ISSN 1440-771X

MONASH University

Australia

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Semiparametric estimation of duration models when the parameters are subject to inequality constraints and the error distribution is unknown

Kulan Ranasinghe and Mervyn J. Silvapulle

June 2008

Working Paper 5/08

Semiparametric estimation of duration models when the parameters are subject to inequality constraints and the error distribution is unknown

By Kulan Ranasinghe and Mervyn J. Silvapulle

Department of Econometrics and Business Statistics, Monash University, P.O. Box 197, Caulfield East, Melbourne, Australia 3145.

kulan.ranasinghe@buseco.monash.edu.au, mervyn.silvapulle@buseco.monash.edu.au,

Abstract

This paper proposes a semiparametric method for estimating duration models when there are inequality constraints on some parameters and the error distribution may be unknown. Thus, the setting considered here is particularly suitable for practical applications. The parameters in duration models are usually estimated by a quasi-MLE. Recent advances show that a semiparametrically efficient estimator[SPE] has better asymptotic optimality properties than the QMLE provided that the parameter space is unrestricted. However, in several important duration models, the parameter space is restricted, for example in the commonly used linear duration model some parameters are non-negative. In such cases, the SPE may turn out to be outside the allowed parameter space and hence are unsuitable for use. To overcome this difficulty, we propose a new constrained semiparametric estimator. In a simulation study involving duration models with inequality constraints on parameters, the new estimator proposed in this paper performed better than its competitors. An empirical example is provided to illustrate the application of the new constrained semiparametric estimator and to show how it overcomes difficulties encountered when the unconstrained estimator of nonnegative parameters turn out to be negative.

Key Words: Adaptive inference; Conditional duration model; Constrained inference; Efficient semiparametric estimation; Order restricted inference; Semiparametric efficiency bound.

JEL Classification: C41, C14.

1 Introduction

The availability of intraday tick-by-tick financial data increased substantially during the past two decades, which in turn has had a phenomenal impact on research in financial market microstructure. Such high frequency data are usually analyzed using essentially two classes of models: generalized autoregressive conditional heteroscedasticity [GARCH] models and duration models. In GARCH type models, the response variable is observed at equally spaced time points. An example is the hourly Dow-Jones index. By contrast, in duration models, the time elapsed between two consecutive events, such as financial transactions, is the response variable. A range of so called duration models has been proposed and studied in the literature to model the data generating process of durations. The class of such models forms an essential tool for the study of market microstructure (Bauwens and Giot 2001).

To introduce the basics of the duration model, let X_i denote the duration between $(i-1)^{th}$ and the i^{th} events, $\mathcal{F}_i$ denote the information up to and including time i, and $\psi_i = E(X_i \mid \mathcal{F}_{i-1})$, the expected duration. A duration model is usually expressed as $X_i = \psi_i \varepsilon_i$ where ε_i is referred to as the *error term* which is assumed to satisfy $E(\varepsilon_i) = 1$ to ensure identifiability of the model. The main objective of duration analysis is to model ψ_i as a function of $\{\ldots, X_{i-2}, X_{i-1}; \ldots, \psi_{i-2}, \psi_{i-1}\}$. For example, a special case of the well-known linear autoregressive conditional duration[ACD] model of Engle and Russell (1998) is the following ACD(1,1) model:

$$\psi_i = \alpha + \beta X_{i-1} + \gamma \psi_{i-1}, \quad \alpha \geq 0, \beta \geq 0, \gamma \geq 0. \tag{1}$$

More generally, the model may take the form $\psi_i = g(\ldots, X_{i-1}; \ldots, \psi_{i-1}; \boldsymbol{\theta})$ where g is a given function and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^T$ is an unknown parameter. Further, g may also depend on exogenous variables.

The objectives of this paper are the following:

1. To propose a method of estimating the unknown parameter $\boldsymbol{\theta}$ in duration models

when the error distribution is unknown and there are inequality constraints on some parameters, for example, some parameters may be nonnegative as in the foregoing ACD model (1).

2. To compare the proposed new constrained estimator with a semiparametrically efficient estimator and the standard quasi-maximum likelihood estimator.

For simplicity, let us temporarily assume that the error terms, $\varepsilon_1, \ldots, \varepsilon_n$, are independently and identically distributed with f denoting their common probability density function[pdf]. If f is known then the model can be estimated by maximum likelihood (for example, see Bauwens and Giot 2000). Since f is usually unknown, the quasi maximum likelihood estimator[QMLE], which is equal to the MLE corresponding to exponential distribution for the error terms, is the standard choice. However, such a QMLE is not necessarily the most efficient if f deviates from the exponential distribution and/or the error terms are not independent. This is important because the time-series nature of $\{X_i\}$ suggests that the error terms $\{\varepsilon_i\}$ are unlikely to be independent and identically distributed with a known density function.

Recently, Drost and Werker (2004) proposed a *semiparametrically efficient* estimator of the unknown parameeter $\boldsymbol{\theta}$ in the duration model when the error distribution is unknown and $\varepsilon_1, \ldots, \varepsilon_n$ may not be independent. In this context, "semiparametrically efficient" essentially means that the estimator has the highest possible asymptotic efficiency in the class of all asymptotically normal estimators. Detailed accounts of this topic are given in Bickel *et al.* (1993), Tsiatis (2006) and Kosorok (2008).

By definition, duration X_i is nonnegative, and hence $\psi_i \geq 0$. Consequently, the parameters α, β and γ in (1) must be nonnegative as well (Nelson and Cao 1992). Further, we also have $\alpha + \beta \leq 1$. However, the Drost-Werker[DW]-estimator does not incorporate such inequality constraints and hence it may turn out to be negative even when the true parameter is known to be nonnegative. If the DW-estimators $\hat{\beta}$ and $\hat{\gamma}$ turn out to be negative, one may be tempted to simply truncate and redefine them as $\hat{\beta} = \hat{\gamma} = 0$. Such a method of truncating estimators is crude, particularly because there is already a well-developed body of statistical theory for incorporating such inequality constraints (Silvapulle and Sen, 2005). The literature on statistical inference under inequality constraints, also known as *order restrictions*, is quite extensive indeed. Some recent relevant references are El Barmi and Mukerjee (2005), El Barmi *et al.* (2006), Peddada *et al.* (2005), Peddada *et al.* (2006), Hwang and Peddada (1994), and Silvapulle and Sen (2005).

In this paper, we propose a new *constrained semiparametric estimator* $\bar{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ when some components of $\boldsymbol{\theta}$ are known to be non-negative, or more generally when there are constraints of the form $\boldsymbol{h}(\boldsymbol{\theta}) \geq \mathbf{0}$ where $\boldsymbol{h}$ is a vector function. A feature of our constrained estimator is that if the DW-estimator satisfies the inequality constraints on the parameters, then the two estimators are the same. Otherwise, the constrained estimator is the point on the boundary of the parameter space that is "closest" to $\hat{\boldsymbol{\theta}}$ in some sense. A theoretical result in section 2.2 provides the asymptotic distribution of our inequality constrained estimator $\bar{\boldsymbol{\theta}}$ and shows that it is closer to the true value than the unconstrained DW-estimator $\hat{\boldsymbol{\theta}}$.

The main findings of a simulation study to compare the foregoing estimators may be summarized as follows:

1. *There are inequality constraints on the parameters of the duration models:* The constrained semiparametric estimator $\bar{\boldsymbol{\theta}}$ introduced in this paper is better than the corresponding unconstrained semiparametrically efficient estimator $\hat{\boldsymbol{\theta}}$.

2. *There are inequality constraints on the parameters and the errors do not satisfy the condition that they are iid with common distribution exp(1):* If the true parameter does not lie in a small particular region of the parameter space, which we shall refer to as A, then our proposed estimator is better than the QMLE and the DW-estimator. In several published empirical studies (see later) we observed that the estimators were not in the region A. Therefore, overall the constrained semiparametric estimator $\bar{\boldsymbol{\theta}}$ is

better than the unconstrained DW-estimator and the constrained QMLE.

3. *The errors are iid and their common distribution is exponential:* In this ideal case, which will serve as a benchmark, the QMLE is equal to the MLE and hence one would expect that the QMLE would be the best. The simulation results are consistent with this, but the differences between QMLE and the semiparametric estimators turned out to be generally small.

We conclude that, when there are constraints on parameters, (i) a theoretical result shows that the estimator proposed in this paper is asymptotically better than the semiparametrically efficient estimator, and (ii) in a large scale simulation study, the estimator proposed in this paper performed better than the 'gold standard' QMLE and the semiparametrically efficient estimator, which corroborates the aforementioned theoretical result. Therefore, the estimator proposed in this paper deserves serious consideration for estimating duration models when there are inequality constraints on parameters.

The plan of the paper is as follows. Section 2 discusses the methodological aspects. In subsection 2.1, we recall some known results on efficient semiparametric inference, and in subsection 2.2 we develop the methodological aspects and propose new inequality constrained semiparametric estimators. Section 3 provides the results of a simulation study, section 4 provides an empirical example to illustrate the new constrained semiparametric estimator, and section 5 concludes.

2 Semiparametric Estimation of Duration Models

As in the previous section, X_i denotes the i^{th} observation of a duration variable X, $\mathcal{F}_i$ denotes the information up to and including the i^{th} observation X_i, $\psi_i = E(X_i \mid \mathcal{F}_{i-1})$ and $\varepsilon_i = X_i/\psi_i$. Fernandes and Grammig (2006) provided a survey of such duration models. Two examples with inequality constraints on parameters are given below.

1. Linear ACD Model: $\quad \psi_i = \alpha + \beta X_{i-1} + \gamma\psi_{i-1}$

2. Power ACD Model: $\psi_i^\lambda = \alpha + \beta X_{i-1}^\lambda + \gamma\psi_{i-1}^\lambda$

Let $\boldsymbol{\theta}$ denote the unknown parameter in the duration model; for example, $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$ for the linear ACD(1,1) model in (1). Within the framework of this paper we do not assume that the error distribution belongs to any known parametric family. Hence $\boldsymbol{\theta}$ does not include parameters of the error distribution. To ensure that the parameters are identified, we assume that $E(\varepsilon_i \mid \mathcal{F}_{i-1}) = 1$. Usually, the errors are assumed to be independently and identically distributed [iid] for simplicity. However, the nature of the durations in practice suggests that this is unlikely to be the case in most practical situations and hence it would be desirable for the method of inference to be robust against violation of the assumption of *iid* errors. To this end, let $\mathcal{H}_{i-1} \subset \mathcal{F}_{i-1}$ and assume that the conditional distribution of ε_i given the past depends only on the information in the set $\mathcal{H}_{i-1}$. Thus, the smaller information set $\mathcal{H}_{i-1}$ contains the relevant past variables that are assumed to affect the distribution of ε_i given the past. Now, with $\psi_i = E(X_i \mid \mathcal{F}_{i-1})$, the semiparametric model is defined formally by

$$X_i = \psi_i\varepsilon_i,\ \psi_i = g(\ldots, X_{i-1}; \ldots, \psi_{i-1}; \boldsymbol{\theta}),\ \text{and } \mathcal{L}(\varepsilon_i \mid \mathcal{F}_{i-1}) = \mathcal{L}(\varepsilon_i \mid \mathcal{H}_{i-1}) \quad (2)$$

where g is a known function and $\mathcal{L}(\varepsilon_i \mid \mathcal{F}_{i-1})$ refers to the distribution of ε_i given $\mathcal{F}_{i-1}$. The special case of independently and identically distributed errors is obtained by setting $\mathcal{H}_i$ equal to the trivial field $\{\phi, \Omega\}$.

The next subsection provides the essentials on semiparametric inference, and states the relevant results in a concise form. For convenience, previously known results are discussed in the next subsection and the new methodological developments are given in subsection 2.2

2.1 Semi-parametric Estimation

Let f_i denote the probability density function [pdf] corresponding to $\mathcal{L}(\varepsilon_i \mid \mathcal{H}_{i-1})$. We shall assume that f_i is smooth, for example, it has continuous first derivative. It follows that the conditional pdf of X_i given $\mathcal{F}_{i-1}$ is $\psi_i^{-1} f_i(x/\psi_i)$ and hence the loglikelihood $\ell(\boldsymbol{\theta})$ is given by $\ell(\boldsymbol{\theta}) = \sum \ell_i(\boldsymbol{\theta}),$ where $\ell_i(\boldsymbol{\theta}) = \ln\{\psi_i^{-1} f_i(X_i/\psi_i)\}$. If f_i were known, then the maximum likelihood estimator [MLE] of $\boldsymbol{\theta}$ would be $argmax_{\boldsymbol{\theta}}\ \ell(\boldsymbol{\theta})$ and it would be asymptotically efficient. In practice, f_i is usually unknown. In this setting, the model is semiparametric and $\boldsymbol{\theta}$ can be estimated consistently by a quasi maximum likelihood estimator[QMLE] obtained by choosing the quasi likelihood equal to the loglikelihood when f_i is the exponential distribution with unit mean (see Bauwens and Giot 2001). Efficient estimation in general semiparametric models has a specialized but a growing literature (see Tsiatis 2006, Kosorok 2008). An important result in this area is that a desirable estimator of an unknown finite dimensional parameter $\boldsymbol{\theta}$ in semiparametric models is the so called, *semiparametrically efficient estimator*, which essentially means that the estimator of $\boldsymbol{\theta}$ is efficient in some sense for the model with the density function of errors treated as an unknown nuisance function. Detailed discussions of such estimators and their relevance for inference are also given in Newey (1990). In this subsection, we shall state the main relevant results, without the technical details or proofs.

To introduce the semiparametrically efficient estimator, first let us suppose that the error density function is known. Let $\dot{g}(\boldsymbol{\theta})$ denote $(\partial/\partial\boldsymbol{\theta})g(\boldsymbol{\theta})$ for any function g, and let $\tilde{\boldsymbol{\theta}}$ denote a $n^{1/2}$-consistent estimator of $\boldsymbol{\theta}$, for example it could be the QMLE introduced in section 1. Let us note that this QMLE is $n^{1/2}$-consistent under a very broad range of conditions, for example the error distribution may not be exp(1). The estimator, $\{\tilde{\boldsymbol{\theta}} + \{n^{-1}\Sigma_{i=1}^{n}\dot{\ell}_i(\tilde{\boldsymbol{\theta}})\dot{\ell}_i(\tilde{\boldsymbol{\theta}})^\top\}^{-1}n^{-1}\Sigma_{i=1}^{n}\dot{\ell}_i(\tilde{\boldsymbol{\theta}})\}$, is called the *one-step estimator*. It is asymptotically equivalent to the MLE, and is obtained by applying a Newton-Raphson type iteration once, starting from any $n^{1/2}$-consistent estimator such as $\tilde{\boldsymbol{\theta}}$ (see Bickel *et al.* 1993).

Now, let us temporarily relax the assumption that the error density function is known. Consequently, $\dot{\ell}_i$ in the foregoing expression for the one-step estimator is also unknown. The main approach in semiparametrically efficient estimation involves replacing this unknown function by $\tilde{\dot{\ell}}_i^*$ which is a 'suitable' estimator of the so called *semiparametrically efficient score function*, which we denote by $\dot{\ell}_i^*(\boldsymbol{\theta})$. The exact form of $\dot{\ell}_i^*(\boldsymbol{\theta})$ would depend on the particular assumptions made about the information set $\mathcal{H}_i$ in (2). In this paper, we shall assume that $\mathcal{H}_i = \mathcal{F}_i$ so that quite minimal assumptions are made about the serial dependence of the error terms; this is particularly suitable for empirical studies. Drost and Werker (2004) referred to this as 'martingale error' structure. In this case,

$$\dot{\ell}_i^*(\boldsymbol{\theta}) = \{(\varepsilon_i - 1)/\mathrm{var}(\varepsilon_i|\mathcal{H}_{i-1})\}(\partial/\partial\boldsymbol{\theta})\log(\psi_i). \tag{3}$$

In this paper, we propose to compute the residual $\tilde{\epsilon}_i$ as $X_i/\psi(\tilde{\boldsymbol{\theta}})$, and define $\tilde{\dot{\ell}}_i^*$ as the sample analogue of $\dot{\ell}_i^*$. This leads to the semiparametrically efficient estimator,

$$\hat{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}} + \left(n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})^\top\right)^{-1} n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}}). \tag{4}$$

We shall refer to this as the Drost-Werker estimator[DW-estimator]. See Drost and Werker (2004) for more details.

Let us note that with the stronger assumptions $\mathcal{H}_i$ equal to $\{\phi, \Omega\}$ and $\sigma(\varepsilon_i)$, which Drost and Werker (2004) referred to as *iid* and *markov* error structures, we obtain different estimators. In our simulation studies we also evaluated these two forms, but they did not perform as well as the aforementioned estimator with martingale errors. The complete simulation results will be reported elsewhere, but for this paper we restrict to the best of the three, namely the one defined by (3, 4).

2.2 Estimation subject to inequality constraints

For the linear ACD(1,1) model $\psi_i = \alpha + \beta X_{i-1} + \gamma\psi_{i-1}$, we have that $\alpha \geq 0$, $\beta \geq 0$ and $\gamma \geq 0$, because $\psi_i \geq 0$ and $X_i \geq 0$ for every i (Nelson and Cao 1992). However, the semiparametrically efficient estimator (4) may not satisfy such inequality constraints. In this section, we modify the approach in Drost and Werker (2004) to ensure that such constraints are satisfied. To this end we adopt results from constrained statistical inference (Silvapulle and Sen 2005). There is no unique way to define suitable constrained estimators. In what follows we propose a suitable method and provide theoretical results to support the proposed method.

Let Θ denote the parameter space of $\boldsymbol{\theta}$. We shall assume that Θ is defined by various combinations of constraints of the form $g(\boldsymbol{\theta}) \geq 0$ and $h(\boldsymbol{\theta}) = 0$ where g and h are continuously differentiable functions of $\boldsymbol{\theta}$. For example, Θ could be of the form $\{\boldsymbol{\theta} : g_1(\boldsymbol{\theta}) \geq 0, \ldots, g_k(\boldsymbol{\theta}) \geq 0, h_1(\boldsymbol{\theta}) = 0, \ldots, h_m(\boldsymbol{\theta}) = 0\}$, where $g_1, \ldots, g_k, h_1, \ldots, h_m$ are continuously differentiable functions. The parameter space for the linear ACD models are of this form (Nelson and Cao 1992). More precisely, we require the parameter space to be *Chernoff Regular*(see Silvapulle and Sen (2005)). Further, we make the mild assumption that $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} \boldsymbol{Z}$ where $\boldsymbol{Z} \sim N(\mathbf{0}, V)$ for some positive definite matrix V where $\hat{\boldsymbol{\theta}}$ is the DW-estimator.

To motivate the ideas underlying the constrained estimator to be introduced, let us temporarily suppose that $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ is distributed exactly as $N(\mathbf{0}, V)$. Therefore, we may treat $\hat{\boldsymbol{\theta}}$ as a single observation from $N(\boldsymbol{\theta}_0, n^{-1}V)$. The corresponding log likelihood is $(-1/2)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top V^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ and the MLE of $\boldsymbol{\theta}_0$ is

$$\bar{\boldsymbol{\theta}}^* = \arg\min_{\boldsymbol{\theta}\in\Theta}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top V^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}). \tag{5}$$

Therefore, $\bar{\boldsymbol{\theta}}^*$ is the projection of $\hat{\boldsymbol{\theta}}$ onto Θ with respect to the inner product $\langle \boldsymbol{x}, \boldsymbol{y}\rangle_V = \boldsymbol{x}^\top V^{-1}\boldsymbol{y}$. The left panel in Figure 1 illustrates this for the simple case of two-dimensions and Θ equal to the first quadrant $\{\theta_1 \geq 0, \theta_2 \geq 0\}$.

Now, let us relax the assumption that $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ is distributed exactly as $N(\mathbf{0}, V)$ and assume that the latter is only the limiting distribution and that V is unknown. Then, motivated by the definition of $\bar{\boldsymbol{\theta}}^*$ in (5), a natural constrained semiparametric estimator is

$$\bar{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\Theta}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top W_n^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \tag{6}$$

where W_n is positive definite. In general, we would choose W_n to be a consistent estimator of V, for example, $W_n = \left(n^{-1}\sum_{i=1}^n \tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})^\top\right)^{-1}$.

Now, to discuss the theoretical results on $\bar{\boldsymbol{\theta}}$, let us briefly recall some definitions. Let $\mathcal{T}(\Theta; \boldsymbol{\theta}_0)$ denote the *tangent cone* (also known as *cone of tangents*) of Θ at $\boldsymbol{\theta}_0$ (see Silvapulle and Sen 2005). Intuitively, the tangent cone $\mathcal{T}(\Theta; \boldsymbol{\theta}_0)$ is constructed as follows: First,

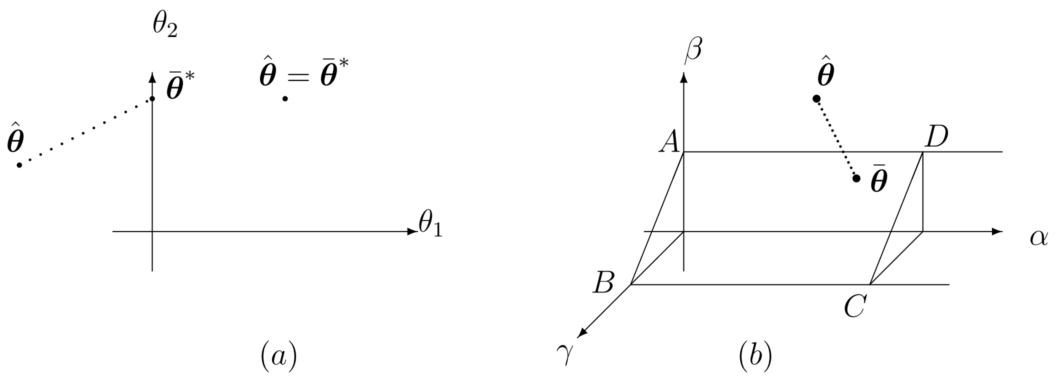

Figure 1: (a) The unconstrained estimator $\hat{\boldsymbol{\theta}}$ and the constrained estimator $\bar{\boldsymbol{\theta}}^*$ of $\boldsymbol{\theta}_0$ subject to $\boldsymbol{\theta} \in \Theta = \{(\theta_1, \theta_2) : \theta_1 \geqslant 0, \theta_2 \geqslant 0\}$ for two possible values of $\hat{\boldsymbol{\theta}}$, one in Θ and the other outside Θ in the second quadrant, when $V = (1, 0.5 \mid 0.5, 1)$. (b) The unconstrained estimator $\hat{\boldsymbol{\theta}}$ and the constrained estimator $\bar{\boldsymbol{\theta}}$ subject to $\boldsymbol{\theta} \in \Theta = \{(\alpha, \beta, \gamma) : \alpha \geq 0, \beta \geq 0, \gamma \geq 0, \beta + \gamma \leq 1\}$ with $\hat{\boldsymbol{\theta}}$ lying outside Θ and $\bar{\boldsymbol{\theta}}$ lying on the face spanned by the rectangle $ABCD$ of the wedge-shaped Θ.

approximate the boundaries of Θ at $\boldsymbol{\theta}_0$ by tangents, and then approximate Θ by the cone, $\mathcal{A}(\Theta; \boldsymbol{\theta}_0)$, formed by these tangents. This is called the *approximating cone* of Θ at $\boldsymbol{\theta}_0$. Now, translate the parameter space so that $\boldsymbol{\theta}_0$ moves to the origin. Consequently, the approximating cone becomes the tangent cone with its vertex at the origin. These are illustrated in Figure 2.

For any $\boldsymbol{x} \in \mathbb{R}^p$, a $p \times p$ positive definite matrix W and a set $\mathcal{C}$, let $\|\boldsymbol{x}\|_W = \{\boldsymbol{x}^\top W^{-1} \boldsymbol{x}\}^{1/2}$ and $\Pi_W\{\boldsymbol{z} \mid \mathcal{C}\} = \arg\min_{\boldsymbol{\theta} \in \mathcal{C}} \|\boldsymbol{z} - \boldsymbol{\theta}\|_W$. Thus, $\Pi_W\{\boldsymbol{z} \mid \mathcal{C}\}$ denotes the projection of $\boldsymbol{z}$ onto $\mathcal{C}$ with respect to the inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_W = \boldsymbol{x}^\top W^{-1} \boldsymbol{y}$. A simple illustration of $\Pi(\hat{\boldsymbol{\theta}} \mid \mathcal{C})$, which is equal to $\bar{\boldsymbol{\theta}}^*$, is given in Figure 1 when $\mathcal{C}$ is the positive orthant in two dimensions.

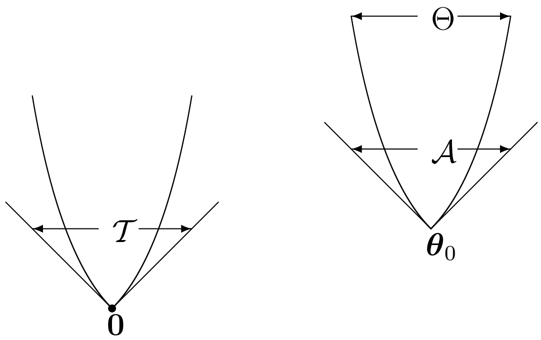

Figure 2: The *approximating cone* $\mathcal{A}$ of Θ at $\boldsymbol{\theta}_0$ and the corresponding *tangent cone* $\mathcal{T}$ with its vertex at the origin, $\mathbf{0}$.

Now, we provide a result about the distribution of $\bar{\boldsymbol{\theta}}$.

Proposition 1. *Suppose that* Θ *is convex,* $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} \boldsymbol{Z}$ *where* $\boldsymbol{Z} \sim N(\boldsymbol{0}, V)$ *for some positive definite matrix* V *and that* $W_n \xrightarrow{p} W$ *where* W *and* W_n *are positive definite. Then*

$$n^{1/2}(\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} \Pi_W\{\boldsymbol{Z} \mid \mathcal{T}(\Theta; \boldsymbol{\theta}_0)\} \tag{7}$$

where $\bar{\boldsymbol{\theta}}$ *is the constrained estimator defined in (6). Further,* $\bar{\boldsymbol{\theta}}$ *is closer to the true value* $\boldsymbol{\theta}_0$ *than* $\hat{\boldsymbol{\theta}}$ *in the following sense:*

$$pr\{\|\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\|_{W_n} \leq \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\|_{W_n}\} = 1, \tag{8}$$

$$pr(\bar{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}) \to pr(\boldsymbol{Z} \in \mathcal{T}(\Theta; \boldsymbol{\theta}_0), \tag{9}$$

$$pr\{\|\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\|_{W_n} < \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\|_{W_n}\} \to pr(\boldsymbol{Z} \notin \mathcal{T}(\Theta; \boldsymbol{\theta}_0). \tag{10}$$

In the rest of this subsection, we shall comment on other possible alternatives to the foregoing approach. The general approach to constructing a constrained estimator exploits the fact that one needs to use only the local behavior of the objective function in an $n^{-1/2}$-neighborhood of the true value $\boldsymbol{\theta}_0$. The foregoing $\bar{\boldsymbol{\theta}}$ adopts this approach. It is also possible to construct other similar estimators. For example, another estimator may be defined as $\hat{\boldsymbol{\theta}}(\lambda_0)$ where $\hat{\boldsymbol{\theta}}(\lambda) = \big[\tilde{\boldsymbol{\theta}} + \lambda\big(n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})^\top\big)^{-1} n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})\big]$ for $0 \leq \lambda \leq 1$ and λ_0 is the maximum value of λ in $[0, 1]$ for which $\hat{\boldsymbol{\theta}}(\lambda)$ lies in Θ. This says that the one-step iteration in (4) moves from $\tilde{\boldsymbol{\theta}}$ in the direction suggested by the DW-estimator but stops before crossing the boundary of Θ.

Another estimator may be defined as $\arg\max_{\boldsymbol{\theta}\in\Theta}\ q(\boldsymbol{\theta})$ where

$$q(\boldsymbol{\theta}) = (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}}) - 2^{-1}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top\Big(n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})^\top\Big)(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}),$$

which may be seen as a pseudo likelihood with score function $n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})$ and information $\big(n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})^\top\big)$. Since the unconstrained maximum of $q(\boldsymbol{\theta})$ is the DW-estimator $\hat{\boldsymbol{\theta}}$, the foregoing estimator $\arg\max_{\boldsymbol{\theta}\in\Theta}\ q(\boldsymbol{\theta})$ can be seen as a constrained version of the DW-estimator. This estimator turns out to be the same as $\bar{\boldsymbol{\theta}}$ in (6) if the W_n in (6) is equal to $\left(n^{-1}\sum_{i=1}^{n}\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})\tilde{\dot{\ell}}_i^*(\tilde{\boldsymbol{\theta}})^\top\right)^{-1}$.

3 Simulation Study

In this section, we report the results of a simulation study conducted to compare $\hat{\boldsymbol{\theta}}$, $\bar{\boldsymbol{\theta}}$ and the standard QMLE; in this study we used the constrained QMLE, which is the maximiser of the quasilikelihood under the inequality constraints.

Design of the study:

We studied the two duration models introduced at the beginning of section 2. For each of these models, the following error distributions were studied:

$$(a)\ \varepsilon_i \sim exp(1), \qquad (b)\ \varepsilon_i \sim \Gamma(\lambda_i^{-2}, \lambda_i^2) \quad \text{and} \quad (c)\ \varepsilon_i \sim LN(-2^{-1}log(1+\lambda_i^2), log(1+\lambda_i^2)),$$

where $\Gamma(a,b)$ is the Gamma distribution with parameters (a,b), and $LN(\mu,\sigma^2)$ is the lognormal distribution. For the purpose of this simulation study, these distributions are particularly relevant. The case $\varepsilon_i \sim exp(1)$ is important because, it is the ideal setting, for example, its role in duration models is similar to that of the normal distribution in linear regression analysis. The gamma distribution was chosen as a more flexible and general alternative to the exponential distribution, and also because this is the most general form for which the semiparametrically efficient estimator is adaptive and hence has the same asymptotic efficiency as the MLE provided that the true parameter is an interior point. The choice $(\lambda_i^{-2}, \lambda_i^2)$ for the parameters of the gamma distribution ensures that ϵ_i has mean 1 as required by the usual standardization for identifiability of the duration model. The lognormal distribution provides a departure from the exponential and gamma distributions so that the performance of the estimators may be evaluated under conditions that are not ideal for the QMLE and the semiparametrically efficient estimator. A dynamic structure on λ_i allows us to depart from the usual desired assumption that the error terms are iid, and evaluate the reliability

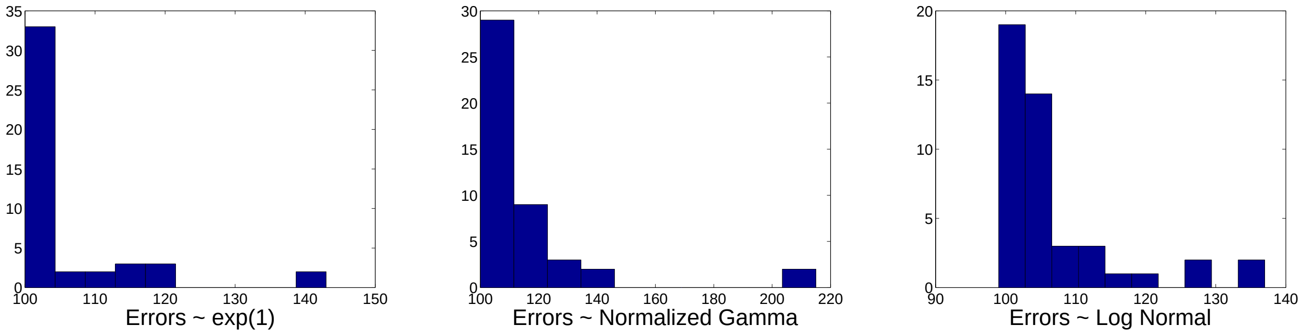

Figure 3: MSE-efficiency of $\bar{\boldsymbol{\theta}}$ relative to $\hat{\boldsymbol{\theta}}$ for the ACD model.

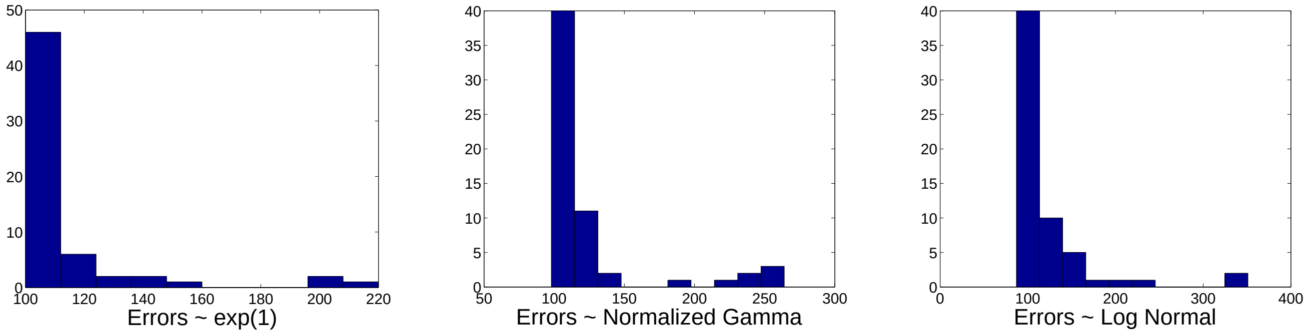

Figure 4: MSE-efficiency of $\bar{\boldsymbol{\theta}}$ relative to $\hat{\boldsymbol{\theta}}$ for the PACD model.

of the semiparametric estimator in the presence of unknown dynamics in the distribution of ϵ_i.

For the gamma and lognormal error distributions in the foregoing settings (b) and (c), we set $\lambda_i^2 = 0.1 + 0.9\varepsilon_{i-1}$. The estimation methods that are compared in this paper do not require the exact form of dependence of λ_i on other variables. This would enable us to compare the estimators when the errors are not *iid*.

Without loss of generality, the unconditional mean of X_i was set equal to 1. All the computations were programmed in MATLAB, and the optimizations were carried out using the optimization toolbox in MATLAB. Since our main objective is to compare the QMLE with the semiparametric estimators, we shall report estimates of relative MSE Efficiency which we define as {MSE of QMLE/ MSE of the estimator}. The results of the simulation study are based on sample size $n = 500$ and 500 repeated samples.

Table 1: MSE-efficiency of $\bar{\boldsymbol{\theta}}$ relative to QMLE for the linear ACD model

True value			$\varepsilon \sim EXP$			$\varepsilon \sim NG$			$\varepsilon \sim LN$		
α_0	β_0	γ_0	α	β	γ	α	β	γ	α	β	γ
0.05	0.30	0.65	103	96	97	179	182	182	153	147	151
0.05	0.05	0.90	99	96	95	156	193	162	143	194	149
0.10	0.20	0.70	106	99	101	174	188	173	144	164	148
*0.25	0.05	0.70	58	96	61	78	162	86	65	212	76
0.10	0.15	0.75	109	99	103	169	195	170	148	174	151
0.05	0.10	0.85	102	97	97	238	207	209	181	184	174
0.20	0.20	0.60	104	101	99	149	168	145	127	155	132
*0.20	0.05	0.75	76	95	76	89	170	98	79	215	91
*0.30	0.10	0.60	76	98	78	86	166	89	78	166	85
0.10	0.10	0.80	104	98	98	147	196	151	138	184	143
0.70	0.20	0.10	87	103	90	107	153	103	111	139	114
0.70	0.25	0.05	88	104	94	150	156	147	122	145	127
0.80	0.10	0.10	82	100	83	106	172	98	101	174	97
0.80	0.12	0.08	86	103	87	120	166	110	106	171	103
0.80	0.15	0.05	89	103	91	143	165	131	112	164	110

MSE-efficiency for θ_i is defined as MSE(QMLE)/MSE($\bar{\boldsymbol{\theta}}$).

Results:

The histograms of the MSE of $\bar{\boldsymbol{\theta}}$ relative to $\hat{\boldsymbol{\theta}}$ and QMLE are shown in Figures 3 - 4. Each figure has three histograms: the one on left, middle and right correspond to ε_i being exp(1), $\Gamma(\lambda_i^{-2}, \lambda_i^2)$ and $LN(-2^{-1}log(1+\lambda_i^2), log(1+\lambda_i^2))$, respectively. Now, let us summarise the main observations.

(A) *Comparison of the constrained semiparametric estimator $\bar{\boldsymbol{\theta}}$ with QMLE:*

First, let us consider the case when the errors are *iid* with common error distribution exp(1). Recall that the QMLE is equal to the MLE in this case. Since this setting is ideal for QMLE, we would expect the QMLE to perform at least as well as, if not better than,

Table 2: MSE-efficiency of $\bar{\boldsymbol{\theta}}$ relative to QMLE for the linear Power ACD model

True Value				$\varepsilon \sim EXP$				$\varepsilon \sim NG$				$\varepsilon \sim LN$			
α_0	β_0	γ_0	λ_0	α	β	γ	λ	α	β	γ	λ	α	β	γ	λ
0.05	0.30	0.65	2	121	94	93	91	1197	136	204	142	498	120	165	110
0.05	0.05	0.9	2	72	84	96	69	352	136	176	101	1108	139	216	93
0.1	0.2	0.70	1.5	107	96	96	92	226	149	200	207	165	128	171	119
*0.25	0.05	0.70	1.5	81	86	82	47	83	117	92	53	90	127	98	59
0.1	0.15	0.75	2	104	89	95	85	579	132	211	136	221	122	179	105
0.05	0.1	0.85	2	73	90	95	83	893	123	189	123	350	129	167	106
0.20	0.2	0.60	1.5	110	97	98	90	182	144	191	198	127	125	141	120
*0.20	0.05	0.75	1.5	89	91	88	56	106	123	115	76	102	146	110	66
*0.3	0.1	0.6	0.5	94	97	95	90	92	123	95	153	82	125	89	92
0.1	0.1	0.8	0.5	115	95	110	85	136	160	140	164	142	150	150	140
0.7	0.2	0.1	0.5	91	99	95	89	107	115	108	136	110	114	113	129
0.7	0.25	0.05	1.5	91	100	96	87	136	111	129	117	111	116	114	99
0.8	0.1	0.1	0.5	91	99	92	90	99	88	93	120	110	119	107	82
0.05	0.05	0.9	0.5	97	92	99	84	158	177	157	123	130	194	142	85
0.8	0.15	0.05	0.5	91	104	92	104	119	97	113	155	113	112	114	101

MSE-efficiency for θ_i is defined as MSE(QMLE)/MSE($\bar{\theta}_i$).

the semiparametric estimators [SPE]. The column with the heading $\varepsilon \sim EXP$ in Tables 1 and 2 show that, as expected, the QMLE performed at least as well as the constrained semiparametric estimator. However, the differences were small in most cases.

Now, let us consider the case when the error distribution is not exponential. The results for these cases are shown under the headings $\varepsilon \sim NG$ and $\varepsilon \sim LN$ in Tables 1 and 2. These results show clearly that the constrained semiparametric estimator performed better, often substantially better, than the QMLE. If the true value of $\boldsymbol{\theta}$ is not in the set A, where $A = \{(\alpha, \beta, \gamma) : \beta$ and (β/α) are close to zero, and α and γ are not close to zero $\}$, (for eg., the rows with '*' in Tables 1 and 2), then $\bar{\boldsymbol{\theta}}$ performs better than QMLE. Even if the true parameter lies in the set A, QMLE does not dominate $\bar{\boldsymbol{\theta}}$; Tables 1 and 2 show that, in region

A, $\bar{\boldsymbol{\theta}}$ is better than QMLE for β, but not for (α, γ). In several empirical studies reported in the literature, for example Engle and Russell (1998), Engle and Russell (1997), Fernandes and Grammig (2006) and Zhang *et al.* (2001), the estimated value of $\boldsymbol{\theta}$ turned out to be away from the aforementioned region A. Therefore, it appears that $\bar{\boldsymbol{\theta}}$ performs better than QMLE in the part of the parameter space that is of practical relevance.

(B) *Comparison of the constrained and the unconstrained semiparametric estimators,* $\bar{\boldsymbol{\theta}}$ *and* $\hat{\boldsymbol{\theta}}$*:*

Figures 3 and 4 show that the relative MSE-efficiencies are at least 100%. Thus, the constrained estimator $\bar{\boldsymbol{\theta}}$ performed at least well as the unconstrained DW-estimator $\hat{\boldsymbol{\theta}}$ for all true parameter values. The cases for which the relative efficiencies are equal to 100% or slightly higher, correspond to the case when the parameter value is away from the boundary and lie deep in the interior of the parameter space. Similarly, relative efficiencies that are substantially higher than 100% correspond to the case when the parameter value is close to the boundary. Therefore, as expected, the constrained estimator $\bar{\boldsymbol{\theta}}$ performed better than the unconstrained estimator $\hat{\boldsymbol{\theta}}$.

Summary of the results:

For the Linear ACD and Power ACD models studied in this paper, for which α, β and γ must be nonnegative and $\beta + \gamma \leq 1$, the new constrained estimator $\bar{\boldsymbol{\theta}}$ performed better than the (unconstrained) semiparametrically efficient $\hat{\boldsymbol{\theta}}$. Further, $\bar{\boldsymbol{\theta}}$ performed better than the QMLE in the part of the parameter space that appears to be practically relevant based on past empirical studies.

4 An empirical example

In this section, we use the IBM transaction data for November 1990, to illustrate the importance of the constrained estimator $\bar{\boldsymbol{\theta}}$. In this example, we do not plan to model the data in order to draw substantive conclusions about IBM transactions, and therefore we do not carry out diagnostics to evaluate goodness of fit. We estimated the parameters in the linear ACD(2,2) model, $\psi_i = \alpha + \beta_1 X_{i-1} + \beta_2 X_{i-2} + \gamma_1 \psi_{i-1} + \gamma_2 \psi_{i-2}$, by QMLE and the semiparametric methods. For this model, the parameter space Θ is given by

$$\Theta = \{\boldsymbol{\theta} : \boldsymbol{\theta} = (\alpha, \beta_1, \beta_2, \gamma_1, \gamma_2)^\top; \alpha \geq 0; 0 \leq \beta_1, \beta_2, \gamma_1, \gamma_2, \beta_1 + \beta_2 + \gamma_1 + \gamma_2 \leq 1\}. \quad (11)$$

It is possible to allow other constraints on $\boldsymbol{\theta}$, but the ones in (11) are sensible from an economic point of view.

The computed values are given in Table 3. To compute the QMLE, we maximized the log likelihood corresponding to the assumption $\varepsilon_i \sim exp(\lambda)$. Since the unconstrained QMLE, given in Table 3, is an interior point of Θ, it is also equal to the QMLE under the constraint $\boldsymbol{\theta} \in \Theta$.

Although the unconstrained QMLE satisfies the constraint $\boldsymbol{\theta} \in \Theta$, the DW-estimator $\hat{\boldsymbol{\theta}}$ is outside the parameter space Θ. This is an example of the type of settings where a constrained estimator such as $\bar{\boldsymbol{\theta}}$ would be essential. Since $\bar{\boldsymbol{\theta}}$ is not asymptotically normal when the true parameter lies on the boundary of the parameter space, it is not particularly meaningful to provide standard errors for $\bar{\boldsymbol{\theta}}$. If a measure of variability is desired, a confidence region can be constructed by inverting an inequality constrained test based on $\bar{\boldsymbol{\theta}}$. This is not a trivial computational task, but possible to do. In any case, it follows from Proposition 1 that the constrained estimator $\bar{\boldsymbol{\theta}}$ in Table 3 is closer to the true value than the unconstrained estimator $\hat{\boldsymbol{\theta}}$.

Note that, the constrained estimation resulted in the unconstrained estimate of β_2 moving from outside its allowed range $(=-0.041)$ to its boundary $\beta_2 = 0$, the estimate of γ_2 moving from outside its allowed range $(=-.082)$ to an interior point (=0.27), and the estimate of γ_1 moving from outside the parameter space (=1.005) to an interior point (=0.616).

This example illustrates that when $\hat{\boldsymbol{\theta}}$ fails to satisfy the constraints imposed by the parameter space Θ, the constrained estimation method introduced in this paper offers a methodologically sound way of obtaining estimators that lie in the parameter space Θ. Not

Table 3: Estimates of parameters for the ACD(2,2) model for the IBM transaction data

	α	β_1	β_2	γ_1	γ_2
	Unconstrained Estimators				
QMLE	0.561	0.098	0.018	0.375	0.492
$\hat{\boldsymbol{\theta}}$	0.321	0.108	**-0.041**	**1.005**	**-0.082**
	Constrained Estimators				
$\bar{\boldsymbol{\theta}}$	0.471	0.099	0.000	0.616	0.270

only does the constrained estimator satisfies the constraints imposed by the parameter space, it is also likely to be closer to the true value than even the unconstrained semiparametrically efficient estimator.

5 Conclusion

We studied estimation of parameters in duration models where the parameter space is restricted. The estimator proposed in this paper is specifically designed for situations when there are constraints on parameters, such as nonnegativity constraints, the error distribution is unknown, and the errors themselves may not be independent. Since such situations are expected to be common in practice and the new method proposed in this paper performed better than its competitors, we conclude that the the proposed method is of significant practical importance.

We used the theoretical results of Drost and Werker (2004) as building blocks, to propose a new semiparametric method of estimation for duration models when some parameters are known to satisfy inequality constraints, for example nonnegativity constraints as in the standard linear ACD model of Engle and Russell (1998). We showed that our proposed constrained estimator is asymptotically better than the unconstrained DW-estimator when there are inequality constraints on parameters.

We carried out a simulation study to compare our estimator with the semiparametrically efficient DW-estimator and the QMLE. In this simulation study, the inequality constrained estimator proposed in this paper performed better than the DW-estimator and the QMLE in most cases of practical interest. Once the unconstrained estimator has been computed, it is straight forward to compute the constrained estimator $\bar{\boldsymbol{\theta}}$.

An empirical application involving the ACD(2,2) model illustrates the relevance and importance of the new method. For example, it illustrates how the new method leads to nonnegative estimates for nonnegative parameters when the unconstrained semiparametrically efficient estimators are negative.

In this paper, we did not discuss about *semiparametric efficiency bound* when there are inequality constraints of the form $\boldsymbol{h}(\boldsymbol{\theta}) \geq \mathbf{0}$. This is because the relevant theory has not been developed yet even for much simpler cases. However, since our constrained estimator $\bar{\boldsymbol{\theta}}$ is based on the building blocks of a semiparametrically efficient estimator, it appears that $\bar{\boldsymbol{\theta}}$ is likely to be 'efficient' in some intuitive sense although it is difficult to formalise.

In summary, the constrained estimator proposed in this paper is better than the corresponding unconstrained estimator and the QMLE when there are inequality constraints.

Acknowledgments

This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP0665710). We are grateful to Feike Drost and Bas Werker for helpful comments and suggestions and for providing a copy of their computer program, Param Silvapulle for constructive comments and suggestions, and Gunky Kim for help with the use of MATLAB and implementation of the DW-approach.

Appendix: Proof of Proposition 1

Main steps only: The technical details of the proof of (7) uses the result that the parameter space Θ can be approximated by its approximating cone at the true value for the purposes of deriving the first order asymptotic properties. For example, the projections of $\hat{\boldsymbol{\theta}}$ onto Θ and onto the approximating cone $\mathcal{A}(\Theta;\boldsymbol{\theta}_0)$ of Θ at θ_0 are asymptotically equivalent: $n^{1/2}(\bar{\boldsymbol{\theta}}-\boldsymbol{\theta}^{\dagger}) = o_p(1)$ where $\boldsymbol{\theta}^{\dagger} = \Pi_{W_n}(\hat{\boldsymbol{\theta}} \mid \mathcal{A}(\Theta;\boldsymbol{\theta}_0))$. Now treating $\boldsymbol{\theta}_0$ as the origin, we have

$$n^{1/2}(\boldsymbol{\theta}^{\dagger}-\boldsymbol{\theta}_0) = \Pi_{W_n}\{n^{1/2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0) \mid \mathcal{A}(\Theta;\boldsymbol{\theta}_0)-\boldsymbol{\theta}_0\} \xrightarrow{d} \Pi_W(\boldsymbol{Z} \mid \mathcal{T}(\Theta;\boldsymbol{\theta}_0)),$$

the last step follows because $\Pi_W(\boldsymbol{z} \mid \mathcal{T})$ is a continuous function of $(\boldsymbol{z}, W)$.

Applying Proposition 3.12.3 on page 114 in Silvapulle and Sen (2005)) for the inner product defined by $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}^{\top} W_n^{-1}\boldsymbol{y}$, we have that $(\bar{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)^{\top} W_n^{-1}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0) \leq 0$. Therefore, $\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|_{W_n} \geq \|\bar{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|_{W_n}$. Since $W_n \xrightarrow{p} W$ and $(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0) = O_p(n^{-1/2})$, we have, by Lemma 4.10.2 on page 216 in Silvapulle and Sen (2005) that $n^{1/2}\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|_{W_n} = n^{1/2}\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|_W + o_p(1)$ and $n^{1/2}\|\bar{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|_{W_n} = n^{1/2}\|\bar{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|_W + o_p(1)$. Now, the proof of (10) follows.

References

Bauwens, L. and Giot, P. (2000). The logarithmic acd model: An application to the bid-ask quote process of three nyse stocks. *Annales d'Economie et de Statistiquie*, **60**, 117–150.

Bauwens, L. and Giot, P. (2001). *Econometric Modelling of Stock Market Intraday Activity.* Kluwer Academic Publishers.

Bickel, P., Klaassen, C., Ritov, Y., and Wellner, J. (1993). *Efficient and Adaptive Estimation in Semiparametric Models.* The Johns Hopkins University Press, Baltimore, Maryland.

Drost, F. C. and Werker, B. J. M. (2004). Semiparametric duration models. *J. Bus. Econom. Statist.*, **22**(1), 40–50.

El Barmi, H. and Mukerjee, H. (2005). Inference under a stochastic ordering constraint: the k-sample case. *Journal of the American Statistical Association*, **100**, 252–261.

El Barmi, H., Kochar, S., and Tsimikas, J. (2006). Likelihood ratio tests for and against ordering of the cumulative incidence functions in multiple competing risks and discrete mark variable models. *Journal of Statistical Planning and Inference*, **136**(5), 1588–1607.

Engle, R. F. and Russell, J. R. (1997). Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model. *Journal of Emprical Finance*, **4**, 187–212.

Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, **66**(5), 1127–1162.

Fernandes, M. and Grammig, J. (2006). A family of autoregressive conditional duration models. *J. Econometrics*, **130**(1), 1–23.

Hwang, J. T. G. and Peddada, S. D. (1994). Confidence interval estimation subject to order restrictions. *Annals of Statistics*, **22**, 67–93.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference.* Springer Series in Statistics, New York.

Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics*, **10**, 229–235.

Newey, W. (1990). Semiparametric efficiency bounds. *Journal of Applied Econometrics*, **5**(2), 99–135.

Peddada, S. D., Dunson, D. B., and Tan, X. (2005). Estimation of order-restricted means from correlated data. *Biometrika*, **92**, 703–715.

Peddada, S. D., Haseman, J., Tan, X., and Travlos, G. (2006). Tests for simple tree order restriction with application to dose-response studies. *Jour. Royal Statist. Soc., Ser - C*, **55**, 493–506.

Silvapulle, M. J. and Sen, P. K. (2005). *Constrained Statistical Inference: inequality, order, and shape restrictions*. John Wiley and Sons, New York.

Tsiatis, A. A. (2006). *Semiparametric theory and missing data*. Springer Series in Statistics. Springer, New York.

Zhang, M. Y., Russel, J. R., and Tsay, R. S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, **104**, 179–207.