2009

A semiparametric approach to a nonlinear ACD Model

Pipat Wongsaart

Jiti Gao

David E. Allen

A Semiparametric Approach to a Nonlinear ACD Model

By

Pipat Wongsaart[1], Jiti Gao[1] and David E. Allen[2]

[1]School of Economics, The University of Adelaide

[2]School of Accounting, Finance and Economics, Edith Cowan University

School of Accounting, Finance and Economics & FEMARC Working Paper Series
Edith Cowan University
June 2009
Working Paper 0903

Correspondence author:

David E. Allen
School of Accounting, Finance and Economics
Faculty of Business and Law
Edith Cowan University
Joondalup, WA 6027
Australia
Phone: +618 6304 5471
Fax: +618 6304 5271
Email: d.allen@ecu.edu.au

ABSTRACT

We introduce in this paper a new semiparametric approach to a nonlinear ACD model, namely the Semiparametric ACD (SEMI-ACD) model. This new model is more flexible in the sense that the data are allowed to speak for themselves, without a hypothetical assumption being imposed arbitrarily on its key component. Moreover, it enables a much more thorough examination of the intertemporal importance of the conditional duration on the ACD process. Our experimental analysis suggests that the new model possesses a sound asymptotic character, while its performance is also robust across data generating processes and assumptions about the conditional distribution of the durations. Furthermore, the empirical analysis in this paper illustrates the advantage of our model over its parametric counterparts. Finally, the paper discusses some important theoretical issues, especially its asymptotic properties, in order to pave the way for a more detailed analysis, which will be presented in a future paper.

Keywords: Dependent point process; Durations; Hazard rates and random measures; Irregularly spaced high frequency data; Semiparametric time series.

1. Introduction

An accurate measure and forecast of the intensity of financial event arrivals is an important tool that makes possible empirical analyzes of a number of important market microstructure issues. Some well known examples of this type of study are Spierdijk [19] who examines the role of trading intensity in information dissemination on a financial market, and Dufour and Engle [5] who investigate the importance of time on the price impact of a trade. In duration studies, an event may refer to a market transaction in general, for example Engle and Russell [7] who study the intensity of transaction arrivals for IBM transaction data. In other cases, it could refer to a particular type of transaction with some specific characteristics, for instance Engle and Russell [6] who employ a thinning algorithm discussed in Engle and Russell [7] to model the intensity of foreign exchange quote arrivals.

To be able to accurately measure and forecast the intensity of financial event arrivals, Engle and Russell [7] develop the so-called Autoregressive Conditional Duration (ACD) model by which arrival times are treated as random variables which follow a conditionally orderly point process with an intensity defined conditionally on the past activity. The ACD model share a number of statistical properties with a GARCH model of Bollerslev [3]. Furthermore, both models rely on a similar empirical motivation, i.e. the clustering of news and financial events in financial markets. As the results, a large number of researchers attempt to extend the model parametrically in a similar manner to those of the GARCH model. Some well known examples are the Logarithmic ACD (Log-ACD) model of Bauwens and Giot [2], Box-Cox ACD (BACD) model of Dufour and Engle [5], Threshold ACD (TACD) model of Zhang, Russell and Tsay [20], and the Augmented ACD (AACD) model of Fernandes and Grammig [8].[2]

[2]See Pacurar [18] for an excellent survey on the theoretical and empirical development of the ACD models.

However, in our view, the applicability and usefulness of the ACD model and its parametric extensions should be carefully scrutinized. Firstly, the scope of nonlinearity offered by the above parametric extensions is still too limited for some, while excessively sophisticated in the others. Furthermore, existing studies have so far failed to examine thoroughly the statistical importance of the conditional duration on the ACD process. Also, the fact that the key components of the models, namely the specification for the expected durations and the conditional distribution of the duration (and hence the shape of the hazard function) should be arbitrarily pre-specified, not only makes the model extremely sensitive to misspecification, but also too restrictive in the sense that the data are not allowed to speak for themselves.

Having taken these issues into consideration, researchers now turn to semiparametrics in search for a more useful generalization to the ACD model. Some interesting studies on semiparametric ACD models are Drost and Werker [4], and Fernandes, Medeiros and Viega [9]. Our proposal is to put forward a new semiparametric method of ACD modeling which is developed to include two important components, namely an iterative estimation algorithm and a semiparametric time series process. In our study, the iterative estimation algorithm is devised in order to address the fact that conditional durations are not observable in practice. As far as nonlinearity is concerned, an excellent starting point in this case is the partially linear additive autoregressive process as extensively discussed in Härdle, Liang and Gao [16].

The main objectives of the current paper are to introduce our newly developed Semiparametric ACD (SEMI-ACD) model and also to illustrate its practicability and usefulness through a number of simulated and real data examples. The paper also discusses a number of important theoretical issues, especially its asymptotic properties when taking into account the fact that the conditional durations themselves are estimates. This is to pave the way for a more detailed analysis, which will be presented in a future paper. The remainder of this paper is organized as follows. Section 2 below develops the statistical underpinning for the model and also presents the basic construction of the above-mentioned estimation algorithm. Section 3 considers a number of illustrative examples, while Section 4 applies the model to a thinned series of quotes arrival times for the USD/EUR exchange rate series. Finally, Section 5 concludes.

2. The SEMI-ACD Model

This section discusses the estimation procedure of the semiparametric partially linear autoregressive model and then presents the computational algorithm adapted within the SEMI-ACD framework in order to obtain estimates of the unobserved process $\{\psi_i\}$. To understand the model, let us consider a stochastic process that is simply a sequence of times $\{t_0, t_1, t_2, \ldots, t_n, \ldots\}$ with $t_0 < t_1 < t_2 < \cdots < t_n < \cdots$. Associated with the arrival times is the counting function $N(t)$, which is the number of events that have occurred by time t. Here, $x_i = t_i - t_{i-1}$ defines the intervals between two arrival times, which is often known as the durations, and ψ_i denotes the expectation of the ith duration, which is given by

$$E\left[x_i | x_{i-1}, \ldots, x_{i-p}, \psi_{i-1}, \ldots, \psi_{i-q}\right] = \psi\left(x_i | x_{i-1}, \ldots, x_{i-p}, \psi_{i-1}, \ldots, \psi_{i-q}; \Theta\right) \equiv \psi_i. \tag{2.1}$$

The ACD class of models consists of parameterizations of (2.1) and the assumption

$$x_i = \psi_i \varepsilon_i, \tag{2.2}$$

where $\{\varepsilon_i\}$ is a sequence of independently and identically distributed (i.i.d) random errors with $E[\varepsilon] = 1$. The basic ACD model as proposed by Engle and Russell [7] relies on a linear parameterizations of (2.1) in which ψ_i depends on p past durations and q past conditional durations as

$$\psi_i = \omega + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{k=1}^{q} \beta_k \psi_{i-k}, \tag{2.3}$$

while assumes that the durations are conditionally exponential. In the literature, this is often referred to as the EACD(p,q) model by which sufficient conditions to ensure the positivity of ψ_i are $\omega > 0$, $\alpha_j > 0$ for $\forall j = 1, \ldots, p$ and $\beta_k \geq 0$ for $\forall k = 1, \ldots, q$.

In spite of the overwhelming evidence of nonlinearity,[3] the question about the most appropriate type of nonlinear ACD models has not been satisfactorily addressed. As an alternative, the current paper proposes the SEMI-ACD model. If the conditional duration was observable in practice, the SEMI-ACD(p,q) model would be a semiparametric parameterization of the form

$$E\left[x_i | x_{i-1}, \ldots, x_{i-p}, \psi_{i-1}, \ldots, \psi_{i-q}\right] = \sum_{j=1}^{p} \gamma_j x_{i-j} + \sum_{k=1}^{q} g_k\left(\psi_{i-k}\right) \equiv \psi_i, \qquad (2.4)$$

where γ_j are unknown parameters and $g_k(\cdot)$ are unknown functions on the real line. An obvious advantage of the SEMI-ACD(p,q) model is the additional flexibility by which the Engle and Russell's linear specification is nested as a special case. Furthermore, the above model enables a much more thorough investigation about the statistical importance of the conditional duration in explaining the duration process. Unlike x_i, the fact that ψ_i is not observable in practice makes imposition of a parametric-intertemporal relationship rather debatable.

In order to present the main idea and methodology without unnecessary complication in our discussion, here attention will be restricted to a special case of (2.4), namely the SEMI-ACD(1,1) model of the form

$$E\left[x_i | x_{i-1}, \psi_{i-1}\right] = \gamma x_{i-1} + g\left(\psi_{i-1}\right) \equiv \psi_i. \qquad (2.5)$$

To derive the estimators of γ and g, we first rewrite (2.2) in its additive form as

$$x_i = \gamma x_{i-1} + g\left(\psi_{i-1}\right) + \eta_i, \qquad (2.6)$$

where $\eta_i = \psi_i\left(\varepsilon_i - 1\right)$ is a sequence of martingale differences and

$$g\left(\psi_{i-1}\right) \equiv E\left[x_i | \psi_{i-1}\right] - \gamma E\left[x_{i-1} | \psi_{i-1}\right] = g_1\left(\psi_{i-1}\right) - \gamma g_2\left(\psi_{i-1}\right). \qquad (2.7)$$

[3] See, for example, Dufour and Engle [5], Zhang, Russell, and Tsay [20], and Fernandes and Grammig [8].

If the parameter γ was known, then the natural estimates of g_j for a given γ would be

$$\hat{g}_{1,h}\left(\psi_{i-1}\right)=\sum_{s=2}^{T} W_{s,h}\left(\psi_{i-1}\right) x_{s}, \tag{2.8}$$

$$\hat{g}_{2,h}\left(\psi_{i-1}\right)=\sum_{s=2}^{T} W_{s,h}\left(\psi_{i-1}\right) x_{s-1}, \tag{2.9}$$

and $g(\psi_{i-1})$ would be estimated by

$$\hat{g}_{h}(\psi_{i-1})=\hat{g}_{1,h}\left(\psi_{i-1}\right)-\gamma \hat{g}_{2,h}\left(\psi_{i-1}\right), \tag{2.10}$$

where $W_{s,h}\left(\psi_{i-1}\right)$ is a probability weight function depending on $\psi_1, \psi_2, \psi_3, \ldots, \psi_{T-1}$ and the number T of observations. Note that here we only consider the case where $W_{s,h}$ is a kernel weight function

$$W_{s,h}(y)=\frac{K_{h}\left(y-\psi_{s-1}\right)}{\sum_{i=2}^{T} K_{h}\left(y-\psi_{i-1}\right)}, \tag{2.11}$$

where $K_h\left(\cdot\right)=h^{-1}K\left(\cdot/h\right)$, K is a real-valued kernel function satisfying Assumptions 4a below and $h=h_T \in H_T=\left[a_1 T^{-1/5-c_1}, b_1 T^{-1/5+c_1}\right]$ in which $0<a_1<b_1<\infty$ and $0<c_1<1/20$.

Now, for a given $\hat{g}_h$ computed based on the model $x_i=\gamma x_{i-1}+\hat{g}_h(\psi_{i-1})+\eta_i$, the kernel weighted least squares estimator of γ can be found by minimizing

$$\sum_{i=2}^{T}\left\{x_{i}-\gamma x_{i-1}-\hat{g}_{h}(\psi_{i-1})\right\}^{2}, \tag{2.12}$$

which implies

$$\hat{\gamma}_{h}-\gamma=\left\{\sum_{i=2}^{T} u_{i}^{2}\right\}^{-1}\left\{\sum_{i=2}^{T} u_{i} \cdot \eta_{i}+\sum_{i=2}^{T} u_{i} \cdot \bar{g}_{h}\left(\psi_{i-1}\right)\right\}, \tag{2.13}$$

where $u_i=x_{i-1}-\hat{g}_{2,h}\left(\psi_{i-1}\right)$ and $\bar{g}_h\left(\psi_{i-1}\right)=g\left(\psi_{i-1}\right)-\hat{g}_h\left(\psi_{i-1}\right)$.

Also, it is important to note that for the case of the SEMI-ACD(1,1)

$$\sigma^2 = E[\eta_i^2] = E[(x_i - \psi_i)^2] = E[\psi_i^2]\sigma_\epsilon^2, \tag{2.14}$$

where $\sigma_\varepsilon^2 = E[\varepsilon_i - 1]^2$. When σ^2 is unknown, it can be estimated by

$$\hat{\sigma}^2(h) = \frac{1}{T-1}\sum_{i=2}^{T}\{x_i - \hat{\gamma}(h)x_{i-1} - \hat{g}_{1,h}(\psi_{i-1}) + \hat{\gamma}(h)\hat{g}_{2,h}(\psi_{i-1})\}^2. \tag{2.15}$$

Moreover, the quality of the proposed estimators can be measured by the average squared error (ASE) of the form

$$D(h) = \frac{1}{T}\sum_{i=2}^{T}\left[\{\hat{\gamma}(h)x_{i-1} + \hat{g}_h^*(\psi_{i-1})\} - \{\gamma x_{i-1} + g(\psi_{i-1})\}\right]^2 \omega(\psi_{i-1}), \tag{2.16}$$

where $\hat{g}_h^*(\psi_{i-1}) = \hat{g}_{1,h}(\psi_{i-1}) - \hat{\gamma}(h)\hat{g}_{2,h}(\psi_{i-1})$ and $\omega(\cdot)$ is a positive weight function.

In practice, we apply a cross-validation (CV) criterion to construct adaptive data-driven estimates for γ and σ. In order to define the CV function, we first introduce the following estimators. For $1 \le n \le N = T-1$ and $j = 1, 2$, let us define

$$\hat{g}_{j,n}(\psi_n) = \frac{1}{N-1}\sum_{s\neq n}\frac{K_h(\psi_n - \psi_s)x_{x+2-j}}{f_{h,n}(\psi_n)} \tag{2.17}$$

and

$$\hat{g}_{h,n}(\psi_n) = \hat{g}_{1,n}(\psi_n) - \gamma\hat{g}_{2,n}(\psi_n), \tag{2.18}$$

where

$$\hat{f}_{h,n}(\psi_n) = \frac{1}{N-1}\sum_{s\neq n} K_h(\psi_n - \psi_s).$$

The leave-one-out estimate $\tilde{\gamma}_h$ of γ can now be founded by minimizing

$$\sum_{n=1}^{N}\left\{x_{n+1} - \gamma x_n - \hat{g}_{h,n}(\psi_n)\right\}^2. \tag{2.19}$$

The CV function in this case can therefore be defined as

$$CV(h) = \frac{1}{N}\sum_{n=1}^{N}\{x_{n+1} - \tilde{\gamma}(h)x_n - \hat{g}_{1,n}(\psi_n) + \tilde{\gamma}(h)\hat{g}_{2,n}(\psi_n)\}^2\omega(\psi_n). \tag{2.20}$$

An optimal value, $\hat{h}_C$, of h is chosen such that

$$CV(\hat{h}_C) = \inf_{h \in H_T} CV(h). \tag{2.21}$$

A data-driven bandwidth $\hat{h}$ is said to be asymptotically optimal if

$$\frac{D(\hat{h})}{\inf_{h \in H_T} D(h)} \xrightarrow{P} 1. \tag{2.22}$$

In the usually nonlinear time series case, Gao [10], Härdle, Liang and Gao [16]; Gao and Yee [12] discuss a number important results about the above estimation procedure. Before we state the main theory of this paper, we need to introduce the following assumption.

ASSUMPTION 1:

(i) Assume that the kernel function, K, is symmetric, Lipschitz continuous and has an absolutely integrable Fourier transform.

(ii) Assume that the weight function ω is bounded and that its support S is compact.

(iii) Assume that the processes $\{\psi_i : i \geq 1\}$ are strictly stationary and α-mixing with mixing coefficient $\alpha(T) \leq Cq^T$ for some $0 < C < \infty$ and $0 < q < 1$.

(iv) Assume that $\{\psi_i; i \geq 1\}$ have a common marginal density f, where f has a compact support containing S, and that g_j, for $j = 1, 2$, and f have two continuous derivatives on the interior of S.

(v) Assume that the K is a bounded probability function with $\int K(u)\, du = 1$, $K(\cdot) \geq 0$ and $\int_{-\infty}^{\infty} u^2 K(u)\, du < \infty$.

(vi) Assume that $P(\psi_i > 0) = 1$ for all $i \geq 1$ and $E[|\psi_i|^k] < \infty$ for all integer $k \geq 1$.

PROPOSITION 1: *Assume that Assumptions 1 holds and that*

$$\sigma_2^2 = E\left\{x_{i-1} - E\left(x_{i-1}|\psi_{i-1}\right)\right\}^2 > 0.$$

Then the following holds uniformly over $h \in H_T$:

$$\sqrt{T}\left\{\hat{\gamma}_h - \gamma\right\} \to N\left[0, \sigma^2/\sigma_2^2\right], \tag{2.23}$$

and

$$\sqrt{T}\left(\hat{\sigma}_h^2 - \sigma^2\right) \longrightarrow N\left[0, Var\left(\eta_1^2\right)\right]. \tag{2.24}$$

PROPOSITION 2: *Assume that Assumption 1 holds. Then the data-driven bandwidth $\hat{h}_C$ of h is asymptotically optimal.*

Hereafter, the current paper takes into account the fact that $\{\psi_i\}$ is not observable in practice and presents the computational algorithm adapted in this research in order to obtain the estimate of the process. Assume that we have a set of data sample $\{x_i; 1 \leq i \leq T\}$, ideally from the generating process described by (2.6). The estimation algorithm is constructed to include five important steps as follows:

Step 2.1: Choose the starting values for the vector of the T conditional durations. Index these values by a zero, i.e. $\{\psi_{i,0}\}$ and set $m = 1$.

Step 2.2: Find the $\hat{h}_m = \hat{h}_{C,m}$ such that $\hat{h}_{C,m} = \arg\min_{h \in H_T} CV_m(h)$, then employ the above estimation procedure to compute $\hat{\gamma}_m(h)$ and $\hat{g}_{h,m}$, based on $\{x_{i-1}; 2 \leq i \leq T\}$ and the estimates of the conditional durations as computed in the previous step, i.e. $\{\hat{\psi}_{i-1,m-1}; 2 \leq i \leq T\}$.

Step 2.3: Compute $\{\hat{\psi}_{i,m}; 2 \leq i \leq T\}$ and also select some sensible value for $\hat{\psi}_{1,m}$, which cannot be computed recursively.

Step 2.4: At $1 \leq m < M$, where M is a pre-specified maximum number of iterations, then increment m and return to Step 2.2.

Step 2.5: At $m = M$, perform the final estimation of x_{i-1} and $\hat{\psi}_{i-1,m}$ to obtain the final estimates of γ and g.

While the method of performing the final estimation in Step 2.5 will be discussed in more detail in Section 3, here let us rewrite the kernel-weighted LS estimators $\hat{\gamma}(h)$ and $\hat{\sigma}^2(h)$ of the above section using the estimates $\hat{\psi}_{i,m}$ as follows

$$\hat{\gamma}_{\hat{\psi}}(h) - \gamma = \left\{ \sum_{i=2}^{T} \hat{u}_i^2 \right\}^{-1} \left\{ \sum_{i=2}^{T} \hat{u}_i \cdot \eta_i + \sum_{i=2}^{T} \hat{u}_i \cdot \bar{\hat{g}}_h \left(\hat{\psi}_{i-1} \right) \right\}, \tag{2.25}$$

and

$$\hat{\sigma}^2_{\hat{\psi}}(h) = \frac{1}{T-1} \sum_{i=2}^{T} \{x_i - \hat{\gamma}_{\hat{\psi}}(h) x_{i-1} - \hat{g}_{1,h}(\hat{\psi}_{i-1,m}) + \hat{\gamma}_{\hat{\psi}}(h) \hat{g}_{2,h}(\hat{\psi}_{i-1,m})\}^2, \tag{2.26}$$

where $\hat{u}_i = x_{i-1} - \hat{g}_{2,h}(\hat{\psi}_{i-1,m})$ and $\bar{\hat{g}}_h = g(\psi_{i-1}) - \hat{g}_h(\hat{\psi}_{i-1,m})$. The $CV_{\hat{\psi}}(h)$ function in this case can be written as

$$CV_{\hat{\psi}}(h) = \frac{1}{N} \sum_{n=1}^{N} \{x_{n+1} - \tilde{\gamma}_{\hat{\psi}}(h) x_n - \hat{g}_{1,n}(\hat{\psi}_{n,m}) + \tilde{\gamma}_{\hat{\psi}}(h) \hat{g}_{2,n}(\hat{\psi}_{n,m})\}^2 \omega(\psi_n). \tag{2.27}$$

Since $\{\psi_i; i \geq 1\}$ are replaced by $\{\hat{\psi}_i; i \geq 1\}$, existing results[4] would need to be significantly modified to show that:

Proposition 3: *Under some regularity conditions on $g(\cdot)$, Propositions 1 and 2 still hold when ψ_i are replaced by the estimates $\hat{\psi}_{i,m}$.*

Mathematical proof of this proposition is tedious and is not the main focus of this paper, but will be included in a future paper instead. The remainder of this paper illustrates the practicability of the above mentioned SEMI-ACD modeling framework through a number of simulated and real data examples.

[4]See, for example, Chapter 6 of Härdle, Liang and Gao [16].

3. Computational Aspects and Illustrative Examples

We present in this section a small sample study for a number of illustrative models, which are specifically designed to demonstrate how that the above procedure works numerically and practically. However, before introducing these models, let us first discuss the computational steps taken in this paper as follows:

Step 3.1: Perform Steps 2.1 to 2.4 of the algorithm as explained in Section 2 to obtain $\hat{\psi}_{i,m}$ for $m = 1, 2, \ldots, M$.

Step 3.2: Averaging over the final K of M iterations to obtain

$$\hat{\psi}_i = \left(\frac{1}{K}\right) \sum_{m=M-K+1}^{M} \hat{\psi}_{i,m}. \tag{3.1}$$

Step 3.3: Compute

$$D_{\hat{\psi}}(h) = \frac{1}{T}\sum_{i=2}^{T}\left[\{\hat{\gamma}_{\hat{\psi}}(h)x_{i-1} + \hat{g}_{1,h}(\hat{\psi}_i) - \hat{\gamma}_{\hat{\psi}}(h)\hat{g}_{2,h}(\hat{\psi}_i)\} - \{\gamma x_{i-1} + g(\psi)\}\right]^2 w(\psi_{i-1})$$

and let $\hat{h}_{D,\hat{\psi}} = \arg\min_{h\in H_T} D_{\hat{\psi}}(h)$, where $H_T = \left[T^{-7/30}, 1.1T^{-1/6}\right]$.

Step 3.4: Find the $\hat{h}_{\hat{\psi}} = \hat{h}_{C,\hat{\psi}}$ such that $\hat{h}_{C,\hat{\psi}} = \arg\min_{h\in H_T} CV_{\hat{\psi}}(h)$.

Step 3.5: For the cases where T=101, 201, 301 and 401, compute

(i) $|\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}|$, $|\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$ and $|\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$.

(ii) $d_1(\hat{h}_C) = \frac{1}{N}\sum_{n=1}^{N}\{\hat{g}^*_{\hat{h}_{C,\hat{\psi}},n}(\hat{\psi}_n) - \hat{g}^*_{\hat{h}_{C,\hat{\psi}},n}(\psi_n)\}^2$,

$d_2(\hat{h}_C) = \frac{1}{N}\sum_{n=1}^{N}\{\hat{g}^*_{\hat{h}_{C,\hat{\psi}},n}(\psi_n) - g(\psi_n)\}^2$

$d_3(\hat{h}_C) = \frac{2}{N}\sum_{n=1}^{N}\{\hat{g}^*_{\hat{h}_{C,\hat{\psi}},n}(\hat{\psi}_n) - \hat{g}^*_{\hat{h}_{C,\hat{\psi}},n}(\psi_n)\}\{\hat{g}^*_{\hat{h}_{C,\hat{\psi}},n}(\psi_n) - g(\psi_n)\}$, and

$d(\hat{h}_C) = d_1(\hat{h}_C) + d_2(\hat{h}_C) + d_3(\hat{h}_C)$, where

$\hat{g}^*_{\hat{h}_{C,\hat{\psi}},n}(\hat{\psi}_n) = \hat{g}_{1,n}(\hat{\psi}_n) - \tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}})\hat{g}_{2,n}(\hat{\psi}_n)$

and

$\hat{g}^*_{\hat{h}_C,n}(\psi_n) = \hat{g}_{1,n}(\psi_n) - \tilde{\gamma}(\hat{h}_C)\hat{g}_{2,n}(\psi_n)$.

With regard to Step 3.2, our experience suggests that using an average as in (3.1), rather than $\hat{\psi}_{i,m}$, can often help in improving the performance of the algorithm. Furthermore, to demonstrate the robustness of the SEMI-ACD procedure, the random variable ε in each example is allowed to follow either a Gamma distribution with $\kappa = 2$ and $\beta = 0.5$, or a Weibull distribution with $\alpha = 3$ and $\beta = 1$. Finally, in the analysis below, we use the quartic kernel function of the form

$$K(u) = \begin{cases} (15/16)(1-u^2)^2 & \text{if } |u| \leq 5 \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

and the weight function

$$w(s) = \begin{cases} 1 & \text{if } |s| \leq 5 \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

We now introduce the illustrative models considered in this paper.

Example 1: The Mackey-Glass ACD (MG-ACD) Model

The MG-ACD model is motivated by the Mackey-Glass model, which can be interpreted as a model for population dynamics.[5] In view of (2.6), the MG-ACD model can be established by specifying

$$\gamma = 0.5 \text{ and } g(\psi) = 0.75\left(\frac{\psi}{1+\psi^2}\right). \tag{3.4}$$

Because of the functional form of g, the fact that the process $\{\psi_i\}$ is strictly stationary follows form Theorem 3.1 of An and Huang [1]. Furthermore, Lemma 3.4.4 and Theorem 3.4.10 of Györfi et al. [13] suggest that the $\{\psi_i\}$ is β-mixing and therefore α-mixing. Finally, it follows from the definitions of K and w above that all the remaining conditions in Assumption 1 are satisfied.

[5] See, for example, Section 4 of Nychka et al. [17] for details.

Example 2: The Logarithmic ACD (Log-ACD) Model

The Log-ACD model of Bauwens and Giot [2] assumes that

$$x_i = \exp(\varphi_i)\,\varepsilon_i, \quad \varphi_i = \omega + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{k=1}^{q} \beta_k \psi_{i-k}, \tag{3.5}$$

where $\{\varepsilon_i\} \sim$ i.i.d. with $E(\varepsilon_i) = \upsilon$. Let us now define

$$\exp(\phi_i) = \upsilon \exp(\varphi_i) \tag{3.6}$$

thereby the model in (3.5) can now be rewritten as

$$x_i = \exp(\phi_i)\,\eta_i, \quad \phi_i = \varpi + \sum_{j=1}^{p} \alpha_j \ln x_{i-j} + \sum_{k=1}^{q} \beta_k \phi_{i-k}, \tag{3.7}$$

where $\varpi = \omega + \ln \upsilon$, $\eta_i = \varepsilon_i / \upsilon$ and $E(\eta_i) = 1$. The so-called Log-ACD(1,1) model imposes the parameterization

$$\phi_i = \varpi + \alpha \ln x_{i-1} + \beta \phi_{i-1} \tag{3.8}$$

such that linearization of (3.10) leads directly to

$$\ln x_i = \varpi + \alpha \ln x_{i-1} + \beta \phi_{i-1} + \mu_i, \tag{3.9}$$

where $\mu_i = \ln \eta_i - 1$, so that $E(\mu_i) = 0$. Below, we illustrate the performance of the SEMI-ACD model in the case where the data generating process for each of the realizations is given by the following Log-ACD(1,1) model

$$x_i = \exp(\phi_i)\,\eta_i, \quad \phi_i = 0.01 + 0.2 \ln x_{i-1} + 0.7 \phi_{i-1}. \tag{3.10}$$

All simulations in this section were performed in S-plus. The means of the results for all four cases, namely the Weibull MG-ACD (WMG-ACD), the Gamma MG-ACD (GMG-ACD), the Weibull Log-ACD (WL-ACD) and the Gamma Log-ACD (GL-ACD) models, are tabulated in Tables 1 to 4, respectively. Note that in these tables N, R and M denote **T-1**, the number of replications and the number of basic iterations, respectively. We will now discuss a number of important findings.

N	100				200			
R	100		500		100		500	
M	3	8	3	8	3	8	3	8
$\lvert\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}\rvert$	0.0123	0.1203	0.1201	0.1200	0.1000	0.0980	0.0855	0.0824
$\lvert\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0774	0.0770	0.0809	0.0814	0.0706	0.0706	0.0636	0.0633
$\lvert\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0773	0.0769	0.0809	0.0813	0.0706	0.0706	0.0636	0.0634
$d_1(\hat{h}_C)$	0.0009	0.0011	0.0008	0.0009	0.0006	0.0005	0.0005	0.0004
$d_2(\hat{h}_C)$	0.0065	0.0070	0.0066	0.0065	0.0046	0.0046	0.0045	0.0045
$d_3(\hat{h}_C)$	-0.0019	-0.0026	-0.0019	-0.0019	-0.0018	-0.0018	-0.0016	-0.0016
$d(\hat{h}_C)$	0.0056	0.0054	0.0055	0.0055	0.0033	0.0033	0.0035	0.0034

N	300				400			
R	100		500		100		500	
M	3	8	3	8	3	8	3	8
$\lvert\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}\rvert$	0.0670	0.0620	0.0638	0.0630	0.0468	0.0466	0.0490	0.0480
$\lvert\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0602	0.0608	0.0561	0.0561	0.0539	0.0540	0.0513	0.0511
$\lvert\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0602	0.0608	0.0561	0.0561	0.0539	0.0540	0.0513	0.0511
$d_1(\hat{h}_C)$	0.0006	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
$d_2(\hat{h}_C)$	0.0038	0.0038	0.0041	0.0041	0.0041	0.0041	0.0036	0.0036
$d_3(\hat{h}_C)$	-0.0014	-0.0013	-0.0016	-0.0016	-0.0015	-0.0015	-0.0014	-0.0014
$d(\hat{h}_C)$	0.0030	0.0029	0.0029	0.0030	0.0030	0.0030	0.0026	0.0026

Table 1: WMG-ACD Model

N	100				200			
R	100		500		100		500	
M	3	8	3	8	3	8	3	8
$\lvert\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}\rvert$	0.0737	0.0858	0.0857	0.0910	0.0673	0.0655	0.0684	0.0668
$\lvert\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.1320	0.1300	0.1280	0.1235	0.1057	0.0959	0.1111	0.1017
$\lvert\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.1321	0.1300	0.1281	0.1236	0.1058	0.0960	0.1112	0.1018
$d_1(\hat{h}_C)$	0.0099	0.0115	0.0100	0.0127	0.0179	0.0236	0.0148	0.0199
$d_2(\hat{h}_C)$	0.0152	0.0150	0.0117	0.0117	0.0096	0.0092	0.0111	0.0118
$d_3(\hat{h}_C)$	-0.0077	-0.0084	-0.0061	-0.0064	-0.0097	-0.0125	-0.0084	-0.0094
$d(\hat{h}_C)$	0.0174	0.0181	0.0156	0.0180	0.0178	0.0203	0.0175	0.0223

N	300				400			
R	100		500		100		500	
M	3	8	3	8	3	8	3	8
$\lvert\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}\rvert$	0.0586	0.0574	0.0577	0.0609	0.0575	0.0594	0.0507	0.0523
$\lvert\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0938	0.0888	0.1083	0.0991	0.0839	0.0781	0.0818	0.0735
$\lvert\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0939	0.0888	0.1084	0.0992	0.0839	0.0782	0.0820	0.0736
$d_1(\hat{h}_C)$	0.0137	0.0170	0.0155	0.0211	0.0191	0.0195	0.0206	0.0206
$d_2(\hat{h}_C)$	0.0087	0.0087	0.0078	0.0095	0.0079	0.0079	0.0044	0.0044
$d_3(\hat{h}_C)$	-0.0082	-0.0097	-0.0087	-0.0098	-0.0079	-0.0088	-0.0102	-0.0114
$d(\hat{h}_C)$	0.0142	0.0160	0.0146	0.0208	0.0191	0.0186	0.0148	0.0136

Table 2: GMG-ACD Model

N	100				200			
R	100		500		100		500	
M	3	8	3	8	3	8	3	8
$\lvert\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}\rvert$	0.1521	0.1506	0.1658	0.1718	0.1233	0.1277	0.1277	0.1279
$\lvert\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0981	0.1031	0.1049	0.1093	0.0790	0.0828	0.0800	0.0832
$\lvert\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0975	0.1027	0.1037	0.1085	0.0789	0.0828	0.0800	0.0832
$d_1(\hat{h}_C)$	0.0018	0.0017	0.0020	0.0018	0.0007	0.0006	0.0004	0.0006
$d_2(\hat{h}_C)$	0.0113	0.0113	0.0152	0.0152	0.0144	0.0144	0.0120	0.0120
$d_3(\hat{h}_C)$	0.0028	0.0026	0.0041	0.0036	0.0027	0.0025	0.0028	0.0026
$d(\hat{h}_C)$	0.0159	0.0156	0.0213	0.0206	0.0178	0.0175	0.0152	0.0152

N	300				400			
R	100		500		100		500	
M	3	8	3	8	3	8	3	8
$\lvert\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}\rvert$	0.1033	0.0983	0.0979	0.0933	0.1151	0.1127	0.1008	0.0980
$\lvert\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0688	0.0712	0.0702	0.0727	0.0621	0.0647	0.0577	0.0597
$\lvert\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma\rvert$	0.0688	0.0713	0.0703	0.0728	0.0622	0.0649	0.0577	0.0598
$d_1(\hat{h}_C)$	0.0006	0.0005	0.0007	0.0007	0.0007	0.0005	0.0007	0.0005
$d_2(\hat{h}_C)$	0.0117	0.0120	0.0131	0.0131	0.0125	0.0125	0.0116	0.0116
$d_3(\hat{h}_C)$	0.0035	0.0030	0.0034	0.0032	0.0031	0.0016	0.0023	0.0022
$d(\hat{h}_C)$	0.0158	0.0155	0.0172	0.0170	0.0163	0.0163	0.0146	0.0143

Table 3: WL-ACD Model

N	100				200			
R	100		500		100		500	
M	3	8	3	8	3	8	3	8
$\lvert \hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}} \rvert$	0.2621	0.2540	0.2682	0.2586	0.1482	0.1430	0.1299	0.1217
$\lvert \hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma \rvert$	0.0970	0.0963	0.0972	0.0970	0.0776	0.0785	0.0709	0.0733
$\lvert \tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma \rvert$	0.0973	0.0969	0.0968	0.0969	0.0778	0.0786	0.0710	0.0733
$d_1(\hat{h}_C)$	0.0430	0.0430	0.0351	0.0351	0.0232	0.0329	0.0348	0.0348
$d_2(\hat{h}_C)$	0.0064	0.0049	0.0060	0.0053	0.0043	0.0019	0.0034	0.0036
$d_3(\hat{h}_C)$	0.0130	0.0107	0.0091	0.0083	0.0097	0.0098	0.0104	0.0093
$d(\hat{h}_C)$	0.0624	0.0586	0.0502	0.0487	0.0372	0.0446	0.0486	0.0477

N	300				400			
R	100		500		100		500	
M	3	8	3	8	3	8	3	8
$\lvert \hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}} \rvert$	0.1105	0.1039	0.0969	0.0939	0.0760	0.0710	0.0810	0.0523
$\lvert \hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma \rvert$	0.0685	0.0656	0.0593	0.0600	0.0584	0.0591	0.0544	0.0935
$\lvert \tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma \rvert$	0.0654	0.0657	0.0594	0.0602	0.0586	0.0593	0.0545	0.0535
$d_1(\hat{h}_C)$	0.0590	0.0509	0.0405	0.0405	0.0312	0.0312	0.0420	0.0420
$d_2(\hat{h}_C)$	0.0030	0.0025	0.0028	0.0025	0.0021	0.0017	0.0028	0.0028
$d_3(\hat{h}_C)$	-0.0004	0.0001	0.0066	0.0061	0.0095	0.0086	0.0147	-0.0114
$d(\hat{h}_C)$	0.0616	0.0535	0.0500	0.0491	0.0429	0.0416	0.0595	0.0334

Table 4: GL-ACD Model

Firstly, the simulation results in Tables 1 to 4 show that in all four cases the absolute error $|\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}|$ has the tendency of approaching zero as N increases. Notice, that $\hat{h}_{C,\hat{\psi}}$ and $\hat{h}_{D,\hat{\psi}}$ shown here are those of the final estimation step. Although the results are not reported here, we also consider $\hat{h}_{C,m}$ and $\hat{h}_{D,m}$ at each of the mth iteration and find that the absolute error $|\hat{h}_{C,m} - \hat{h}_{D,m}|$ has the tendency of approaching zero in all cases. Note also that here $\hat{h}_{C,\hat{\psi}}$ and $\hat{h}_{D,\hat{\psi}}$ are selected within the interval H_T of relatively small values, as the results of the absolute error $|\hat{h}_{C,\hat{\psi}} - \hat{h}_{D,\hat{\psi}}|$ might not provide the most accurate representation of the distance between the two estimates. Therefore, if we were to consider $N \geq 500$, the relative measure of the form $|\hat{h}_{C,\hat{\psi}}/\hat{h}_{D,\hat{\psi}} - 1|$ should be more appropriate.

Secondly, our estimation method is able to provide estimates for the parameter γ with comparable degree of accuracy to those of a one-step partially linear autoregressive estimation reported in Gao and Yee [12]. In all four cases, the absolute errors $|\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$ and $|\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$ have the tendency of approaching zero as $N \to \infty$ at a similar rate as those reported in Tables 1 and 2 of Gao and Yee [12]. Furthermore, these results are quite stable and not significantly affected by the increases in the number of replications, R. However, it is interesting to report that our estimation method seems to perform better, with respect to $|\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$ and $|\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$, at a smaller number of basic iterations, M, when applied to the WMG-ACD and the WL-ACD models, while performs better at a larger number of iterations when applied to the GMG-ACD and the GL-ACD models. Moreover, switching from the Weibull to the Gamma distributed standardized duration seems to have affected the results, with respect to both $|\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$ (and $|\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$ and $d_2(\hat{h}_C)$ significantly. In all aspects, $|\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$ and $|\tilde{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}}) - \gamma|$ of the GMG-ACD model are much larger than those of the WMG-ACD at a smaller number of observations, while the results become more comparable as N increases.

It is obvious that here $d_2(\hat{h}_C)$ is equivalent to that of a case where ψ_i was observable. Therefore, it should not be surprising to see that our results are quite comparable to those reported in Tables 1 and 2 of Gao and Yee [12]. Notice in Table 3 that $d_2(\hat{h}_C)$ for the WL-ACD model are relatively large when compared to those of the WMG-ACD model. Again, this should not be surprising given the linear nature of the Log-ACD model. Usually we expect the above semiparametric procedure to perform better with the Mackey-Glass style model.

We will now turn our attention to the results of $d_1(\hat{h}_C)$, which represents the estimation error due to the fact that the conditional duration is an estimate. The simulation results in Table 1 and 3 indicate that, for cases of the Weibull based models, the values of $d_1(\hat{h}_C)$ are significantly smaller than those of $d_2(\hat{h}_C)$ and have the tendency of approaching zero as N increases. The results in Table 2 and 4 show that $d_1(\hat{h}_C)$ of the Gamma based models are relatively large compared with those of their Weibull counterpart. The highest $d_1(\hat{h}_C)$ in Tables 2 and 4 is 0.0590 compared to only 0.0018 in Tables 1 and 3. In addition, it is apparent in Tables 2 and 4 that in this case $d_1(\hat{h}_C)$ has less tendency of approaching zero. Though, further investigation indicates that similar results to those of Tables 1 and 3 can also be obtained for the Gamma based models when the number of observations, N, is increased to more than 1,000. The question of how changes in the distributional assumption of ε may affect the simulation results is the subject of further investigation.

4. Real Example

This section applies the SEMI-ACD model to a thinned series of quotes arrival times for the USD/EUR exchange rate series.[6] The foreign exchange market operates around the clock 7 days a week, while the complete data set is one whole week covering March 11, 2001 through to March 17, 2001. The current study assumes that the business week is periodic, i.e. 5 days, beginning Sunday 22:00 GMT to Friday 22:00 GMT. Therefore, the weekend observations of the USD/EUR data are filtered out. To eliminate the problem of bid-ask bounce, this paper defines the current price as the midpoint of the bid-ask spread, i.e. the midprice of the form

$$p_i = \frac{bid_i + ask_i}{2} \tag{4.1}$$

such that bid_i and ask_i are the current bid and ask prices associated with transaction at time t_i. Then the dependent thinning is performed, so that only the points at which prices have changed significantly since the occurrence of the last price change are kept. Formally we retain point $i > 1$ if

$$|p_i - p_j| > c, \tag{4.2}$$

where j is the index of the most recent retained point and the constant c represents a threshold value. Clearly the value of c is what characterizes a significant price change such that if $c = 0$, then we would count every single movement in the midpoint as a price change. However, to better capture movements in the price at which transactions are occurring and also to minimize the impact of asymmetric quote setting due to portfolio adjustment by individual banks, this study sets $c = 0.0005$, i.e. five pips. [7] This choice of c yields a sample size of 1,663 or 1.5% of the original sample. The average price duration for the sample is 258 seconds (or just over 4 minutes), while the minimum and the maximum are zero and 25,765 seconds (or just over 7 hours), respectively.

[6]The foreign exchange data are provided by Olsen and Associates (O&A).

[7]See also Engle and Russell [6].

The fact that for currency trading there are clear periods of high and low activity as markets around the world open and close leads us to believe that the intraday durations may consist of not only the stochastic, but also the deterministic components. In the discussion that follows we assume that the deterministic effect of time can be formulated as

$$x_i = \phi(t_{i-1})\nu_i, \tag{4.3}$$

where $\phi(\cdot)$ denotes the diurnal factor of the calendar time t_{i-1} at which the ith duration begins, and such that

$$\nu_i = \frac{x_i}{\phi(t_{i-1})} \backsim \text{ i.i.d.} \tag{4.4}$$

represents the diurnally adjusted data. Given (4.3) and (4.4), this study defines the expected price duration as

$$E_{i-1}(x_i) = E[\nu_i|\mathcal{F}_i]\phi(t_{i-1}) = \psi_i\phi(t_{i-1}), \tag{4.5}$$

where $\mathcal{F}_i$ is the σ-field representing the past information set. Engle and Russell [7] assume that the seasonal factor can be approximated by a cubic spline, while the parameters in both the deterministic and the stochastic components can be jointly estimated using maximum likelihood estimation.

A simple linear transformation of (4.3) into an additive noise of the form

$$x_i = \phi(t_{i-1}) + \xi_i, \tag{4.6}$$

where $\xi_i = \phi(t_{i-1})(\nu_i - 1)$ is a martingale difference series, leads to an alternative approach which is to initially estimate the diurnal factor and then model the ratio of actual to fitted value

$$\tilde{x}_i = \frac{x_i}{\hat{\phi}(t_{i-1})} \tag{4.7}$$

as the diurnally adjusted series of durations. [8] In the current paper, the diurnal factor is estimated using the kernel regression smoothing technique such that the

[8]See also Engle and Russell [6] and [7].

smoother is defined as

$$\hat{\phi}_h(t_{i-1}) = \sum_{s=2}^{T} W_{s,h}(t_{s-1}) x_s, \tag{4.8}$$

where $W_{s,h}(y) = \frac{K_h(y-t_{s-1})}{\sum_{i=2}^{T} K_h(y-t_{i-1})}$ and $K_h(\cdot)$ is as defined in the earlier section. An asymptotically optimal bandwidth parameter is selected using the leave-one-out cross validation based selection criteria such that

$$H_T = \left\{ h = h_{\max} a^k : h \geq h_{\min}, k = 0, 1, 2, \ldots \right\},$$

where $0 < h_{\min} < h_{\max}$ and $0 < a < 1$. Now, let J_T denotes the number of elements of H_T, then we have in this case $J_T \leq \log_{1/a}(h_{\max}/h_{\min})$.[9] Figure 1 presents the seasonal component for the price durations.

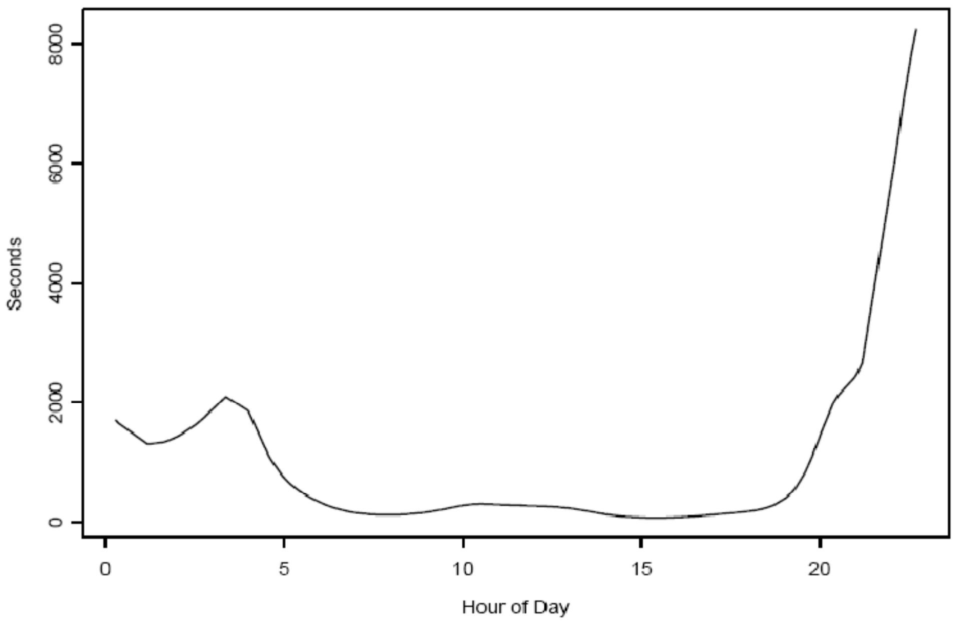

Figure 1: Expected price duration on hour of day, where 0 denotes 24:00GMT.

[9]See also Gao, Hawthorne and Bag [11].

There is enough evidence in the figure to suggest that that price movements change characteristics as a business day around the world starts and ends. A moderate fluctuation, which occurs between hours 01:00 and 08:00GMT, marks the beginning and the end of a business period in Tokyo by which the sudden slow down during hours 03:00 and 04:00GMT belongs to the Japanese lunch hours. Furthermore, the period of high price volatility, which occurs between hours 14:00 and 16:00GMT, takes place when both London and New York markets are active. Finally, it is clear that price change occurs much less frequently between hours 22:00 and 24:00GMT, which is the period when business activity in London, New York and Tokyo is relatively less intense. Engle and Russell [6] report similar intraday seasonal pattern in their study on the price intensity of USD/Deutschmark exchange rate. Figure 2 shows the diurnally adjusted price duration, $\tilde{x}_i$.

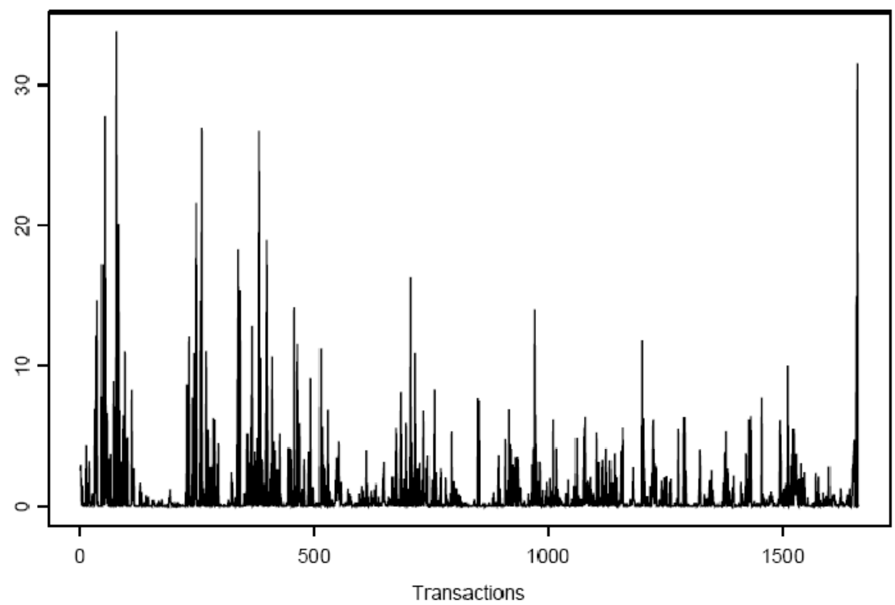

Figure 2: Diurnal adjusted price durations.

The ACD model is proposed as a model for intertemporally correlated event arrival times. Hence, to examine the dependence, we calculate the autocorrelations and partial autocorrelations in the waiting times between price changes. In Table 5, it is obvious that, while most are positive, the autocorrelations and partial autocorrelations are far from zero. The Ljung-Box statistic is then conducted to formally test the null hypothesis that the first 15 autocorrelations are 0. The test statistic is distributed as a χ^2_{15} with 5% critical value of 25. Therefore, the test provides us with some evidence against the null hypothesis of no autocorrelation up to order 15.

	Raw Price Durations		Diurnal Adj. Durations	
Lag	acf	pacf	acf	pacf
1	0.153	0.153	0.082	0.082
2	0.195	0.175	0.148	0.143
3	0.064	0.013	0.087	0.067
4	0.082	0.040	0.037	0.006
5	0.043	0.015	0.129	0.108
6	0.067	0.040	0.059	0.034
7	0.096	0.074	0.029	(0.011)
8	0.143	0.107	0.024	(0.006)
9	0.015	(0.050)	0.017	0.005
10	0.084	0.040	0.058	0.040
11	0.008	(0.016)	0.043	0.025
12	0.035	0.002	0.071	0.053
13	0.044	0.035	0.018	(0.006)
14	0.028	(0.006)	0.060	0.039
15	0.013	(0.022)	0.092	0.071
Ljung-Box(15)	199.93(0.00)		137.60(0.00)	
Sample Size	1663		1663	

Table 5: Autocorrelations and partial autocorrelations of price durations.

We now apply the above mentioned SEMI-ACD(1,1) model to model the diurnally adjusted price duration, $\tilde{x}_i$. A number of previous studies in the field have suggested that the choice of the kernel function is much less critical than that of the bandwidth.[10] To study the current problem, we employ the normal kernel function of the form

$$K(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \tag{4.9}$$

while computation of the CV function for the m iteration follows directly from (2.27) namely

$$CV_m = \frac{1}{N} \sum_{n=1}^{N} \{x_{n+1} - \tilde{\gamma}_m(h) x_n - \hat{g}_{1,n}(\hat{\psi}_{n,m-1}) + \tilde{\gamma}_m(h) \hat{g}_{2,n}(\hat{\psi}_{n,m-1})\}^2 \omega(\psi_n). \tag{4.10}$$

To specify the most appropriate bandwidth interval for each of the nth iteration, we follow a similar procedure to that suggested in Härdle, Hall and Marron [15]. The first step is to compute the score for each of the CV functions among one hundred sample values of h drawn sequentially from the set

$$H_0 = \left\{ \hat{h}_s : 0.01 < \hat{h}_s \leq 4 \right\},$$

where $s = 1, 2, \ldots, 100$. The results show that the interval $H_T = [0.0532, 0.3486]$ is the smallest possible bandwidth interval by which $CV_m(h)$ can attain their smallest values. The above step is then repeated, except that $\hat{h}_s$ are now drawn sequentially from H_T.

With regard to the maximum number of basic iterations, initially it is set at M=15. However, it is found that the average squared error of $\hat{\psi}_{i,m-1}$ and $\hat{\psi}_{i,m}$ at $M \geq 7$ is virtually zero, which indicates that no further improvement can be obtained. Therefore, the analysis that follows is based on $M = 6$.

[10] See, for example, Gao and Yee [12].

Iteration	$\hat{\gamma}_m(\hat{h}_{C,m})$	$\hat{h}_{C,m}$
1st	0.08182	0.2000
2nd	0.07997	0.2775
3rd	0.07929	0.2775
4th	0.07926	0.2775
5th	0.07927	0.2775
6th	0.07926	0.2775
Semiparametric Model	$\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}})$	$\hat{h}_{C,\hat{\psi}}$
SEMI-ACD I	0.07234	0.2775
SEMI-ACD II	0.07017	0.3175
Parametric Model	$\hat{\gamma}$	-
EACD Model	0.1288	-

Table 6: Estimation results for the SEMI-ACD(1,1) and EACD(1,1) price model, where those of SEMI-ACD I and II were computed based on Step 2.5 and Step 3.2, respectively.

Table 6 shows the estimation results of $\hat{\gamma}_m(\hat{h}_{C,m})$, $\hat{h}_{C,m}$, $\hat{\gamma}_{\hat{\psi}}(\hat{h}_{C,\hat{\psi}})$ and $\hat{h}_{C,\hat{\psi}}$ for the SEMI-ACD(1,1) model of price durations, and also the parameter estimate for the EACD as a comparison. While the semiparametric estimates of 0.07234 and 0.07017 are quite consistent with that of Engle and Russell [6], who report an estimate of 0.07315 in their study on the price intensity of USD/Deutschmark exchange rate, they are significantly lower than the parametric estimate of 0.1288.

To further investigate the source of such an inconsistency, let us now empirically examine the intertemporal importance of the conditional duration on the ACD process. Figure 3 presents the partial plot of the nonparametric estimate of the unknown real valued function g in the SEMI-ACD model, while Figure 4 compares this estimate with that of β computed based on the EACD model.

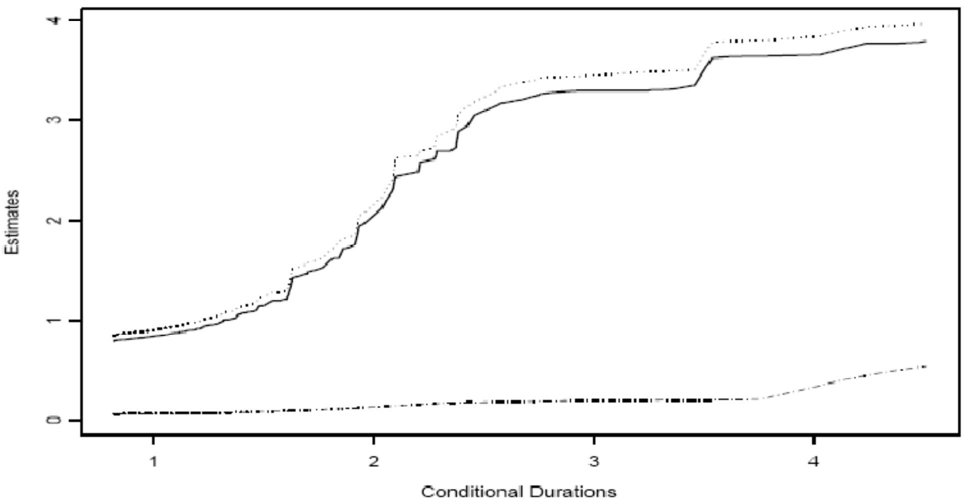

Figure 3: The solid line displays the partial plot of the nonparametric estimates of g, while the dots-and-dashes and the dotted lines show those of g_1 and g_2, respectively.

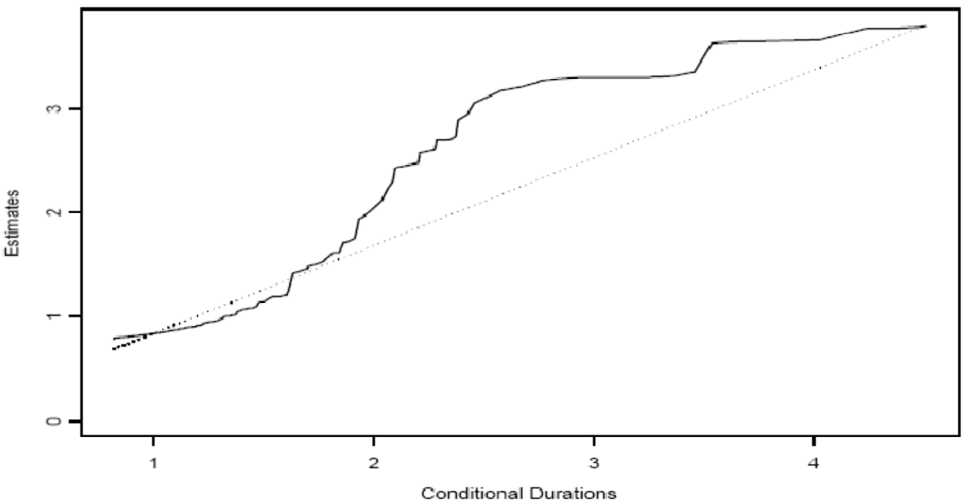

Figure 4: SEMI-ACD(1,1): $\hat{g}_h(\cdot)$ Vs. EACD(1,1): $\hat{\beta} = 0.8438$.

The solid line in Figure 3 displays the partial plot of the nonparametric estimates of g, while the dots-and-dashes and the dotted lines show those of g_1 and g_2, respectively. It is quite clear that the shape of g is significantly determined by that of g_2. Moreover, there is enough evidence in the figure to suggest that g is in fact nonlinear with an intertemporal asymmetry occurring between the conditional duration of above and below 2. In particular, the empirical estimate of the function suggests that g is convex for all points below 2, while is concave for the remaining points above it. Furthermore, it is this kind of asymmetric-intertemporal impact that makes a linear parameterization, e.g. the basic Engle and Russell [7], inappropriate. The slope of the dotted line in Figure 4 represents the empirical estimate of the unknown parameter β based on the EACD model. For the current study of the price durations, clearly the EACD model slightly overestimates the intertemporal impact of conditional duration at the points below 2 second, while significantly underestimates it for all the remaining points above it. Failure to capture this asymmetric-intertemporal impact clearly is the key reason why the EACD model significantly overestimate the unknown parameter γ in Table 6. Finally, Figure 5 presents the empirical estimate of the expected price duration in (4.5) computed based on the SEMI-ACD(1,1) model with nonparametrically estimated diurnal component.

To obtain the estimate of the baseline hazard, let us first define the empirical estimate of the standardized duration as

$$v_i = \frac{\tilde{x}_i}{\psi_i} \tag{4.11}$$

with density p_v and the associated survival function S_v. In a parametric ACD study, for example Engle and Russell [7], a stochastic transformation of the data, such as that in (4.11), is often assumed i.i.d.. Nonetheless, an advantage of semiparametrics in general is its flexibility in the sense that such a statistically restrictive property is not usually required. In a future paper, we intend to show that the above-mentioned SEMI-ACD estimation does also enjoy a similar benefit. However, for the sake of completion, we present here results of the Ljung-Box test statistics with 15 lags on υ_i. Even though in this case the Ljung-Box test statistics reduces to 85.143 compared to 199.93 and 137.60 in Table 5, the null hypothesis is still rejected at the 5% significance level.

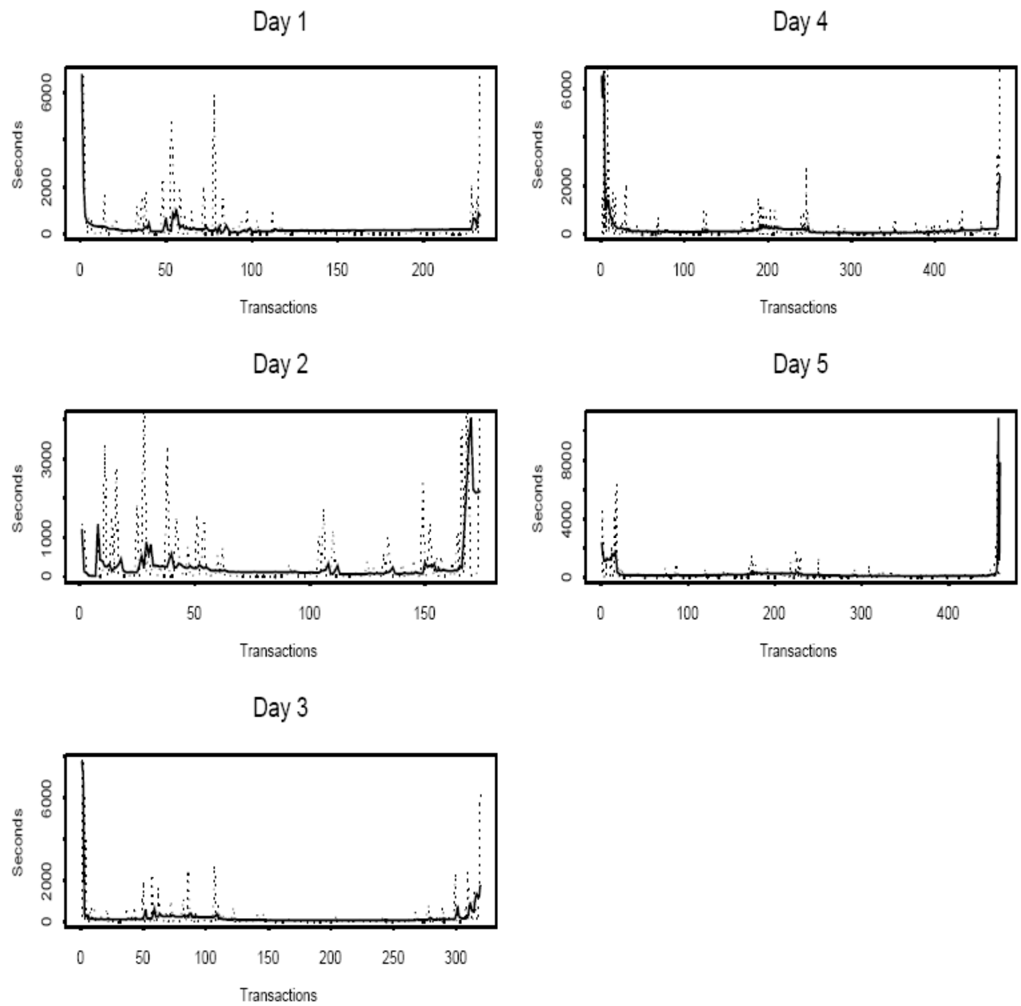

Figure 5: The dotted line displays the observed price durations, while the solid curve shows the one-step forecast of price durations computed based on the SEMI-ACD(1,1) for each of the five days considered.

There are numerous suggestions in the duration literature on how the baseline hazard for the price durations can be empirically estimated. An alternative approach we consider in this paper is to (i) estimate the density of the empirical standardized duration using kernel density estimation, (ii) compute the associated survival function and (iii) take the ratio of the two to obtain the baseline hazard. We will now explain the first two steps in more details.

The survival function of v is the function S_v defined by

$$S_v(e) = Pr(v > e) \tag{4.12}$$

for all e. If the cumulative distribution function F_v is known, then generally S_v can be computed as

$$S_v(e) = 1 - F_v(e). \tag{4.13}$$

Otherwise, $\hat{F}_v$ can be estimated by

$$\hat{F}_v(e) = \int_{-\infty}^{e} \hat{p}_v(y)dy, \tag{4.14}$$

where in this case $\hat{p}_v(y)$ is the nonparametric kernel density estimate of the form

$$\hat{p}_v(y) = \frac{1}{Th}\sum_{i=1}^{T} K\left(\frac{e_i - y}{h}\right), \tag{4.15}$$

and h is the bandwidth parameter. We can now write (4.14) using the estimate in (4.15) as

$$\hat{F}_v(e) = \frac{1}{Th}\sum_{i=1}^{T} \int_{-\infty}^{e} K\left(\frac{e_i - y}{h}\right) dy. \tag{4.16}$$

Defining $z = \frac{e_i - y}{h}$ such that $dy = (-h)dz$, then performing the change of variable should lead immediately to

$$\hat{F}_v(e) = \frac{1}{T}\sum_{i=1}^{T} - \int_{-\infty}^{\frac{e_i - e}{h}} K(z)dz = \frac{1}{T}\sum_{i=1}^{T} \int_{\frac{e_i - e}{h}}^{\infty} K(z)dz. \tag{4.17}$$

In (4.17), if $K(z)$ is the normal kernel function, then we immediately have

$$\hat{F}_{\upsilon}(\epsilon) = \frac{1}{T}\sum_{i=1}^{T}\left[\Phi(\infty) - \Phi\left(\frac{e_i - e}{h}\right)\right] = \frac{1}{T}\sum_{i=1}^{T}\left[1 - \Phi\left(\frac{e_i - e}{h}\right)\right] \tag{4.18}$$

giving way to

$$\hat{S}_{\upsilon}(e) = \frac{1}{T}\sum_{i=1}^{T}\Phi\left(\frac{e_i - e}{h}\right) \tag{4.19}$$

where

$$\Phi(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u}\exp^{-u^2/2}du.$$

Now, in order to estimate (4.15), we compute the bandwidth parameter $\hat{h}_0$ based on the following rule of thumb

$$\hat{h}_0 = 1.06\min\left(\hat{\sigma}_{\upsilon}, \frac{\hat{R}}{1.34}\right)T^{-1/5}, \tag{4.20}$$

where $\hat{R}$ is the inter-quartile range defined as[11]

$$\hat{R} = \upsilon_{0.75T} - \upsilon_{0.25T}. \tag{4.21}$$

We will now present the empirical estimates of (4.17) and (4.19). Figure 6 presents the kernel density estimates of p_{υ}. Also, to give some idea about the kind of distribution υ may follow, the figure compares these estimates with those of a Gamma distribution. The dashed line in the figure displays the Kernel density estimates of p_{υ}, while the solid line shows the density of the Gamma(1,1/2) distribution. The fact that the two curves are quite similar in shape suggests that the standardized duration may have a Gamma distribution. However, a more formal testing is required. Finally, Figures 7 and 8 present the empirical estimates of the survival function S_{υ} and the associated cumulative distribution function, respectively.

[11]See, for example, Härdle [14].

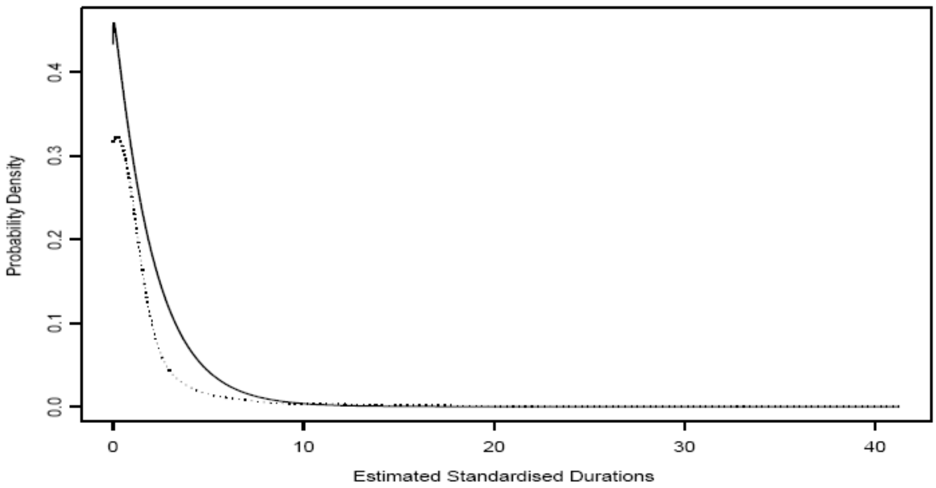

Figure 6: The dashed line displays the Kernel density estimates of p_v, while the solid line show the density of the Gamma(1,1/2) distribution.

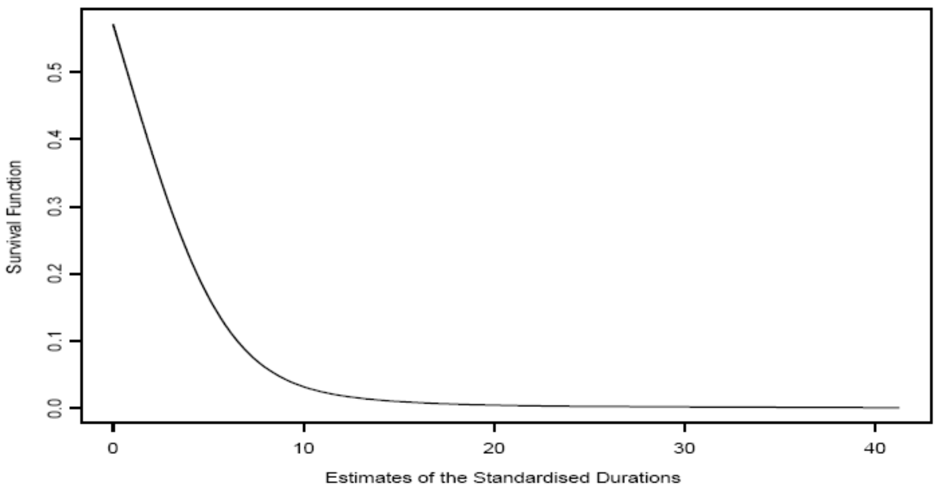

Figure 7: Empirical estimate of the survival function S_v

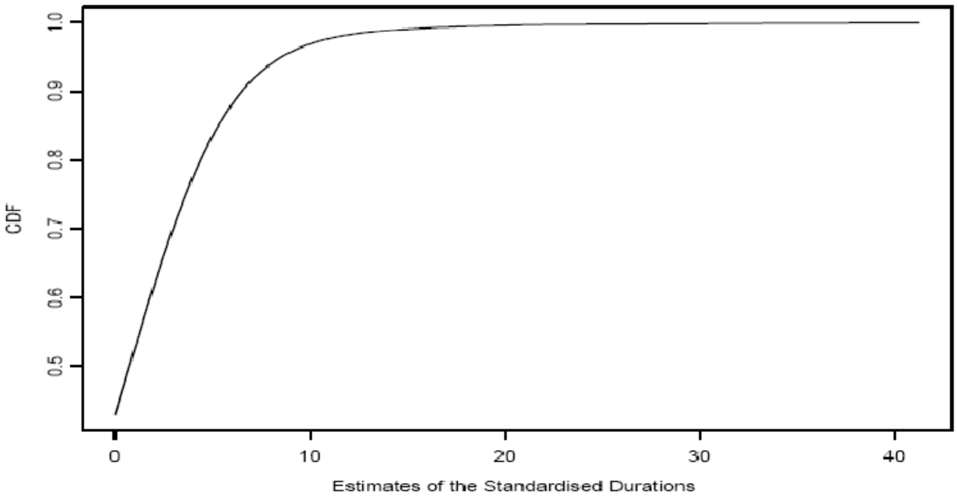

Figure 8: Empirical estimate of the cumulated distribution function F_{υ}.

5. Conclusions

Having concluded that thus far the question about the most appropriate type of nonlinear ACD model has not been satisfactorily answered, the current paper introduce a new Semiparametric ACD modeling method, namely the SEMI-ACD model. The SEMI-ACD model is developed to consist of two important components, i.e. the iterative estimation algorithm established in this paper to address the latency problem arises because of the fact that conditional durations are not observable in practice and the adaptive estimation of the partially linear additive autoregressive process. Our experimental analysis indicates that the SEMI-ACD model possesses a sound asymptotic character, while its performance is also robust across data generating processes and hypotheses about the conditional distribution of durations. Not only the model produces estimation results which are in agreement with previous studies when applied to a thinned series of quotes arrival times for the USD/EUR exchange rate, but it also provides valuable information and evidence about (i) the asymmetric-intertemporal impact of the conditional duration on the ACD process,

(ii) the conditional distribution of the price durations and (iii) how the results of a parametric ACD, e.g. the EACD model, could be unreliable. Overall, the analysis in this paper indicates that our new semiparametric approach to a nonlinear ACD model performs reasonably well in practice.

References

[1] H.Z. An and F.C. Huang. The geometrical ergodicity of nonlinear autoregressive models. *Statist. Sinica*, 6(4):943–956, 1996.

[2] L. Bauwens and P. Giot. The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks. *Annales d'Economie et de Statistique*, 60:117–149, 2000.

[3] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 31(3):307–327, 1986.

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

[5] A. Dufour and R.F. Engle. Time and the Price Impact of a Trade. *The Journal of Finance*, 55(6):2467–2498, 2000.

[6] R.F. Engle and J.R. Russell. Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model. *Journal of Empirical Finance*, 4(2-3):187–212, 1997.

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

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

[9] M. Fernandes, M.C. Medeiros, and A. Veiga. A (semi-) parametric functional coefficient autoregressive conditional duration model. *Texto para Discussão*, 2006.

[10] J. Gao. *Nonlinear time series: semiparametric and nonparametric methods.* Chapman & Hall/CRC, London, 2007.

[11] J. Gao and K. Hawthorne. Semiparametric estimation and testing of the trend of temperature series. *Econometrics Journal*, 9(2):332–355, 2006.

[12] J. Gao and T. Yee. Adaptive estimation in partially linear autoregressive models. *Canad. J. Statist.*, 28(3):571–586, 2000.

[13] L. Györfi, W. Härdle, P. Sarda, and P. Vieu. *Nonparametric curve estimation from time series*, volume 60 of *Lecture Notes in Statistics*. Springer-Verlag, Berlin, 1989.

[14] W. Härdle. *Smoothing techniques: with implementation in S*. Springer, 1991.

[15] W. Härdle, P. Hall, and S. Marron. How far are automatically chosen regression smoothing parameter from their optimum? *Journal of the American Statistical Association*, 83:86–101, 1988.

[16] W. Härdle, H. Liang, and J. Gao. *Partially linear models*. Springer-Verlag, Berlin, 2000.

[17] D. Nychka, S. Ellner, D. McCaffrey, and A.R. Gallant. Finding chaos in noisy systems. *J. Roy. Statist. Soc. Ser. B*, 54(2):399–426, 1992.

[18] M. Pacurar. Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. *Journal of Economic Surveys*, 22(4):711–751, 2008.

[19] L. Spierdijk. An empirical analysis of the role of the trading intensity in information dissemination on the NYSE. *Journal of Empirical Finance*, 11(2):163–184, 2004.

[20] M. Y. Zhang, J. R. Russell, and R.S. Tsay. A nonlinear autoregressive conditional duration model with applications to financial transaction data. *J. Econometrics*, 104(1):179–207, 2001.