

Ridder, Geert; Woutersen, Tiemen

Working Paper

The singularity of the efficiency bound of the mixed proportional hazard model

Research Report, No. 2001-9

Provided in Cooperation with:

Department of Economics, University of Western Ontario

Suggested Citation: Ridder, Geert; Woutersen, Tiemen (2001) : The singularity of the efficiency bound of the mixed proportional hazard model, Research Report, No. 2001-9, The University of Western Ontario, Department of Economics, London (Ontario)

This Version is available at:

<https://hdl.handle.net/10419/70365>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

THE SINGULARITY OF THE EFFICIENCY BOUND OF THE MIXED PROPORTIONAL HAZARD MODEL*

BY GEERT RIDDER AND TIEMEN WOUTERSEN[†]
University of Southern California and University of Western Ontario

November 6, 2001

KEYWORDS: Duration, Semi-parametric Efficiency Bound, Mixed Proportional Hazard.

1 Introduction

WE RECONSIDER THE EFFICIENCY BOUND for the semi-parametric Mixed Proportional Hazard (MPH) model with parametric baseline hazard and regression function. This bound was first derived by Hahn (1994). One of his results is that if the baseline hazard is Weibull, the efficiency bound is singular, even if the model is semi-parametrically identified¹. This implies that neither the Weibull parameter nor the regression coefficients can be estimated at the \sqrt{N} rate² (Van der Vaart (1998), Theorem 25.32).

Hahn's result has had an impact on the use of MPH models in empirical research. The singularity of the efficiency bound seems to confirm the results of simulation studies, see e.g. Baker and Melino (2000), that suggest that it is difficult to estimate both the baseline hazard and the distribution of the random effects (or unobserved heterogeneity) with a sufficient degree of accuracy with the sample sizes that one encounters in practice. Indeed Honoré's (1990) estimator for the parameters of a semi-parametric Weibull MPH model converges at a rate that is (much) slower than \sqrt{N} . Altogether these results seem to imply that although the

*We would like to thank Guido Imbens who urged us to write this paper. All errors are ours.

[†]Mailing address: University of Southern California, Department of Economics, Kaprielian Hall, University Park Campus, Los Angeles, CA-90089, USA, and University of Western Ontario, Department of Economics, Social Science Centre, London, Ontario, N6A 5C2, Canada. Email: ridder@usc.edu and twouters@uwo.ca.

MPH model is semi- and even non-parametrically identified, the estimation of the parameters in a semi-parametric MPH model requires a larger dataset than usual.

In this note we show that this impression is false. In particular, we show that the efficiency bound is singular if and only if the parametric model of the (integrated) baseline hazard is closed under the power transformation. The Weibull baseline hazard is the most prominent member of this class of models. All models that are closed under the power transformation have a baseline hazard that is either 0 or ∞ in 0. The restriction that the baseline hazard in 0 is bounded away from 0 and ∞ rules out that the baseline hazard model is closed under the power transformation and this implies that the efficiency bound is nonsingular with this restriction.

We also show that the MPH model is semi-parametrically identified if we restrict the baseline hazard in 0 to be bounded away from 0 and ∞ . Hence, there are (at least) two restrictions that are sufficient for semi-parametric identification: (i) the restriction that the mean of the unobserved random effect is finite (Elbers and Ridder (1982)³), and (ii) the restriction that the baseline hazard in 0 is bounded from 0 and infinity. The first restriction does not preclude that the efficiency bound is singular, the second restriction does. Hence, estimators that impose the second restriction can be \sqrt{N} consistent. Indeed, the rank estimator that has been proposed by Bijwaard and Ridder (2001) and the quantile censoring estimator of Ridder and Woutersen (2001) both are \sqrt{N} consistent for the baseline hazard and regression parameters in the semiparametric MPH model with parametric baseline hazard and regression function. The finite mean assumption is not sufficient to obtain \sqrt{N} consistent estimators in the semi-parametric MPH model. In particular, Ishwaran (1996) has shown that the rate of convergence approaches \sqrt{N} if all moments of the distribution of the unobserved random effect are finite.

The plan of the paper is as follows. In section 2 we give the semi-parametric MPH model and its efficiency bound as obtained by Hahn (1994). We also give an example that shows that if we change the Weibull baseline hazard slightly so that it is bounded from 0 and ∞ in 0, then the information bound becomes nonsingular. Section 3 contains the main result and

section 4 concludes.

2 The Semi-parametric MPH Model: Identification and Efficiency Bound

2.1 The semi-parametric MPH model

We consider the semi-parametric MPH model for the conditional distribution of T given a vector of nonconstant covariates X

$$(1) \quad \theta(t \mid X, \gamma) = \lambda(t, \alpha) e^{\beta' X} e^U$$

with parametric baseline hazard $\lambda(t, \alpha)$ and regression function $e^{\beta' X}$ and (α, β) in a parameter space that is the product of a parameter space for α , A , and a parameter space for β , B , both of which assumed to be open subsets of the Euclidean space. The following assumptions are sufficient for the efficiency bound as derived by Hahn (see Hahn (1994), p. 610)

(A1) $\lambda(t, \alpha)$ and $\Lambda(t, \alpha)$ are assumed to be continuously differentiable with respect to α on A .

(A2) $E(X'X) < \infty$ and there exist non-negative functions $\zeta_i(T, X)$, $i = 1, 2, 3$ such that

$$\begin{aligned} \left| \frac{\partial \ln \lambda(T, \alpha)}{\partial \alpha} \right| &\leq \zeta_1(T) \\ \left| e^{\beta' X} \frac{\partial \ln \Lambda(T, \alpha)}{\partial \alpha} \right| &\leq \zeta_2(T, X) \\ \left| X e^{\beta' X} \Lambda(T, \alpha) \right| &\leq \zeta_3(T, X) \end{aligned}$$

with $E(\zeta_1(T)^2) < \infty$, $E(e^{2U} \zeta_i(T, X)^2) < \infty$, $i = 2, 3$.

The unobserved covariates are captured by the random effect U . Note that our notation deviates somewhat from that used by Hahn (1994): the parameter vector is γ and the baseline hazard is λ (and the integrated baseline hazard Λ). We also take the regression function to be loglinear in X^4 .

If we define the unconditional integrated baseline hazard at the population values of the parameters as

$$(2) \quad S = \Lambda(T, \alpha_0) e^{\beta_0' X}$$

then it is not difficult to show that

$$(3) \quad S \stackrel{d}{=} \frac{W}{e^U}$$

with W a standard exponential random variable that is independent of U and $\stackrel{d}{=}$ means that the random variables on both sides have the same distribution.

2.2 Semi-parametric identification

Elbers and Ridder (1982) have shown that this MPH model is semi-parametrically identified if the following assumptions hold.

(B1) $\Lambda(t_0, \alpha_0) = 1$ for some $t_0 > 0$, $\Lambda(\infty, \alpha_0) = \infty$ with $\Lambda(t, \alpha_0) = \int_0^t \lambda(s, \alpha_0) ds$.

(B2) U and X are independent and $E(e^U) < \infty$.

(B3) There are x_1, x_2 in the support of X with $\beta'_0 x_1 \neq \beta'_0 x_2$.

(B4) If $\lambda(t, \alpha_0) = \lambda(t, \tilde{\alpha}_0)$ for all $t > 0$, then $\alpha_0 = \tilde{\alpha}_0$, and if $\beta'_0 x = \tilde{\beta}'_0 x$ for all x in the support of X , then $\beta_0 = \tilde{\beta}_0$.

Because we can multiply the baseline hazard by a positive constant and subtract the log of that constant from U without changing the model (the mean of U is arbitrary), the first part of assumption (B1) is a normalization. For the same reason, there is no constant in $\beta'X$. Assumption (B4) ensures parametric identification of α_0, β_0 .

We propose an alternative set of assumptions that is sufficient for the semi-parametric identification of the MPH model.

(C1) $\Lambda(t_0, \alpha_0) = 1$ for some $t_0 > 0$, $\Lambda(\infty, \alpha_0) = \infty$ with $\Lambda(t, \alpha_0) = \int_0^t \lambda(s, \alpha_0) ds$. Moreover $0 < \lim_{t \downarrow 0} \lambda(t, \alpha_0) = \lambda(0, \alpha_0) < \infty$.

(C2) U and X are independent and $\Pr(U = -\infty) = \Pr(U = \infty) = 0$.

(C3) There are x_1, x_2 in the support of X with $\beta'_0 x_1 \neq \beta'_0 x_2$.

(C4) If $\lambda(t, \alpha_0) = \lambda(t, \tilde{\alpha}_0)$ for all $t > 0$, then $\alpha_0 = \tilde{\alpha}_0$, and if $\beta'_0 x = \tilde{\beta}'_0 x$ for all x in the support of X , then $\beta_0 = \tilde{\beta}_0$.

The difference between these two sets of assumptions is that the finite mean assumption on the random individual effect $E(e^U) < \infty$ in (B2) is replaced by the assumption that the baseline hazard in 0 is bounded away from 0 and ∞ in (C1).

Proposition 1

If the conditional distribution of T given X has a distribution with a (conditional) hazard as in (1) and if assumptions (C1)-(C4) are satisfied, then α_0, β_0 and the distribution of U are identified, i.e. there are no observationally equivalent $\tilde{\alpha}_0, \tilde{\beta}_0$.

Proof: See Appendix A.

Although both sets of conditions ensure that the semi-parametric MPH model is identified, they have different implications for the information bound of this model. In particular, with the finite mean assumption the information bound can be singular, while with the assumption that the baseline hazard in 0 is bounded from 0 and ∞ this cannot be the case.

2.3 The information bound of the MPH model

Hahn (1994), Theorem 1 derives the efficient score of the MPH model. The variance matrix of the efficient score is the information bound. The efficient score is

$$(4) \quad l = \begin{bmatrix} l_\alpha \\ l_\beta \end{bmatrix} = \begin{bmatrix} a_{11} - a_{12}S \cdot E[e^U|S] \\ a_2 - a_2S \cdot E[e^U|S] \end{bmatrix}.$$

with

$$(5) \quad \begin{aligned} a_{11} &= \frac{\partial \ln \lambda(T, \alpha)}{\partial \alpha} - E \left[\frac{\partial \ln \lambda(T, \alpha)}{\partial \alpha} | S \right] \\ a_{12} &= \frac{\partial \ln \Lambda(T, \alpha)}{\partial \alpha} - E \left[\frac{\partial \ln \Lambda(T, \alpha)}{\partial \alpha} | S \right] \\ a_2 &= X - E(X|S) = X - E(X). \end{aligned}$$

Without loss of generality we assume that $E(X) = 0$.

For the Weibull baseline hazard $\lambda(t, \alpha) = \alpha t^{\alpha-1}$ we have

$$(6) \quad a_{11} = a_{12} = \ln T - E(\ln T|S)$$

and by (2) $\ln T = \frac{\ln S - \beta'_0 X}{\alpha_0}$ so that

$$(7) \quad a_{11} = a_{12} = -\frac{\beta'_0}{\alpha_0} X.$$

Substitution in (4) yields

$$(8) \quad l = (1 - SE(e^U|S)) \begin{bmatrix} -\frac{\beta'_0}{\alpha_0} X \\ X \end{bmatrix}$$

and because the first component of this vector is a linear combination of the other components with nonrandom coefficients, the distribution of the efficient score is singular as is its variance matrix. This is the argument given by Hahn (1994), p. 614.

Note that this argument is not restricted to the Weibull baseline hazard. It applies to all integrated baseline hazards of the form $\lambda(t, \alpha) = h(t)^\alpha$ with h a known strictly increasing function of t with $h(0) = 0$. However, a small change in the Weibull baseline hazard gives a nonsingular information bound. In particular, consider the translated Weibull baseline hazard with integrated baseline hazard $\Lambda_\varepsilon(t, \alpha) = (t + \varepsilon)^\alpha - \varepsilon^\alpha$ with $\varepsilon > 0$ a known constant. The component for β in the efficient score (4) is still equal to X . The only change is in the component for α . To be specific for $\alpha = \alpha_0$

$$(9) \quad a_{11} = \frac{1}{\alpha_0} \ln(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0}) - \frac{1}{\alpha_0} E_X \left[\ln(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0}) \right]$$

$$(10) \quad a_{12} = \frac{e^{\beta'_0 X}}{\alpha_0 S} (e^{-\beta'_0 X} S + \varepsilon^{\alpha_0}) \ln(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0}) - \\ - E_X \left[\frac{e^{\beta'_0 X}}{\alpha_0 S} (e^{-\beta'_0 X} S + \varepsilon^{\alpha_0}) \ln(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0}) \right] - \\ - \frac{e^{\beta'_0 X} \varepsilon^{\alpha_0} \ln \varepsilon}{S} + E_X \left[\frac{e^{\beta'_0 X} \varepsilon^{\alpha_0} \ln \varepsilon}{S} \right].$$

To see that the distribution of the efficient score is nonsingular, consider the special case $U \equiv 0$ so that $E[e^U|S] = 1$. Then a necessary condition for singularity is that

$$(11) \quad a_{11} - a_{12} S = \frac{1}{\alpha_0} \ln(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0}) - \frac{1}{\alpha_0} E_X \left[\ln(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0}) \right] -$$

$$\begin{aligned}
& -\frac{e^{\beta'_0 X}}{\alpha_0} \left(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0} \right) \ln \left(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0} \right) + \\
& + \mathbb{E}_X \left[\frac{e^{\beta'_0 X}}{\alpha_0} \left(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0} \right) \ln \left(e^{-\beta'_0 X} S + \varepsilon^{\alpha_0} \right) \right] + \\
& + \frac{e^{\beta'_0 X} \varepsilon^{\alpha_0} \ln \varepsilon}{S} - \mathbb{E}_X \left[\frac{e^{\beta'_0 X} \varepsilon^{\alpha_0} \ln \varepsilon}{S} \right]
\end{aligned}$$

is constant in S for all x in the support of X , and this is true if and only if $\varepsilon = 0$. We conclude that the efficiency bound for the ε translated Weibull baseline hazard is nonsingular. Note that this class of integrated baseline hazard models is not closed under the power transformation. Also the baseline hazard of this model is $\lambda_\varepsilon(t, \alpha) = \alpha(t + \varepsilon)^{\alpha-1}$ with $\lambda_\varepsilon(0, \alpha) = \varepsilon^{\alpha-1}$ which is bounded away from 0 and ∞ if $\varepsilon > 0$.

3 Necessary and Sufficient Conditions for the Singularity of the Efficiency Bound

We first rewrite the efficient score in (4) and (5) to reflect the dependence on T, X, S and the parameters α_0, β_0 . For that purpose we denote $\Lambda_0(t) = \Lambda(t, \alpha_0)$, $\lambda_0(t) = \lambda(t, \alpha_0)$, and $Z = \beta'_0 X$ so that $T = \Lambda_0^{-1}(S e^{-Z})$. Also $V = e^U$.

$$(12) \quad l = \begin{bmatrix} l_\alpha \\ l_\beta \end{bmatrix} = \begin{bmatrix} a_{11}(T, S, \alpha_0) - a_{12}(T, S, \alpha_0) H_V(S) \\ X(1 - H_V(S)) \end{bmatrix}$$

with $H_V(S) = SE(V|S) = -\frac{SL_V''(S)}{L_V'(S)}$ and $L_V(s) = \mathbb{E}(e^{-sV})$ the Laplace transform of V . Note that H_V does not depend on the parameters. Because S and Z are independent we have

$$\begin{aligned}
a_{11}(T, S, \alpha_0) &= \frac{\partial \ln \lambda_0(T)}{\partial \alpha} - \mathbb{E}_Z \left[\frac{\partial \{ \ln \lambda_0(\Lambda_0^{-1}(S e^{-Z})) \}}{\partial \alpha} \right] \\
a_{12}(T, S, \alpha_0) &= \frac{\partial \ln \Lambda_0(T)}{\partial \alpha} - \mathbb{E}_Z \left[\frac{\partial \{ \ln \Lambda_0(\Lambda_0^{-1}(S e^{-Z})) \}}{\partial \alpha} \right].
\end{aligned}$$

By (2) the variables are related by

$$(13) \quad \ln \Lambda_0(T) + Z = \ln S.$$

By assumptions (A1) and (A2) the information bound is continuous in α_0 . We first consider the case that α is a scalar. If the information bound has a rank equal to the number of regressors in X , i.e. one less than full rank, for some value of α_0 , then it has the same rank for population parameters in a small neighborhood of that value of α_0 by continuity in α_0 . We keep β_0 fixed. Note that T depends on X only through $\beta_0'X$. Because of (B4) or (C4) the linear combination that makes the score singular must contain l_α . Because l_α depends on X only through $\beta_0'X$, loss of rank occurs if and only if l_α is proportional to $\beta_0'X$, i.e. there is a $c(\alpha_0) \neq 0$ such that

$$(14) \quad c(\alpha_0)a_{11}(T, S, \alpha_0) - c(\alpha_0)a_{12}(T, S, \alpha_0)H_V(S) = Z(1 - H_V(S))$$

for all α_0 in some open interval, $S \geq 0$ and Z, T that satisfy (13). From (14) it follows that

$$(15) \quad a_{11}(t, s, \alpha_0) = a_{12}(t, s, \alpha_0)$$

for if this equality does not hold for some α_0 , it does not hold on some open interval, because of (A1) and (A2). Moreover, there is a t such that $a_{11}(t, s, \alpha_0), a_{12}(t, s, \alpha_0)$ are not constant in α_0 on that interval by (B4) or (C4). Only if the equality holds we can find $c(\alpha_0)$ such that the left-hand side does not depend on α_0 .

Substitution in (15) gives that for all α_0 in some open interval and for $s \geq 0$ and t that satisfies (13) for some z in the support of Z

$$(16) \quad \frac{\partial \ln \lambda(t, \alpha_0)}{\partial \alpha} - E_Z \left[\frac{\partial \ln \lambda(\Lambda_0^{-1}(se^{-Z}), \alpha_0)}{\partial \alpha} \right] - \frac{\partial \ln \Lambda(t, \alpha_0)}{\partial \alpha} - E_Z \left[\frac{\partial \ln \Lambda(\Lambda_0^{-1}(se^{-Z}), \alpha_0)}{\partial \alpha} \right] = 0.$$

Note that both a_{11} and a_{12} are identically equal to 0 if Z takes only one value. If Z takes two (or more) values, then (16) holds if and only if for α_0 in some open interval and $t > 0$

$$(17) \quad \frac{\partial \ln \lambda(t, \alpha_0)}{\partial \alpha} - \frac{\partial \ln \Lambda(t, \alpha_0)}{\partial \alpha} = e(\alpha_0).$$

Integration with respect to α_0 and t gives (using $\Lambda(t_0, \alpha_0) = 1$)

$$(18) \quad \ln \Lambda(t, \alpha_0) = e^{f(\alpha_0)} \int_{t_0}^t e^{k(s)} ds$$

with $\int_0^{t_0} e^{k(s)} ds = -\infty$ and $\int_{t_0}^{\infty} e^{k(s)} ds = \infty$. If we define $\ln h(t) = \int_{t_0}^t e^{k(s)} ds$ and $d(\alpha_0) = e^{f(\alpha_0)} > 0$, we find

$$(19) \quad \Lambda_0(t, \alpha_0) = h(t)^{d(\alpha_0)}$$

with h an increasing function with $h(0) = 0$ and $h(\infty) = \infty$.

We have proved the following proposition.

Proposition 2

If the assumptions (A1), (A2) and (C1)-(C4) hold except for the assumption that the baseline hazard is bounded from 0 and ∞ in 0, then the information bound is singular if and only if the integrated baseline hazard is of the form (19) for some strictly increasing continuous function h with $h(0) = 0$, $h(\infty) = \infty$ and $d(\alpha_0) > 0$.

Proposition 2 is for the case that the baseline hazard depends on one parameter α_0 . If α_0 is a vector, (17) becomes

$$(20) \quad c(\alpha_0)' \frac{\partial \ln \lambda(t, \alpha_0)}{\partial \alpha} - c(\alpha_0)' \frac{\partial \ln \Lambda(t, \alpha_0)}{\partial \alpha} = e(\alpha_0)$$

for some vector $c(\alpha_0)$. Consider the case that α_0 has two parameters. Then from (20)

$$(21) \quad \frac{\partial \ln \lambda(t, \alpha_0)}{\partial \alpha_1} - \frac{\partial \ln \Lambda(t, \alpha_0)}{\partial \alpha_1} = \frac{e(\alpha_0)}{c_1(\alpha_0)} - \frac{c_2(\alpha_0)}{c_1(\alpha_0)} \left(\frac{\partial \ln \lambda(t, \alpha_0)}{\partial \alpha_2} - \frac{\partial \ln \Lambda(t, \alpha_0)}{\partial \alpha_2} \right).$$

Integration with respect to α_1 and t yields the representation

$$(22) \quad \Lambda_0(t, \alpha_0) = h(t, \alpha_0)^{d(\alpha_0)}$$

with $d(\alpha_0) = e^{\int \frac{e(\alpha_0)}{c_1(\alpha_0)} d\alpha_1}$ and $\ln h(t, \alpha_0) = \int_{t_0}^t e^{l(s, \alpha_1) - \frac{c_2(\alpha_0)}{c_1(\alpha_0)} \left(\frac{\partial \ln \lambda(s, \alpha_0)}{\partial \alpha_2} - \frac{\partial \ln \Lambda(s, \alpha_0)}{\partial \alpha_2} \right)} ds$

so that Proposition 2 still holds with an obvious modification.

The baseline hazard that corresponds to (19) is

$$(23) \quad \lambda(t, \alpha_0) = d(\alpha_0) h(t)^{d(\alpha_0)-1} h'(t).$$

Note that the proposition only restricts $d(\alpha_0)$ to be positive. In particular, it can be either smaller or larger than 1. If $d(\alpha_0) < 1$, then by (23) $\lambda(0, \alpha_0) = \infty$. If $d(\alpha_0) > 1$, then

$\lambda(0, \alpha_0) = 0$. Only if $d(\alpha_0) = 1$, in which case the baseline hazard is known, the baseline hazard in 0 can be bounded from 0 and ∞ . Hence we have

Theorem

If the assumptions for Proposition 2 hold, then $0 < \lambda(0, \alpha_0) < \infty$ implies that the efficiency bound of the semi-parametric MPH model in (1) is nonsingular.

4 Conclusion

By Proposition 1 the condition that the baseline hazard in 0 is bounded away from 0 and ∞ is sufficient for semi-parametric identification. This condition is also sufficient for a non-singular efficiency bound. Hence, there may be (regular) estimators in the semi-parametric MPH model with parametric baseline hazard and regression function that are \sqrt{N} consistent.

Appendix A: Proof of Proposition 1

By (2) and (3) we have for all $t > 0$

$$(24) \quad \Pr(T \leq t | X) = F_V \left(\Lambda(t, \alpha_0) e^{\beta_0 X} \right)$$

where $V = \frac{W}{e^\sigma}$ is distributed as a mixture of exponential distributions and hence has a strictly increasing cdf F_V . Without loss of generality we can assume that $\Lambda(t, \alpha_0)$ is strictly increasing in t^5 . If $\tilde{\alpha}_0, \tilde{\beta}_0, \tilde{U}$ are observationally equivalent, then for all $t > 0$

$$(25) \quad F_V \left(\Lambda(t, \alpha_0) e^{\beta_0 X} \right) = F_{\tilde{V}} \left(\Lambda(t, \tilde{\alpha}_0) e^{\tilde{\beta}_0 X} \right)$$

We denote $\Lambda(t, \alpha_0) = \Lambda(t)$, $\Lambda(t, \tilde{\alpha}_0) = \tilde{\Lambda}(t)$, $e^{-\beta'_0 x_1} = \phi_1$, $e^{-\beta'_0 x_2} = \phi_2$, $e^{-\tilde{\beta}'_0 x_1} = \tilde{\phi}_1$, $e^{-\tilde{\beta}'_0 x_2} = \tilde{\phi}_2$ with x_1, x_2 as in (B3) and without loss of generality $1 = \phi_1 > \phi_2$, $1 = \tilde{\phi}_1 > \tilde{\phi}_2$.

Because M is strictly increasing its inverse M^{-1} exists and from (25) for all $t > 0$

$$(26) \quad F_V \left(\Lambda \left(\tilde{\Lambda}^{-1}(t\tilde{\phi}_2) \right) \frac{1}{\phi_2} \right) = F_{\tilde{V}}(t) = F_V \left(\Lambda \left(\tilde{\Lambda}^{-1}(t) \right) \right)$$

If we denote $K = \Lambda \left(\tilde{\Lambda}^{-1}(t) \right)$ with K is strictly increasing and $K(0) = 0$, then (26) implies that

$$(27) \quad K(t\tilde{\phi}_2) = \phi_2 K(t)$$

and by iteration for all $n \geq 1$

$$(28) \quad K(t\tilde{\phi}_2^n) = \phi_2^n K(t)$$

If we take the derivative of (27) we obtain

$$(29) \quad \frac{\phi_2}{\tilde{\phi}_2} K'(t) = K'(\tilde{\phi}_2 t)$$

and by iteration for all $n \geq 1$

$$(30) \quad \left(\frac{\phi_2}{\tilde{\phi}_2} \right)^n K'(t) = K'(\tilde{\phi}_2^n t)$$

Taking the ratio of (30) and (28) we obtain because $K'(t) = \frac{\lambda(\tilde{\Lambda}^{-1}(t))}{\lambda(\tilde{\Lambda}^{-1}(t))}$ with $\lambda(t) = \lambda(t, \alpha_0)$, $\tilde{\lambda}(t) = \lambda(t, \tilde{\alpha}_0)$

$$(31) \quad \frac{K'(t)}{K(t)} = \frac{1}{t} \lim_{n \rightarrow \infty} \frac{K'(\tilde{\phi}_1^n t)}{\frac{K(\tilde{\phi}_1^n t)}{\tilde{\phi}_1^n t}} = \frac{1}{t} \frac{\frac{\lambda(\tilde{\Lambda}^{-1}(\tilde{\phi}_1^n t))}{\tilde{\lambda}(\tilde{\Lambda}^{-1}(\tilde{\phi}_1^n t))}}{\frac{K(\tilde{\phi}_1^n t)}{\tilde{\phi}_1^n t}} = \frac{1}{t}$$

by assumption (B1). Because $K(0) = 0$ this implies that $K(t) = t$ and hence $\lambda(t, \alpha_0) = \lambda(t, \tilde{\alpha}_0)$ for $t > 0$ so that $\alpha_0 = \tilde{\alpha}_0$ by (B4). By (27) $\beta'_0 x_2 = \tilde{\beta}_0 x_2$ for all x_2 in the support of X and hence $\beta_0 = \tilde{\beta}_0$ by (B4).

5 References

- Baker, M. and A. Melino (2000): “Duration Dependence and Nonparametric Heterogeneity: A Monte Carlo Study Source,” *Journal of Econometrics*, 96, 357-93.
- Bijwaard, G. and G. Ridder (2001): “Correcting for Selective Compliance in a Re-employment Bonus Experiment”, CLEO Working paper, University of Southern California.
- Elbers, C. and G. Ridder (1982): “True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model,” *Review of Economic Studies*, 49, 402-409.
- Hahn, J. (1994): “The Efficiency Bound of the Mixed Proportional Hazard Model,” *Review of Economic Studies*, 61, 607-629.
- Heckman, J. J. and B. Singer (1984): “The Identifiability of the Proportional Hazard Model,” *Review of Economic Studies*, 60, 231-243.
- Honoré, B.E. (1990): “Simple Estimation of a Duration Model with Unobserved Heterogeneity,” *Econometrica*, 58, 453-473.
- Ishwaran, H. (1996): “Uniform Rates of Estimation in the Semiparametric Weibull Mixture Model,” *The Annals of Statistics*, 24, 1572-1585.
- Jewell, N. (1982): “Mixtures of Exponential Distributions,” *The Annals of Statistics*, 10, 479-484.
- Ridder, G. (1990): “The Non-Parametric Identification of Generalized Accelerated Failure-Time Models, *Review of Economic Studies*, 57, 167-182.
- Ridder, G. and T. Woutersen (2001): “Method of Moments Estimation of Duration Models with Exogenous and Endogenous Regressors,” , in preparation.
- Van der Vaart, A.W. (1998): *Asymptotic Statistics*, Cambridge University Press, Cambridge, UK.

Notes

¹The singularity holds if we have single-spell duration data. Hahn shows that the efficiency bound is nonsingular, if we have two or more spells for the same individual provided that the individual random effect is the same for both spells.

²That is by a regular estimator sequence (for a definition see Van der Vaart (1998), p. 115).

³See also Jewell (1982) and Heckman and Singer(1984) who consider an alternative identifying assumption that allows for an infinite mean, but assumes that the power transformation is fixed

⁴The class of loglinear regression functions is closed under the power transformation. The semi-parametric MPH model can be identified by choosing a class of parametric regression functions that is not closed under the power transformation. However, this is identification by arbitrary functional form assumptions. The loglinear specification imposes no (identifying) restrictions.

⁵If not then there are intervals that have zero probability. These intervals can be omitted and by redefining T accordingly we have that $\lambda(t, \alpha_0) > 0$ for $t > 0$.