gamma_\tau \in \mathbb{R}^2_+, \text{ for some specific } \tau \in (0,s) \qquad \text{(initial conditions)}, \tag{10}$$

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

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

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 identifed (see the first paragraph) . Also, the rest quantities on the right hand side of the above equation are identified by the first step of the current proof (see the fist paragraph). Furthermoer, the quantity $\mathcal{Q}_j$ is observed from the data. By making use of Lemma 1, the above system of difefrential 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 imlpies that the quantities Υ_1 and Υ_2 are uniquely identified. By definition, the latter yields identification of Δ_1 and Δ_2. The proof is complete.

■

4 Discussion

The identification strategy is characterized by two main steps. More precisely, the first step exploits the fact that the three durations up to the first exit are characterized by a mixed proportional hazard competing risks model (see Abbring and Van den Berg, 2003b). This argument is also used in Abbring and Van den Berg (2003a).

On the other hand, the second step, which is concerned with the identification of Δ_1 and Δ_2, is different and actually more complicated than the respective step in the proof of Abbring and Van den Berg (2003a) who derive an identification result for the case with a single risk. Their identification result regarding the single treatment effect function directly follows by considering the quantity $\frac{\partial}{\partial s}\mathbb{P}(Y > y, S > s, Y > S|x)$, where Y in their model denotes the time to the occurence of the exit due to this single risk. In particular, the above partial derivative is expressed as a strictly decreasing function of a quantity which has one-to-one correspondence with the treatement effect function.

However, in the more general case with two risks that we consider here it is not possible to employ a similar argument. More precisely, our model implies that for all $y \in \mathbb{R}_+$, almost every $s \in \mathbb{R}_+$ such that $s < y$, and all $x \in \mathbb{X}$,

$$\begin{aligned} \frac{\partial}{\partial s} Q_S(y, s|x) &:= \frac{\partial}{\partial s}\mathbb{P}(Y_1 > y, Y_2 > y, S > s, Y_1 > S, Y_2 > S|x) \\ &= \mathcal{L}_G^3\left(\phi_1(x)(\Lambda_1(y) + \Upsilon_1(y|s,x)), \phi_2(x)(\Lambda_2(y) + \Upsilon_2(y|s,x)), \phi_S(x)\Lambda_S(s)\right) \\ &\quad \times \lambda_S(s)\phi_S(x), \end{aligned} \tag{11}$$

where $\mathcal{L}_G^3$ denotes the partial derivative with rerspect to the third argument of the trivariate Laplace Transform of the random vector $(V_1\ V_2\ V_S)'$. The quantity on the left hand side is nonparametrically observed from the data. But, in contrast to the setup of Abbring and Van den Berg (2003a) there are two unknowns on the right hand side and therefore we cannot adopt their identification strategy. To overcome this difficulty, we consider i) the partial derivatives $\frac{\partial Q_{Y_j}(y|x)}{\partial y}$, and ii) the mixed derivatives $\frac{\partial^2 Q_{Y_j,S}(y,s|x)}{\partial s \partial y}$ such that $s < y$. The latter derivatives yield a system of two differential equations of first order which can be solved with the help of Lemma 1. Finally, note that our model deals with two competing risks. Obviously, the proof can be easily modified such that to study any mixed proportional hazard competing risks model with more than two risks.

5 Conclusion

The contribution of this note is to show the identifiability of the mixed proportional hazard competing risks model with a dynamic treatment effect in the presence of selection on unobservables. A crucial feature of our model is that it allows to identify different effects of the treatment on different competing exit risks. In applied work, this model could be used, for instance, to evaluate the effects of labor market programs on different competing exit risks such as 'finding a part-time job' vs. 'finding full-time employment' vs. 'exiting the labor force. In this paper we show that the different effects of an endogenous dynamic treatment in the mixed proportional hazard competing risks model can be identified from single-spell data.

Appendix

The appendix presents a technical result which 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 depends also on the variable t. The next Lemma establishes existence and uniqueness of a solution for the (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 $\mathbb{T} = [\tau, \tau + a]$ 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 $(t, \mathcal{H}), (t, \mathcal{H}^*) \in \mathbb{S}$

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

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 $\mathcal{H} \in \mathbb{H}$, it will hold for some particular positive constant $\mathcal{C}_1 < \infty$

$$\sup_{(t,\mathcal{H})\in\mathbb{S}} \left| \frac{\partial r_\rho(t, \mathcal{H})}{\partial \mathcal{H}_\rho} \right| < \mathcal{C}_1. \tag{A.4}$$

This implies for $t \in \mathbb{T}$

$$\sup_{\mathcal{H}\in\mathbb{H}} \left| \frac{\partial r_\rho(t, \mathcal{H})}{\partial \mathcal{H}_\rho} \right| < \mathcal{C}_1. \tag{A.5}$$

Hence, by the mean value theorem, we get for $(t, \mathcal{H}), (t, \mathcal{H}^*) \in \mathbb{S}$ and some positive constant $\mathcal{C}_2 < \infty$

$$\|r(t, \mathcal{H}) - r(t, \mathcal{H}^*)\| \leq \mathcal{C}_2 \, \|\mathcal{H} - \mathcal{H}^*\| \, . \tag{A.6}$$

Therefore, combining the inequalities (A.3) and (A.6), 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 measurability and integrability of the functions $\mathcal{Q}_1(t)$ and $\mathcal{Q}_2(t)$ over $\mathbb{T}$, imply measurability and integrability of $l(t)$ over $\mathbb{T}$, too.

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

References

Abbring, J.H. and G.J. van den Berg (2003a), "The Nonparametric Identification of Treatment Effects in Duration Models", *Econometrica*, 71, 1491–1517.

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

Arni, P., R. Lalive and Jan C. van Ours. (2009), "How Effective Are Unemployment Benefit Sanctions? Looking Beyond Unemployment Exit", *Forthcoming in Journal of Applied Econometrics.*

Elbers, C. and G. Ridder (1982), "True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model", *The Review of Economic Studies*, 22, 211–241.

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

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, 211–241.

Walter, W. (1998), "Ordinary Differential Equations", *Springer-Verlag, New York.*