

MEMORANDUM

No 20/2000

**Identification of Structural Duration Dependence and Unobserved
Heterogeneity with Time-varying Covariates**

By
Christian Brinch

ISSN: 0801-1117

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: <http://www.sv.uio.no/sosoek/>
e-mail: econdep@econ.uio.no

In co-operation with
**The Frisch Centre for Economic
Research**

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: <http://www.frisch.uio.no/>
e-mail: frisch@frisch.uio.no

List of the last 10 Memoranda:

| | |
|-------|---|
| No 19 | By Knut Røed and Morten Nordberg: Have the Relative Employment Prospects for the Low-Skilled Deteriorated After All? 21 p. |
| No 18 | By Jon Vislie: Environmental Regulation under Asymmetric Information with Type-dependent outside Option. 20 p. |
| No 17 | By Tore Nilssen and Lars Sørgard: Strategic Informative Advertising in a TV-Advertising Duopoly. 21 p. |
| No 16 | By Michael Hoel and Perry Shapiro: Transboundary Environmental Problems with a Mobile Population: Is There a Need for Central Policy? 19 p. |
| No 15 | By Knut Røed and Tao Zhang: Labour Market Transitions and Economic Incentives. 21 p. |
| No 14 | By Dagfinn Rime: Private or Public Information in Foreign Exchange Markets? An Empirical Analysis. 50 p. |
| No 13 | By Erik Hernæs and Steinar Strøm: Family Labour Supply when the Husband is Eligible for Early Retirement. 42 p. |
| No 12 | By Erik Bjørn: The rate of capital retirement: How is it related to the form of the survival function and the investment growth path? 36 p. |
| No 11 | By Geir B. Asheim and Wolfgang Buchholz: The Hartwick rule: Myths and facts. 33 p. |
| No 10 | By Tore Nilssen: Risk Externalities in Payments Oligopoly. 31p. |

A complete list of this memo-series is available in a PDF® format at:
<http://www.sv.uio.no/sosoek/memo/>

Identification of Structural Duration Dependence and Unobserved Heterogeneity with Time-varying Covariates*

Christian Brinch

Department of Economics, University of Oslo[†]

July 13, 2000

Abstract

Known results on the identification of structural duration dependence in the presence of unobserved heterogeneity depend crucially on the proportional hazards assumption. Here, I show 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.

JEL Classification: C41.

*Thanks to John K. Dagsvik and Kåre Bævre for valuable comments that have substantially improved the paper.

[†]P.O. Box 1095 Blindern, N - 0317 Oslo, Norway, telephone: (47) 22 85 51 59, fax: (47) 22 85 50 35, E-mail: c.n.brinch@econ.uio.no.

1 Introduction

In the analysis of duration data, the econometrician is usually interested in some behavioral structure governing the distributions of duration spells. Such behavioral structure often takes the form of optimal stopping decisions in stochastic frameworks. Durations then follow from sequences of decisions of whether to end e.g. a strike “now” or whether to continue striking. With a behavioral structure of this sort, it is a fruitful and common approach to represent the distributions of durations through the hazard rate, which may be interpreted as proportional to the probability of a spell to end in a short interval following t , given a duration of at least t . There is then a direct link between the hazard rate and the probability of ending a spell from the perspective of an individual economic agent.

Duration dependence is defined as dependence of the hazard rate on the elapsed duration. In economic applications, such as in the study of unemployment durations, duration dependence will usually be of both theoretical interest and of more direct practical interest to policymakers. Here, I will consider the problem of recovering duration dependence from duration data. This is a well known and far from trivial problem, a general discussion is given in Heckman (1991), and a more comprehensive survey is given in Heckman and Taber (1994). The main obstacle to recovering duration dependence is heterogeneity in the underlying distributions.

A decrease over time in the relative frequency of durations ending at time

t , given survival to at least t , may be given several interpretations. Such sample duration dependence may be due to duration dependence in the underlying distributions, or “structural duration dependence.” However, sample duration dependence may also be due to heterogeneity in the sample. Units with low hazard rates tend to have longer durations than units with high hazard rates. Consequently, units with low hazard rates constitute an increasing proportion of the sample over time, causing the average hazard rate in the sample to decrease over time. In the study of unemployment durations, one will usually find decreasing hazard rates over time in the sample. Thus, one cannot in general conclude whether this is because long term unemployment cause low hazard rates or because individuals with low hazard rates become long term unemployed?

While it may be possible to control for some heterogeneity through observed covariates, there are usually still reasons to expect some sample heterogeneity to remain unobserved. My purpose here is to show how variation in covariates over time may be used to relax some of the assumptions that has until now been considered necessary for identification of structural duration dependence in the presence of unobserved heterogeneity. The following section provides the necessary background for describing the results.

2 Preliminaries

The hazard rate, $\theta(t)$, associated with a random variable T with support on the positive scale is defined as

$$\theta(t) = \lim_{\Delta t \rightarrow 0^+} \frac{\Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \quad (1)$$

The hazard rate is an exhaustive description of the distribution of the random variable in question. The hazard rate is often specified in some way conditional on a vector of covariates x , e.g. as in the proportional hazards model in Cox (1972), where the hazard rate as a function of x is specified as

$$\theta(t; x) = \psi(t) \phi(x), \quad (2)$$

where $\psi(t)$ and $\phi(x)$ are functions taking values on the positive real line. To the econometrician, elements of the vector x may be unobserved. Lancaster (1979) extended the proportional hazards model to take into account the possibility of such unobserved heterogeneity by the mixed proportional hazard model (MPH),

$$\theta(t; x, v) = v\psi(t) \phi(x), \quad (3)$$

where v is a random variable with an unknown distribution function $F(v)$, with support on the positive real line, interpreted as an analogue to the residual term

in standard regression analysis.

Lancaster and Nickel (1980) observed that without variation in x , it is impossible to distinguish between different combinations of ψ and F from a sample of durations. However, Elbers and Ridder (1982) and Heckman and Singer (1984) showed under different regularity conditions that variation over observations in the covariates is sufficient to identify the MPH model. Ridder (1990) and Heckman and Honoré (1989), in a slightly different context, have also shown that this holds for the larger class of generalized accelerated failure-time (GAFT) models. The problem with these results for applied work is that all of them are crucially dependent on the proportional hazards assumption, for which there is, in general, no theoretical basis. As the proportional hazards assumption has been considered necessary for identification, it has not in general been testable. However, see McCall (1994b) for a test of the proportional hazards assumption within an identified model that nests versions of the MPH model.

My purpose here is to demonstrate how variation over time (within durations) in observed covariates, combined with variation in covariates across observations, may help us identify structural duration dependence in the presence of unobserved heterogeneity without invoking the proportional hazards assumption. In particular, I demonstrate how nonparametric identification is partially achieved in a GAFT-like model and how full nonparametric identification is achieved in a MPH-like model without assuming proportional hazards.

3 The GGAF T class and the MH model

The class of models that I will call generalized GAFT (GGAF T) can be described in the following way. Denote the “structural hazard function” by $\lambda(t; x) > 0$. The “integrated structural hazard function” is now given by

$$\Lambda(t; x) = \int_0^t \lambda(s; x) dx. \quad (4)$$

In a model without unobserved heterogeneity, the survival function of T is given by

$$G(t; x) = \exp(-\Lambda(t; x)). \quad (5)$$

This is not generally the case in the presence of unobserved heterogeneity. The GGAF T model involves a generalization of the equation above. The survival function may now be any continuously differentiable strictly decreasing function, K , in the “integrated structural hazard function,”

$$G(t; x) = K(\Lambda(t; x)), \quad (6)$$

such that $K(0) = 1$ and $K(\infty) = 0$. However, without further specification, the argument of K does not in general have the interpretation of an integrated structural hazard function. Models in the GGAF T class belong to the class of

GAFIT models if there exist functions $\Psi(t)$ and $\phi(x)$ that satisfy

$$\Lambda(t; x) = \Psi(t) \phi(x) \tag{7}$$

for all x and $t \geq 0$. Models in the GAFIT class are under mild regularity conditions identified, see Ridder (1990). This is not in general the case for models in the wider GGAFIT class.

I will further consider a particular model in the GGAFIT class, the Mixed Hazard (MH) model, discussed in Lancaster (1990), Heckman (1991) and McCall (1994a). The hazard rate of the MH model is specified as

$$\theta(t; x, v) = v\lambda(t, x), \tag{8}$$

$\lambda(t, x)$ is any positive function and v is a random term with a distribution function $F(v)$ with support on the positive real line. The MH model is obviously a generalization of the MPH model. The survival function of the model is given by

$$G(t; x) = \mathcal{L}(\Lambda(t; x)), \tag{9}$$

where \mathcal{L} is the Laplace transform of F . The MH model provides interpretability to the GGAFIT class. In the GGAFIT class, the “integrated structural hazard rate” is in general not interpretable, while in the MH model, it is interpretable as an integrated structural hazard rate. The MH model also provides restrictions

in that \mathcal{L} must be the Laplace transform of a distribution function.

The introduction of time-varying covariates in these models is straightforward.

Just substitute (4) with

$$\Lambda(t; \tilde{x}) = \int_0^t \lambda(s; x(s)) ds, \quad (10)$$

where $\tilde{x} = \{x(t), t \geq 0\}$. Regularity conditions are required as the covariate processes must be exogenous to the durations they are supposed to explain, see Yashin and Arjas (1988).

In the identification results below, I have chosen a particular specification of the covariate process. The covariates are specified as a jump process such that a single value is realized in the interval $t \in [0, t_1)$, where t_1 is a constant, and then another single value is realized in the interval $t \in [t_1, \infty)$. This specification is chosen to emphasize how little variation over time is needed for the results. The results can easily be generalized to include further jumps after t_1 . It also seems straightforward to extend the results to covariate jump processes with random waiting times.

4 Identification results

In this section two identification results will be provided. Both identification results show how variation over time in covariates may substitute for the propor-

tional hazards assumption, in the sense that the models considered would not be identified in the absence of time-varying covariates, but would be identified in the absence of time-varying covariates if we were willing to assume proportional hazards. Theorem 1 considers identification in the GGAFIT class.

For the discussion of identification, it is as usual assumed that we observe an unlimited number of (T, \tilde{x}) , such that $G(t; \tilde{x})$ can be considered observable. The identification problem then amounts to recovering K and Λ from G .

Assumption 1: *Let $\tilde{x} = \{x(t), t > 0\}$ be a covariate process such that $x(t) = x_1$ for $t \in [0, t_1)$ and $x(t) = x_2$ for $t \in [t_1, \infty)$, $t_1 > 0$. (x_1, x_2) is drawn from a probability distribution that has a density on an open set $S^2 = S \times S \subset \mathbb{R}^2$.*

Assumption 2: *$\lambda(s; x) > 0$ for all $(t, x) \in \mathbb{R}_+ \times S$, and is continuously differentiable and non-constant in x .*

Assumption 3: *(GGAFIT class) The durations T are generated by some distribution function such that the conditional survival function of T is given by*

$$G(t; \tilde{x}) = K(\Lambda(t; \tilde{x})), \quad (11)$$

where K is continuously differentiable and strictly decreasing with $K(0) = 1$, $K(\infty) = 0$ and

$$\Lambda(t; \tilde{x}) = \int_0^t \lambda(s; x) dx. \quad (12)$$

Theorem 1 Let $A = \{(t, \tilde{x}) : \tilde{x} \in S^2, t \in R_+, G(t, \tilde{x}) \leq \max_{x_1 \in S} G(t_1, \tilde{x})\}$. Under assumptions 1, 2 and 3, $\lambda(t; x)$ is identified up to a scale parameter $C_1 > 0$ for $(t, \tilde{x}) \in A$. Λ is identified up to a location parameter $C_0 > 0$ and a scale parameter C_1 for $(t, \tilde{x}) \in A$. K is identified up to a linear transformation of the argument for $(t, \tilde{x}) \in A$.

Proof. For any $(x_1, x_2) \in S^2$ and $t \geq t_1$,

$$\frac{\partial G(t; \tilde{x}) / \partial t}{\partial G(t; \tilde{x}) / \partial x_1} = \frac{\lambda(t, x_2)}{\int_0^{t_1} \lambda'_2(s, x_1) ds}. \quad (13)$$

By normalizing the denominator in (13) for a particular value of x_1 , we can trace out $\lambda(t; x)$ for $t \geq t_1$, by varying x_2 and t . Thus, $\lambda(t, x)$ is identified up to a scale parameter for $t \geq t_1$.

Let $x_1^0 = \arg \max_{x_1 \in S} G(\Lambda(t, \tilde{x}))$. For any (t', \tilde{x}') and (t'', \tilde{x}'') such that $t' \geq 0$, $t'' \geq t_1$, $\tilde{x}' \in S^2$, $\tilde{x}'' \in S^2$, $x_1'' = x_1^0$ and $G(\Lambda(t', \tilde{x}')) = G(\Lambda(t'', \tilde{x}''))$, we obtain

$$\frac{\partial G(t', \tilde{x}') / \partial t}{\partial G(t'', \tilde{x}'') / \partial t} = \frac{\lambda(t', x'(t'))}{\lambda(t'', x''_2)}. \quad (14)$$

Since $t'' \geq t_1$, the denominator is identified subject to a normalization of the scale. Thus, the numerator is also identified for all $(t', x'(t'))$ such that the requirements for observing this ratio can be satisfied. The requirements can be

satisfied for any (t', \tilde{x}') such that $G(t', \tilde{x}') \leq \max_{x_1 \in S} G(t_1, \tilde{x})$, that is, for all $(t', \tilde{x}') \in A$. Denote the identified version of $\lambda(t; x)$ by $\lambda^*(t; x)$. Thus

$$\lambda(t; x) = C_1 \lambda^*(t; x), \quad (15)$$

where C_1 is a positive scale parameter.

Define a function $t_c(x_1)$, $x_1 \in S$, implicitly by the equation

$$G(t_c(x_1); x_1) = G(t_1, x_1^0). \quad (16)$$

This is possible since G is strictly decreasing in its first argument. Let $C_0 = \Lambda(t_1; x_1^0)$. Now, we can write

$$\Lambda(t; \tilde{x}) = C_0 + C_1 \int_{t_c(x_1)}^t \lambda^*(s, x(s)) ds, \quad (t, \tilde{x}) \in A.$$

Given a choice of C_0 and C_1 , K is trivially identified from (11) for arguments in the part of the domain of K where Λ is identified. ■

I will now proceed with Theorem 2 before commenting on the results. Theorem 2 considers identification in the MH model.

Assumption 3': (*Mixed Hazard model*) *The durations T are generated from*

a conditional probability distribution characterized by a hazard function given by

$$\theta(t; \tilde{x}, v) = v\lambda(t, x(t)), \quad (17)$$

where v is a random term, independent of \tilde{x} , with an unknown distribution function $F(v)$ with support on the positive scale.

Assumption 4: $E(v) = 1$.

The last assumption serves to identify the scale parameter in the structural hazard function.

Theorem 2 *Under assumptions 1, 2, 3' and 4, the functions $\lambda(t; x)$ and $F(v)$ are fully identified from $G(t; \tilde{x})$.*

Proof. In the mixed hazards model, the function K given in (11) is the Laplace transform of F , and Assumption 3' implies that Assumption 3 is satisfied. We can then use the identification results from Theorem 1 directly.

Standard results on Laplace transforms, see Feller (1966), include their analyticity. In other words, the Laplace transform \mathcal{L} possesses derivatives of all orders. From Theorem 1 above, the Laplace transform of F is identified (up to a scale and location parameter on the argument) in an open interval. From analyticity, the Laplace transform is then identified in its full domain. The scale and location parameters are identified by $\mathcal{L}(0) = 1$, which is a requirement on

the Laplace transform of a distribution function and $\mathcal{L}'(0) = -1$, which follows from Assumption 4.

When the Laplace transform of $F(v)$ is identified, identification of $F(v)$ and $\lambda(t; x)$ is trivial. ■

Note that some of the Assumptions can easily be relaxed with respect to Theorem 2, as the only results we require from Theorem 1 is the identification of the structural hazard rate in any arbitrarily small open interval. In particular, the requirements on the joint distribution of (x_1, x_2) in Assumption 1 and the differentiability and nonconstantness of $\lambda(t, x)$ in assumption 2 need only be satisfied in an arbitrarily small open subset of S^2 .

Even though identification is only partially achieved for the GGAFT class, Theorem 1 is fairly strong in that only a minimal variation over time is required. It is instructive to see how variation over time in covariates substitutes for the proportional hazards assumption in the GAFT model. In models in the GGAFT class without variation in covariates over time, the analogue of (13) is

$$\frac{\partial G(t; \tilde{x}) / \partial t}{\partial G(t; \tilde{x}) / \partial x} = \frac{\lambda(t, x)}{\int_0^t \lambda_2'(s, x) ds}, \quad (18)$$

where both the numerator and the denominator depends on both t and x , so that the ratio will not be helpful for identification purposes. In the case of the GAFT

class, this equation looks like

$$\frac{\partial G(t; \tilde{x}) / \partial t}{\partial G(t; \tilde{x}) / \partial x} = \frac{\psi(t) \phi(x)}{\int_0^t \psi(s) ds \phi'(x)}, \quad (19)$$

and by appropriate normalizations, identification of λ and ϕ is straightforward.

I will provide further interpretation on how identification is achieved through variation over time in the context of the MH model. Consider the “observed hazard function,” the relative frequency of spells with duration at least t that end at t ,

$$\tilde{\theta}(t, x) = \frac{\partial G(t; \tilde{x}) / \partial t}{G(t; \tilde{x})} = E(v | T \geq t, \tilde{x}) \lambda(t; x(t)). \quad (20)$$

For $t > t_1$, there are observed covariates, x_1 , that only affect the observed hazard rate through the conditional expectation of v . Since

$$E(v | T \geq t, \tilde{x}) = \frac{\mathcal{L}'(\Lambda(t; \tilde{x}))}{\mathcal{L}(\Lambda(t; \tilde{x}))}, \quad (21)$$

there is a one-to-one relationship between the integrated structural hazard rate, Λ , at t and the conditional expectation of v at t . There is of course also a one-to-one relationship between the integrated structural hazard rate at t and the observed survival probability to t . By varying t in (20), both the conditional expectation of v and the structural hazard rate $\lambda(t, x)$ may change. However, by varying x_1 appropriately along with t , we can keep the probability of survival

to t and thus the conditional expectation of v constant, and find the changes in $\tilde{\theta}(t, \tilde{x})$ due to changes in the structural hazard rate. Equation (20) also suggests a simple way of testing for unobserved heterogeneity by testing for effects of past covariates in the “observed hazard function.”

The identification result presented here may be viewed as an analogue to the instrumental variable approach. By assuming that past realizations of the covariates do not directly affect the structural hazard rate, we can use the observed relationship between past covariates and the present “observed hazard rate” to deduce the effects of the past on the present.

As the observant reader will note, this interpretation only covers the part of the identification results that are given in Theorem 1. It seems much harder to give a simple interpretation of the identification results in Theorem 2, as these are due to the analyticity of the Laplace transform. A suggestion may be that the analyticity property of the survival function provides us with a theoretical foundation for extrapolation of the survival function. However, any number of derivatives may still be necessary for a sufficiently close Taylor approximation.

Even though Theorem 2 may seem to of most direct relevance to applied duration analysis, Theorem 1 is also of some importance. First, Theorem 1 provides identifiability for a wider class of models, and as such provides us with testable restrictions for the MH model. Secondly, and more important, Theorem 1 teaches us that identification of structural duration dependence in the MH model is, in a potentially large region, identified independently of our ability to identify

the distribution of unobserved heterogeneity. This is important for practical purposes, as we will rarely be able to find good estimates on the distribution of unobserved heterogeneity even if it is identified.

Identification in the mixed hazards model has previously been studied by Heckman (1991) and McCall (1994a). In Heckman (1991) it is showed that in the MH model without time-varying covariates, the structural hazard function is identified for $t = 0$, as

$$\tilde{\theta}(t; x) = E(v\lambda(t; x); T \geq t; x) = \lambda(t; x), \quad (22)$$

for $t = 0$.

In McCall (1994a) the result above is generalized to the case with time-varying covariates. His results rely on the following argument: If (i) $\tilde{x} = \{x(t), t \geq 0\}$ 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) $\lambda(t; x) \rightarrow 0$ for some x , then

$$\tilde{\theta}(t; \tilde{x}) = E(v\lambda(t; x); T \geq t; \tilde{x}) = \lambda(t; x),$$

for all \tilde{x} such that $\lambda(s; x) \rightarrow 0$ for all $s < t$. It is only necessary to use this argument for very small t , as the identification of $\lambda(t, x)$ in an arbitrarily small interval is sufficient for full identification through the analyticity of the Laplace

transform, as in Theorem 2 above. It should also be noted that McCall (1996) has provided further identification results for a particular version of the MH model in absence of time-varying covariates.

The results provided in this paper are stronger than the results in McCall (1994a) in the sense that less variation over time is required and that λ is not required to approach zero for any x , though the results are not nested as McCall (1994a) does not require differentiability of the function λ . The results provided here also show that the identification of structural duration dependence does not rely on identification of the distribution of unobserved heterogeneity in a potentially large region.

5 Concluding discussion

The results here establish how variation over time in covariates may be used to relax the proportional hazards assumption that pervades identification results in the literature on structural duration dependence and unobserved heterogeneity. This is important for applied research as the proportional hazards assumption is usually introduced without theoretical support while it is crucial to the results obtained. The results provided here should have clear implications. For decomposition of sample duration dependence into structural duration dependence and unobserved heterogeneity, relevant time-varying covariates are crucial as they allow us to relax the proportional hazards assumption.

References

- [1] 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.
- [2] Feller, W., 1971, *An Introduction to Probability Theory and Its Applications*, Vol. II, New York: John Wiley.
- [3] Heckman, J. J., 1991, "Identifying the Hand of Past: Distinguishing State Dependence from Heterogeneity", *American Economic Review*, May 1991, 75-79.
- [4] Heckman, J. J. and B. Honore, 1989, "The Identifiability of The Competing Risks Model", *Biometrika*, 76, 325-330.
- [5] Heckman, J. J. and B. Singer, 1984, "The Identifiability of the Proportional Hazards Model", *Review of Economic Studies* 51, 231-241.
- [6] Heckman, J. J. and C. Taber, 1994, "Econometric Mixture Models and More General Models for Unobservables in Duration Analysis", NBER Technical working paper no.157.
- [7] Lancaster, T., 1979, "Econometric methods for the duration of unemployment.", *Econometrica* 47(4): 939-956.

- [8] Lancaster, T., 1990, *The Analysis of Transition Data*. New York: Cambridge University Press.
- [9] Lancaster, T. and S. Nickel, 1980, "The analysis of reemployment probabilities for the unemployed", *Journal of the Royal Statistical Society A*, 143, 2:141-165.
- [10] McCall, B. P., 1994a, "Identifying State Dependence in Durations Models," *American Statistical Association 1994, Proceedings of the Business and Economics Section*, 14-17.
- [11] McCall, B. P., 1994b, "Testing the Proportional Hazards Assumption in the Presence of Unmeasured Heterogeneity", *Journal of Applied Econometrics*, 9, 321-334.
- [12] McCall, B. P., 1996, "The Identifiability of the Mixed Proportional Hazards Model with Time-varying Coefficients", *Econometric Theory*, 12, 733-738.
- [13] Ridder, G., 1990, "The Non-Parametric Identification of Generalized Accelerated Failure Time Models", *Review of Economic Studies*, 57, 167-182.
- [14] Yashin, A. I. and A. Arjas, 1988, "A Note on Random Intensities and Conditional Survivor Functions", *Journal of Applied Probability*, 25, 630-635.