Econometric Theory, **23**, 2007, 349–354. Printed in the United States of America.
DOI: 10.1017/S0266466607070144

NONPARAMETRIC IDENTIFICATION OF THE MIXED HAZARDS MODEL WITH TIME-VARYING COVARIATES

CHRISTIAN N. BRINCH
University of Oslo

Most nonparametric identification results for the mixed proportional hazards model for single spell duration data depend crucially on the proportional hazards assumption. Here, it is shown that variation in covariates over time, combined with variation across observations, is sufficient to ensure identification without the proportional hazards assumption. The required variation over time is minimal, and the mixed hazards model is identified without the proportional hazards assumption in the presence of standard time-varying covariates.

1. INTRODUCTION

The hazard function is central in modeling of econometric duration data. Economic theory often has more direct implications for the hazard function than for other descriptions of the duration distribution, as the hazard function may reflect decisions taken over time, conditional on past events. For example, in unemployment duration data, the hazard function may reflect decisions to accept or reject job offers appearing to unemployed workers in a random flow.

Inference about the hazard function using duration data is nontrivial in the presence of unobserved heterogeneity. Estimators that do not take heterogeneity explicitly into account tend to give a negative bias of the slope of the hazard function; see Heckman and Singer (1984a).

One common class of models taking unobserved heterogeneity explicitly into account is mixed hazards (MH) models. In MH models, unobserved heterogeneity is modeled in terms of individual specific independent random variables. Let V be such a random variable with a distribution function F, with support on the positive real line. Then, the conditional hazard function θ (conditional on $V = v$) is defined as

$$\theta(t;x,v) = v\lambda(t,x), \tag{1}$$

where λ is the structural hazard function, t is duration, and x is a vector of covariates.

Thanks to Kåre Bævre, Zhiyang Jia, Tore Schweder, Rolf Aaberge, and John K. Dagsvik for useful comments. Address correspondence to Christian N. Brinch, Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, P.O. Box 1050 Blindern, N-0316 Oslo, Norway; e-mail: c.n.brinch@bio.uio.no.

© 2007 Cambridge University Press 0266-4666/07 $12.00

https://doi.org/10.1017/S0266466607070144 Published online by Cambridge University Press

In the literature, MH models usually appear in the form of mixed proportional hazards (MPH) models, where it is also required that the structural hazard function factors into separate functions of duration and covariates. The MPH model was introduced in econometrics by Lancaster (1979), as a generalization of the proportional hazards model from Cox (1972). Lancaster and Nickel (1980) observed that without variation in x, it may be impossible to distinguish between different combinations of λ and F from a sample of durations. However, Elbers and Ridder (1982) and Heckman and Singer (1984b) showed under different conditions on F that variation over observations in the covariates is sufficient to identify the MPH model. For surveys of nonparametric identification results for duration models, see Heckman and Taber (1994) or van den Berg (2001).

From the perspective of applied work, there are two important problems with these results. First, the proportional hazards assumption does not usually have any firm theoretical basis (see van den Berg, 2001) and is empirically questionable, at least for applications to unemployment duration data (see McCall, 1994b). Second, the models are only identified under the assumption that the distribution function F has finite mean or under an assumption about the thickness of the tail of F, and these assumptions give rise to different identified models; see Ridder (1990).

One alternative to the proportional hazards assumption is to rely on semiparametric identification results, such as McCall (1996). Another alternative is to base identification on data with multiple spells per individual; see Honoré (1993). The result presented in the next section provides an alternative requirement on the data for identification of MH models. It is shown that in the presence of covariates that not only vary across observations but also over time within individual observations, the MH model is nonparametrically identified with single spell duration data.

2. THE IDENTIFICATION RESULT

Define a covariate path as a function $x : \mathbb{R}_+ \to X$. Let $S \subset (\mathbb{R}_+ \times X)$ be a family of such paths.

The role of a structure of the MH model is to specify the distribution of a nonnegative random variable T as a function on S. The survival function of T, $G : (\mathbb{R}_+ \times S) \to [0,1]$ is given by

$$G(t;x) = \phi(\Lambda(t,x)), \tag{2}$$

where ϕ is the Laplace transform of another nonnegative random variable V and

$$\Lambda(t;x) = \int_0^t \lambda(s;x(s))\, ds, \tag{3}$$

where $\lambda : (\mathbb{R}_+ \times X) \to \mathbb{R}_+$.

https://doi.org/10.1017/S0266466607070144 Published online by Cambridge University Press

A *structure of the MH model* is defined as a pair (ϕ, λ) that satisfies the preceding requirements. The *MH model* is defined as the set of all such admissible pairs. The *regular continuous time MH model* is defined as the subset of the MH model with continuous and strictly decreasing survival functions.

The MH model is *identified* if the structure of the MH model is uniquely determined by G or equivalently, if no two different structures imply the same survival function. The MH model is *identified up to a trivial normalization* if two structures $\{\gamma, \pi\}$ and $\{\phi, \lambda\}$ that imply the same survival function must satisfy $\pi(t,x) = \alpha\lambda(t,x)$, $\gamma(\alpha s) = \phi(s)$ for some $\alpha > 0$.

Identification depends on the domain of the survival function G, that is, the variation in covariate paths. The following condition on this domain will be sufficient for the identification result.

Condition 1. There are two elements $x_1 \in S$ and $x_2 \in S$, such that $x_1(t) = x_2(t)$ for $t \in (t_0, t_1)$, for some $t_0, t_1 \in \mathbb{R}_+$, with $t_0 < t_1$ and $G(t_0; x_1) \neq G(t_0; x_2)$.

THEOREM 1. *Under Condition 1, the regular continuous time MH model is identified up to a trivial normalization.*

Proof. Assume that $G(t_0;x_1) < G(t_0;x_2)$. It is clear that $\Lambda(t,x_1) = \Lambda(t,x_2) + C$ on (t_0,t_1), with $C > 0$. Fix C to an arbitrary value.

Define the function $g(s) = \phi(\phi^{-1}(s) + C)$ for $s \in (0,1]$. There is a one-to-one relationship between g and ϕ. That g is uniquely determined by ϕ is trivial. The fact that ϕ is uniquely determined by g requires a deeper argument.

A completely monotone function (such as a Laplace transform) is uniquely determined by its values at a sequence of points such that the sum of the inverses of the points diverges. This follows from the Muntz theorem; see Feller (1968) or the footnote at page 430 in Feller (1971). As $\phi^{-1}(1) = 0$, $\phi(C) = g(1)$. Further, $\phi(2C) = g(g(1))$, $\phi(3C) = g(g(g(1)))$, etc. Now, g uniquely determines ϕ at $C, 2C, \ldots$, and thus uniquely determines ϕ everywhere as $\sum_{j=1}^{\infty}(jC)^{-1}$ diverges.

For $t \in (t_0, t_1)$,

$$G(t;x_1) = \phi(\Lambda(t,x_1)) = \phi(\Lambda(t,x_2) + C)$$
$$= \phi(\phi^{-1}(G(t;x_2)) + C) = g(G(t;x_2)). \quad \textbf{(4)}$$

The term $G(t;x_2)$ takes on values on an open set as t ranges over (t_0,t_1). Thus, equation (4) holds with the argument of g taking on values on an open set. As g is analytic, g is then uniquely determined. Using the one-to-one relationship between g and ϕ, the Laplace transform ϕ is then uniquely determined.

With the Laplace transform uniquely determined, the integrated structural hazard function follows as

$$\Lambda(t,x) = \phi^{-1}(G(t;x)). \quad \textbf{(5)}$$

https://doi.org/10.1017/S0266466607070144 Published online by Cambridge University Press

Different choices of C lead to trivially different models. If $C = C_1$ gives the structure $\{\phi, \lambda\}$, and $C = C_2 = \alpha C_1$ for some $\alpha > 0$ gives $\{\gamma, \pi\}$, then $\pi(t, x(t)) = \alpha\lambda(t, x(t))$ for all $t > 0$, $x \in S$, and $\gamma(\alpha r) = \phi(r)$. ■

There are two main features to note from Theorem 1, as compared to the known identification results for MH models from Elbers and Ridder (1982), Heckman and Singer (1984b), and Ridder (1990). First, the proportional hazards assumption is not necessary for identification of the model. Second, no assumptions on the moments or the tail of the distribution of the random variable are necessary for identification. It should be noted that Heckman and Taber (1994) presented a result from a working paper by Honoré (1991), showing that the latter kind of assumption is dispensable with time-varying covariates. The result is proved in an MPH model, and the proof requires the proportional hazards assumption.

3. DISCUSSION

To get more intuition for the result given previously, it is worthwhile to study the unconditional hazard function, derived from the unconditional survival function,

$$\tau(t, x) = \frac{\phi'(\Lambda(t, x))}{\phi(\Lambda(t, x))}\lambda(t, x(t)), \tag{6}$$

where ϕ' denotes the derivative of the Laplace transform ϕ. The ratio in front of the structural hazard function expresses the conditional expectation of V, given survival to t. This expectation is completely determined by, and nondecreasing in, the survival probability.

The shape of ϕ is identified through the effect of survival probability on the unconditional hazard function. To separate the effects of covariates through the conditional expectation of heterogeneity and through the structural hazard function, we need an exclusion restriction: Some covariates that affect the survival probability to t (for some t) must be excluded from affecting the structural hazard rate at t. Such an exclusion restriction is implicit in time-varying covariates. As a simple example, unconditional hazard functions that are independent of past covariates (that affect survival probability) imply no unobserved heterogeneity.

Nonparametric identification in the MH model without the proportional hazards assumption has previously been studied by Heckman (1991) and McCall (1994a). Heckman (1991) shows that the structural hazard function is identified in the neighborhood of $t = 0$ as

$$\tau(t; x) = E(V\lambda(t; x); T \geq t; x) \sim \lambda(t; x), \tag{7}$$

https://doi.org/10.1017/S0266466607070144 Published online by Cambridge University Press

for $t = 0$, where $\sim$ means "is proportional to." The intuition is simple. At $t = 0$, there has been no selection of the sample, and the conditional expectation of V equals the unconditional expectation.

In McCall (1994a), the argument in Heckman (1991) is generalized to the case with time-varying covariates. The following argument is used: If (i) $x = \{x(t), t \geq 0\}$ is a stochastic jump process with random waiting times, such that the probability for a jump in any period of positive length is positive, and (ii) $\lambda(t;x) \to 0$ for some x, then

$$\tau(t;x) = E(V\lambda(t;x);T \geq t;x) \sim \lambda(t;x), \tag{8}$$

for all x such that $\lambda(s;x) \to 0$ for all $s < t$. Identification of the structural hazard function is thus based on the subset of covariate paths with survival probability approaching 1. It is only necessary to use this argument for small t, as the identification of $\lambda(t,x)$ in an arbitrarily small interval is sufficient for full identification of the model through the analyticity of the Laplace transform.

It is useful to highlight the differences from Theorem 1. In McCall (1994a) covariates are required to change (with some probability) along the continuous time scale. In Theorem 1, covariates are required to change only once. In McCall (1994a), covariates must affect the structural hazard rate such that the structural hazard rate approaches zero as the covariates approach some limit. In Theorem 1, the only requirement is that the covariates take on at least two values that give different structural hazard rates.

Instead of explicitly specifying a covariate process, Theorem 1 is given with a condition on the covariate process involved. The crucial feature is that there is, for some time t, variation in $x(s)$ for some $s < t$ over observations with the same $x(t)$. This excludes three types of covariate processes: constant processes (such as gender), covariate processes that do not vary over the population but vary only over time (such as aggregate unemployment rates), and covariate processes where past covariate values are functions of current covariate values (such as age). In addition, covariate processes that do not affect the survival probabilities will not satisfy Condition 1.

The minimalist covariate process that will satisfy Condition 1 is a process taking on two values, constant until a jump time t, when some, but not all, observations change value. This process illustrates how small the required variation actually is. More variation will not lead to problems of identification. Identification, in consistency with the literature, is defined as whether the structure of the model is derivable from the unconditional survival function (unconditional on unobserved heterogeneity, conditional on covariates). Thus, whether the unconditional survival function is identified from data is not considered.

With this in mind, it is hard to come up with covariate processes varying over time and observations that do not satisfy Condition 1. I will mention two examples of time-varying covariates from unemployment duration analysis that could satisfy Condition 1. First, for reduced form analysis, indicators of local

https://doi.org/10.1017/S0266466607070144 Published online by Cambridge University Press

labor market conditions will typically vary over time and individuals, assuming data for more than one locality. Second, more structural analyses could use the present value of future unemployment benefits as a covariate. Such a variable would vary over time and individuals as required if, for example, individuals with different benefit entitlements approach an expiry date for the benefits.

Nonparametric identification results play an important role as a heuristic tool for model specification in complex models of duration data. The absence of nonparametric identification implies that parametric specification, which is often ad hoc, plays a crucial role in inference. With variation over time in covariates, the proportional hazards assumption is not crucial for separating structural duration dependence and unobserved heterogeneity.

REFERENCES

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

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

Feller, W. (1968) On Muntz' theorem and completely monotone functions. *American Mathematics Monthly* 75, 342–359.

Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*, vol. 2. Wiley.

Heckman, J.J. (1991) Identifying the hand of past: Distinguishing state dependence from heterogeneity. *American Economic Review* 81(2), 75–79.

Heckman, J.J. & B. Singer (1984a) Econometric duration analysis. *Journal of Econometrics* 24, 63–132.

Heckman, J.J. & B. Singer (1984b) The identifiability of the proportional hazards model. *Review of Economic Studies* 51, 231–241.

Heckman, J.J. & C. Taber (1994) Econometric mixture models and more general models for unobservables in duration analysis. *Statistical Methods in Medical Research* 3, 279–299.

Honoré, B. (1991) Identification results for duration models with multiple spells or time-varying covariates. Working paper, Northwestern University.

Honoré, B. (1993) Identification results for duration models with multiple spells. *Review of Economic Studies* 60, 241–246.

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

Lancaster, T. & S. Nickel (1980) The analysis of reemployment probabilities for the unemployed. *Journal of the Royal Statistical Society, Series A* 143, 141–165.

McCall, B.P. (1994a) Identifying state dependence in durations models. In *American Statistical Association 1994, Proceedings of the Business and Economics Section*, 14–17. American Statistical Association.

McCall, B.P. (1994b) Testing the proportional hazards assumption in the presence of unmeasured heterogeneity. *Journal of Applied Econometrics* 9, 321–334.

McCall, B.P. (1996) The identifiability of the mixed proportional hazards model with time-varying coefficients. *Econometric Theory* 12, 733–738.

Ridder, G. (1990) The non-parametric identification of generalized accelerated failure time models. *Review of Economic Studies* 57, 167–182.

van den Berg, G. (2001) Duration models: Specification, identification, and multiple durations. In J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, vol. 5. North-Holland.

https://doi.org/10.1017/S0266466607070144 Published online by Cambridge University Press