Semiparametric autoregressive conditional proportional hazard models

Frank Gerhard [a,*] and Nikolaus Hautsch [b]

[a] Nuffield College, Oxford

[b] Center of Finance and Econometrics, University of Konstanz

December 18, 2001

Abstract

A new semiparametric proportional hazard rate model is proposed which extends standard models to include a dynamic specification. Two main problems are resolved in the course of this paper. First, the partial likelihood approach to estimate the components of a standard proportional hazard model is not available in a dynamic model involving lags of the log integrated baseline hazard. We use a discretisation approach to obtain a semiparametric estimate of the baseline hazard. Second, the log integrated baseline hazard is not observed directly, but only through a threshold function. We employ a special type of observation driven dynamic which allows for a computationally simple maximum likelihood estimation. This specifications approximates a standard ARMA model in the log integrated baseline hazard and is identical if the baseline hazard is known.

It is shown that this estimator is quite flexible and easily extended to include unobserved heterogeneity, censoring and state dependent hazard rates. A Monte Carlo study on the approximation quality of the model and an empirical study on BUND future trading at the former DTB complement the paper.

JEL classification: C22, C25, C41, G14

Keywords: autoregressive duration models, dynamic ordered response models, generalised residuals, censoring

*Corresponding author: Tel. ++44-1865-2-78692, fax ++44-1865-2-78621, email frank.gerhard@nuffield.oxford.ac.uk.

1 Introduction

In this paper, we propose a dynamic extension to semiparametric proportional hazard models. This can be seen as an alternative to the recently proposed ACD model by Engle and Russell (1998) and Engle (2000) that captures the dynamic structure of durations in the context of special accelerated failure time models and the approach of Gamerman and West (1987) building on the dynamic generalized linear model framework suggested by West, Harrison, and Migon (1985).[1] Our approach is a special type of modulated renewal process in the sense of Cox (1972b) and thereby a special type of proportional hazard model. The latter class of models is widely employed in microeconometrics when the waiting times of individuals, households or firms in particular states is of primary interest.[2] Hence, in this sense, the proposed approach is particularly useful to analyze data which have typical duration characteristics, like censoring structures or unobservable heterogeneity effects but which also reveal serial dependencies. The model dynamic is specified based on an observation driven autoregressive process for the log integrated baseline hazard and we call it semiparametric autoregressive conditional proportional hazard (ACPH) model.

Since this specification involves dynamics in terms of the log integrated baseline hazard, the partial likelihood approach proposed by Cox (1972a), (1975) and Oakes and Cui (1994) leading to a two step estimation of covariates and the baseline hazard function,

[1]For a general exposition of durations models, see e.g. the recent survey by Oakes (2001) or Cox and Oakes (1984), Kalbfleisch and Prentice (1980) or Lancaster (1997).

[2]A well known example is the analysis of the length of unemployment spells which is studied by a wide range of theoretical and empirical papers, see e.g. van den Berg and van der Klaauw (2001), McCall (1996), Gritz (1993), Meyer (1990), Moffitt (1985), Heckmann and Singer (1984), Nickell (1979), Lancaster (1979) among many others.

is not available. A further problem is that the log integrated baseline hazard is not observable directly, and thus the usual prediction error decomposition, see e.g. Harvey (1990, chapter 3.5), is not applicable.

To circumvent these problems, we use a discretisation approach based on the work of Han and Hausman (1990). A categorisation of the durations allows us to estimate the autoregressive parameters and discrete points of the baseline hazard function simultaneously. Instead of a parameter driven dynamic model, in the sense of Cox (1981), which would involve a computationally cumbersome maximum likelihood (ML) estimation, we propose an observation driven model. For this specification computationally simple ML estimators of the dynamic parameters and of the underlying distribution function at a finite number of points are directly available. Given that there are many ways to specify an observation driven dynamic, we specify a model which approximates the properties, especially the autocorrelation function, of a corresponding parameter driven model. This approximation is better, the finer the chosen categorization. It is shown that this model is equivalent to a simple ARMA model for the log integrated baseline hazard, if the baseline hazard is known.

Subsequently, we illustrate that higher order AR(p) and MA(q) terms are easily incorporated in the model, yielding a semiparametric ACPH(p,q) model. Moreover, we show that further extensions, like the inclusion of exogenous regressors, the consideration of censoring mechanisms, unobserved heterogeneity or state dependent hazard rates are easily accommodated in the model framework.

Some small sample evidence on the quality of this approximation is provided on the basis of a Monte Carlo study for the semiparametric ACPH(1,0) and ACPH(0,1) model. Results indicate that the approximation works nicely for a reasonable number of categories and a moderate sample size. Finally, we provide a brief demonstration of the estimator in the context BUND future trading at the former DTB in Frankfurt.

The outline of the rest of the paper is as follows: in section 2 we propose the semiparametric ACPH model. Section 3 outlines some possible extensions. Small sample evidence on the approximation quality is provided in section 4. Section 5 reports on an application to BUND future trading and section 6 concludes.

2 Semiparametric autoregressive conditional proportional hazard models

2.1 Semiparametric proportional hazard models

Consider a sequence of events taking place at times $\vartheta_0, \vartheta_1, \ldots, \vartheta_T$ with $\vartheta_0 = 0$ and $\vartheta_0 < \vartheta_1 < \ldots < \vartheta_T$ being a stochastic process. Associated with this process of arrival times is a process of the durations between individual events, defined as the first difference of the original process, i.e. $\tau_t := \vartheta_t - \vartheta_{t-1}$, $t = 1, \ldots, T$. Denote the information generated by this process and potential covariates x_t as $\mathcal{F}_t = \sigma(\tau_t, \tau_{t-1}, \ldots, \tau_1, x_{t+1}, x_t, \ldots, x_1)$[3].

To capture the information contained in the process of durations τ_t, one way to proceed is to specify a stochastic model based on the hazard rate of τ_t, following the approach suggested by Cox (1972a), (1975).

The hazard rate is defined as the limit of the conditional probability to observe an event in a small time interval, given that the event has not occurred until the beginning

[3]The inclusion of x_{t+1} into the information set $\mathcal{F}_t$ simplifies the notation in the rest of the paper since it allows us to condition on past durations and contemporaneous covariates. This notation will be used in section 2.4.

of the small interval, i.e.

$$\lambda_\tau(s) := \lim_{\Delta \to 0} \frac{\text{Prob}\left[s \leq \tau < s + \Delta \,|\tau \geq s\right]}{\Delta} \tag{1}$$

The relationship to the distribution function Q_τ of τ is simply

$$\lambda_\tau(s) = -\frac{\partial \log(1 - Q_\tau(s))}{\partial s}. \tag{2}$$

In order to include explanatory variables x_t in a cross-sectional context, frequently a proportional hazard model is specified as the product of a baseline hazard $\lambda_0(s) > 0$ and a strictly positive function of x_t with coefficients β as

$$\lambda_\tau(s|x_t) = \lambda_0(s) \exp(-x_t'\beta), \qquad t = 1, \ldots, T. \tag{3}$$

The baseline hazard $\lambda_0(s)$ corresponds to $\lambda_\tau(s|x_t = 0)$, i.e. if the regressors are centered, $\lambda_0(s)$ has an interpretation as the hazard function for the mean values of x_t. Note that the information generated by past covariates x_s, $s < t$, and past durations does not enter this specification.

Much insight can be gained, if one transforms the dependent variable τ_t to another random variable $\tilde{\tau}_t$ which has a flat, i.e. constant, hazard rate, implying that the time since the start of the spell is uninformative. This is achieved by taking the integral of (3) and using the fact that this integral is a random variable described by a unit exponential distribution, i.e. $\int_0^{\tau_t} \lambda_\tau(s|x_t)ds \sim Exp(1)$, see e.g. Lancaster (1997, p. 19),

$$\tilde{\tau}_t = \exp(x_t'\beta)\xi_t, \quad \xi_t \sim Exp(1), \quad \text{with } \tilde{\tau}_t := \int_0^{\tau_t} \lambda_0(s)ds. \tag{4}$$

Note that the exponential distribution implies indeed a flat hazard rate, which is easily verified from (2). Taking the log of (4) yields a model linear in the transformed duration $\tau_t^* = \ln \tilde{\tau}_t$,

$$\tau_t^* = x_t'\beta + \epsilon_t^*, \tag{5}$$

where ϵ_t^* is an identically independently type II extreme value distributed error term with mean $\varsigma \equiv \mathrm{E}[\epsilon^*] = -0.577216$, variance $\mathrm{Var}[\epsilon^*] = \pi^2/6$ (see e.g. Johnson, Kotz, and Balakrishnan (1994)), and density function

$$f(s) = \exp(s - \exp(s)). \tag{6}$$

Due to the one-to-one nature of the transformation applied to the durations τ_t, the information sets generated by either dependent variable τ_t or τ_t^* are equal, i.e. $\mathcal{F}_t = \mathcal{F}_t^* \equiv \sigma(\tau_t^*, \tau_{t-1}^*, \ldots, \tau_1^*, x_{t+1}, x_t, \ldots, x_1)$. This rather peculiar setup, is chosen since the distribution of the durations τ_t contains valuable information on the point process generating the observations, beyond the marginal effects of regressors x_t. The latter is quite different from a standard linear regression model where one is content with a model having i.i.d. errors, which implies that no information is missed by the model. In the context of point processes however, even an i.i.d. error still contains valuable information on the DGP in the form of the slope of the hazard rate. Thus, a good regression model should imply not only an i.i.d. error structure but a flat hazard rate for the error term of the transformed model (4) as well.

By having made the transition from τ_t to τ_t^*, the information on the shape of the hazard rate of τ_t is actually captured by a transformation of the dependent variable. In the framework of Cox (1972a), λ_0 is an arbitrary unspecified baseline hazard function leading to a model which is sufficiently flexible for many applications. In this context the transformation from τ_t to τ_t^* is unknown, thus (5) corresponds to a latent model driving the observable outcomes of τ_t.

For the estimation of this model, different solutions have been proposed. Cox (1975) proposed the partial likelihood approach which is based on the fact that the baseline hazard λ_0 and the coefficients β can be estimated separately. The estimation of the baseline hazard follows from a modification of the estimator by Kaplan and Meier (1958) proposed by Breslow (1972), (1974). An alternative proceeding relies on the relationship between conventional binary response models and proportional hazard specifications for grouped durations models (see Sueyoshi, 1995). Based on this framework Han and Hausman (1990) introduce a model which allows for a simultaneous estimation of β and λ_0 at discrete points.[4] A further strand of econometric duration literature is based on Bayesian estimation methods, see, for example, Campolieti (2001), (2000) or Ruggiero (1994). A fourth method to estimate the semiparametric proportional hazard model uses recent results on the estimation of (censored) regression models with an unknown transformation of the dependent variable. Based on the work of Horowitz (1996), Gorgens and Horowitz (1999) introduce an approach for a general class of censored regression models where the left hand variable is an unknown increasing function of an observable variable[5]. An important special case of this class of models is the (censored) semiparametric proportional hazard model with unobserved heterogeneity. In this context the unknown distribution function as well as the unknown transformation of the left hand variable are estimated nonparametrically.

2.2 Dynamic proportional hazards models

Using the terminology of Cox (1981) there are two ways to incorporate a dynamic in the proportional hazards model outlined in (4). On the one hand, there are observation driven models, which are characterised by a mean function m_t measurable with respect to the observable information set $\mathcal{F}_{t-1}$. On the other hand, there are parameter driven models

[4] A related approach has been introduced simultaneously by Meyer (1990). This class of specifications is extended by Romeo (1999) in order to test hypotheses regarding the shape of the hazard function.

[5] For a related study see Abrevaya (1999).

whose mean function is measurable with respect to the latent information set $\mathcal{F}^*_{t-1}$. At first glance the distinction seems point-less since $\tau^*_t = \log \Lambda_0(\tau_t)$ is a one-to-one mapping from τ_t to τ^*_t and the information sets coincide. The importance of the distinction will become clear once both approaches have been outlined in greater detail.

An observation driven dynamic model of τ^*_t can be depicted as

$$\tau^*_t = m(\mathcal{F}_{t-1}, \theta) + \epsilon^*_t, \tag{7}$$

where m is a possibly nonlinear function of past observed durations and present and past exogeneous regressors x_t and parameters θ. The estimations of this type of model uses again the partial likelihood procedure proposed by Cox (1975), which has been shown by Oakes and Cui (1994) to be available, even in the dynamic case. Therefore, the two-step estimation of parameters θ and baseline hazard λ_0 is still possible. A simple example would be to specify $m(\mathcal{F}_{t-1}, \theta)$ in terms of lagged durations. However, it turns out, that the dynamic properties of such models are non-trivial and that the selection of a specific functional form in empirical applications is a complicated task. See e.g. Hautsch (1999). An alternative is to specify the model dynamics directly in terms of the log integrated baseline hazard τ^*_t leading to a parameter driven dynamic proportional hazard model. The important difference of this specification of the general form

$$\tau^*_t = m(\mathcal{F}^*_{t-1}, \theta) + \epsilon^*_t, \tag{8}$$

is that the latter type of models does not allow for a two step estimation. This becomes obvious, if one considers an AR(1) process for τ^*_t as a special case of (8)

$$\tau^*_t = \rho\tau^*_{t-1} + \epsilon^*_t = \rho \log \Lambda_0(\tau_{t-1}) + \epsilon^*_t. \tag{9}$$

The baseline hazard enters explicitly the mean function and the partial likelihood approach is unfeasible. This is simply because τ^*_t is a function of the baseline hazard λ_0

and the latter is left unspecified. Since τ^* is not observable directly, the usual prediction error decomposition, again, is not available. To circumvent the problems induced by a dynamic extension of the standard proportional hazard model, we propose a model which embodies characteristics of observation driven specifications as well as parameter driven models. In particular, we specify an observation driven dynamic process for the log integrated baseline hazard. Since this model prevents a two step estimation of θ and λ_0, in the following subsection we present a discretisation approach which allows for a simultaneous estimation of discrete points of the baseline hazard function and of the parameter vector. In section 2.4 we introduce a dynamic extension of the semiparametric proportional hazard model which is quite easy to implement and allows for a standard (conditional) maximum likelihood estimation of the unknown parameters.

Other avenues might be worthwhile pursuing in order to obtain specifications which are more flexible in a dynamic framework. On the one hand there are parametric extensions, which allow for more flexible distributions. This naturally leads to a type of semiparametric time series model as it is suggested by Drost and Klaassen (1997) for the case of GARCH models and Drost and Werker (2000) for ACD models. On the other hand, one might think of alternative nonparametric procedures along the lines of Gorgens and Horowitz (1999) or of specifications based on a kernel estimator of the baseline hazard, as it was recently proposed by Horowitz (1999). The latter is however not easily extended to account for dynamics as in the given context.

2.3 A discretisation approach

Since in a dynamic framework involving lags of the log integrated baseline hazard λ_0, the estimation of λ_0 and the dynamic parameters are not separable, we use a discretisation approach for the estimation of the baseline hazard rate, see e.g. Oakes (1972), Breslow (1972), Kalbfleisch and Prentice (1973) or Han and Hausman (1990). We assume that the baseline hazard function is constant over certain intervals of the support of the durations τ_t. In doing so we follow the approach suggested by Han and Hausman (1990) and build on a categorization of the durations τ_t leading to a semiparametric estimator of the baseline hazard, λ_0, at a number of discrete points.

The discretisation rule $d(\cdot)$ by which a categorical random variable d_t is derived from the observable durations τ_t is as follows

$$d(\tau_t) := \begin{cases} 1 & \text{if } \tau_t \in (-\infty, \mu_1], \\ 2 & \text{if } \tau_t \in (\mu_1, \mu_2], \\ \vdots & \\ K & \text{if } \tau_t \in (\mu_{K-1}, \infty). \end{cases} \tag{10}$$

where $\mu_k,\ k = 1, \ldots, K-1$, denote the boundaries of the K categories. For the sake of a simple exposition, the K categories are denoted by integers, however the values denoting individual categories can be chosen arbitrarily without having an influence on the estimated dynamic. Let $\mathcal{F}_t^d = \sigma(d(\tau_t), d(\tau_{t-1}), \ldots, d(\tau_1), x_{t+1}, x_t, \ldots, x_1)$ the information set generated by the categorized durations.

The discretisation rule in (10) describes the recoding of the data, while the random variable d_t is described by a categorical response type framework, i.e. we use τ_t^* as a latent variable, which is only observed through a step function $g(\cdot)$ defined using parameters μ^*,

$$g(\tau_t^*, \mu^*) = \begin{cases} 1 & \text{if } \tau_t^* \in (-\infty, \mu_1^*], \\ 2 & \text{if } \tau_t^* \in (\mu_1^*, \mu_2^*], \\ \vdots & \\ K & \text{if } \tau_t^* \in (\mu_{K-1}^*, \infty). \end{cases} \tag{11}$$

By this, the observation rule of the latent variable τ_t^* is such that a duration within the category $(\mu_{k-1}; \mu_k]$ is observed if the latent variable τ_t^* lies between the two thresholds $(\mu_{k-1}^*; \mu_k^*]$.

The discretisation approach implies of course a certain loss in information, whereby we have for the information generated by the discretized durations $\mathcal{F}_t^d \subset \mathcal{F}_t$. This is a standard approach in the analysis of grouped durations, see e.g. Thompson (1977), Prentice and Gloeckler (1978), Meyer (1990), Kiefer (1988), Han and Hausman (1990), Sueyoshi (1995) and Romeo (1999).

By having introduced the discretisation of the data (10) and the observation rule (11) of the latent process τ_t^*, the model admits an interpretation as a special type of ordered response model based on an extreme value distribution. The functional relationship between the duration τ_t and the log integrated baseline hazard τ_t^* is estimated in a semiparametric fashion. Note the important difference between μ and μ^*. The former being a set of thresholds governing the precision of the semiparametric approximation. The latter being the parameters of the model, which are directly related to the distribution function of the errors in the linear model (5).

The relationship between the baseline hazard λ_0 in (3) and the baseline distribution function Q_0 is just

$$Q_0(s) = 1 - \exp\left[-\int_0^s \lambda_0(u)du\right] \tag{12}$$

This relationship is exploited to relate the thresholds μ^*, which are parameters of the model, to the baseline hazard λ_0. The unknown distribution function Q_0 can be estimated at $K-1$ discrete points by a nonlinear function of the estimated thresholds μ_k^* using

$$Q_0(\mu_k) = 1 - \exp(-\exp(\mu_k^*)), \quad k = 1, \ldots, K-1. \tag{13}$$

Based on the discrete points of the baseline distribution function we can estimate a discrete baseline hazard $\tilde{\lambda}_0$, corresponding to the conditional failure probability, given the elapsed time since the last event, given by

$$\tilde{\lambda}_0(\mu_k) := \text{Prob}\left[\tau_t \in [\mu_k, \mu_{k+1}) \,|\tau_t > \mu_k\right] \tag{14}$$

$$= \frac{Q_0(\mu_{k+1}) - Q_0(\mu_k)}{1 - Q_0(\mu_k)}, \quad k = 0, \ldots, K-2 \tag{15}$$

with $\mu_0 \equiv 0$. This formulation serves as an approximation to the baseline hazard defined in (1) if divided by the length of the discretisation interval, i.e. if the length of the intervals goes to zero, the approximation converges to the original definition, thus

$$\lambda_0(\mu_k) \approx \frac{\tilde{\lambda}_0(\mu_k)}{\mu_{k+1} - \mu_k}. \tag{16}$$

Although, this approximation yields results comparable to a nonparametric estimation procedure, there are important differences. The most obvious is the missing theoretical link between the number of categories K and the sample size T, as it is the case between the bandwidth parameter and the sample size in a nonparametric estimation. Therefore, the approximation of the baseline hazard (16) should not be termed nonparametric.

2.4 Specification and estimation of a semiparametric ACPH(p,q) model

Using the discretisation approach, another obstacle is the estimation problem incurred if a dynamic is attached to the latent of a limited dependent variable like (11). In general this leads to a maximum likelihood estimator involving a T-fold integral, as has been amply discussed in the literature, see e.g. the surveys by Hajivassiliou and Ruud (1994) and Manrique and Shephard (1998). A viable alternative poses the observation driven dynamic proposed by Gerhard (2001).

Therefore, we propose an observation driven dynamic for the conditional mean function of τ_t^*. The main idea is to specify the dynamic on the basis of conditional expectations of the error ϵ_t^*

$$\epsilon_t := \mathrm{E}\left[\epsilon_t^* \middle| \mathcal{F}_t^d\right]. \tag{17}$$

Note that ϵ_t relates to a concept known in econometrics as generalised residuals, see Gourieroux, Monfort, Renault, and Trognon (1987), or in statistics as Bayesian residuals, see Albert and Chib (1995).

By using for expositional ease the simplest case of an AR(1) like dynamic, the ACPH(1,0) model is given by

$$\tau_t^* = m_t + \epsilon_t^*, \tag{18}$$

where m_t is defined through a recursion, conditioned on an initial m_0,

$$m_t = \phi(m_{t-1} - x_{t-1}'\beta + \epsilon_{t-1}) + x_t'\beta. \tag{19}$$

Note that the computation of m_t allows to use a conditioning approach which exploits the observation of m_{t-1} and thus prevents us from computing cumbersome T-fold integrals.

Therefore, the semiparametric ACPH(p,q) model is built on an ARMA structure based on the conditional expectations of the past latent variables, given the observable categorized durations which correspond to generalized residuals in the sense of Gourieroux, Monfort, Renault, and Trognon (1987). This specification poses a dynamic extension to the approach of Han and Hausman (1990), i.e. the autoregressive structure is built on past values of a transformation of the baseline hazard which is assumed to be piecewise constant.

Thus, the semiparametric ACPH model can be written in terms of the hazard rate as

$$\lambda_\tau(s|\mathcal{F}_{t-1}^d) = \lambda_0(s)\exp(-m_t). \tag{20}$$

This specification has two major advantages. First, it is shown that this model is equivalent to a simple ARMA model for the log integrated baseline hazard, if the baseline hazard was known. In this case, the step function $g(\cdot)$ is replaced by a Borel measurable one-to-one function and $d(\cdot)$ is reduced to the identity mapping, so that $\mathcal{F}_t^d = \mathcal{F}_t$, see e.g. Davidson (1994, theorem 10.3). Therefore, $\epsilon_t = \epsilon_t^*$ and $m_t = \mathrm{E}[\tau_t^*|\mathcal{F}_t^*] - \varsigma$, where $\varsigma \equiv \mathrm{E}[\epsilon^*]$ and thus the semiparametric ACPH(1,0) model would collapse to a standard model of an ARMA(1,0) process of the log integrated hazard. Hence, the parameter ϕ can be interpreted as an approximation to the parameter ρ in the ARMA(1,0) model

$$\tau_t^* = \rho(\tau_{t-1}^* - x_{t-1}'\beta) + x_t'\beta + \epsilon_t^*. \tag{21}$$

Note that this approximation is the better the finer the chosen categorization (see also section 4).

The second major advantage is the straightforward computation of the likelihood of the model, which does not necessitate the use of extensive simulation methods. Thus, a computationally simple maximum likelihood estimator of the dynamic parameter ϕ and the parameters of the baseline hazard approximation μ^* is directly available.

The mean function of the latent case, given in (19) for the case of an ACPH(1,0) model can easily accommodate higher order AR(p) and MA(q) terms to yield an ACPH(p,q) model. Define the matrices F and H and the dimension r of a generic state space model as

$$F = \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_{r-1} & \phi_r \\ 1 & 0 & \dots & & 0 \\ 0 & 1 & 0 & \dots & \vdots \\ \vdots & & \ddots & & 0 \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}, \quad H' = \begin{bmatrix} 1 & \theta_1 & \dots & \theta_{r-1} \end{bmatrix}, \quad r = \max(p, q+1), \quad (22)$$

where we have for the AR parameters $\phi_i = 0$ for $i > p$ and the MA parameters $\theta_i = 0$ for $i > q$. See e.g. Hamilton (1994, chap. 13.1). Thereby, the mean function m_t can be given as the function of the mean function s_t in a state space context as

$$m_t = H' s_t + x_t' \beta, \quad (23)$$

where the mean of the state process is assumed to follow

$$s_t = F(s_{t-1} + e_1 \epsilon_{t-1}), \qquad \text{where } e_1' = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}, \quad (24)$$

conditional on some initial s_0.

Once the system is cast in state space form, weakly exogenous variables can be included not only statically in (23), corresponding to (19), but also in the dynamic (24) leading to an infinite lag structure. For the sake of brevity the explicit model is omitted here.

2.5 Computation of the log likelihood

Since the observation driven dynamic enables us to use the standard prediction error decomposition, see e.g. Harvey (1990), the likelihood is evaluated in a straightforward iterative fashion:

(a) The mean function s_t of the state process is initialised with its unconditional expectation

$$s_0 := \mathrm{E}[s_t]. \tag{25}$$

(b) The likelihood contribution of observation t given the observation rule (11) under the given distributional assumptions is

$$\mathrm{Prob}\left[d_t \,\middle|\, \mathcal{F}^d_{t-1}\right] = \begin{cases} Q(\mu_1^* - m_t) & \text{if } d_t = 1, \\ Q(\mu_2^* - m_t) - Q(\mu_1^* - m_t) & \text{if } d_t = 2, \\ \vdots & \\ 1 - Q(\mu_{K-1}^* - m_t) & \text{if } d_t = K, \end{cases} \tag{26}$$

where $Q(.)$ denotes the c.d.f. of the extreme value distribution, given by

$$Q(s) = \exp(-\exp(s)).$$

(c) The generalized error ϵ_t, which drives the mean function is a (conditionally) deterministic function of the observations, which make up the conditional expectation

m_t and the present observation d_t and is given by

$$\epsilon_t = \mathrm{E}\left[\epsilon_t^* \middle| \mathcal{F}_t^d\right] = \mathrm{E}\left[\epsilon_t^* \middle| d_t, m_t\right]$$
$$= \begin{cases} \frac{q(\eta_{t,1})}{Q(\eta_{t,1})} & \text{if } d_t = 1, \\ \frac{q(\eta_{t,k}) - q(\eta_{t,k-1})}{Q(\eta_{t,k}) - Q(\eta_{t,k-1})} & \text{if } d_t \in \{2, \ldots, K-1\}, \\ \frac{\varsigma - q(\eta_{t,K})}{1 - Q(\eta_{t,K})} & \text{if } d_t = K, \end{cases} \tag{27}$$

where $\eta_{t,k} = \mu_k^* - m_t$, $q(s) \equiv \int_{-\infty}^{s} uf(u)du$ and $f(\cdot)$ denotes the p.d.f. of the extreme value distribution, eq. (6). See the original paper by Gourieroux, Monfort, Renault, and Trognon (1987) for an extended discussion of generalized errors in the context of non-dynamic models.

(d) The calculation of the conditional expectation of the future latent variable given the present information and the definition of the mean processes m_t and s_t in (23) and (24).

(e) Steps 2 through 4 are repeated for all d_t, $t = 1, \ldots, T$.

(f) The observable log likelihood is given by

$$\log \mathcal{L}(\bar{d}_T | \phi, \theta, \beta, \mu^*) = \sum_{t=1}^{T} \sum_{k=1}^{K} \delta_{t,k} \log \mathrm{Prob}\left[d_t = k \middle| \mathcal{F}_{t-1}^d\right], \tag{28}$$

where

$$\delta_{t,k} = \begin{cases} 1 & \text{if } \ d_t = k, \\ 0 & \text{else}, \end{cases}$$

and $\bar{d}_T$ collects all observations of d_t, $t = 1, \ldots, T$.

3 Extensions to the semiparametric ACPH model

3.1 Censoring

A typical property of economic duration data is the occurrence of censoring, leading to incomplete spells. Therefore, a wide strand of econometric duration literature focusses on the consideration of censoring mechanisms, see, for example, Orbe, Ferreira, and Nunez-Anton (2000), Gorgens and Horowitz (1999), the survey by Neumann (1997) or Horowitz and Neumann (1987), (1989). In the context of serial dependent arrival rates, censoring occurs if some of the arrival times ϑ_t, $t = 1, \ldots, T$, cannot be observed. Assume that for each arrival time ϑ_t there exists a censoring time ϑ_t^c which is associated with the beginning of a time period between ϑ_{t-1} and ϑ_{t+1} in which the occurrence of an event is not observable.

If ϑ_t occurs after ϑ_t^c, then ϑ_t is not observable and the observed duration is not $\tau_t = \vartheta_t - \vartheta_{t-1}$ but $\tau_t^c = \vartheta_t^c - \vartheta_{t-1}$ with $\tau_t^c \leq \tau_t$. In this case τ_t^c is called 'right-censored' because the exact stopping time is unknown. If the stopping time ϑ_t is also the starting time of the next spell with duration τ_{t+1}, then instead of the exact duration $\tau_{t+1} = \vartheta_{t+1} - \vartheta_t$, only the 'left-censored' duration $\tau_{t+1}^c = \vartheta_{t+1} - \vartheta_t^c$ is observable.

In this framework the observed duration process is $\{\tau_t, c_t\}$, $t = 1, \ldots, T$, where

$$c_t = \begin{cases} 1 & \text{if observation } t \text{ is censored,} \\ 0 & \text{else} \end{cases}$$

and the corresponding (observable) durations in t and $t+1$ are $\tau_t = \min\{\tau_t, \tau_t^c\}$ and $\tau_{t+1} = \min\{\tau_{t+1}, \tau_{t+1}^c\}$.

In the following we assume independent censoring, i.e. the censored durations τ_t^c, and the true (but in the case of censoring, unobservable) durations are independent.

This assumption is fulfilled if the censoring mechanism is determined exogenously and is not driven by the duration process. For a detailed exposition and different types of censoring mechanisms see e.g. Neumann (1997). Under this assumption the likelihood can be decomposed into

$$\mathcal{L}(\bar{d}_T, \bar{c}_T|\phi, \theta, \beta, \mu^*) = \mathcal{L}(\bar{d}_T|\bar{c}_T, \phi, \theta, \beta, \mu^*)\mathcal{L}(\bar{c}_T), \tag{29}$$

where $\bar{d}_T$ and $\bar{c}_T$ collects all observations of d_t and c_t, $t = 1, \ldots, T$. Note that in the case of an exogenously given censoring scheme, the second factor does not depend on the parameters of the model, thus the parameters ϕ, θ, β and μ^* are estimated by maximizing the first factor of (29) and the corresponding conditional likelihood function is given by

$$\log \mathcal{L}(\bar{d}_T|\bar{c}_T, \phi, \theta, \beta, \mu^*) = \sum_{t=1}^{T}\sum_{k=1}^{K}(1-c_t)(1-c_{t-1}) \cdot \delta_{t,k} \log \int\limits_{\mu^*_{k-1}-m_t}^{\mu^*_k-m_t} f(s)ds \tag{30}$$

$$+ \sum_{t=1}^{T}(1-c_{t-1})c_t \cdot \log \int\limits_{\mu^*_{l,r}}^{\infty} f(s)ds + \sum_{t=1}^{T} c_{t-1}(1-c_t) \cdot \log \int\limits_{\mu^*_{l,l}}^{\infty} f(s)ds + \sum_{t=1}^{T} c_{t-1}c_t \cdot \log \int\limits_{\mu^*_{l,lr}}^{\infty} f(s)ds,$$

where $\mu^*_{l,r}$, $\mu^*_{l,l}$ and $\mu^*_{l,lr}$, respectively, denote the corresponding thresholds of the lower boundary of the censored duration τ^c_t in the case of right-censoring ($\tau^c_t = \vartheta^c_t - \vartheta_{t-1}$), left-censoring ($\tau^c_t = \vartheta_t - \vartheta^c_{t-1}$) and left-right-censoring ($\tau^c_t = \vartheta^c_t - \vartheta^c_{t-1}$). Hence, the first term in (30) is the probability for observing a (noncensored) duration in category k while the following terms are the probability for the duration τ_t to be determined by the particular censoring scheme.

The maximum likelihood estimation proceeds along the lines described for the standard ACPH(p,q) model in section 2.4, where the likelihood in (30) accounting for the pseudo-censoring takes the place of the original likelihood in (28). The derivation of generalised residuals needs to be slightly modified in order to account for censoring. In (27) the list of cases needs to be amended to accommodate for the conditional expectation of the error given a censored observation, which is computed as

$$\epsilon_t = \frac{\varsigma - q(\eta_{t,l})}{(1 - Q(\eta_{t,l}))} \quad \text{if } c_t{=}1 \text{ or } c_{t-1} = 1, \tag{31}$$

where

$$\eta_{t,l} = \begin{cases} \mu^*_{l,r} - m_t & \text{if } (1 - c_{t-1}) \cdot c_t = 1, \\ \mu^*_{l,l} - m_t & \text{if } c_{t-1} \cdot (1 - c_t) = 1, \\ \mu^*_{l,lr} - m_t & \text{if } c_{t-1} \cdot c_t = 1. \end{cases}$$

3.2 Unobserved heterogeneity

Because the ACPH model is based on the traditional proportional hazard specification, it is relatively easy to control for unobservable heterogeneity. From an econometric point of view, accounting for unobserved heterogeneity can be interpreted as an additional degree of freedom where Lancaster (1997) illustrates that the inclusion of a heterogeneity variable can capture errors in the variables. From an economic point of view, in the context of financial transaction data, unobservable heterogeneity effects can be associated with different kinds of traders or different states of the market. The consideration of unobserved heterogeneity is an ongoing topic in duration analysis. Lancaster (1979) or Heckmann and Singer (1984) illustrated that ignoring unobserved heterogeneity can lead to biased estimates of the hazard function. While in former studies either the underlying duration distribution or the heterogeneity component was specified nonparametrically (see Heckmann and Singer, 1984, Honore, 1990, Bearse, Canals, and Rilstone (1996), 1996, Han and Hausman, 1990 or Meyer, 1990), recent specifications, like Horowitz (1996) or Gorgens and Horowitz (1999), allow for a nonparametric estimation of both components.[6]

[6]For a Monte Carlo study which analyzes the impact of misspecifications in both components, see Baker and Melino (2000).

Since in our approach a nonparametric estimation of unobserved heterogeneity effects would be quite cumbersome, we specify such effects, following Han and Hausman (1990), by a random variable which enters the hazard function multiplicatively leading to a mixed ACPH model. The standard procedure to account for unobserved heterogeneity in the semiparametric ACPH model is to introduce a random variable ν_t in the specification (20) to obtain

$$\lambda_\tau(s|\mathcal{F}_{t-1}^d, \nu_t) = \lambda_0(s) \cdot \nu_t \cdot \exp(-m_t), \qquad t = 1, \ldots, T. \tag{32}$$

We assume for the random variable ν_t a Gamma distribution with mean one and variance ζ^{-1}, which is standard for this type of mixture models, see e.g. Lancaster (1997). The distribution function of a compounded model $\tilde{Q}$ is obtained by integrating out ν_t

$$\tilde{Q}(s|\mathcal{F}_{t-1}^d) = 1 - \left[1 + \zeta^{-1} \exp(-m_t) \int_0^s \lambda_0(u) du\right]^{-\zeta}. \tag{33}$$

Note that this is identical to the distribution function of a BurrII model under appropriate parametrisation.

The latter gives rise to an analogue model based on the discretisation approach outlined in 2.3. By augmenting the log linear model of the integrated baseline hazard rate by a compounder, we obtain an ACPH(p,q) model including unobserved heterogeneity based on the modified latent process

$$\tau_t^* = \ln(\zeta) + m_t + \epsilon_t^{**}, \tag{34}$$

where the error term ϵ_t^{**} follows a BurrII(ζ) distribution with density function

$$\tilde{f}(s) = \frac{\zeta \exp(s)}{[1 + \exp(s)]^{\zeta+1}}. \tag{35}$$

It is easy to show that the BurrII(ζ) distribution includes the extreme value distribution as a limiting case for $\zeta^{-1} = \text{Var}[\nu_t] \to 0$, i.e. if no unobservable heterogeneity effects exist, the model corresponds to the basic ACPH model.

The estimation procedure is similar to the procedure described in section 2.4. The difference is just that the model is now based on a BurrII(ζ) distribution. Apart from an obvious adjustment to the generalised errors, the relationship between the estimated thresholds and the estimation of the distribution function of the error term given in (13), which relates directly to the baseline hazard, is slightly modified, to

$$\tilde{Q}_0(\mu_k) = 1 - \frac{1}{[1 + \exp(\mu_k^* - \ln(\zeta))]^\zeta}, \qquad k = 1, \ldots, K-1. \tag{36}$$

Thus, the semiparametric ACPH model can also accommodate for unobserved heterogeneity, which can be of considerable importance, when some of the imminent market microstructure effects are not modelled explicitly.

3.3 State dependent hazard rates

The standard proportional hazard model underlies the assumption that, for any two sets of explanatory variables x_1 and x_2, the hazard functions are related by

$$\lambda(s|x_1) \propto \lambda(s|x_2).$$

Sometimes there exist important economic factors causing a violation of this relationship. To obtain more flexibility and to relax the proportionality assumption we stratify the data set and define state dependent baseline hazard functions $\lambda_{0,s}(u)$, $s = 1, \ldots, S$, where the index s denotes the particular state. In the context of transaction data the estimation of state dependent hazard rates and survivor functions can be used as valuable tools to obtain state dependent liquidity or volatility measures.

Assume a state defining variable z_t, which is at least weakly exogenous and which determines the functional relationship between τ_t^* and the baseline hazard $\lambda_{0,s}(u)$. Hence, the transformation from τ_t to τ_t^*, as outlined in section 2.3, is now state-dependent and is given by

$$\tau_t^* = \sum_{s=1}^{S} \mathbf{1}_{\{\underline{z}_{s-1} < z_t \leq \underline{z}_s\}} \ln \int_0^{\tau_t} \lambda_{0,s}(u) du, \tag{37}$$

where $\underline{z}_0 < \underline{z}_1 < \ldots < \underline{z}_S$ denote the state defining values of the variable z_t. The assumption of a model with S distinct hazard rates translates to S sets of distinct threshold parameters $\mu_{k,s}^*$, where $k = 1, \ldots, K-1$ and $s = 1, \ldots, S$. Correspondingly to (13) we obtain S distribution functions evaluated at the $K-1$ thresholds[7]

$$Q_{0,s}(\mu_{k,s}) = 1 - \exp(-\exp(\mu_{k,s}^*)), \;\; k = 1, \ldots, K-1, \;\; s = 1, \ldots, S. \tag{38}$$

Therefore, the generalized residuals are also state-dependent

$$\epsilon_t := \mathrm{E}\left[\epsilon_t^* | \mathcal{F}_t^{d,z}\right] = \sum_{s=1}^{S} \mathbf{1}_{\{\underline{z}_{s-1} < z_t \leq \underline{z}_s\}} \epsilon_{t,s},$$

where the information set $\mathcal{F}_t^d$ is extended to $\mathcal{F}_t^{d,z}$ in an obvious way and $\epsilon_{t,s}$ is computed according to (27) based on the corresponding set of threshold parameters $\mu_{k,s}^*$.

Hence, a semiparametric ACPH model with state dependent hazard rates is defined as

$$\lambda(s|\mathcal{F}_{t-1}^{d,z}) = \sum_{s=1}^{S} \lambda_{0,s}(s) \mathbf{1}_{\{\underline{z}_{s-1} < z_t \leq \underline{z}_s\}} \exp(-m_t). \tag{39}$$

The calculation of the log likelihood is based on the procedure proposed in section 2.4, therefore, we obtain the log likelihood function by

$$\log \mathfrak{L}(\bar{d}_T | \bar{z}_T, \phi, \theta, \beta, \mu^*) = \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{s=1}^{S} \delta_{t,k} \mathbf{1}_{\{\underline{z}_{s-1} < z_t \leq \underline{z}_s\}} \log \int_{\mu_{k-1,s}^* - m_t}^{\mu_{k,s}^* - m_t} f(s) ds. \tag{40}$$

---

[7]Note that it is also possible to use different categorizations for the duration τ_t within each state, thus, the number of thresholds estimated for each state could actually differ.

4 The approximation error in small samples

The main advantages of the given approach using a ACPH dynamic in contrast to a dynamic in the log integrated hazard rate τ_t^*, eq. (21), are the straightforward and fast maximum likelihood estimation and the fact that the generalized errors ϵ_t employed in the approximation are equal to the true errors ϵ_t^*, if the baseline hazard λ_0 is known. Yet, as it is at the core of our model that the baseline hazard λ_0 is not known, the question to be answered in this section is how large is the approximation error, if the true baseline hazard λ_0 is approximated at $K-1$ discrete points using $\tilde{\lambda}_0$ given by (14).

To obtain some evidence on the bias incurred by the discretisation approach a small Monte Carlo study is performed for an AR(1) process, like (21). The two specifications considered are based on two ($K_1 = 2$) and 11 ($K_2 = 11$) categories. The former being the worst possible approximation of the true baseline hazard λ_0 one could possible think of in the context of the given model and the latter being a more realistic case of using a moderate number of thresholds in the estimation.

Table 1: Monte Carlo study of semiparametric ACPH(p,q) model with K categories and T observations

p	q	K	T	bias	MSE	MAE
1	0	2	50	-0.0062	0.0294	0.1188
1	0	10	50	0.0045	0.0096	0.0734
1	0	2	500	0.0053	0.0024	0.0372
0	1	2	50	-.00795	0.0743	0.2025
0	1	10	50	-.00535	0.0189	0.0939
0	1	2	500	.01589	0.0200	0.0959

The K_1 model is replicated for two sample sizes $T_{11} = 50$ and $T_{12} = 500$. The model with more thresholds is only estimated for a small sample size $T_2 = 50$. This set-up allows to compare the improvement achieved by increasing the number of observations versus the benefit of a better approximation of the baseline hazard. Parameter estimation is based on the semiparametric ACPH(1,0) model given by (19) and the semiparametric ACPH(0,1). Since the focus is here on the bias of the dynamic parameters, the threshold parameters are fixed to their true values. A range of parameter values for ϕ and θ are covered in the simulations, concisely, $\phi, \theta \in Q = \{-0.9, -0.8, -0.7, \ldots, 0.9\}$ providing $N_i = 1,000$, $i \in Q$, replications for each value. The errors ϵ_t^* are drawn from the extreme value distribution as in the assumed DGP.

Overall results for all $N = 19,000$ replications are reported in table 1. It gives descriptive statistics of the difference between true parameters and estimates, $\phi_i - \hat{\phi}_i$, and $\theta_i - \hat{\theta}_i$, for $i = 1, \ldots, N$. Although aggregated over all parameter values, the small sample properties match the expectation build from the asymptotic results, i.e. the variance decreases over an increasing sample size and likewise do the interquantile ranges. The results indicate that even a moderately sized sample of 50 observations is quite sufficient to obtain reasonable results. The results indicate that for the ACPH(1,0) model the asymptotic properties seem to hold quite nicely. The performance of the MA based ACPH(0,1) model seems be worse than the corresponding ACPH(1,0) models.

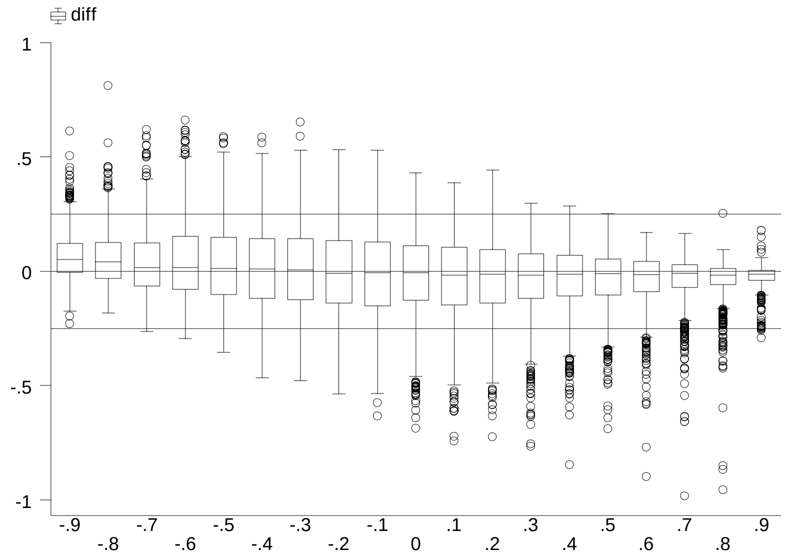

Figure 1: ACPH(1,0), $K_1 = 2$, $T_{11} = 50$: Box plots of $\phi_i - \hat{\phi}_i$ for 19 values of the parameter ϕ_i in a Monte Carlo study.

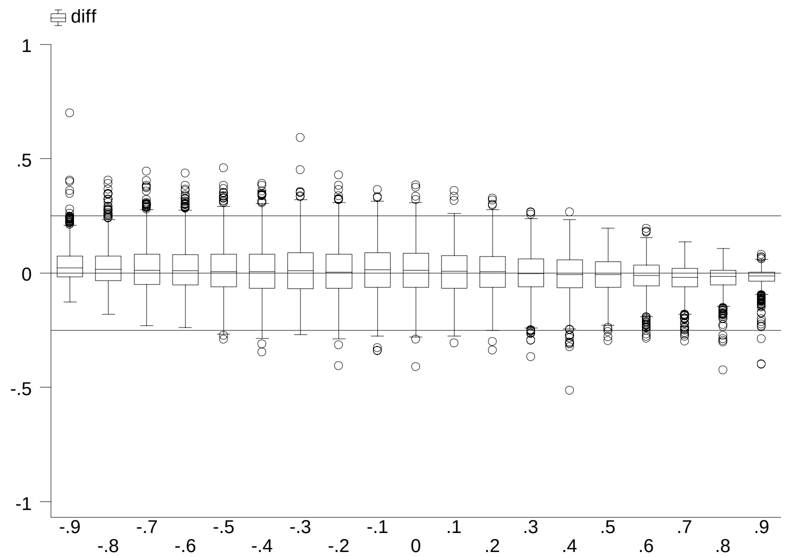

Figure 2: ACPH(1,0), $K_2 = 10$, $T_2 = 50$: Box plots of $\phi_i - \hat{\phi}_i$ for 19 values of the parameter ϕ_i in a Monte Carlo study.

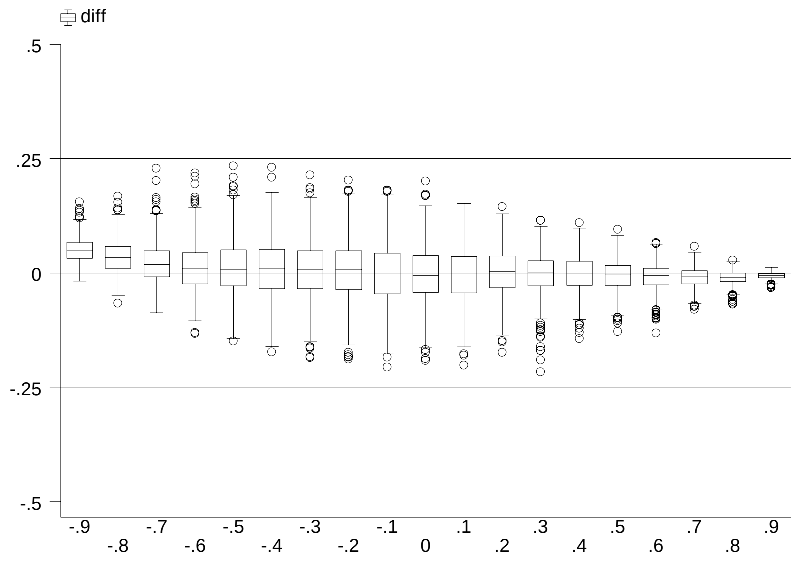

Figure 3: ACPH(1,0), $K_1 = 2$, $T_{12} = 500$**:** Box plots of $\phi_i - \hat{\phi}_i$ for 19 values of the parameter ϕ_i in a Monte Carlo study.

To gain more insight into the consequences the discretisation grid of the durations bears for the estimation, the results of the Monte Carlo experiment are scrutinized with respect to the parameters of the model, ϕ and θ, respectively. Simulation results for each of the 19 considered values of the true parameter in the DGP are illustrated in Box plots reported in figure 1-3. The results are quite encouraging, indicating that the quite considerable bias incurred for an ACPH(1,0) based on $K = 2$ categories is reduced considerably once a more realistic model based on $K = 11$ categories is employed. For a reasonable sample size (T_{12}) even for the categories the performance of the estimator is quite encouraging over all parameter values considered.

Figures 4-6 give the corresponding results for ACPH(0,1) models. Although, qualitattively similar, it is evident from the study that the ACPH(0,1) performs worse than the corresponding AR model. After an increase in the number of categories from $K = 2$ to $K = 11$ the approximation reaches about the quality of the ACPH(1,0) process with

$K = 2$ categories, except for the parameter value $\theta = 0.9$ The reason for this can be found in the differing ACF of an AR(1) and an MA(1). The relatively bad performance of the ACPH(0,1) process for parameters θ with large absolute value is due to the flattening out of the ACF towards the limits of the invertible region. See also the similar results for Gaussian models in Gerhard (2001).

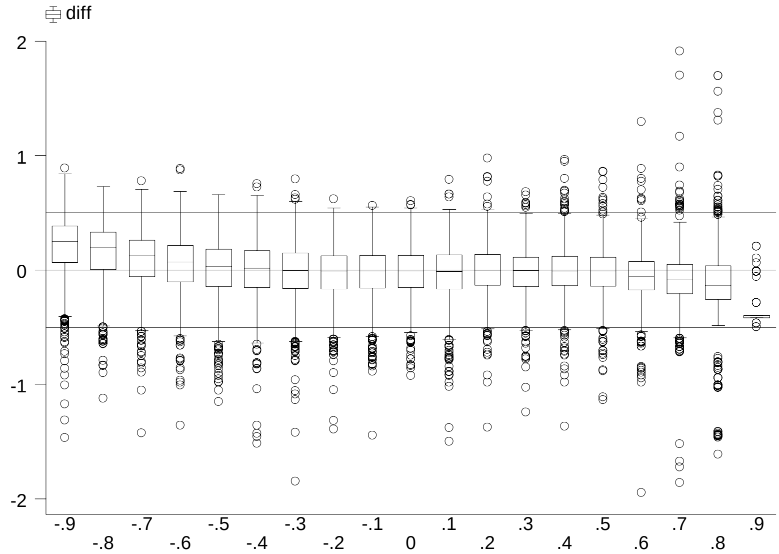

Figure 4: ACPH(0,1), $K_1 = 2$, $T_{11} = 50$: Box plots of $\theta_i - \hat{\theta}_i$ for 19 values of the parameter θ_i in a Monte Carlo study.

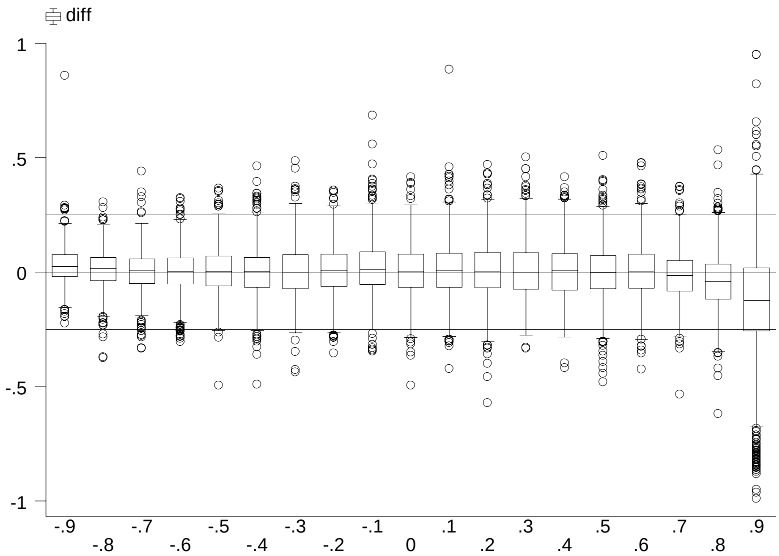

Figure 5: ACPH(0,1), $K_2 = 10$, $T_2 = 50$: Box plots of $\theta_i - \hat{\theta}_i$ for 19 values of the parameter θ_i in a Monte Carlo study.

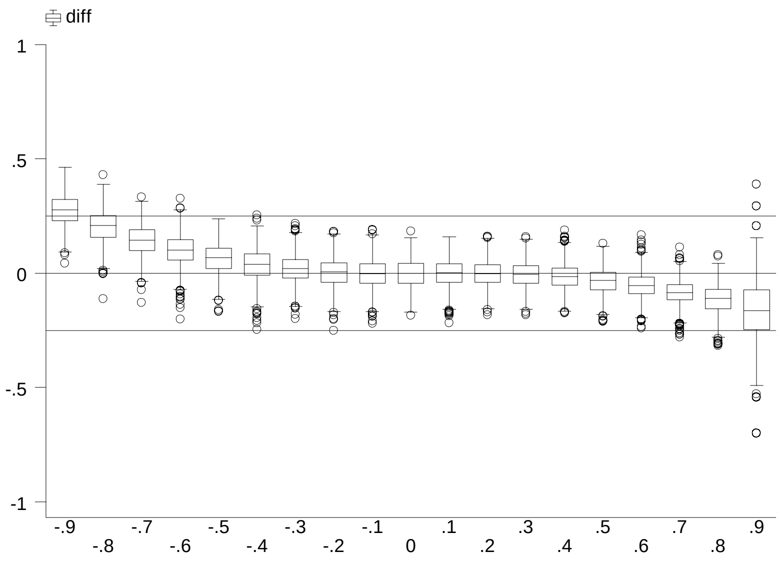

Figure 6: ACPH(0,1), $K_1 = 2$, $T_{12} = 500$: Box plots of $\theta_i - \hat{\theta}_i$ for 19 values of the parameter θ_i in a Monte Carlo study.

5 An application to BUND Future trading

To illustrate a small application of the model we estimate hazard rates of inter-trade durations based on intra-day transaction data from the Bund Future trading at the electronic trading system of the former Deutsche Terminbörse (DTB)[8], Frankfurt. The sample contains data from 01/30/95 to 02/24/95, corresponding to 20 trading days. Within this period the Bund-Future was one the most liquid futures in Europe and corresponded to a 6% German government bond of DEM 250,000 face value. The Bund Future had a maturity of 8.5 years and four contract maturities per year, March, June, September and December. In the sampling interval prices were denoted in basis points of face value, thus, one tick was equivalent to a value of DEM 25. The data set consists of 44,810 observations, where the overnight durations are omitted. Furthermore, we refrain from using the first 10 minutes of a trading day to avoid the opening phase. The descriptive statistics of

[8]Now EUREX.

the inter-trade durations are given in table 2. To account for intraday seasonalities we use the flexible Fourier form proposed by Andersen and Bollerslev (1998) based on Gallant (1981) which is given by[9]

$$s(\delta, t^*, P) = \delta_1 \cdot t^* + \sum_{p=1}^{P} \left(\delta_{c,p} \cos(t^* \cdot 2\pi p) + \delta_{s,p} \sin(t^* \cdot 2\pi p)\right), \tag{41}$$

where p is identical with the order of the term, $t^* \in [0,1]$ is defined by

$$t^* = \frac{\text{seconds since 8:40 a.m.}}{\text{seconds between 8:40 a.m. and 5:15 p.m.}} \tag{42}$$

and $\delta_{c,p}$, $\delta_{s,p}$ and δ denote the corresponding coefficients[10].

To illustrate the serial dependency of the inter-trade durations we computed the autocorrelation function (ACF) based on the raw durations and seasonal adjusted durations (figure 7). The ACF indicates the typical shape of a process with high persistence indicated by a relatively slowly decaying rate of dependence. Furthermore it is shown, that the serial dependence is slightly reduced when deterministic intraday seasonality patterns are taken into account.

Table 2: Descriptive statistics of inter-trade durations. Based on BUND futures trading at DTB, Frankfurt, from 01/30/95 to 02/24/95. 44810 observations.

Mean	Std. Dev.	Quantiles						
		5%	10%	25%	50%	75%	90%	95%
14.16	26.56	1	1	2	6	16	35	54

[9] For an application of the flexible Fourier form to intra-day and inter-day volatility estimation, see e.g. Gerhard and Hautsch (2001).

[10] Within the observation period at the DTB trading took place between 8:30 a.m. and 5.15 p.m.

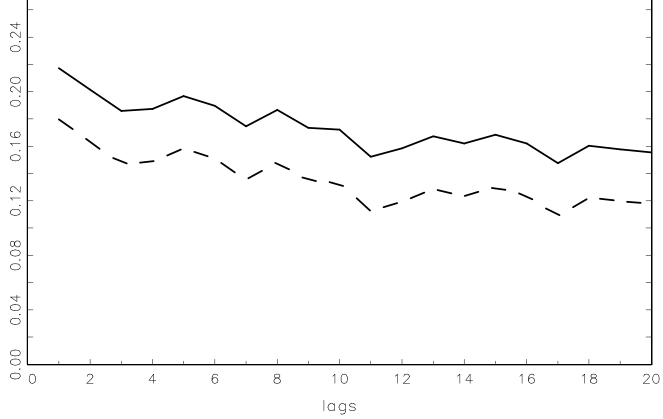

Figure 7: Autocorrelation functions of inter-trade durations. Based on BUND futures trading at DTB, Frankfurt, from 01/30/95 to 02/24/95. 44810 observations. Solid line: Raw durations. Broken line: Seasonal adjusted durations (residuals of an OLS regression of the durations on a flexible Fourier form with order $p = 4$).

Table 5 shows the estimation results of the basic ACPH(1,1) model based on different categorizations. The strong similarity of the corresponding estimates in the particular specifications indicates the robustness of the results against the choice of the categorization. Thus, the estimates of the autoregressive parameters and the coefficients associated with the explanatory variables seem not to be affected by the underlying categorization. Confirming the descriptive statistics, the estimated ARMA coefficients indicate a duration process with strong persistence which is a typical result for inter-trade durations. Figure 8 depicts the daily pattern of survivor probabilities, especially the probability to observe at least 30 seconds without a trade, computed based on the flexible Fourier coefficients. The figure shows the typical intra-day seasonality pattern with a high market activity in the morning, corresponding to low survivor probabilities, a significant dip around lunch time and a minimum after the opening of the American trading at 2:30 p.m.

Table 3: Estimates of ACPH(1,1) models based on different categorizations. Based on BUND futures trading at DTB, Frankfurt, from 01/30/95 to 02/24/95. 44810 observations. P-values based on asymptotic t-statistics.

Variables	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
			Thresholds			
μ_1 $(\tau_t = 1)$	-3.109	0.000	-3.095	0.000	-3.106	0.000
μ_2 $(\tau_t = 5)$	-2.043	0.000				
μ_3 $(\tau_t = 10)$	-1.532	0.000	-1.519	0.000		
μ_4 $(\tau_t = 15)$	-1.254	0.000			-1.256	0.000
μ_5 $(\tau_t = 20)$	-1.052	0.000	-1.040	0.000		
μ_6 $(\tau_t = 25)$	-0.889	0.000				
μ_7 $(\tau_t = 30)$	-0.763	0.000	-0.751	0.000	-0.766	0.000
μ_8 $(\tau_t = 35)$	-0.648	0.000				
μ_9 $(\tau_t = 40)$	-0.546	0.000	-0.535	0.000		
μ_{10} $(\tau_t = 45)$	-0.463	0.000			-0.467	0.000
μ_{11} $(\tau_t = 50)$	-0.386	0.000	-0.375	0.000		
μ_{12} $(\tau_t = 55)$	-0.313	0.000				
μ_{13} $(\tau_t = 60)$	-0.254	0.000	-0.242	0.017	-0.258	0.001
μ_{14} $(\tau_t = 65)$	-0.196	0.000				
μ_{15} $(\tau_t = 70)$	-0.142	0.000	-0.131	0.125		
μ_{15} $(\tau_t = 75)$	-0.093	0.000			-0.097	0.134
μ_{16} $(\tau_t = 80)$	-0.043	0.000	-0.032	0.388		
μ_{17} $(\tau_t = 85)$	-0.003	0.256				
μ_{18} $(\tau_t = 90)$	0.035	0.000	0.047	0.339	0.031	0.360

Table 3 continued: Estimates of ACPH(1,1) models based on different categorizations. Based on BUND futures trading at DTB, Frankfurt, from 01/30/95 to 02/24/95. 44810 observations. P-values based on asymptotic t-statistics.

Variables	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
			Intra-day Seasonalities			
trend	0.230	0.002	0.242	0.003	0.241	0.002
$\delta_{c,1}$	-0.215	0.000	-0.211	0.000	-0.217	0.000
$\delta_{c,2}$	0.187	0.000	0.187	0.000	0.182	0.000
$\delta_{c,3}$	-0.076	0.001	-0.076	0.000	-0.066	0.002
$\delta_{c,4}$	0.044	0.031	0.044	0.026	0.044	0.026
$\delta_{s,1}$	0.210	0.000	0.209	0.000	0.205	0.000
$\delta_{s,2}$	0.063	0.027	0.070	0.006	0.076	0.003
$\delta_{s,3}$	-0.102	0.000	-0.102	0.000	-0.106	0.000
$\delta_{s,4}$	0.076	0.001	0.079	0.000	0.073	0.000
			ARMA Parameters			
AR1	0.978	0.000	0.978	0.000	0.978	0.000
MA1	-0.917	0.000	-0.916	0.000	-0.916	0.000
			BIC and Mean Log Likelihood			
BIC	-92681		-69705		-57679	
LL	-2.064		-1.553		-1.285	

The three plots in figure 9 show the corresponding discrete baseline hazard functions based on the estimated thresholds. In general, the hazard function depicts a decreasing shape, i.e. the longer the last trade dates back the lower the probability for observing a further trade in the next instant of time. Furthermore, these graphs illustrate the loss of information induced by the choice of larger categories. Thus, while a categorization based on 5 secs provides a quite exact computation of the baseline hazard function, a categorization based on 15 secs allows only for a relatively coarse computation.

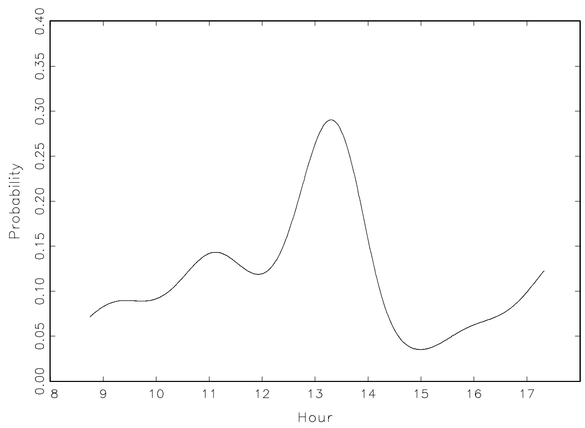

Figure 8: Intra-day seasonality pattern of the probability to observe 30 seconds without a trade. Based on BUND futures trading at DTB, Frankfurt, from 01/30/95 to 02/24/95. 44810 observations.

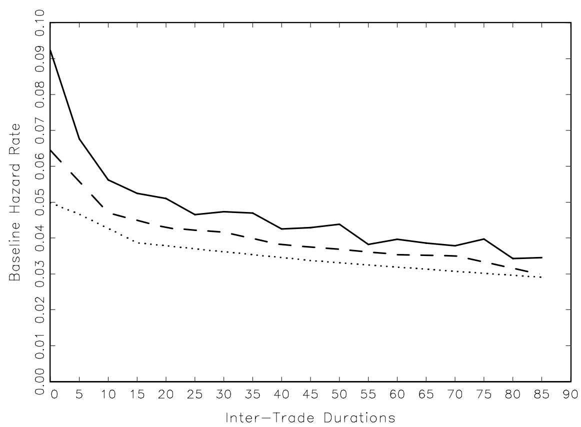

Figure 9: Estimated baseline hazard rates λ_0 based on different categorizations (left: 5,10,. . .,90 secs, middle: 10,20,. . .,90 secs, right: 15,30,. . .,90 secs.) Based on BUND futures trading at DTB, Frankfurt, from 01/30/95 to 02/24/95. 44810 observations.

Note that more sophisticated investigations, especially concerning the application of state dependent hazard rates, are beyond the scope of the paper and are provided in Gerhard and Hautsch (2000).

6 Conclusions

In this paper we have proposed semiparametric ACPH(p,q) models as a special class of semiparametric proportional hazard models. This approach has the virtue of being relatively easy to estimate and yielding a model with an unspecified baseline hazard and a dynamic in the log integrated baseline hazard. The latter is in contrast to standard ACD type models which focus usually on the conditional mean of the duration process. A thorough comparison of the implications of either approach is well beyond this paper which concentrates on the semiparametric estimation aspect and is to be pursued in future research. It is demonstrated, that the model can be easily extended to accommodate unobserved heterogeneity, censoring and state dependent hazard rates. In a small Monte Carlo study the quality of the approximation involved by using a discrete approximation to the unknown baseline hazard is assessed. An empirical study based on Bund future data of the former DTB demonstrated the flexibility of the approach.

Thanks

For valuable comments we would like to thank Luc Bauwens, Sir David Cox, Robert Engle, David Veredas, Joachim Inkmann, Bent Nielsen, Neil Shephard and Winfried Pohlmeier. An earlier version of this paper was presented at the 8th World Congress of the Econometric Society, Seattle and the CoFE Conference on Intertemporal Finance, Konstanz. The authors would like to thank the Center of Finance and Econometrics (CoFE) for financial support. The first author would also like to thank the UK's Economic and Social Research Council which supported him through the grant 'Econometrics of trade-by-trade price dynamics', which is coded R00023839. All remaining errors are our sole responsibility.

References

Abrevaya, J. (1999): "Leapfrog Estimation of a Fixed-Effects Model with Unknown Transformation of the Dependent Variable," *Journal of Econometrics*, 93, 203–228.

Albert, J., and S. Chib (1995): "Bayesian Residual Analysis for Binary Response Regression Models," *Biometrika*, 82, 747–759.

Andersen, T. G., and T. Bollerslev (1998): "Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies," *Journal of Finance*, 53, 219–265.

Baker, M., and A. Melino (2000): "Duration dependence and nonparametric heterogeneity: A Monte Carlo study," *Journal of Econometrics*, 96, 357–393.

Bearse, P., J. Canals, and P. Rilstone (1996): "Semiparametric Maximum Likelihood Estimation of Duration Models with Unobserved Heterogeneity," Discussion paper, University of Tennessee and University of Virginia and York University.

Breslow, N. (1972): "Contribution to the Discussion of Paper the Paper by D. R. Cox," *Journal of the Royal Statistical Society, Series B*, 34, 216–17.

Breslow, N. E. (1974): "Covariance Analysis of Censored Survival Data," *Biometrics*, 30, 89–100.

Campolieti, M. (2000): "Bayesian estimation and smoothing of the baseline hazard in discrete time duration models," *Review of Economics and Statistics*, 82(4), 685–701.

——— (2001): "Bayesian semiparametric estimation of discrete duration models: An application of the dirichlet process prior," *Journal of Applied Econometrics*, 16, 1–22.

Cox, D. R. (1972a): "Regression Models and Life Tables," *Journal of the Royal Statistical Society, Series B*, 34, 187–220.

——— (1972b): "The Statistical Analysis of Dependencies in Point Processes," in *Stochastic Point Processes*, ed. by P. A. W. Lewis, pp. 55–66. John Wiley, New York.

——— (1975): "Partial Likelihood," *Biometrika*, 62, 269.

——— (1981): "Statistical Analysis of Time Series: Some Recent Developments," *Scandinavian Journal of Statistics*, 8, 93–115.

Cox, D. R., and D. Oakes (1984): *Analysis of Survival Data*. Chapman and Hall, London.

Davidson, J. (1994): *Stochastic Limit Theory*. Oxford University Press, Oxford, New York.

Drost, F. C., and C. A. J. Klaassen (1997): "Efficient Estimation in Semiparametric GARCH Models," *Journal of Econometrics*, 81, 193–221.

Drost, F. C., and B. J. M. Werker (2000): "Efficient Estimation in Semiparametric Time Series: The ACD Model," Discussion paper, CentER, Tilburg University.

Engle, R. F. (2000): "The Econometrics of Ultra-High-Frequency Data," *Econometrica*, 68, 1, 1–22.

Engle, R. F., and J. Russell (1998): "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127–1162.

Gallant, R. A. (1981): "On the Bias in Flexible Functional Forms and an Essential Unbiased Form: The Fourier Flexible Form," *Journal of Econometrics*, 15, 211–245.

Gamerman, D., and M. West (1987): "An Application of Dynamic Survival Models in Unemployment Studies," *The Statistician*, 36, 269–274.

Gerhard, F. (2001): "A Simple Dynamic for Limited Dependent Variables," Discussion Paper W11, Nuffield College, Oxford.

Gerhard, F., and N. Hautsch (2000): "Determinants of Inter-Trade Durations and Hazard Rates Using Proportional Hazard ARMA Models," Discussion Paper 00/20, Center of Finance and Econometrics.

——— (2001): "Volatility Estimation on the Basis of Price Intensities," *forthcoming Journal of Empirical Finance.*

Gorgens, T., and J. L. Horowitz (1999): "Semiparametric Estimation of a Censored Regression Model with an Unknown Transformation of the Dependent Variable," *Journal of Econometrics*, 90, 155–191.

Gourieroux, C., A. Monfort, E. Renault, and A. Trognon (1987): "Generalised Residuals," *Journal of Econometrics*, 34, 5–32.

Gritz, R. M. (1993): "The Impact of Training on the Frequency and Duration of Employment," *Journal of Econometrics*, 57, 21–51.

Hajivassiliou, V. A., and P. A. Ruud (1994): "Classical Estimation Methods for LDV Models Using Simulation," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. L. McFadden, chap. 40, pp. 2383–2441. Elsevier Science B.V.

Hamilton, J. D. (1994): *Time Series Analysis.* Princeton University Press.

Han, A., and J. A. Hausman (1990): "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 5, 1–28.

Harvey, A. C. (1990): *The Econometric Analysis of Time Series.* Philip Allan, Hemel Hempstead, 2nd. edn.

Hautsch, N. (1999): "Analyzing the Time Between Trades with a Gamma Compounded Hazard Model. An Application to LIFFE Bund Future Transctions," Discussion Paper 99/03, CoFE, Konstanz.

Heckmann, J. J., and B. Singer (1984): "Econometrics Duration Analysis," *Journal of Econometrics*, 24, 63–132.

Honore, B. E. (1990): "Simple Estimation of Duration Model with Unobserved Heterogeneity," *Econometrica*, 58, 453–473.

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

Horowitz, J. L. (1999): "Semiparametric Estimation of a Proportional Hazard Model with Unobserved Heterogeneity," *Econometrica*, 67, 1001–1028.

Horowitz, J. L., and G. R. Neumann (1987): "Semiparametric Estimation of Employment Duration Models," *Econometric Review*, 6(1), 5–40.

——— (1989): "Specification Testing in Censored Regression Models: Parametric and Semiparametric Methods," *Journal of Applied Econometrics*, 4, 61–86.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994): *Distributions in Statistics: Continuous Univariate Distributions (2. Vol.).* Wiley, New York.

Kalbfleisch, J. D., and R. L. Prentice (1973): "Marginal Likelihoods Based on Cox's Regression and Life Model," *Biometrika*, 60, 267–278.

Kalbfleisch, J. D., and R. L. Prentice (1980): *The Statistical Analysis of Failure Time Data.* Wiley.

Kaplan, E. L., and P. Meier (1958): "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 53, 457–481.

Kiefer, N. M. (1988): "Economic Duration Data and Hazard Functions," *Journal of Economic Literature*, 26, 646–679.

Lancaster, T. (1979): "Econometric Methods for the Duration of Unemployment," *Econometrica*, 47 (4), 939–956.

——— (1997): *The Econometric Analysis of Transition Data.* Cambridge University Press.

Manrique, A., and N. Shephard (1998): "Simulation-Based Likelihood Inference for Limited Dependent Processes," *Econometrics Journal*, 1, C174–C202.

McCall, B. P. (1996): "Unemployment Insurance Rules, Joblessness and Part-Time Work," *Econometrica*, 64 (3), 647–682.

Meyer, B. D. (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58 (4), 757–782.

Moffitt, R. (1985): "Unemployment Insurance and the Distribution of Unemployment Spells," *Journal of Econometrics*, 28, 85–101.

Neumann, G. (1997): *Search Models and Duration Data,* chap. 7, pp. 300–351, Handbook of Applied Econometrics. M.H. Pesaran, ed. (Basil Blackwell), Oxford.

Nickell, S. (1979): "Estimating the Probability of Leaving Unemployment," *Econometrica*, 47(5), 1249–1266.

Oakes, D. (1972): "Contribution to the Discussion of the Paper by D. R. Cox," *Journal of the Royal Statistical Society, Series B*, 34, 208.

Oakes, D. (2001): "Biometrika Centenary: Survival Analysis," *Biometrika*, 88, 99–142.

Oakes, D., and L. Cui (1994): "On Semiparametric Inference for Modulated Renewal Processes," *Biometrika*, 81, 83–90.

Orbe, J., E. Ferreira, and V. Nunez-Anton (2000): "A Censored Partial Regression Model for Lifetime Data Analysis," Discussion paper, University del Pais Vasco, Spain.

Prentice, R. L., and L. A. Gloeckler (1978): "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data," *Biometrics*, 34, 57–67.

Romeo, C. J. (1999): "Conducting inference in semiparametric duration models und inequality restrictions on the shape of the hazard implied by job search theory," *Journal of Applied Econometrics*, 14, 587–605.

Ruggiero, M. (1994): "Bayesian Semiparametric Estimations of Proportional Hazards Models," *Journal of Econometrics*, 62, 277–300.

Sueyoshi, G. T. (1995): "A Class of Binary Response Models for Grouped Duration Data," *Journal of Applied Econometrics*, 10, 411–431.

Thompson, W. A. (1977): "On the treatment of grouped durations in life studies," *Biometrics*, 33, 463–470.

van den Berg, G. J., and B. van der Klaauw (2001): "Combining micro and macro unemployment duration data," *Journal of Econometrics*, 102, 271–309.

West, M., P. J. Harrison, and H. S. Migon (1985): "Dynamic Generalized Linear Models and Bayesian Forecasting," *Journal of the American Statistical Association*, 80, 73–83.