CentER

Discussion Paper

No. 2011-135

Regular Variation and the Identification of Generalized Accelerated Failure-Time Models

By

Jaap H. Abbring, Geert Ridder

December 2011

ISSN 0924-7815

TILBURG ◆ UNIVERSITY

Regular Variation and the Identification of Generalized Accelerated Failure-Time Models*

Jaap H. Abbring† Geert Ridder‡

December 11, 2011

Abstract

Ridder (1990) provides an identification result for the Generalized Accelerated Failure-Time (GAFT) model. We point out that Ridder's proof of this result is incomplete, and provide an amended proof with an additional necessary and sufficient condition that requires that a function varies regularly at 0 and ∞. We also give more readily interpretable sufficient conditions on the tails of the error distribution or the asymptotic behavior of the transformation of the dependent variable. The sufficient conditions are shown to encompass all previous results on the identification of the Mixed Proportional Hazards (MPH) model. Thus, this paper not only clarifies, but also unifies the literature on the non-parametric identification of the GAFT and MPH models.

*We thank Áureo de Paula, seminar participants at Tilburg University, and participants in the 10th World Congress of the Econometric Society in Shanghai and the 21st $(\mathrm{EC})^2$ Conference in Toulouse for comments.

†CentER, Department of Econometrics & OR, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: jaap@abbring.org. Web: jaap.abbring.org.

‡Department of Economics, University of Southern California, Los Angeles, USA. E-mail: ridder@usc.edu. Financial support for this research was generously provided through NSF SES 0819612 and 081963.

Keywords: duration analysis, identifiability, Mixed Proportional Hazards model, regular variation.
JEL codes: C14, C41.

1 Introduction

The Generalized Accelerated Failure-Time (GAFT) model introduced by Ridder (1990) specifies the cumulative distribution function $F(\cdot|x)$ of a positive random time T given a q-vector of covariates x as

(A-1). $F(t|x) = G\left[\phi(x)\Lambda(t)\right]; \qquad t \in (0,\infty),\ x \in \mathcal{X} \subseteq \mathbb{R}^q;$

where

(A-2). $\Lambda : (0,\infty) \to (0,\infty)$ can be written as $\Lambda(t) = \int_0^t \lambda(u)du$, $t \in (0,\infty)$, for some $\lambda : (0,\infty) \to (0,\infty)$ that is integrable on bounded intervals, and $\lim_{t\to\infty} \Lambda(t) = \infty$;

(A-3). $G : (0,\infty) \to (0,1)$ is a cumulative distribution function that is absolutely continuous with respect to the Lebesgue measure with density $g : (0,\infty) \to (0,\infty)$; and

(A-4). $\phi : \mathcal{X} \to (0,\infty)$ is such that $\phi(x_0) \neq \phi(x_1)$ for some $x_0, x_1 \in \mathcal{X}$.

If Λ is linear, the GAFT model reduces to the Accelerated Failure-Time model of Cox (1972) (pp. 200–01) with baseline distribution G. If $G(s) = 1 - \int_0^\infty \exp(-sv)\, \mathrm{d}H(v)$ for some cumulative distribution function H on $(0,\infty)$, then it is Lancaster (1979)'s Mixed Proportional Hazards (MPH) model, with baseline hazard λ and mixing distribution H.

Assumptions (A-1)–(A-4) are equivalent to Ridder (1990)'s Assumptions (A-1)–(A-4). Ridder required that $F(\cdot|x)$ has a positive density with respect to the Lebesgue measure and that Λ is non-decreasing and left-continuous. For expositional convenience, we have directly assumed that Λ is absolutely continuous with respect to the Lebesgue measure and increasing, and that G has a positive Lebesgue density. From Ridder's analysis, it is clear that this is without loss of generality relative to his assumptions.

Ridder (1990) studied the identifiability of the GAFT model. Section 2 provides a new proof of his main identification result (Theorem 1) for the GAFT model with a new necessary and sufficient condition, and shows that without it his proof is incomplete. Sufficient conditions are provided that encompass the various assumptions that are made in the literature to ensure the non-parametric identification of the MPH model. Section 3 gives a unification of the MPH identification literature based on these sufficient conditions. Section 4 concludes. Three appendices provide further proofs and results.

2 Identification Results

Suppose that the data provide us with $F(t|x)$ for all $t \in (0, \infty)$ and all $x \in \mathcal{X}$.[1] In the GAFT model, this cumulative distribution function is fully determined by the triplet (Λ, ϕ, G). Conversely, two GAFT triplets may imply the same cumulative distribution function F. In this case, we say that the triplets are observationally equivalent. Assumption (A-1) implies the following, more formal definition.

Definition 1. Two GAFT triplets (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ are *observationally equivalent* if $G[\phi(x)\Lambda(t)] = F(t|x) = \tilde{G}\left[\tilde{\phi}(x)\tilde{\Lambda}(t)\right]$ for all $t \in (0, \infty)$ and all $x \in \mathcal{X}$.

We will study the GAFT model's *identification* by characterizing the relation between observationally equivalent GAFT triplets. A GAFT triplet is identified if no other triplets are observationally equivalent. The GAFT model is identified if all GAFT triplets are. A feature of the GAFT model, such as the sign of the effect $\phi(x_1) - \phi(x_0)$ of changing the covariates from x_0 to x_1, is identified if it does not vary across observationally equivalent GAFT triplets.

2.1 Preliminary Results

First, note that the *signs* of the covariates' effects are identified. For future reference, we formalize this result in a lemma.

[1] Section 4 and Appendix C discuss an alternative setup with discretely observed durations.

Lemma 1. *Let* (Λ, ϕ, G) *and* $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ *be observationally equivalent GAFT triplets that satisfy (A-1)–(A-3). Then; for* $x_0, x_1 \in \mathcal{X}$; $\phi(x_0) < \phi(x_1)$ *if and only if* $\tilde{\phi}(x_0) < \tilde{\phi}(x_1)$.

Proof. Pick some $t \in (0, \infty)$. Observational equivalence implies that $G[\phi(x_0)\Lambda(t)] < G[\phi(x_1)\Lambda(t)]$ if and only if $\tilde{G}\left[\tilde{\phi}(x_0)\tilde{\Lambda}(t)\right] < \tilde{G}\left[\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]$. Moreover, because G and $\tilde{G}$ are strictly increasing by (A-2) and $\Lambda(t) > 0$ and $\tilde{\Lambda}(t) > 0$ by (A-3), $G[\phi(x_0)\Lambda(t)] < G[\phi(x_1)\Lambda(t)]$ if and only if $\phi(x_0) < \phi(x_1)$ and $\tilde{G}\left[\tilde{\phi}(x_0)\tilde{\Lambda}(t)\right] < \tilde{G}\left[\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]$ if and only if $\tilde{\phi}(x_0) < \tilde{\phi}(x_1)$. □

Next, we present an implication of observational equivalence that is key to both our main result (Section 2.2's Theorem 1) and Ridder (1990)'s Theorem 1. Denote the composition of two functions f and g with $f \circ g$; that is, for all s, $f \circ g(s) \equiv f(g(s))$.

Lemma 2. *Let* (Λ, ϕ, G) *and* $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ *be observationally equivalent GAFT triplets that satisfy (A-1)–(A-4). Define* $K \equiv \Lambda \circ \tilde{\Lambda}^{-1}$, *with derivative* $K' : (0, \infty) \to (0, \infty)$ *almost everywhere.*[2] *Then; for* $x_0, x_1 \in \mathcal{X}$ *and almost every* $s \in (0, \infty)$;

$$\frac{sK'(s)}{K(s)} = \frac{\tilde{\beta}^n s\, K'\left(\tilde{\beta}^n s\right)}{K\left(\tilde{\beta}^n s\right)}, \quad n \in \mathbb{Z}; \tag{1}$$

and

$$\frac{sK'(s)}{K(s)} = \lim_{n \to -\infty} \frac{\tilde{\beta}^n s\, K'\left(\tilde{\beta}^n s\right)}{K\left(\tilde{\beta}^n s\right)} = \lim_{n \to \infty} \frac{\tilde{\beta}^n s\, K'\left(\tilde{\beta}^n s\right)}{K\left(\tilde{\beta}^n s\right)}; \tag{2}$$

where $\tilde{\beta} \equiv \tilde{\phi}(x_0)/\tilde{\phi}(x_1)$ *and* $\mathbb{Z}$ *is the set of integers.*

Proof. See Appendix A. □

By Assumption (A-4), we can take x_0 and x_1 in Lemma 2 such that $\tilde{\beta} \neq 1$. Then, for given $s \in (0, \infty)$, $\tilde{\beta}^n s \to \infty$ in one of the limits in (2) and $\tilde{\beta}^n s \to 0$ in the other limit.

[2] Here and in the sequel, the exceptional sets have Lebesgue measure 0.

Intuitively, with conditions on the tail behavior of K' and K at 0 and ∞, $sK'(s)/K(s)$ can be determined almost everywhere from the limits in the right-hand side of (2). In turn, because by definition $\Lambda = K \circ \tilde{\Lambda}$, this characterizes the relation between the observationally equivalent GAFT triplets. Our main result gives such a characterization based on conditions on the tail behavior of K' and K.

2.2 Main Result

The statement of our main result requires Karamata's concepts of regular and slow variation (Feller, 1971, Section VIII.8).

Definition 2. A function $k : (0,\infty) \to (0,\infty)$ *varies regularly* with exponent $\tau \in \mathbb{R}$ at 0 (at ∞) if $k(\alpha s)/k(s) \to \alpha^\tau$ as $s \to 0$ ($s \to \infty$) for every $\alpha \in (0,\infty)$.

A function that varies regularly with exponent 0 is also said to be *slowly varying*. Any function that has a positive (and finite) limit varies slowly; but slowly varying functions may converge to 0 or diverge, such as $s \mapsto |\ln(s)|$ and $s \mapsto 1/|\ln(s)|$. If k varies regularly with exponent τ, then $k(s) = s^\tau k_0(s)$ for some slowly varying function k_0. The function k varies regularly at 0 with exponent τ if and only if $s \mapsto k(1/s)$ varies regularly at ∞ with exponent $-\tau$. By Feller (1971), Section VIII.8, a function k that varies regularly with exponent τ at ∞ (at 0) asymptotically satisfies $s^{\tau-\varepsilon} < k(s) < s^{\tau+\varepsilon}$, for any given $\varepsilon > 0$ ($\varepsilon < 0$).

With these definitions in place, we can state our main result. Here and in the sequel statements that involve functions hold on their domain; for example, $\Lambda = c\tilde{\Lambda}^\rho$ means that $\Lambda(t) = c\tilde{\Lambda}(t)^\rho$ for all $t \in (0,\infty)$.

Theorem 1. *Let* (Λ, ϕ, G) *and* $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ *be observationally equivalent GAFT triplets that satisfy (A-1)–(A-4). Define* $K \equiv \Lambda \circ \tilde{\Lambda}^{-1}$*, with derivative* $K' : (0, \infty) \to (0, \infty)$ *almost everywhere. Let* $\rho \in (0, \infty)$*. Then,*

(i). for some $c, d \in (0, \infty)$*,*

$$\Lambda = c\tilde{\Lambda}^{\rho},$$

$$\phi = d\tilde{\phi}^{\rho}, \text{ and}$$

$$\tilde{G}(s) = G(cds^{\rho}) \text{ for all } s \in (0, \infty)$$

if and only if

(ii). K' *varies regularly at* 0 *and* ∞ *with exponent* $\rho - 1$*.*

A sufficient condition for (ii) or, equivalently, (i) is that

(iii). $\rho = (\overline{\tau} + 1)/(\underline{\tau} + 1)$ *with at least one of the following true:*

(a) λ *and* $\tilde{\lambda}$ *vary regularly at* 0*, with exponents* $\overline{\tau} \in (-1, \infty)$ *and* $\underline{\tau} \in (-1, \infty)$*;*

(b) λ *and* $\tilde{\lambda}$ *vary regularly at* ∞*, with exponents* $\overline{\tau} \in (-1, \infty)$ *and* $\underline{\tau} \in (-1, \infty)$*;*

(c) g *and* $\tilde{g}$ *vary regularly at* 0*, with exponents* $\underline{\tau} \in (-1, \infty)$ *and* $\overline{\tau} \in (-1, \infty)$*; or*

(d) g *and* $\tilde{g}$ *vary regularly at* ∞*, with exponents* $\underline{\tau} \in (-\infty, -1)$ *and* $\overline{\tau} \in (-\infty, -1)$*.*

Proof. The proof proceeds in three steps. It first (A) provides an alternative characterization of (ii); subsequently (B) uses this to prove that (i) and (ii) are equivalent; and finally (C) shows that (iii) is sufficient for (ii).

A Alternative Characterization of (ii)

By Karamata's theorem (Feller, 1971, Section VIII.9, Theorem 1), (ii) is equivalent to

$$\lim_{s \to \infty} \frac{sK'(s)}{K(s)} = \lim_{s \to 0} \frac{sK'(s)}{K(s)} = \rho. \tag{3}$$

Specifically, by Theorem 1(b) in Feller (1971, Section VIII.9), the first limit in (3) holds if and only if K' varies regularly at ∞ with exponent $\rho - 1$. For the second limit, define $K^\star(s) \equiv K'(1/s)$. Using that $K(0) = 0$,

$$\lim_{s\to 0} \frac{sK'(s)}{K(s)} = \lim_{s\to 0} \frac{sK'(s)}{\int_0^s K'(u)\,du} = \lim_{s\to\infty} \frac{s^{-1}K'(1/s)}{\int_0^{1/s} K'(u)\,du} = \lim_{s\to\infty} \frac{s^{-1}K^\star(s)}{\int_s^\infty u^{-2}K^\star(u)\,du} = \rho. \quad (4)$$

By Theorem 1(a) in Feller (1971, Section VIII.9), (4) is equivalent to regular variation of $K^\star$ at ∞ with exponent $-\rho + 1$. Consequently, the second limit in (3) holds if and only if K' varies regularly at 0 with exponent $\rho - 1$.

B Equivalence of (i) and (ii)

First, suppose that (i) holds. Then; $K(s) = cs^\rho$, so that $K'(s) = c\rho s^{\rho-1}$; $s \in (0, \infty)$; and (ii) holds.

Next, we will prove that, conversely, (ii) implies (i), by showing that (3) implies (i). Recall that, by Lemma 2, observational equivalence implies (2). Let $x_0, x_1 \in \mathcal{X}$ be such that $\tilde{\beta} \equiv \tilde{\phi}(x_0)/\tilde{\phi}(x_1) \neq 1$ (Assumption (A-4) ensures that x_0 and x_1 exist). Then, for given $s \in (0, \infty)$, $\tilde{\beta}^n s \to \infty$ in one of the limits in (2) and $\tilde{\beta}^n s \to 0$ in the other limit. Now suppose that (3) holds. Then, the limits in (2), and therefore $sK'(s)/K(s)$, $s \in (0, \infty)$, equal ρ. In turn, this implies that $K(s) = cs^\rho$, $s \in (0, \infty)$, for some $c \in (0, \infty)$. Using the definition of K, we conclude that $\Lambda = c\tilde{\Lambda}^\rho$. Substituting this into (13) in the proof of Lemma 2 in Appendix A gives $\tilde{G}(s) = G(cds^\rho)$, $s \in (0, \infty)$, with $d \equiv \phi(x_0)/\tilde{\phi}(x_0)^\rho \in (0, \infty)$. Finally, observational equivalence (Definition 1) implies $\phi = d\tilde{\phi}^\rho$. Consequently, (i) holds. This establishes that (i) and (ii) are equivalent.

C Sufficiency of (iii)

The final step is to prove that (iii) is sufficient for (ii). We will do so by showing that each of (iii)a–(iii)d implies

$$\text{either} \quad \lim_{s\to\infty} \frac{sK'(s)}{K(s)} = \rho \quad \text{or} \quad \lim_{s\to 0} \frac{sK'(s)}{K(s)} = \rho, \tag{5}$$

so that the corresponding limit in the right-hand side of (2) equals ρ. Then, Lemma 2 implies that $sK'(s)/K(s) = \rho$ for all $s \in (0,\infty)$, so that (3) and, equivalently, (ii) hold.

We first consider regular variation of $\lambda, \tilde{\lambda}$. Because $K = \Lambda \circ \tilde{\Lambda}^{-1}$, by Assumption (A-2), we have $K'(s) = \lambda\left[\tilde{\Lambda}^{-1}(s)\right]/\tilde{\lambda}\left[\tilde{\Lambda}^{-1}(s)\right]$, so that[3]

$$\frac{sK'(s)}{K(s)} = \frac{\tilde{\Lambda}^{-1}(s)\,\lambda\left[\tilde{\Lambda}^{-1}(s)\right]}{\Lambda\left[\tilde{\Lambda}^{-1}(s)\right]} \cdot \frac{\tilde{\Lambda}\left[\tilde{\Lambda}^{-1}(s)\right]}{\tilde{\Lambda}^{-1}(s)\,\tilde{\lambda}\left[\tilde{\Lambda}^{-1}(s)\right]}, \qquad s \in (0,\infty). \tag{6}$$

Suppose that (iii)a holds: λ and $\tilde{\lambda}$ vary regularly at 0 with exponents $\overline{\tau} \in (-1,\infty)$ and $\underline{\tau} \in (-1,\infty)$, respectively. Then, by Theorem 1(a) in Feller (1971, Section VIII.9) and an argument like that for K' around (4), $\lim_{t\to 0} t\lambda(t)/\Lambda(t) = \overline{\tau}+1$ and $\lim_{t\to 0} t\tilde{\lambda}(t)/\tilde{\Lambda}(t) = \underline{\tau}+1$. Because, by (A-2), $\tilde{\Lambda}^{-1} : (0,\infty) \to (0,\infty)$ and $\lim_{s\to 0}\tilde{\Lambda}^{-1}(s) = 0$, this implies that the first factor in the right-hand side of (6) converges to $\overline{\tau}+1$ and the second factor to $1/(\underline{\tau}+1)$, as $s \to 0$. Consequently, $\lim_{s\to 0} sK'(s)/K(s) = \rho$, and (5) holds, with $\rho = (\overline{\tau}+1)/(\underline{\tau}+1)$.

For the case that (iii)b holds— λ and $\tilde{\lambda}$ vary regularly at ∞ with exponents $\overline{\tau} \in (-1,\infty)$ and $\underline{\tau} \in (-1,\infty)$, respectively— Theorem 1(b) in Feller (1971, Section VIII.9) implies that $\lim_{t\to\infty} t\lambda(t)/\Lambda(t) = \overline{\tau}+1$ and $\lim_{t\to\infty} t\tilde{\lambda}(t)/\tilde{\Lambda}(t) = \underline{\tau}+1$. Now using that $\lim_{s\to\infty}\tilde{\Lambda}^{-1}(s) = \infty$, we conclude from (6) that $\lim_{s\to\infty} sK'(s)/K(s) = \rho$, and (5) holds, with $\rho = (\overline{\tau}+1)/(\underline{\tau}+1)$.

[3]Here and in the sequel, we omit the qualifier "almost every" from "almost every $s \in (0,\infty)$".

Next, we consider regular variation of $g, \tilde{g}$. With (A-3), (13) implies that

$$K(s) = \frac{1}{\phi(x_0)} G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right] \quad \text{and} \quad K'(s) = \frac{\tilde{\phi}(x_0)\tilde{g}\left[\tilde{\phi}(x_0)s\right]}{\phi(x_0)g\left\{G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right]\right\}}$$

so that

$$\frac{sK'(s)}{K(s)} = \frac{\tilde{\phi}(x_0)s\,\tilde{g}\left[\tilde{\phi}(x_0)s\right]}{\tilde{G}\left[\tilde{\phi}(x_0)s\right]} \cdot \frac{G\left\{G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right]\right\}}{G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right] g\left\{G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right]\right\}} \tag{7}$$

and

$$\frac{sK'(s)}{K(s)} = \frac{\tilde{\phi}(x_0)s\,\tilde{g}\left[\tilde{\phi}(x_0)s\right]}{1-\tilde{G}\left[\tilde{\phi}(x_0)s\right]} \cdot \frac{1-G\left\{G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right]\right\}}{G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right] g\left\{G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right]\right\}}, \qquad s \in (0,\infty). \tag{8}$$

Suppose that (iii)c holds: g and $\tilde{g}$ vary regularly at 0 with exponents $\underline{\tau} \in (-1,\infty)$ and $\overline{\tau} \in (-1,\infty)$, respectively. Then, by Theorem 1(a) in Feller (1971, Section VIII.9) and an argument like that for K' around (4), $\lim_{s\to 0} s\tilde{g}(s)/\tilde{G}(s) = \overline{\tau}+1$ and $\lim_{s\to 0} sg(s)/G(s) = \underline{\tau}+1$. Because, by (A-2) and (A-4); $\tilde{\phi}(x_0) \in (0,\infty)$, $G^{-1}\left[\tilde{G}(\tilde{\phi}(x_0)s)\right] \in (0,\infty)$ for $s \in (0,\infty)$, and $\lim_{s\to 0} G^{-1}\left[\tilde{G}(\tilde{\phi}(x_0)s)\right] = 0$; this implies that the first factor in the right-hand side of (7) converges to $\overline{\tau}+1$ and the second factor to $1/(\underline{\tau}+1)$, as $s \to 0$. Consequently, $\lim_{s\to 0} sK'(s)/K(s) = \rho$, and (5) holds, with $\rho = (\overline{\tau}+1)/(\underline{\tau}+1)$.

For the case that (iii)d holds— g and $\tilde{g}$ vary regularly at ∞ with exponents $\underline{\tau} \in (-\infty,-1)$ and $\overline{\tau} \in (-\infty,-1)$, respectively— Theorem 1(a) in Feller (1971, Section VIII.9) implies that $\lim_{s\to\infty} s\tilde{g}(s)/\left[1-\tilde{G}(s)\right] = \overline{\tau}+1$ and $\lim_{s\to\infty} sg(s)/\left[1-G(s)\right] = \underline{\tau}+1$. Now using that $\lim_{s\to\infty} G^{-1}\left[\tilde{G}(\tilde{\phi}(x_0)s)\right] = \infty$, we conclude from (8) that $\lim_{s\to\infty} sK'(s)/K(s) = \rho$, and (5) holds, with $\rho = (\overline{\tau}+1)/(\underline{\tau}+1)$. □

Theorem 1 is an amended version of Ridder (1990)'s Theorem 1. It shows that GAFT triplets are identified, up to obvious normalizations, if and only if an additional condition, (ii), is satisfied. Because $K'(s) = \lambda\left[\tilde{\Lambda}^{-1}(s)\right]/\tilde{\lambda}\left[\tilde{\Lambda}^{-1}(s)\right]$, this necessary and sufficient condition requires that the ratio $\lambda(t)/\tilde{\lambda}(t)$ behaves like $\left[\tilde{\Lambda}(t)\right]^{\rho-1}$ for t near 0 and ∞. This is made explicit in the more readily interpretable conditions (iii)a–(iii)d. Each one of these conditions is sufficient for (i) to hold, but none of them is necessary. For example, suppose that $(\Lambda, \phi, G) = (\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$. Then, trivially, $K(s) = s$ for all $s \in (0, \infty)$ and K' varies slowly at 0 and ∞, even if λ and $\tilde{\lambda}$ satisfy neither (iii)a nor (iii)b and g and $\tilde{g}$ satisfy neither (iii)c nor (iii)d.

Within the context of the literature on the identification of the (single spell) MPH model, the new sufficient conditions are mild, because they are implied by each of the identifying assumptions made in that literature (see Section 3). The restrictions on the ranges of $\overline{\tau}$ and $\underline{\tau}$ in Theorem 1(iii)a–(iii)d are implied by Assumptions (A-2) and (A-3) and constitute no additional restrictions, except for exclusion of the boundary cases that $\overline{\tau}$ and/or $\underline{\tau}$ equal -1 (which we will further discuss below). In particular, by the Lemma in Section VIII.9 of Feller (1971), the requirement that $\lim_{t\to\infty} \Lambda(t) = \infty$ implies that λ cannot vary regularly with exponent $\overline{\tau} < -1$ at ∞, and $\lim_{s\to\infty} G(s) = 1$ implies that g cannot vary regularly with exponent $\underline{\tau} > -1$ at ∞. By that same Lemma, and an argument like that for K' around (4), existence of $\Lambda(t) = \int_0^t \lambda(u)du < \infty$ and $G(s) = \int_0^s g(u)du < \infty$ for finite t and s implies that λ and g cannot vary regularly with respective exponents $\overline{\tau} < -1$ and $\underline{\tau} < -1$ at 0. Obviously, the same restrictions hold for $\tilde{\lambda}$ and $\tilde{g}$.

Ridder (1990)'s Theorem 1 does not impose conditions on the asymptotic behavior of the transformation or the baseline distribution, but his proof implicitly relies on such a condition. In particular, condition (17) in that proof states that $0 < K(s) < \infty$ and $0 < K'(s) < \infty$ for all $s \in (0, \infty)$. His limit result (21) essentially claims that, because of these bounds on K and K', the right-hand side of our equation (2) converges to a constant $\rho \in (0, \infty)$ as $n \to \infty$. Our Theorem 1 shows that this requires that K' varies regularly

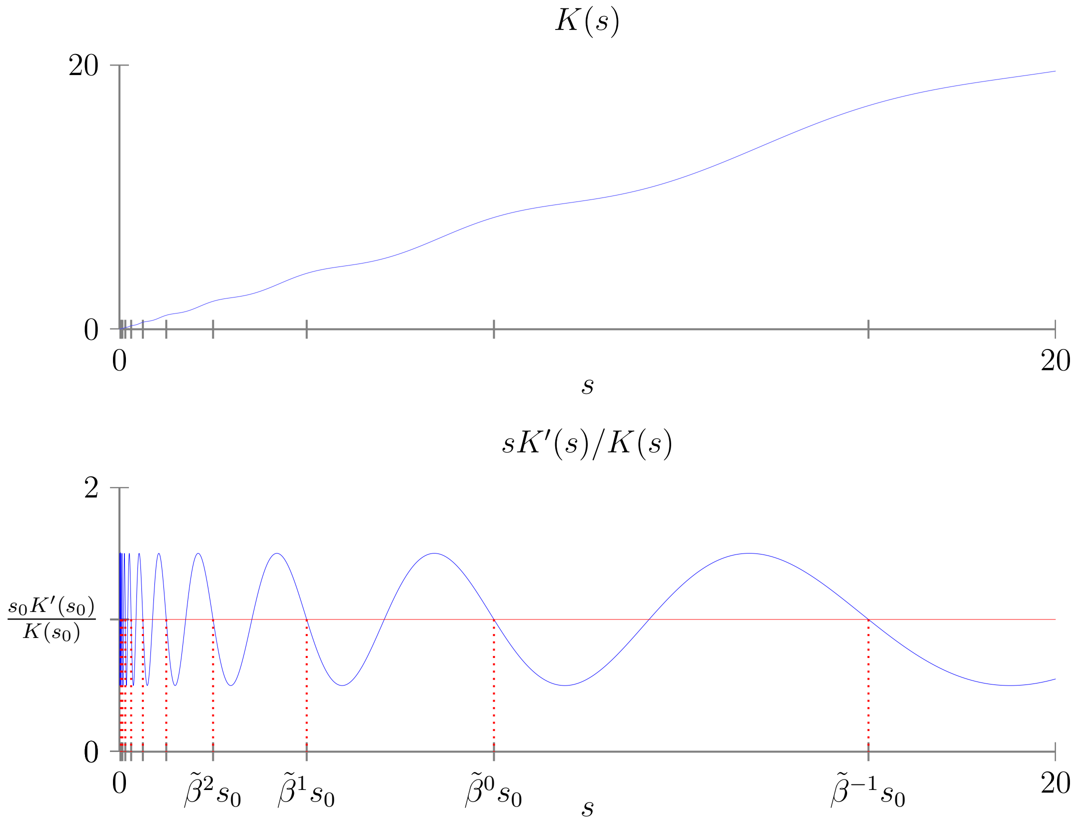

Figure 1: A function $K \equiv \Lambda \circ \tilde{\Lambda}^{-1}$ such that, for given $\tilde{\beta} \in (0,1)$ and all $s_0 \in (0,\infty)$, $s \mapsto sK'(s)/K(s)$ is constant on $\{\tilde{\beta}^n s_0; n \in \mathbb{Z}\}$ but not on $(0,\infty)$ (plotted for $\tilde{\beta} = \frac{1}{2}$, $s_0 = 8$)

with exponent $\rho - 1 \in (-1,\infty)$, which is not guaranteed by Ridder's conditions.

The top panel of Figure 1 plots a counterexample to Ridder (1990)'s claims,

$$K(s) = s^2 \exp\left\{-\frac{\ln(\tilde{\beta})}{2\pi} \cos\left[\frac{2\pi}{\ln(\tilde{\beta})} \ln(s)\right]\right\}, \tag{9}$$

for the case that $\tilde{\beta} = \frac{1}{2}$. In this example, Ridder's bounds on K and K' hold. Moreover, Lemma 2's implication of observational equivalence (1) is satisfied:

$$\begin{aligned} \frac{sK'(s)}{K(s)} &= 2 + \sin\left[\frac{2\pi}{\ln(\tilde{\beta})} \ln(s)\right] \\ &= 2 + \sin\left[\frac{2\pi}{\ln(\tilde{\beta})} \ln(s) + 2\pi n\right] = \frac{\tilde{\beta}^n s K'\left(\tilde{\beta}^n s\right)}{K\left(\tilde{\beta}^n s\right)}, \quad n \in \mathbb{Z}. \end{aligned} \tag{10}$$

However, $sK'(s)/K(s)$ is not a constant and the right-hand side of (10) does not converge to a constant as $n \to \infty$ or $n \to -\infty$. Figure 1's bottom panel illustrates this by plotting $sK'(s)/K(s)$ and, for given $s_0 = 8$, its value $s_0K'(s_0)/K(s_0) = 1$ on the set $\{\tilde{\beta}^n s_0; n \in \mathbb{Z}\}$. It follows that K is not uniquely determined (up to two constants) by (2). Moreover, $\lim_{s\to 0} sK'(s)/K(s)$ and $\lim_{s\to\infty} sK'(s)/K(s)$ do not exist so that, by Karamata's theorem, K' does not vary regularly (see Part A of Theorem 1's proof). We conclude that observationally equivalent triplets are not necessarily related as in Theorem 1(i), so that the GAFT model is not non-parametrically identified, under Ridder (1990)'s conditions alone. Also, the corresponding GAFT triplets do not satisfy any of the sufficient conditions in Theorem 1(iii)a–(iii)d.

This counterexample can also be applied to Theorem 1(iii)'s excluded boundary cases. Consider, for example, Theorem 1(iii)b. Let K be specified as in (9) and suppose that $\tilde{\Lambda}(t) = \ln(t+1)$. Then, both λ and $\tilde{\lambda}$ vary regularly at ∞ with exponent -1, but $\Lambda = K \circ \tilde{\Lambda}$ and Λ are not related as in Theorem 1(i). Clearly, regular variation of λ and $\tilde{\lambda}$ at ∞ with exponent -1 is not sufficient for Theorem 1(i) to hold.[4] Appendix B provides further discussion of the boundary cases.

In the special case of the MPH model, $G(s) = 1 - \int_0^\infty \exp(-sv)\, dH(v)$ for some cumulative distribution function H on $(0, \infty)$. Then, because

$$K(s) = \frac{1}{\phi(x_0)} \cdot G^{-1}\left[\tilde{G}\left(\tilde{\phi}(x_0)s\right)\right],$$

$s \mapsto sK'(s)/K(s)$ is real analytic. One may wonder whether this additional structure on K is sufficient to prove Theorem 1 without reference to conditions on the model's tails. This is not the case: The counterexample in (9) is real analytic and thus continues to be valid in the analytic case. It is nevertheless instructive to develop the argument based on real analyticity as far as it goes. For a given $s_0 \in (0, \infty)$, Lemma 2 tells

[4] We have not been able to find such counterexamples for all boundary cases. Thus, it is possible that some of the other boundary cases imply Theorem 1(i).

us that $s \mapsto sK'(s)/K(s)$ equals the constant $\rho \equiv s_0 K'(s_0)/K(s_0) \in (0,\infty)$ on a set $\{\tilde{\beta}^n s_0; n \in \mathbb{Z}\}$ that is dense near 0. From this, one may want to conclude from an analytic extension result, like Krantz and Parks (2002), Corollary 1.2.7, that $sK'(s)/K(s) = \rho$ everywhere. However, such an analytic extension result does not readily apply, because the accumulation point 0 is on the boundary of the domain of $s \mapsto sK'(s)/K(s)$. This suggests that we need further conditions on the tail of $s \mapsto sK'(s)/K(s)$. Theorem 1(ii) and (iii)a–(iii)d are such conditions.

3 Application to the Mixed Proportional Hazards Model

Ridder (1990) provides an extensive discussion of his Theorem 1's implications for, in particular, the empirical content of the MPH model; that is, a GAFT model with $G(s) = 1 - \int_0^\infty \exp(-sv)\, dH(v)$ for some cumulative distribution function H on $(0,\infty)$. This discussion remains valid for our amended version of his theorem, and can be extended to more recent results in the literature on the identification of the MPH model from single spell data, because our new sufficient conditions on the tails of λ and g, Theorem 1(iii), nest all related assumptions made in this literature. We list these assumptions and their connection to our Theorem 1.

- Elbers and Ridder (1982) and Kortram et al. (1995) achieve point identification of the MPH model (up to the obvious scale normalizations) under the *finite-mean assumption* that

$$\lim_{s \to 0} g(s) = \int_0^\infty v dH(v) < \infty.$$

Because H has no support outside $(0,\infty)$, we also have that $\lim_{s\to 0} g(s) > 0$.[5]

[5]This continues to be true if, more generally, H is a cumulative distribution function on $[0,\infty)$, because (A-3) precludes the case that H is concentrated at 0.

Consequently, Elbers and Ridder's finite-mean assumption implies that g varies slowly at 0. This is equivalent to setting $\underline{\tau} = \overline{\tau} = 0$ in Theorem 1(iii)c.

- Heckman and Singer (1984) instead make assumptions that guarantee that $v \mapsto M(v) \equiv \int_0^v s dH(s)$ varies regularly at ∞ with an a priori given exponent $-\tau \in (0,1)$. The Laplace transform of M is $g(s) = \int_0^\infty v \exp(-sv) dH(v)$. Therefore, by an Abelian-Tauberian theorem (Feller, 1971, Section XIII.5, Theorem 2), their assumption is equivalent to the assumption that g varies regularly at 0 with a priori given exponent $\tau \in (-1,0)$. This corresponds to setting $\underline{\tau} = \overline{\tau} = \tau$, for given $\tau \in (-1,0)$, in Theorem 1(iii)c. Note that, in contrast to Elbers and Ridder (1982)'s finite-mean assumption, Heckman and Singer (1984)'s assumption implies that $\lim_{s \to 0} g(s) = \lim_{v \to \infty} M(v) = \infty$.

- Ridder and Woutersen (2003) take a different angle, and assume that $0 < \lim_{s \to 0} \lambda(s) < \infty$. This implies that λ varies slowly at 0. In turn, this corresponds to setting $\overline{\tau} = \underline{\tau} = 0$ in Theorem 1(iii)a.

Note that, in all three cases, point identification is obtained by not only assuming that λ and g vary regularly in one of their tails (as in Theorem 1(iii)), but also by a priori fixing the corresponding exponent of regular variation. In terms of Theorem 1, in each case, $\underline{\tau} = \overline{\tau}$ is set to a known constant, so that $\rho = 1$.

4 Conclusion and Extensions

Our main result corrects a flaw in proof of the non-parametric identification of the GAFT model in Ridder (1990). We obtain a new necessary and sufficient condition under which the GAFT model is non-parametrically identified up to obvious normalizations. The GAFT model is not identified if we can find observationally equivalent GAFT triplets that are not related by these normalizations. We also provide novel sufficient conditions for non-parametric identification in terms of the GAFT model's primitives. Section 3 uses this to clarify and unify the previous results on the non-parametric identification of the MPH model for single spell data, which is a special case of the GAFT model.

Our results have relevance beyond the MPH model for single spell data. They can easily be extended to a competing-risks setting, and used to interpret and extend the identification results of Heckman and Honoré (1989) and Abbring and Van den Berg (2003). They can also be applied to Honoré and de Paula (2010)'s recent analysis of an optimal stopping game. All three papers study multivariate extensions of the MPH and GAFT models, and rely on multivariate versions of restrictions on the behavior of g near 0, as in Theorem 1(iii)c. Finally, Abbring (2011) shows that many of the identification arguments for MPH and GAFT models, including those in this paper, can be adapted to a class of mixed hitting-time models that specify durations as the first time a Lévy process hits a threshold that may depend on both observed covariates and an unobserved heterogeneity factor.

The GAFT model for duration analysis is closely related to transformation models for the analysis of general continuous variables. Horowitz (1996) and Chiappori and Komunjer (2009) analyzed the semiparametric and nonparametric identification of transformation models. They do not rely on tail conditions like our sufficient conditions, but instead assume continuous variation in the covariates. For example, suppose that $\mathcal{X}$ is an interval in $\mathbb{R}$. Let (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ be observationally equivalent GAFT triplets that satisfy (A-1)–(A-3). Moreover, instead of (A-4), assume that ϕ and $\tilde{\phi}$ are continuously differentiable, with $\phi'(x_0) \neq 0$ and $\tilde{\phi}'(x_0) \neq 0$ for some $x_0 \in \mathcal{X}$. Then, in the spirit of Horowitz (1996),

$$\frac{\phi'(x)}{\phi(x)} \cdot \frac{\Lambda(t)}{\lambda(t)} = \frac{\partial F(t|x)/\partial x}{\partial F(t|x)/\partial t} = \frac{\tilde{\phi}'(x)}{\tilde{\phi}(x)} \cdot \frac{\tilde{\Lambda}(t)}{\tilde{\lambda}(t)}; \quad t \in (0, \infty),\ x \in \mathcal{X};$$

implies that

$$\frac{\lambda}{\Lambda} = \rho\frac{\tilde{\lambda}}{\tilde{\Lambda}} \quad \text{and} \quad \frac{\phi'}{\phi} = \rho\frac{\tilde{\phi}'}{\tilde{\phi}}, \quad \text{with } \rho \equiv \frac{\phi'(x_0)}{\phi(x_0)} \cdot \frac{\tilde{\phi}(x_0)}{\tilde{\phi}'(x_0)} \in (0,\infty).$$

In turn, this implies the characterization of the observationally equivalent GAFT triplets in Theorem 1(i). In this argument, direct continuous variation with the covariates substitutes for Theorem 1's evaluation of $F(\cdot|x_0)$ and $F(\cdot|x_1)$ near the common limits of their supports.

The present paper's analysis requires data for only two covariate values. It also applies in that case that x is discrete, where the results for continuously varying x cannot be used. The binary valued case shows that the GAFT model with continuously varying x is (heavily) overidentified. If we fix x_0 and let x_1 take values in $\mathcal{X}$ then we can identify Λ and G for each x_1 such that $\phi(x_0) \neq \phi(x_1)$. The binary case only requires that Λ and G are the same in the subpopulations with covariate values x_0 and x_1; different x_1 could identify different Λ and G.

The binary covariate case can also be applied to the common situation in which durations are discretely observed, but continuous regressor variation is available. Ridder (1990) shows that, without further assumptions, identification breaks down when durations are only observed in intervals; and that identification can be restored by exploiting continuous and parametric variation with the covariates. In the context of the MPH model, Brinch (2011) noted that Elbers and Ridder (1982)'s results for continuously observed durations and discrete covariate values can be applied to this problem by simply exchanging the roles of time and the covariates.[6] This idea extends to our analysis.

For example, suppose that $F(t|x)$ is known for only two values t_0 and t_1 of t such that $0 < t_0 < t_1 < \infty$. Consider two GAFT triplets (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ that are

[6] Ridder (1990) essentially focuses on a GAFT model in which $F(t|x)$ is only observed for *one* value of t in $(0,\infty)$ (and where it is in addition known that $\lim_{t\to 0} F(t|x) = 0$ and $\lim_{t\to\infty} F(t|x) = 1$). For this case, he provides both a nonidentification result and a semiparametric identification result that relies on parametric structure on ϕ. For the MPH special case, Brinch (2011) points out that such parametric structure is not needed when $F(t|x)$ is known for at least *two* values of t in $(0,\infty)$.

observationally equivalent with such data: $G[\phi(x)\Lambda(t)] = F(t|x) = \tilde{G}\left[\tilde{\phi}(x)\tilde{\Lambda}(t)\right]$ for all $t \in \{t_0, t_1\}$ and all $x \in \mathcal{X}$. Suppose that both these triplets satisfy (A-1)–(A-3) (note that this implies that $0 < \Lambda(t_0) < \Lambda(t_1) < \infty$ and $0 < \tilde{\Lambda}(t_0) < \tilde{\Lambda}(t_1) < \infty$). For the sake of simplicity, let $\mathcal{X} = (0, \infty)$, and assume that ϕ and $\tilde{\phi}$ satisfy the conditions on Λ in (A-2). In particular, this requires that $\phi(x)$ and $\tilde{\phi}(x)$ attain all values in $(0, \infty)$ if x varies over $\mathcal{X}$. It also requires that ϕ and $\tilde{\phi}$ are strictly increasing functions of a scalar covariate, but this can easily be relaxed, as shown in Appendix C. Then, Theorem 1 applies directly; with $\Lambda(t_0)$, $\Lambda(t_1)$, $\tilde{\Lambda}(t_0)$, and $\tilde{\Lambda}(t_1)$ taking the roles of $\phi(x_0)$, $\phi(x_1)$, $\tilde{\phi}(x_0)$, and $\tilde{\phi}(x_1)$; and ϕ and $\tilde{\phi}$ substituting for Λ and $\tilde{\Lambda}$. Therefore, up to obvious normalizations we can non-parametrically identify the regression function and G and the transformation is identified at the interval boundaries.

Appendices

A Proof of Lemma 2

By Definition 1, observational equivalence of (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ implies that

$$\begin{aligned} \tilde{G}\left[\tilde{\phi}(x_0)\tilde{\Lambda}(t)\right] &= G\left[\phi(x_0)\Lambda(t)\right] \text{ and} && (11) \\ \tilde{G}\left[\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right] &= G\left[\phi(x_1)\Lambda(t)\right] && (12) \end{aligned}$$

for all $t \in (0, \infty)$. Because $\tilde{\Lambda} : (0, \infty) \to (0, \infty)$ is bijective by (A-2), and $\tilde{\phi}(x_0) > 0$ and $\tilde{\phi}(x_1) > 0$ by (A-4), changing variables to $s = \tilde{\phi}(x_0)\tilde{\Lambda}(t)$ in (11) and to $s = \tilde{\phi}(x_1)\tilde{\Lambda}(t)$ in (12) gives

$$\begin{aligned} \tilde{G}(s) &= G\left\{\phi(x_0)\Lambda\left[\tilde{\Lambda}^{-1}\left(\frac{s}{\tilde{\phi}(x_0)}\right)\right]\right\} \text{ and} && (13) \\ \tilde{G}(s) &= G\left\{\phi(x_1)\Lambda\left[\tilde{\Lambda}^{-1}\left(\frac{s}{\tilde{\phi}(x_1)}\right)\right]\right\} \end{aligned}$$

for $s \in (0, \infty)$. In turn, because G is strictly increasing by (A-3), this implies that

$$\phi(x_0)\Lambda\left[\tilde{\Lambda}^{-1}\left(\frac{s}{\tilde{\phi}(x_0)}\right)\right] = \phi(x_1)\Lambda\left[\tilde{\Lambda}^{-1}\left(\frac{s}{\tilde{\phi}(x_1)}\right)\right], \quad s \in (0, \infty).$$

Changing variables to $t = \tilde{\Lambda}^{-1}(s/\tilde{\phi}(x_0))$ and rearranging, using that $\tilde{\Lambda}^{-1}$ is bijective and $\tilde{\phi}(x_0) > 0$, and that Λ is invertible by (A-2), gives

$$\Lambda^{-1}\left(\frac{\phi(x_0)}{\phi(x_1)}\Lambda(t)\right) = \tilde{\Lambda}^{-1}\left(\frac{\tilde{\phi}(x_0)}{\tilde{\phi}(x_1)}\tilde{\Lambda}(t)\right), \quad t \in (0, \infty). \tag{14}$$

With $\beta \equiv \phi(x_0)/\phi(x_1)$ and $\tilde{\beta} \equiv \tilde{\phi}(x_0)/\tilde{\phi}(x_1)$, we can write (14) more succinctly as

$$\Lambda^{-1} \circ \beta \cdot \Lambda = \tilde{\Lambda}^{-1} \circ \tilde{\beta} \cdot \tilde{\Lambda} \tag{15}$$

or, by inverting the left- and right-hand sides,

$$\Lambda^{-1} \circ \beta^{-1} \cdot \Lambda = \tilde{\Lambda}^{-1} \circ \tilde{\beta}^{-1} \cdot \tilde{\Lambda}. \tag{16}$$

By composing the left- and right-hand sides of (15) and (16) n times with themselves, we obtain the equivalent relation

$$\Lambda^{-1} \circ \beta^{n} \cdot \Lambda = \tilde{\Lambda}^{-1} \circ \tilde{\beta}^{n} \cdot \tilde{\Lambda}, \qquad n \in \mathbb{Z} \tag{17}$$

(note that this equation is trivially satisfied if $n = 0$). Using that $K = \Lambda \circ \tilde{\Lambda}^{-1}$, and a change of variables to $s = \tilde{\Lambda}(t)$, (17) implies

$$\beta^n K(s) = K(\tilde{\beta}^n s), \qquad s \in (0, \infty),\ n \in \mathbb{Z}. \tag{18}$$

Assumption (A-2) implies that $K(s) > 0$ for $s \in (0, \infty)$. So, we can take the derivative of the logarithm of (18) for $s \in (0, \infty)$. Multiplying the result of this by s gives (1).

Finally, because (1) holds for all integer n, it should hold in the limit as $n \to -\infty$ or $n \to \infty$. This gives (2). □

B Boundary Cases

The boundary cases that $\overline{\tau} = -1$ and $\underline{\tau} = -1$ are excluded from Theorem 1(iii)a–(iii)d because, in them, $\rho = (\overline{\tau} + 1)/(\underline{\tau} + 1)$ is not defined and Part C of Theorem 1's proof breaks down. Intuitively, in these cases, the tail behavior of $F(\cdot|x)$ provides comparatively little information about the model primitives.

To gain some intuition for this, let (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ be observationally equivalent GAFT triplets that satisfy (A-1)–(A-4). For any $x_0, x_1 \in \mathcal{X}$; observational equivalence implies that

$$\frac{1-\tilde{G}\left[\frac{\tilde{\phi}(x_0)}{\tilde{\phi}(x_1)}\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]}{1-\tilde{G}\left[\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]} = \frac{1-F(t|x_0)}{1-F(t|x_1)} = \frac{1-G\left[\frac{\phi(x_0)}{\phi(x_1)}\phi(x_1)\Lambda(t)\right]}{1-G\left[\phi(x_1)\Lambda(t)\right]}, \qquad t \in (0,\infty), \tag{19}$$

Now consider, for example, Theorem 1(iii)d. Suppose that g and $\tilde{g}$ vary regularly at ∞ with exponents $\underline{\tau} \in (-\infty,-1]$ and $\overline{\tau} \in (-\infty,-1]$. Then, again by Feller's Lemma, $1-G$ and $1-\tilde{G}$ vary regularly at ∞ with exponents $\underline{\tau}+1 \in (-\infty,0]$ and $\overline{\tau}+1 \in (-\infty,0]$. By the definition of regular variation (Definition 2), (19) converges to

$$\left[\frac{\tilde{\phi}(x_0)}{\tilde{\phi}(x_1)}\right]^{\overline{\tau}+1} = \lim_{t\to\infty}\frac{1-F(t|x_0)}{1-F(t|x_1)} = \left[\frac{\phi(x_0)}{\phi(x_1)}\right]^{\underline{\tau}+1} \tag{20}$$

as $t \to \infty$ (note that $\tilde{\phi}(x_1)\tilde{\Lambda}(t) \to \infty$ and $\phi(x_1)\Lambda(t) \to \infty$ as $t \to \infty$). Without loss of generality, by (A-4), take x_0 and x_1 such that $\phi(x_0) < \phi(x_1)$. Then, Lemma 1 implies that $\tilde{\phi}(x_0) < \tilde{\phi}(x_1)$ as well.

If both $\underline{\tau} \in (-\infty,-1)$ and $\overline{\tau} \in (-\infty,-1)$, as in Theorem 1(iii)d, then

$$\lim_{t\to\infty}\frac{1-F(t|x_0)}{1-F(t|x_1)} \in (1,\infty). \tag{21}$$

Conversely, (21) and regular variation of g and $\tilde{g}$ imply that $\underline{\tau},\overline{\tau} \in (-\infty,-1)$. In this case, ϕ and $\tilde{\phi}$ are related as in Theorem 1(i).

If $\underline{\tau} = \overline{\tau} = -1$, then

$$\lim_{t\to\infty}\frac{1-F(t|x_0)}{1-F(t|x_1)} = 1. \tag{22}$$

Conversely, (22) and regular variation of g and $\tilde{g}$ imply that $\underline{\tau} = \overline{\tau} = -1$. In this case, ϕ and $\tilde{\phi}$ may be related as in Theorem 1(i), for example if $(\Lambda,\phi,G) = (\tilde{\Lambda},\tilde{\phi},\tilde{G})$. However, (20) does not imply that they are; in particular, in this case, (20) is satisfied for all ϕ and $\tilde{\phi}$.

The other boundary case of Theorem 1(iii)d occurs if $\underline{\tau} = \overline{\tau} = -\infty$. In this case, g and $\tilde{g}$ *vary rapidly* at ∞. Rapidly varying functions (De Haan, 1970, Section 1.1) satisfy Definition 2 with $\tau = -\infty$ or $\tau = \infty$ (De Haan, 1970, Section 1.1) if we define, for $\alpha \in (0, \infty)$,

$$\alpha^{\infty} \equiv \begin{cases} 0 & \alpha < 1, \\ 1 \quad \text{if} & \alpha = 1, \\ \infty & \alpha > 1; \end{cases} \qquad \text{and} \qquad \alpha^{-\infty} \equiv \begin{cases} \infty & \alpha < 1, \\ 1 \quad \text{if} & \alpha = 1, \text{ and} \\ 0 & \alpha > 1. \end{cases}$$

This gives the following generalization of regular variation.

Definition 3. A function $k : (0, \infty) \to (0, \infty)$ is τ*-varying* at 0 (at ∞), $\tau \in [-\infty, \infty]$, if $k(\alpha s)/k(s) \to \alpha^{\tau}$ as $s \to 0$ ($s \to \infty$) for every $\alpha \in (0, \infty)$.

If g and $\tilde{g}$ are $-\infty$-varying, and still $\phi(x_0) < \phi(x_1)$, then

$$\lim_{t \to \infty} \frac{1 - F(t|x_0)}{1 - F(t|x_1)} = \infty. \tag{23}$$

Conversely, if $g, \tilde{g}$ are $\underline{\tau}, \overline{\tau}$-varying and (23) holds, then $\underline{\tau} = \overline{\tau} = -\infty$. As in the other boundary case, ϕ and $\tilde{\phi}$ may be related as in Theorem 1(i), but (20) does not guarantee that they are.

Taken together, this implies that GAFT triplets that satisfy the sufficient condition in Theorem 1(iii)d cannot be observationally equivalent to GAFT triplets that satisfy a boundary case of this same condition. The following lemma summarizes this result and extends it to the other boundary cases.

Lemma 3. *Let* (Λ, ϕ, G) *and* $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ *be observationally equivalent GAFT triplets that satisfy (A-1)–(A-4). Let* $x_0, x_1 \in \mathcal{X}$ *be such that* $\phi(x_0) \neq \phi(x_1)$.

(i). If $\lambda, \tilde{\lambda}$ *are* $\overline{\tau}, \underline{\tau}$*-varying at 0; with necessarily* $\overline{\tau}, \underline{\tau} \in [-1, \infty]$*; then*

$$\lim_{s \to 0} \frac{F^{-1}(s|x_0)}{F^{-1}(s|x_1)} \in \begin{cases} \{0, \infty\} & \iff \quad \overline{\tau} = \underline{\tau} = -1, \\ (0,1) \cup (1, \infty) & \iff \quad \overline{\tau}, \underline{\tau} \in (-1, \infty), \ \textit{and} \\ \{1\} & \iff \quad \overline{\tau} = \underline{\tau} = \infty. \end{cases}$$

(ii). If $\lambda, \tilde{\lambda}$ *are* $\overline{\tau}, \underline{\tau}$*-varying at* ∞*; with necessarily* $\overline{\tau}, \underline{\tau} \in [-1, \infty]$*; then the same result holds if we take the limit* $s \to \infty$.

(iii). If $g, \tilde{g}$ *are* $\underline{\tau}, \overline{\tau}$*-varying at 0; with necessarily* $\underline{\tau}, \overline{\tau} \in [-1, \infty]$*; then*

$$\lim_{t \to 0} \frac{F(t|x_0)}{F(t|x_1)} \in \begin{cases} \{1\} & \iff \quad \underline{\tau} = \overline{\tau} = -1, \\ (0,1) \cup (1, \infty) & \iff \quad \underline{\tau}, \overline{\tau} \in (-1, \infty), \ \textit{and} \\ \{0, \infty\} & \iff \quad \underline{\tau} = \overline{\tau} = \infty. \end{cases}$$

(iv). If $g, \tilde{g}$ *are* $\underline{\tau}, \overline{\tau}$*-varying at* ∞*; with necessarily* $\underline{\tau}, \overline{\tau} \in [-\infty, -1]$*; then*

$$\lim_{t \to 0} \frac{1 - F(t|x_0)}{1 - F(t|x_1)} \in \begin{cases} \{0, \infty\} & \iff \quad \underline{\tau} = \overline{\tau} = -\infty, \\ (0,1) \cup (1, \infty) & \iff \quad \underline{\tau}, \overline{\tau} \in (-1, \infty), \ \textit{and} \\ \{1\} & \iff \quad \underline{\tau} = \overline{\tau} = -1. \end{cases}$$

Proof. (i). First, suppose that $\lambda, \tilde{\lambda}$ are $\overline{\tau}, \underline{\tau}$-varying at 0 with $\overline{\tau}, \underline{\tau} \in [-1, \infty]$. Define $\lambda^\star(s) \equiv \lambda(1/s)$. Let

$$\Lambda^\star(s) \equiv \Lambda(1/s) = \int_0^{1/s} \lambda(u) du = \int_s^\infty u^{-2} \lambda^\star(u) du$$

and note that $s \mapsto s^{-2}\lambda^\star(s)$ is $(-\overline{\tau} - 2)$-varying at ∞. By Lemma 1.2.2 in De Haan

(1970)[7], $\Lambda^\star$ is $(-\overline{\tau}-1)$-varying at ∞. Consequently, Λ is $(\overline{\tau}+1)$-varying at 0 and, by Corollary 2.2.1 in De Haan (1970), Λ^{-1} is $(\overline{\tau}+1)^{-1}$-varying at 0 (here, we take $(\overline{\tau}+1)^{-1}=\infty$ if $\overline{\tau}=-1$ and $(\overline{\tau}+1)^{-1}=0$ if $\overline{\tau}=\infty$). Similarly, it follows that $\tilde{\Lambda}^{-1}$ is $(\underline{\tau}+1)^{-1}$-varying at 0. By observational equivalence

$$\frac{\tilde{\Lambda}^{-1}\left[\frac{\tilde{\phi}(x_1)}{\tilde{\phi}(x_0)}\tilde{\phi}(x_1)^{-1}\tilde{G}(s)\right]}{\tilde{\Lambda}^{-1}\left[\tilde{\phi}(x_1)^{-1}\tilde{G}(s)\right]}=\frac{F^{-1}(s|x_0)}{F^{-1}(s|x_1)}=\frac{\Lambda^{-1}\left[\frac{\phi(x_1)}{\phi(x_0)}\phi(x_1)^{-1}G(s)\right]}{\Lambda^{-1}\left[\phi(x_1)^{-1}G(s)\right]},\qquad s\in(0,\infty); \tag{24}$$

so that

$$\begin{aligned}\left[\frac{\tilde{\phi}(x_1)}{\tilde{\phi}(x_0)}\right]^{\frac{1}{\underline{\tau}+1}}&=\lim_{s\to 0}\frac{\tilde{\Lambda}^{-1}\left[\frac{\tilde{\phi}(x_1)}{\tilde{\phi}(x_0)}\tilde{\phi}(x_1)^{-1}\tilde{G}(s)\right]}{\tilde{\Lambda}^{-1}\left[\tilde{\phi}(x_1)^{-1}\tilde{G}(s)\right]}\\&=\lim_{s\to 0}\frac{F^{-1}(s|x_0)}{F^{-1}(s|x_1)}=\lim_{s\to 0}\frac{\Lambda^{-1}\left[\frac{\phi(x_1)}{\phi(x_0)}\phi(x_1)^{-1}G(s)\right]}{\Lambda^{-1}\left[\phi(x_1)^{-1}G(s)\right]}=\left[\frac{\phi(x_1)}{\phi(x_0)}\right]^{\frac{1}{\overline{\tau}+1}}.\end{aligned}$$

With $\phi(x_0)\neq\phi(x_1)$ and Lemma 1, this gives the desired result.

(ii). Next, let $\lambda,\tilde{\lambda}$ be $\overline{\tau},\underline{\tau}$-varying at ∞ with $\overline{\tau},\underline{\tau}\in[-1,\infty]$. By Lemma 1.2.2 and Corollary 2.2.1 in De Haan (1970), Λ^{-1} is $(\overline{\tau}+1)^{-1}$-varying at ∞ and $\tilde{\Lambda}^{-1}$ is $(\underline{\tau}+1)^{-1}$-varying at ∞. With observational equivalence, in particular (24), this implies that

$$\begin{aligned}\left[\frac{\tilde{\phi}(x_1)}{\tilde{\phi}(x_0)}\right]^{\frac{1}{\underline{\tau}+1}}&=\lim_{s\to \infty}\frac{\tilde{\Lambda}^{-1}\left[\frac{\tilde{\phi}(x_1)}{\tilde{\phi}(x_0)}\tilde{\phi}(x_1)^{-1}\tilde{G}(s)\right]}{\tilde{\Lambda}^{-1}\left[\tilde{\phi}(x_1)^{-1}\tilde{G}(s)\right]}\\&=\lim_{s\to \infty}\frac{F^{-1}(s|x_0)}{F^{-1}(s|x_1)}=\lim_{s\to \infty}\frac{\Lambda^{-1}\left[\frac{\phi(x_1)}{\phi(x_0)}\phi(x_1)^{-1}G(s)\right]}{\Lambda^{-1}\left[\phi(x_1)^{-1}G(s)\right]}=\left[\frac{\phi(x_1)}{\phi(x_0)}\right]^{\frac{1}{\overline{\tau}+1}}.\end{aligned}$$

With $\phi(x_0)\neq\phi(x_1)$ and Lemma 1, this gives the desired result.

(iii). Now, suppose that $g,\tilde{g}$ are $\underline{\tau},\overline{\tau}$-varying at 0 with $\underline{\tau},\overline{\tau}\in[-1,\infty]$. By an argument

[7] Lemma 1.2.2 in De Haan (1970) is an extension to τ-varying functions of the Lemma in Feller (1971, Section VIII.9) used in the main text.

as that for Λ and $\tilde{\Lambda}$ in (i), G is $\underline{\tau}+1$-varying at 0 and $\tilde{G}$ is $\overline{\tau}+1$-varying at 0. With observational equivalence; in particular

$$\frac{\tilde{G}\left[\frac{\tilde{\phi}(x_0)}{\tilde{\phi}(x_1)}\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]}{\tilde{G}\left[\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]} = \frac{F(t|x_0)}{F(t|x_1)} = \frac{G\left[\frac{\phi(x_0)}{\phi(x_1)}\phi(x_1)\Lambda(t)\right]}{G\left[\phi(x_1)\Lambda(t)\right]}, \qquad t \in (0,\infty);$$

this implies that

$$\begin{aligned}\left[\frac{\tilde{\phi}(x_0)}{\tilde{\phi}(x_1)}\right]^{\overline{\tau}+1} &= \lim_{t\to 0}\frac{\tilde{G}\left[\frac{\tilde{\phi}(x_0)}{\tilde{\phi}(x_1)}\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]}{\tilde{G}\left[\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]}\\ &= \lim_{t\to 0}\frac{F(t|x_0)}{F(t|x_1)} = \lim_{t\to 0}\frac{G\left[\frac{\phi(x_0)}{\phi(x_1)}\phi(x_1)\Lambda(t)\right]}{G\left[\phi(x_1)\Lambda(t)\right]} = \left[\frac{\phi(x_0)}{\phi(x_1)}\right]^{\underline{\tau}+1}.\end{aligned}$$

With $\phi(x_0) \neq \phi(x_1)$ and Lemma 1, this gives the desired result.

(iv). Finally, suppose that $g, \tilde{g}$ are $\underline{\tau}, \overline{\tau}$-varying at ∞ with $\underline{\tau}, \overline{\tau} \in [-\infty, 1]$. By Lemma 1.2.2 in De Haan (1970), $1-G$ is $\underline{\tau}+1$-varying at ∞ and $1-\tilde{G}$ is $\overline{\tau}+1$-varying at ∞. With observational equivalence, in particular (19), this implies that

$$\begin{aligned}\left[\frac{\tilde{\phi}(x_0)}{\tilde{\phi}(x_1)}\right]^{\overline{\tau}+1} &= \lim_{t\to \infty}\frac{1-\tilde{G}\left[\frac{\tilde{\phi}(x_0)}{\tilde{\phi}(x_1)}\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]}{1-\tilde{G}\left[\tilde{\phi}(x_1)\tilde{\Lambda}(t)\right]}\\ &= \lim_{t\to \infty}\frac{1-F(t|x_0)}{1-F(t|x_1)} = \lim_{t\to \infty}\frac{1-G\left[\frac{\phi(x_0)}{\phi(x_1)}\phi(x_1)\Lambda(t)\right]}{1-G\left[\phi(x_1)\Lambda(t)\right]} = \left[\frac{\phi(x_0)}{\phi(x_1)}\right]^{\underline{\tau}+1}.\end{aligned}$$

With $\phi(x_0) \neq \phi(x_1)$ and Lemma 1, this gives the desired result. □

C Discrete Duration Data

Suppose that, in contrast to the setup in Section 2, the data provide us with $F(t|x)$ for only two distinct values $t_0, t_1 \in (0,\infty)$ of t and all $x \in \mathcal{X}$. Without loss of generality, let $t_0 < t_1$. With such data, two GAFT triplets (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ are *observationally* *equivalent* if $G[\phi(x)\Lambda(t)] = F(t|x) = \tilde{G}\left[\tilde{\phi}(x)\tilde{\Lambda}(t)\right]$ for all $t \in \{t_1, t_2\}$ and all $x \in \mathcal{X}$.

Consider the following alternatives for (A-2) and (A-4):

(A-2*). $\phi : \mathcal{X} \to (0,\infty)$ is such that, for some *covariate path* $\xi : (0,\infty) \to \mathcal{X}$, $\Psi \equiv \phi \circ \xi$ can be written as $\Psi(s) = \int_0^s \psi(u)du$, $s \in (0,\infty)$, for some $\psi : (0,\infty) \to (0,\infty)$ that is integrable on finite intervals, and $\lim_{s\to\infty} \Psi(s) = \infty$.

(A-4*). $\Lambda(t_0) \in (0,\infty)$ and $\Lambda(t_1) \in (0,\infty)$ are such that $\Lambda(t_0) < \Lambda(t_1)$.

Assumption (A-2*) requires that there exists a covariate path ξ such that $\Psi \equiv \phi \circ \xi$ satisfies the conditions on Λ in (A-2): Ψ is absolutely continuous on bounded intervals and strictly increasing, $\lim_{s\to 0} \Psi(s) = 0$, and $\lim_{s\to\infty} \Psi(s) = \infty$. The following result shows that such a covariate path can be determined from the data.

Lemma 4. *Let (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ be observationally equivalent GAFT triplets that satisfy (A-1), (A-2*), (A-3), and (A-4*). Let $\xi : (0,\infty) \to \mathcal{X}$ be a covariate path such that $\Psi \equiv \phi \circ \xi$ satisfies the conditions on Λ in (A-2). Then, $\tilde{\Psi} \equiv \tilde{\phi} \circ \xi$ satisfies these same conditions.*

Proof. Observational equivalence implies that

$$G[\Psi(s)\Lambda(t_0)] = F(t_0|x) = \tilde{G}\left[\tilde{\Psi}(s)\tilde{\Lambda}(t_0)\right], \quad s \in (0,\infty). \tag{25}$$

Because Ψ is absolutely continuous on bounded intervals and nondecreasing; and, by (A-3), G is absolutely continuous; the left-hand side of (25), as a function of s, is absolutely continuous on bounded intervals (and on $(0,\infty)$, because it is monotone and bounded). Moreover, (A-2*) and (A-3) imply that it is strictly increasing, converges to 0 as $s \to 0$, and converges to 1 as $s \to \infty$.

The right-hand side of (25) should have these same properties. Because, by (A-3), $\tilde{G}$ is absolutely continuous and strictly increasing, with $\lim_{s\to 0} \tilde{G}(s) = 0$ and $\lim_{s\to\infty} \tilde{G}(s) = 1$; this requires that $\tilde{\Psi}$ is absolutely continuous on bounded intervals and strictly increasing, $\lim_{s\to 0} \tilde{\Psi}(s) = 0$, and $\lim_{s\to\infty} \tilde{\Psi}(s) = \infty$. □

Lemma 4's assumption that ϕ satisfies (A-2⋆) ensures that the covariate path ξ exists. Lemma 4 shows that, in (A-2⋆), the *same* covariate paths can be used across all observationally equivalent GAFT triplets. More constructively, such covariate paths can be identified with the paths ξ such that $s \mapsto F[t_0|\xi(s)]$ is absolutely continuous and strictly increasing, with $\lim_{s\to 0} F[t_0|\xi(s)] = 0$ and $\lim_{s\to\infty} F[t_0|\xi(s)] = 1$. For example, if $F(t_0|x) = h(\theta' x)$ with h increasing, then a sufficient condition is that one of the components of θ is nonzero, the corresponding x has 'large' support and h is 0 and 1 at the boundary of that support.

Assumption (A-4⋆) ensures that $F(t_0|x) \neq F(t_1|x)$. Note that, because $t_0 < t_1$, (A-2) is sufficient for (A-4⋆). Because we do not need assumptions on $\Lambda(t)$ for $t \notin \{t_1, t_2\}$, (A-2) is not necessary.

Together, Assumptions (A-2⋆) and (A-4⋆) ensure that Theorem 1 can be applied to data on $F[t|\xi(s)]$ for $t \in \{t_0, t_1\} \subset (0,\infty)$ and $s \in (0,\infty)$, with $\Psi \equiv \phi \circ \xi$ taking the role of Λ and $\Lambda(t_0)$ and $\Lambda(t_1)$ taking the roles of $\phi(x_0)$ and $\phi(x_1)$. To this end, consider observationally equivalent GAFT triplets (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$. Define $\tilde{\Psi} \equiv \tilde{\phi} \circ \xi$ and $K \equiv \Psi \circ \tilde{\Psi}^{-1}$. Then, Theorem 1 gives $\Psi = c\tilde{\Psi}^{\rho}$, $\Lambda(t_0) = d\tilde{\Lambda}(t_0)^{\rho}$, $\Lambda(t_1) = d\tilde{\Lambda}(t_1)^{\rho}$, and $\tilde{G}(s) = G(cds^{\rho})$ for all $s \in (0,\infty)$; for some $c, d \in (0,\infty)$; if and only if K' varies regularly at 0 and ∞ with exponent $\rho - 1 \in (-1,\infty)$. With these relations between (Λ, ϕ, G) and $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ in hand, and observational equivalence, it is easy to show that $\phi = c\tilde{\phi}^{\rho}$.

We summarize this result as a corollary to Theorem 1.

Corollary 1. *Let* (Λ, ϕ, G) *and* $(\tilde{\Lambda}, \tilde{\phi}, \tilde{G})$ *be observationally equivalent GAFT triplets that satisfy (A-1), (A-2*⋆*), (A-3), and (A-4*⋆*). Let* $\xi : (0, \infty) \to \mathcal{X}$ *be a covariate path, as in (A-2*⋆*), such that* $\Psi \equiv \phi \circ \xi$ *satisfies the conditions on* Λ *in (A-2). Define* $\tilde{\Psi} \equiv \tilde{\phi} \circ \xi$ *and* $K \equiv \Psi \circ \tilde{\Psi}^{-1}$*, with derivative* $K' : (0, \infty) \to (0, \infty)$ *almost everywhere. Let* $\rho \in (0, \infty)$*. Then,*

(i). for some $c, d \in (0, \infty)$*,*

$$\begin{aligned}
&\phi = c\tilde{\phi}^{\rho}, \\
&\Lambda(t_0) = d\tilde{\Lambda}(t_0)^{\rho}, \\
&\Lambda(t_1) = d\tilde{\Lambda}(t_1)^{\rho}, \text{ and} \\
&\tilde{G}(s) = G(cds^{\rho}) \text{ for all } s \in (0, \infty)
\end{aligned}$$

if and only if

(ii). K' *varies regularly at* 0 *and* ∞ *with exponent* $\rho - 1$*.*

Conditions Theorem 1(iii)c and (iii)d continue to be sufficient for Corollary 1(ii). Theorem 1(iii)a and (iii)b can be straightforwardly adapted to sufficient conditions in terms of the tails of ψ and $\tilde{\psi}$ (which in turn require conditions on $\mathcal{X}$ and ϕ).

References

Abbring, J. H. (2011). Mixed hitting-time models. *Econometrica*. forthcoming. 14

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 Series B 65*, 701–710. 14

Brinch, C. N. (2011). Non-parametric identification of the mixed proportional hazards model with interval-censored durations. *Econometrics Journal 14*(2), 343–350. 15

Chiappori, P.-A. and I. Komunjer (2009, March). Correct specification and identification of nonparametric transformation models. Mimeo, Department of Economics, Columbia University, New York, NY. 14

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society 34*, 187–220. 1

De Haan, L. (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*. Amsterdam: Mathematical Centre. 20, 21, 22, 23

Elbers, C. and G. Ridder (1982). True and spurious duration dependence: The identifiability of the proportional hazards model. *Review of Economic Studies 49*, 402–411. 12, 13, 15

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications (2nd ed.), Volume II*, Volume 2. New York: New York: Wiley. 4, 5, 6, 7, 8, 9, 13, 19, 22

Heckman, J. J. and B. E. Honoré (1989). The identifiability of the competing risks model. *Biometrika 76*, 325–330. 14

Heckman, J. J. and B. Singer (1984). The identifiability of the proportional hazard model. *Review of Economic Studies 51*, 231–241. 13

Honoré, B. E. and A. de Paula (2010). Interdependent durations. *Review of Economic Studies 77*(3), 1138–1163. 14

Horowitz, J. (1996). Semiparametric estimation of a regression model with an unknown transformation of the dependent variable. *Econometrica 64*, 103–137. 14

Kortram, R., A. Lenstra, G. Ridder, and A. van Rooij (1995). Constructive identification of the mixed proportional hazard model. *Statistica Neerlandica 49*, 269–281. 12

Krantz, S. G. and H. R. Parks (2002). *A Primer of Real Analytic Functions.* Birkhäuser. 12

Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrica 47*, 939–956. 1

Ridder, G. (1990). The non-parametric identification of generalized accelerated failure-time models. *Review of Economic Studies 57*, 167–182. 1, 3, 9, 10, 11, 12, 13, 15

Ridder, G. and T. Woutersen (2003). The singularity of the information matrix of the mixed proportional hazard model. *Econometrica 71*, 1579–1589. 13