

CORE DISCUSSION PAPER
2001/XX

A FAMILY OF AUTOREGRESSIVE CONDITIONAL DURATION MODELS

Marcelo Fernandes, ¹Joachim Grammig²

August 16, 2001

Abstract

This paper develops a family of autoregressive conditional duration (ACD) models that encompasses most specifications in the literature. The nesting relies on a Box-Cox transformation with shape parameter λ to the conditional duration process and a possibly asymmetric shocks impact curve. We establish conditions for the existence of higher-order moments, strict stationarity, geometric ergodicity and β -mixing property with exponential decay. We next derive moment recursion relations and the autocovariance function of the power λ of the duration process. Finally, we assess the practical usefulness of our family of ACD models using NYSE price duration data on the IBM stock. While the in-sample results warrant the extra flexibility provided either by the Box-Cox transformation or by the asymmetric response to shocks, we find no specification that entails satisfactory out-of-sample performance.

Keywords: Asymmetry, Box-Cox transformation, mixing property, price duration, shocks impact curve, stationarity.

JEL classification: C22, C41.

¹Graduate School of Economics, Fundação Getulio Vargas, Praia de Botafogo 190, 22953-900 Rio de Janeiro, Brazil. E-mail: mfernand@fgv.br

²CORE, Voie du Roman Pays, 34, B-1348 Louvain-La-Neuve, Belgium. Email: grammig@core.ucl.ac.be.

The authors acknowledge the financial support from the Jean Monnet fellowship at the European University Institute and from the Landeszentralbank Hessen, respectively. We are grateful to Pierre Giot for providing access to the NYSE database. Helpful comments were provided by Luc Bauwens and David Veredas. The scientific responsibility is assumed by the authors.

1 Introduction

The seminal work of Engle and Russell (1998) hoisted a great interest in the implications of price and trade durations in empirical finance. For instance, the modeling of price duration processes hinges the approaches to option pricing and intraday risk management recently proposed by Pringent, Renault, and Scaillet (1999) and Giot (2000), respectively. Although Engle and Russell's (1998) autoregressive conditional duration (ACD) model is the starting point of such analyses, the literature carries several extensions. Bauwens and Giot (2000) work with a logarithmic version of the ACD model that avoids the nonnegativeness constraints implied by the original specification so as to facilitate the testing of market microstructure hypotheses. Bauwens and Veredas (1999) propose the stochastic conditional duration process, leaning upon a latent stochastic factor to capture the unobserved random flow of information in the market. Ghysels, Gouriéroux, and Jasiak (1997) introduce the stochastic volatility duration model to cope with higher order dynamics in the duration process. Zhang, Russell, and Tsay (1999) argue for a nonlinear version based on self-exciting threshold autoregressive processes.

This paper develops a family of ACD models encompassing most of the existing models in the literature, such as the nonlinear ACD specifications recently put forward by Dufour and Engle (2000). For that purpose, we exploit the common features shared by the ACD and GARCH processes and follow a similar approach taken by Hentschel (1995) to build a family of asymmetric GARCH models. The nesting relies on a Box and Cox's (1964) transformation with shape parameter $\lambda \geq 0$ to the conditional duration process and on an asymmetric response to shocks. The motivation for the latter stems from Engle and Russell (1998), who show that standard ACD models applied to financial data tend to overpredict after extreme (very long or very short) durations.

We establish sufficient conditions for the existence of higher order moments, strict stationarity, geometric ergodicity and β -mixing property with exponential decay in this class of augmented ACD models. Although there are no analytical solutions for the autocorrelation function and moments of the duration process, we show that it is possible to derive the autocovariance function and moment recursion relations for the power λ of the duration process. Alternatively, one must restrict attention to particular subclasses, e.g. $\lambda \rightarrow 0$ and $\lambda = 1$, in order to work out expressions for any arbitrary moment and the autocovariance function.

We then demonstrate the practical usefulness of our ACD family modeling IBM price durations. Our findings clearly reject the restrictions imposed by the existing models in the literature. Further, we show that allowing for a concave shocks impact curve is paramount when fitting IBM price durations, because it mitigates the problem of overpredicting short durations. It is thus no wonder that we find some sort of substitutability between the Box-Cox transformation and the asymmetric effects given that both may lead to concavity of the shocks impact curve. In line with the findings of Bauwens, Giot, Grammig, and Veredas (2000), we unfortunately find no autoregressive duration process with good out-of-sample performance.

The remainder of the paper is organized as follows. Section 2 outlines the

statistical properties of the family of augmented ACD processes. Section 3 collects the findings of the empirical application to IBM price durations. Section 4 summarizes the main results and offers some concluding remarks.

2 The augmented ACD model

2.1 Model formulation

Let $x_i = t_i - t_{i-1}$ denote the time spell between two events occurring at times t_i and t_{i-1} . For example, price durations correspond to the time interval needed to observe a certain cumulative change in the stock price, whereas trade durations stand for the time elapsed between two consecutive transactions. To account for the serial dependence that is common to financial duration data, Engle and Russell (1998) formulate the accelerated time process $x_i = \psi_i \epsilon_i$, where the conditional duration process $\psi_i \propto E(x_i | \Omega_{i-1})$ is stochastically independent of iid sequence formed by ϵ_i and Ω_{i-1} is the set including all information available at time t_{i-1} . As in Hentschel (1995), we generalize the ACD processes by applying a Box-Cox transformation with parameter $\lambda \geq 0$ to the conditional duration process ψ_i , giving way to

$$\frac{\psi_i^\lambda - 1}{\lambda} = \omega_* + \alpha_* \psi_{i-1}^\lambda \left[|\epsilon_{i-1} - b| - c(\epsilon_{i-1} - b) \right]^v + \beta \frac{\psi_{i-1}^\lambda - 1}{\lambda}. \quad (1)$$

The shape parameter λ determines whether the Box-Cox transformation is concave ($\lambda \leq 1$) or convex ($\lambda \geq 1$).

The augmented autoregressive conditional duration (AACD) model then ensues by rewriting (1) as

$$\psi_i^\lambda = \omega + \alpha \psi_{i-1}^\lambda \left[|\epsilon_{i-1} - b| + c(\epsilon_{i-1} - b) \right]^v + \beta \psi_{i-1}^\lambda, \quad (2)$$

where $\omega = \lambda\omega_* - \beta + 1$ and $\alpha = \lambda\alpha_*$. The AACD model provides a flexible functional form that permits the conditional duration process $\{\psi_i\}$ to respond in distinct manners to small and large shocks. The shocks impact curve $g(\epsilon_i) = [|\epsilon_i - b| + c(\epsilon_i - b)]^v$ incorporates such asymmetric responses through the shift and rotation parameters b and c , respectively. As ϵ_i is nonnegative, the shift parameter b is key to the identification of the asymmetric response implied by the shocks impact curve. In turn, the parameter c determines whether rotation is clockwise ($c < 0$) or counterclockwise ($c > 0$). Interestingly, it is not necessarily the case that shift and rotation reinforce each other. Indeed, the shift parameter affects mostly small shocks, whereas rotation is dominant for large shocks. The shape parameter v plays a similar role to λ , inducing either concavity ($v \leq 1$) or convexity ($v \geq 1$) to the shocks impact curve. Figure 1 illustrates the behavior of the shocks impact curve $g(\cdot)$ according to the values of the shift, rotation and shape parameters.

The original ACD model of Engle and Russell (1998) is recovered by imposing $\lambda = v = 1$ and $b = c = 0$, whereas letting $\lambda \rightarrow 0$ and $b = c = 0$ renders the Box-Cox ACD specification put forward by Dufour and Engle (2000). Further, (1) reduces to Bauwens and Giot's (2000) logarithmic ACD models either if $\lambda \rightarrow 0$,

$v = 1$ and $b = c = 0$ (Type I) or if $\lambda, v \rightarrow 0$ and $b = c = 0$ (Type II). Following the GARCH literature, one may build other conditional duration models by imposing restrictions on (1). The examples we consider in the sequel include the asymmetric logarithmic ACD ($\lambda \rightarrow 0$ and $v = 1$), asymmetric power ACD ($\lambda = v$), asymmetric ACD ($\lambda = v = 1$), and power ACD ($\lambda = v$ and $b = c = 0$). Dufour and Engle (2000) independently propose a version of the asymmetric logarithmic ACD model with $b = 1$ under the name of exponential ACD model. We keep our notation because the linear ACD model with exponential distribution is sometimes referred to as the exponential ACD model. Table 1 summarizes the typology of ACD models under consideration.

2.2 Properties

In this section, we build heavily on Carrasco and Chen's (2000) general results to establish sufficient conditions that ensure β -mixing and finite higher order moments for (conditional) duration processes belonging to the augmented ACD family. The first step consists in casting (2) into a generalized polynomial random coefficient autoregressive model

$$X_{i+1} = A(e_i)X_i + B(e_i), \quad i = 0, 1, 2, \dots \quad (3)$$

where $\{e_i\}$ forms an iid sequence. Next, we apply Mokkadem's (1990) result for polynomial autoregressive models to derive the mixing properties of $\{\psi_i\}$. For the duration process $\{x_i\}$, we take advantage of Carrasco and Chen's result on the mixing properties of a process $Y_i = X_i + \varepsilon_i$, where X_i is a β -mixing homogeneous Markov process and ε_i is an iid noise with a continuous density. These two results are collected in Propositions 2 and 4 of Carrasco and Chen (2000), respectively.

Proposition 1: *Let $x_i = \psi_i \epsilon_i$, where ψ_i satisfies (2) and ϵ_i is an iid random variable that is stochastically independent of ψ_i . Assume further that the probability distribution of ϵ_i is absolutely continuous with respect to the Lebesgue measure on $(0, \infty)$ and such that the density is positive almost everywhere. Suppose that $|\beta| < 1$ and*

$$E |\beta + \alpha[|\epsilon_i - b| + c(\epsilon_i - b)]^v|^m < 1, \quad (4)$$

for some integer $m > 1$. Then, $\{\psi_i\}$ is a geometrically ergodic Markov process and, if initialized from its ergodic distribution, is also strictly stationary and β -mixing with exponential decay. Further, $E(\psi_i^{\lambda m}) < \infty$ and $E(x_i^{\lambda m}) < \infty$. Condition (4) with $m = 2$ is also necessary to entail geometric ergodicity of $\{\psi_i\}$ and $E(\psi_i^{2\lambda}) < \infty$. Lastly, if initialized from its ergodic distribution, $\{x_i\}$ is strictly stationary and β -mixing with exponential decay.

Proof: The first two results follow immediately from Carrasco and Chen's Proposition 2 with $X_i = \psi_i^\lambda$, $A(e_i) = \beta + \alpha[|\epsilon_i - b| + c(\epsilon_i - b)]^v$, and $B(e_i) = \omega$, where $e_i = (|\epsilon_i - b|, \epsilon_i)'$. The need for condition (4) with $m = 2$ stems from Lemma 2 of Pham (1986). The last result follows from Carrasco and Chen's Proposition 4 with $Y_i = \log x_i$, $X_i = \log \psi_i$, and $\varepsilon_i = \log \epsilon_i$. ■

If the interest were only in deriving sufficient conditions for the duration processes x_i to be nondegenerate and covariance stationary, one could alternatively

use the tools provided by Nelson (1990, Theorems 1 to 3) as in Hentschel (1995). Actually, for most of the models in the family spanned by the augmented ACD process, the conditions in the proposition above are both necessary and sufficient. The exceptions are formed by the models that ascertain a positive conditional duration even when at least one of the following restrictions are violated: $\omega > 0$, $\alpha > 0$, $\beta > 0$, and $|c| \leq 1$ for some odd integer v . For instance, letting $\lambda \rightarrow 0$ ensures nonnegativeness of the duration process without imposing further restrictions.

For the sake of completeness, we establish similar properties for the ACD models belonging to the family of augmented ACD processes. At first glance, it seems that it suffices to consider the parametric restrictions implied by each model in condition (4) to extract the corresponding result. That is not the case, though. To derive (4), one must impose restrictions on $A(\cdot)$ and $B(\cdot)$, which vary according to the specification of the model. More specifically, Carrasco and Chen's (2000) results require that $|A(0)| < 1$ and that, for some integer $m \geq 1$, $E|A(e_i)|^m < 1$ and $E|B(e_i)|^m < \infty$.

The generalized polynomial random coefficient autoregressive representation of the asymmetric logarithmic ACD process ensues from $X_i = \log \psi_i$, $A(e_i) = \beta$, and $B(e_i) = \omega + \alpha[|\epsilon_i - b| + c(\epsilon_i - b)]$, implying that condition (4) becomes $E(\epsilon_i^m) < \infty$. For the asymmetric power ACD model, $A(e_i) = \beta + \alpha[|\epsilon_i - b| + c(\epsilon_i - b)]^\lambda$ and $B(e_i) = \omega$, so that it suffices to impose that $E|\beta + \alpha[|\epsilon_i - b| + c(\epsilon_i - b)]^\lambda|^m < 1$. The latter condition also holds for the asymmetric ACD model with $\lambda = 1$ and for the power ACD specification with $b = c = 0$. While the Box-Cox ACD process asks for $E(\epsilon_i^m) < \infty$, the logarithmic ACD models of Bauwens and Giot (2000) require either that $E(\epsilon_i^m)$ exists (Type I) or that $|\alpha + \beta| < 1$ and $E|\log \epsilon_i|^m < \infty$ (Type II). As advanced by Carrasco and Chen (2000), in the linear ACD model, condition (4) reduces to $E|\beta + \alpha\epsilon_i|^m < 1$, which is equivalent to assuming that $\alpha + \beta < [E(\epsilon_i^m)]^{-1/m}$ is finite.

2.3 Higher-order moments and autocovariance function

In general, there is no analytical solution for the moments and autocorrelation function of duration processes belonging to the augmented ACD family. Nonetheless, it is possible to derive moment recursion relations and the autocovariance function of the power λ of the duration process by extension of He and Teräsvirta's (1999) results for the family of GARCH models.

To derive the λm -th moment $\mu_{\lambda m}$ of the duration process, we write (2) in its generalized polynomial random coefficient autoregressive representation

$$\psi_i^\lambda = A_{i-1}\psi_{i-1}^\lambda + B, \quad (5)$$

where $B = \omega$ and $A_i = \beta + \alpha g(\epsilon_i)$. Raising both sides to the power $m > 0$ and then applying recursions give

$$\begin{aligned} \psi_i^{\lambda m} &= \left(A_{i-1}\psi_{i-1}^\lambda + B \right)^m \\ &= A_{i-1}^m \psi_{i-1}^{\lambda m} + \sum_{j=1}^m \binom{m}{j} B^j A_{i-1}^{m-j} \psi_{i-1}^{\lambda(m-j)} \end{aligned}$$

$$= \psi_{i-n}^{m\lambda} \prod_{k=1}^n A_{i-k}^m + \sum_{j=1}^m \binom{m}{j} \sum_{k=1}^n B^j A_{i-k}^{m-j} \psi_{i-k}^{\lambda(m-j)} \prod_{\ell=1}^{k-1} A_{i-\ell}^m. \quad (6)$$

We are now ready to state the next proposition that documents moment recursion relations for the augmented ACD class of processes.

Proposition 2: *Let $x_i = \psi_i \epsilon_i$, where ψ_i satisfies (5) with $E|A_i^m| < 1$ and $\{\epsilon_i\}$ is an iid process stochastically independent of $\{\psi_i\}$. Assume further that the process started at some finite value infinitely many periods ago. It then follows that*

$$\mu_{\lambda m} = \frac{E\epsilon_i^{\lambda m}}{1 - EA_i^m} \sum_{j=1}^m \binom{m}{j} \frac{B^j EA_i^{m-j}}{E\epsilon_i^{\lambda(m-j)}} \mu_{\lambda(m-j)}. \quad (7)$$

for some integer $m \geq 1$ and $\mu_0 = 1$.

Proof: Because the process started at some finite value infinitely many periods ago, we can let $n \rightarrow \infty$ in (6) and take expectations, resulting in

$$\begin{aligned} E\psi_i^{\lambda m} &= E\psi_i^{m\lambda} (EA_i^m)^n + \sum_{j=1}^m \binom{m}{j} B^j EA_i^{m-j} E\psi_i^{\lambda(m-j)} \sum_{k=1}^n (EA_i^m)^{k-1} \\ &= \frac{1}{1 - EA_i^m} \sum_{j=1}^m \binom{m}{j} B^j EA_i^{m-j} E\psi_i^{\lambda(m-j)}. \end{aligned}$$

To complete the proof, it suffices to observe that ψ_i and ϵ_i are stochastically independent, and hence $\mu_{\lambda m} = E\epsilon_i^{\lambda m} E\psi_i^{\lambda m}$. ■

Before moving to the autocovariance function of the power λ of the duration process, two remarks are in order. First, assuming that $E|A_i^m| < 1$ is analogous to imposing condition (4) in Proposition 1. Second, the moment recursion relation in (7) involves moments that are possibly of fractional order. Unfortunately, it is not possible to derive expressions for a moment of an arbitrary order for such a general family of processes without restricting the shape parameter λ of the Box-Cox transformation of the conditional duration process. For instance, imposing linearity ($\lambda = 1$) suffices to extract a recursion relation involving moments of any integer order. Alternatively, one could also consider the subclass of conditional duration processes determined by the limiting case $\lambda \rightarrow 0$. We follow the latter approach in the end of this section in view that the log-transformation of the duration process is quite convenient for avoiding nonnegativeness constraints.

Proposition 3: *Let $x_i = \psi_i \epsilon_i$, where ψ_i satisfies (5) with $E|A_i^m| < 1$ for some integer $m \geq 2$. Let ϵ_i form an iid sequence stochastically independent of $\{\psi_i\}$ such that both $E\epsilon_i^\lambda$ and $EA_i\epsilon_i^\lambda$ are finite. It then follows that the autocovariance function $\gamma_{\lambda,n} = Ex_i^\lambda x_{i-n}^\lambda - \mu_\lambda^2$ of order $n \geq 1$ reads*

$$\gamma_{\lambda,n} = \frac{B^2 E\epsilon_i^\lambda}{1 - EA_i} \left[\sum_{j=0}^{n-1} (EA_i)^j + \frac{(EA_i)^{n-1} EA_i \epsilon_i^\lambda (1 + EA_i)}{1 - EA_i^2} - \frac{E\epsilon_i^\lambda}{1 - EA_i} \right]. \quad (8)$$

Proof: Multiplying both sides of (5) by ψ_{i-n}^λ yields

$$\begin{aligned}\psi_i^\lambda \psi_{i-n}^\lambda &= (A_{i-1} \psi_{i-1}^\lambda + B) \psi_{i-n}^\lambda \\ &= [A_{i-1} (A_{i-2} \psi_{i-2}^\lambda + B) + B] \psi_{i-n}^\lambda \\ &= \left(B + B \sum_{j=2}^n \prod_{k=1}^{j-1} A_{i-k} + \prod_{k=1}^n A_{i-k} \psi_{i-n}^\lambda \right) \psi_{i-n}^\lambda.\end{aligned}$$

Multiplying now both sides by $\epsilon_i \epsilon_{i-n}$ and then taking expectations give

$$\begin{aligned}E x_i^\lambda x_{i-n}^\lambda &= E \epsilon_i^\lambda B \left[1 + \sum_{j=2}^n (E A_i)^{j-1} \right] E \psi_i^\lambda + E \epsilon_i^\lambda (E A_i)^{n-1} E A_i \epsilon_i^\lambda E \psi_i^{2\lambda} \\ &= E \epsilon_i^\lambda \left[B E \psi_i^\lambda \sum_{j=0}^{n-1} (E A_i)^j + (E A_i)^{n-1} E A_i \epsilon_i^\lambda E \psi_i^{2\lambda} \right].\end{aligned}$$

The result then ensues from the fact that equation (6) implies that the first and second moments of ψ_i^λ are respectively $E \psi_i^\lambda = B/(1 - E A_i)$ and $E \psi_i^{2\lambda} = B^2(1 + E A_i)/[(1 - E A_i)(1 - E A_i^2)]$, whereas the moment recursion relation in (7) gives $\mu_\lambda = B E \epsilon_i^\lambda / (1 - E A_i)$. $\blacksquare$

As an example, consider the linear ACD process with an exponential noise introduced by Engle and Russell (1998), which results from $\lambda = 1$, $A_i = \beta + \alpha \epsilon_i$ and $B = \omega$. Proposition 2 then implies that

$$\mu_m = \frac{\Gamma(m+1)}{1 - E(\beta + \alpha \epsilon_i)^m} \sum_{j=1}^m \binom{m}{j} \frac{E(\beta + \alpha \epsilon_i)^{m-j}}{\Gamma(m-j)} \omega^j \mu_{m-j}$$

provided that $\alpha + \beta < 1$. Solving for $m = 1$ and $m = 2$ yields the first two moments as derived in Engle and Russell (1998). In turn, it follows from Proposition 3 that the autocovariance function of order n reads

$$\gamma_n = \omega^2 \left\{ \frac{1 - (\alpha + \beta)^n}{1 - (\alpha + \beta)} + \frac{(\alpha + \beta)^{n-1} (1 + \alpha + \beta) (2\alpha + \beta)}{[1 - (\alpha + \beta)] [1 - (\alpha + \beta)^2 - \alpha^2]} \right\}.$$

This expression provides a sharper result than Bauwens and Giot's (2000) recursive formula for computing the autocovariance function of a linear ACD process with exponential errors.

We now focus on a particular subclass of the augmented ACD family that permits working out expressions for any arbitrary moment as well as the autocorrelation function. This subclass is determined by shrinking the Box-Cox shape parameter to zero ($\lambda \rightarrow 0$), yielding

$$\log \psi_i = \omega + \alpha g(\epsilon_{i-1}) + \beta \log \psi_{i-1}. \quad (9)$$

This subclass is particularly interesting for ensuring that the duration process is always positive regardless of the sign and magnitude of the parameters. In particular, it nests the asymmetric logarithmic ACD model, Bauwens and Giot's (2000)

logarithmic ACD specifications, and the Box-Cox ACD process put forward by Dufour and Engle (2000). He, Teräsvirta, and Malmsten (1999) derive analogous results for a class of exponential GARCH models.

To derive the m -th moment $\mu_m = E x_i^m$ of the duration process, it is convenient to write equation (9) in the exponential form. Raising both sides to the power $m > 0$ and then applying recursions give

$$\begin{aligned}\psi_i^m &= \exp [m \omega + m \alpha g(\epsilon_{i-1})] \psi_{i-1}^{m \beta} \\ &= \exp \left(m \omega \sum_{k=0}^{n-1} \beta^k \right) \prod_{k=1}^n \exp [m \alpha \beta^{k-1} g(\epsilon_{i-k})] \psi_{i-n}^{m \beta^n}.\end{aligned}$$

Assuming that $E \{ \exp [\kappa g(\epsilon_i)] \} < \infty$ for $\kappa \in (0, \infty)$ and that $|\beta| < 1$, yields

$$E \psi_i^m = \exp \left(m \omega \frac{1 + \beta^n}{1 - \beta} \right) \prod_{k=1}^n E \left\{ \exp [m \alpha \beta^{k-1} g(\epsilon_i)] \right\} E \psi_i^{m \beta^n}. \quad (10)$$

We are now ready to state the next result that reports the m -th moment of the duration process defined in (9).

Corollary 1: *Let $x_i = \psi_i \epsilon_i$, where ψ_i satisfies (9) with $|\beta| < 1$ and $\{\epsilon_i\}$ is an iid process stochastically independent of $\{\psi_i\}$. Assume that the process started at some finite value infinitely many periods ago. If both $E \epsilon_i^m$ and $E \{ \exp [m \alpha g(\epsilon_i)] \}$ are finite for some integer $m \geq 1$, it then follows that*

$$\mu_m = E \epsilon_i^m \exp [m \omega (1 - \beta)^{-1}] \prod_{k=1}^{\infty} E \exp \left\{ [m \alpha \beta^{k-1} g(\epsilon_i)] \right\}. \quad (11)$$

Proof: Because the process started at some finite value infinitely many periods ago, we can let $n \rightarrow \infty$ in (10), resulting in

$$E \psi_i^m = \exp [m \omega (1 - \beta)^{-1}] \prod_{k=1}^{\infty} E \left\{ \exp [m \alpha \beta^{k-1} g(\epsilon_i)] \right\}.$$

The result then follows from the fact that ψ_i and ϵ_i are stochastically independent. ■

Next we move to the autocovariance function of duration processes in the $(\lambda \rightarrow 0)$ -subclass of augmented ACD models. As before, the exponential form of (9) facilitates the task.

Corollary 2: *Let $x_i = \psi_i \epsilon_i$, where ψ_i satisfies (9) with $|\beta| < 1$ and is stochastically independent of the iid process $\{\epsilon_i\}$. Assume further that both $E \{ \exp [\alpha g(\epsilon_i)] \}$ and $E \{ \epsilon_i \exp [\alpha g(\epsilon_i)] \}$ are finite. It then follows that the autocovariance function $\gamma_n = E x_i x_{i-n} - \mu_1^2$ of order $n \geq 1$ reads*

$$\begin{aligned}\gamma_n &= E \epsilon_i E \left\{ \epsilon_i \exp [\alpha \beta^{n-1} g(\epsilon_i)] \right\} \prod_{k=1}^{n-1} E \left\{ \exp [\alpha \beta^{k-1} g(\epsilon_i)] \right\} \\ &\quad \times \prod_{k=1}^{\infty} E \left\{ \exp [\alpha (1 + \beta^n) \beta^{k-1} g(\epsilon_i)] \right\} \exp \left(\frac{2\omega}{1 - \beta} \right) - \mu_1^2. \quad (12)\end{aligned}$$

Proof: Consider the exponential form of the conditional duration process (9), then

$$\psi_i \psi_{i-n} = \exp \left(\omega \sum_{k=0}^{n-1} \beta^k \right) \prod_{k=1}^n \exp \left[\alpha \beta^{k-1} g(\epsilon_{i-k}) \right] \psi_{i-n}^{\beta^n + 1},$$

which means that

$$x_i x_{i-n} = \epsilon_i \epsilon_{i-n} \exp \left(\omega \sum_{k=0}^{n-1} \beta^k \right) \prod_{k=1}^n \exp \left[\alpha \beta^{k-1} g(\epsilon_{i-k}) \right] \psi_{i-n}^{\beta^n + 1}.$$

Taking expectations in both sides yields (12). ■

3 Empirical application

In this section, we estimate different ACD specifications using IBM price durations at the New York Stock Exchange (NYSE) from September to November 1996. Data were kindly provided by Luc Bauwens and Pierre Giot, who have formed a broad data set using the NYSE's Trade and Quote database. We define price duration as the time interval needed to observe a cumulative change in the mid-price of at least \$0.125 as suggested by Giot (2000). Price durations are closely tied to the instantaneous volatility of the mid-quote price process (Engle and Russell, 1997 and 1998); hence it is not surprising that they may have serious implications to option pricing (Pringent, Renault, and Scaillet, 1999) and intra-day risk management (Giot, 2000).

Apart from an opening auction, NYSE trading is continuous from 9:30 to 16:00. Overnight spells, as well as durations between events recorded outside the regular opening hours of the NYSE, are removed. As documented by Giot (2000), durations feature a strong time-of-the-day effect. We therefore consider diurnally adjusted durations $x_i = D_i / \varrho(t_i)$, where D_i is the plain duration in seconds and $\varrho(\cdot)$ denotes the diurnal factor determined by first averaging the durations over thirty minutes intervals for each day of the week and then fitting a cubic spline with nodes at each half hour. The resulting (diurnally adjusted) durations serve as input for the remainder of the analysis.

Table 2 describes the main statistical properties of the IBM price durations as well as, for the purpose of comparison, trade durations. The latter are the time elapsed between two consecutive transactions, gauging to some extent the degree of market activity. For each duration, we compute descriptive statistics for both plain and diurnally adjusted data. On average, the market witnesses a trade involving IBM stocks every 24 seconds, whereas it takes on average 3.5 minutes for a cumulative price change of \$0.125 to take place. The overdispersion that characterizes both price and trade durations is robust to the time-of-the-day effect, thus it is not an artifact due to data seasonality. Sample autocorrelations reveal that persistence is slightly reduced when we account for the diurnal factor.

Next, we estimate by maximum likelihood the ACD models listed in Table 1 assuming that ϵ_i is iid with Burr density

$$f_B(\epsilon_i; \theta_B) = \frac{\kappa \mu_{B,1}^\kappa \epsilon_i^{\kappa-1}}{\left(1 + \gamma \mu_{B,1}^\kappa \epsilon_i^\kappa\right)^{1+1/\gamma}}, \quad (13)$$

where $\kappa > \gamma > 0$ and

$$\mu_{B,m} \equiv \frac{\Gamma(1 + m/\kappa) \Gamma(1/\gamma - m/\kappa)}{\gamma^{1+m/\kappa} \Gamma(1 + 1/\gamma)}$$

denotes the m -th moment, which exists for $m < \kappa/\gamma$. The Burr family encompasses both the Weibull ($\gamma \rightarrow 0$), exponential ($\gamma \rightarrow 0$ and $\kappa = 1$), and log-logistic ($\gamma \rightarrow 1$) distributions.

Before discussing the estimation results, two remarks are in order. First, we assess both in-sample and out-of-sample performances by splitting the sample. More precisely, we reserve the last third for out-of-sample evaluation. Second, asymptotic standard errors are based on the outer-product-of-the-gradient (OPG) estimator of the information matrix since the absolute value function in the shocks impact curve makes Hessian-based estimates tricky to compute due to numerical problems. Nonetheless, for the models without the asymmetric effect, we have also computed robust standard errors rooted in the sandwich form involving both Hessian and OPG terms. In comparison to the OPG counterpart, the robust standard errors indicate a significantly lower degree of accuracy in the estimation of the parameters of the duration process, while a slightly better precision for the Burr parameters. The results are qualitatively similar and are therefore omitted.

Tables 3 and 4 report respectively the in-sample results for the existing models in the literature and the novel specifications. It is interesting to notice that the estimates of the Burr parameters κ and γ are quite robust regardless of the specification of the duration process. They imply that the baseline hazard rate function is nonmonotonic and that there are at most three finite moments in view that $\hat{\kappa}/\hat{\gamma} \in [2.7173, 3.0438]$. The parameter estimates of the linear and logarithmic ACD models are very much in line with the previous results in the literature (see columns ACD, LACD I and LACD II, respectively). Interestingly, the log-likelihood value of the logarithmic ACD Type I model substantially differs from the values of the linear and logarithmic ACD Type II specifications. The asymmetric logarithmic ACD model with $b = 1$ introduced by Dufour and Engle (2000) palpably increases the log-likelihood value (-4,920.5 versus -4,950.5), suggesting that asymmetry may play a role (see column EXACD). The last column BCACD shows however that letting the power v of ϵ_{i-1} free to vary in the logarithmic ACD processes amplifies even more the log-likelihood value than introducing asymmetric effects. Indeed, in the Box-Cox ACD model, $\hat{v}$ is significantly different from both zero and one, lending support against the logarithmic ACD Type I and II models, respectively.

In the power ACD specification, we notice that the shape parameter λ of the Box-Cox transformation is also significantly different from both zero and one (see column PACD). This indicates that the restrictions imposed by the linear and the logarithmic ACD Type I models seem inconsistent with the data, even though the latter is only marginally inferior to the power ACD model in terms of log-likelihood value. Introducing an asymmetric effect to the power ACD specification ameliorates only marginally the fit of the model (see column A-PACD). Despite the fact that b is significantly different from zero, the standard error of c is quite large, showing that the shocks impact curve features no rotation. Although both shift and rotation parameters are significant in the asymmetric ACD specification

(see column A-ACD), it violates the constraints usually imposed to ensure the nonnegativeness of the duration process, namely $\alpha > 0$ and $|c| < 1$. The A-LACD column shows that all parameters are significantly different from zero in the asymmetric logarithmic ACD model. In particular, given that $\hat{\alpha}$ is negative, the shift and rotation effects are such that the shocks impact curve is concave.

The figures displayed in the column AACD demonstrate that the double Box-Cox transformation ($\lambda \neq v$) brings about further improvements as indicated by the value of the log-likelihood of the augmented ACD model. The difference between $\hat{v}$ and $\hat{\lambda}$ is striking. Indeed, there is strong evidence supporting that λ converges to zero (i.e. the log transformation), whereas $0.1310 < \hat{v} < 0.5178$ with 99% of confidence. The fact that $\hat{\lambda}$ is close to zero also explains why the estimate of α is not statistically different from zero. From equations (1) and (2), it happens that $\alpha = \alpha_* \lambda$ only if $\lambda > 0$, while α and α_* are equivalent in the limiting case $\lambda \rightarrow 0$. Table 4 reports $\hat{\alpha} = \hat{\alpha}_* \hat{\lambda}$ and the corresponding standard error as computed by the delta method. It is therefore straightforward to retrieve the estimate of α_* from the figures in Table 4: Indeed, $\hat{\alpha}_* = 0.3898$ with standard error equal to 0.1839.

To have a better idea about the in-sample fit of the models, we undertake an informal log-likelihood comparison that accounts for overparameterization. We do not pursue a formal analysis based on log-likelihood ratio tests because, due to the presence of inequality constraints in the parameter space, the limiting distribution of the test statistic is a mixing of chi-square distributions with probability weights depending on the variance of the parameter estimates (Wolak, 1991). Accordingly, it is extremely difficult to obtain empirically implementable asymptotically exact critical values. As an alternative, Wolak suggests applying asymptotic bounds tests. However, bounds are in most instances quite slack, often yielding inconclusive results. We therefore compute the Bayesian information criterion $BIC = -(2 \log \mathcal{L} - k \log T)/T$, where $\log \mathcal{L}$ denotes the value of the log-likelihood, k the number of parameters and T the number of in-sample observations. In terms of BIC values, the horse race winners are the Box-Cox ACD, power ACD, the asymmetric logarithmic ACD, the logarithmic ACD Type I models and, to a lesser extent, the augmented ACD process. Indeed, the rewards of the extra flexibility granted by these specifications are in contrast to the poor performance of the linear and logarithmic ACD Type II models. Further, letting λ free to vary and accounting for asymmetric effects seem operate as substitute sources of flexibility. For instance, the power ACD specification fits better the data than the asymmetric power ACD model according to the Bayesian information criterion, whereas it is very rewarding to introduce asymmetric responses to shocks in specifications with fixed λ .

Figure 2 portrays the effective shocks impact curves of each specification by depicting the variation of the conditional duration $\Delta\psi_i \equiv \psi_i - \psi_{i-1}$ in response to a shock ϵ_{i-1} at time t_{i-1} . We fix the conditional duration process ψ_{i-1} at time t_{i-1} to one, while we vary the shock ϵ_{i-1} from zero to five.¹ It is striking that, in all instances, $\Delta\psi_i$ reacts in a very similar fashion to the shock. In particular, it seems that the concavity of the shocks impact curve is the most important

¹ We refrain from plotting the shocks impact curve for larger shocks because it is merely a byproduct of the assumed specification of the duration process, without necessarily representing some meaningful property of the data (Hentschel, 1995).

feature to account for when modeling IBM price durations, alleviating the problem of overpredicting short durations. In the sequel, we argue that the apparently substitutability between the Box-Cox transformation and the asymmetric effects is chiefly caused by the need to achieve concavity of the shocks impact curve.

The asymmetric linear and logarithmic ACD's shocks impact curves are concave only for certain values of the shift and rotation parameters, namely $b > 0$ and $c < -1$. From this perspective, the parameter estimates reported in the column EXACD in Table 3 and columns A-ACD and A-LACD in Table 4 are not surprising. The estimates of the shift and rotation parameters are significantly different from zero and inferior to minus one, respectively. In contrast, if the shape parameter v is inferior to one, both the Box-Cox and augmented ACD models produce concave shocks impact curves. In the case of the AACD, this holds regardless of the shift and rotation parameters, hence it comes with no wonder that the corresponding estimates are not jointly significant in the AACD specification. As the power ACD model imposes $\lambda = v$, the estimate of $\hat{\lambda}$ sets in towards the estimates of v in the Box-Cox and augmented ACD models so as to entail a concave shocks impact curve. The same happens with the asymmetric power ACD model, despite the fact that, at first glance, one could also induce concavity through the shift and rotation parameters. It turns out, however, that to ensure a concave shocks impact curve the absolute value of the rotation parameter must exceed one, running counter to the nonnegativeness constraint.² Indeed, the augmented ACD model avoids this problem by letting λ converge to zero, thereby mimicking the asymmetric logarithmic ACD specification. All in all, Figure 2 illustrates some of the pitfalls from the specific to general modeling approach: There are various ways to achieve a concave shocks impact curve that the data call for and failing to start from a sufficiently general specification may point to quite misleading directions.

We now infer about the statistical properties of the duration processes by checking whether they satisfy the sufficient conditions for strict stationarity derived in Proposition 1. The aim is to illustrate how to use Proposition 1 for testing purposes. Maximum likelihood requires strict stationarity of the duration process to ensure consistency, hence estimates that violate either $|\beta| < 1$ or (4) are not very reliable. In the linear ACD model, this is equivalent to verifying whether $|\alpha + \beta| < \mu_{B,m}^{-1/m} < \infty$ for some integer $m > 1$. The second inequality poses no problem as $\mu_{B,m}$ exists for $m < \hat{\kappa}/\hat{\gamma} = 3.0139$. However, $\hat{\alpha} + \hat{\beta} = 0.9915$, whereas $m = 2$ yields $\hat{\mu}_{B,2}^{-1/2} = 0.4348$. In contrast, all other specifications seem to satisfy the sufficient conditions put forth in Proposition 1. For both versions of the logarithmic ACD model satisfy condition (4) since $|\hat{\alpha} + \hat{\beta}| < 1$ in Type I and $|\hat{\beta}| < 1$ in Type II. Further, $|\hat{\beta}| < 1$ guarantees that both restricted (EXACD) and unrestricted (A-LACD) versions of the asymmetric logarithmic ACD model as well as the Box-Cox ACD process are strictly stationary. The power ACD model requires that $E|\beta + \alpha \epsilon_i^\lambda|^2 < 1$ for some integer $m > 1$, which reduces to $|\alpha + \beta| < \mu_{B,2\lambda}^{-1/2}$ for $m = 2$. The latter inequality is empirically satisfied as the parameter estimates are such that $0.9738 = \hat{\alpha} + \hat{\beta} < \hat{\mu}_{B,2\hat{\lambda}}^{-1/2} = 1.0058$. Numerical results based on 10,000

² Unlike what occurs in the asymmetric ACD case, the estimation of the asymmetric power ACD model depends heavily on this constraint, since if the shocks impact curve is negative complex numbers would arise disrupting the maximum likelihood algorithm.

Monte Carlo simulations also show that (4) holds for the asymmetric ACD and asymmetric power ACD models. As $\lambda \rightarrow 0$ in the augmented ACD model, strict stationarity follows from the fact that $|\hat{\beta}| < 1$.

To check the assumptions on the error term, we first apply Fernandes and Grammig's (2000) D-test to assess the goodness-of-fit of the Burr distribution. The D-test statistic gauges the closeness between the parametric and nonparametric estimates of the density function. While the parametric density estimate is consistent only under the null of correct specification of the distribution, the kernel estimate is consistent even under the alternative. The test statistic is thus presumably close to zero under the null, whereas it should be large under the alternative. The D-test is attractive for three reasons. First, although it is slightly conservative, the D-test entails excellent power against both fixed and local alternatives. Second, it is nuisance parameter free in that there is no asymptotic cost in replacing errors with estimated residuals. Third, the asymptotic normality of the D-test is derived under mixing conditions, keeping the procedure clear of the assumption of serial independence.

To avoid boundary effects in the kernel density estimates due to the nonnegativeness of standardized durations, we work with log-residuals rather than plain residuals. All nonparametric density estimates use a Gaussian kernel, whereas the bandwidths are chosen according to an adjusted-version of Silverman's (1986) rule of thumb. The adjustment is necessary because the asymptotic theory of the D-test requires a slight degree of undersmoothing so as to avoid additional bias terms (see Fernandes and Grammig, 2000). Despite the fact that the p-values of the D-test tend to decrease with the degree of smoothing, the results are qualitatively robust to the bandwidth values.

The D-test results illustrate the rewards of the extra flexibility provided by the AACD family of models. There is no standard specification that performs well as seen in Table 3. At the 1% level of significance, we soundly reject the linear and logarithmic Type I ACD models, whereas we find a borderline result for the asymmetric logarithmic ACD model with $b = 1$ proposed by Dufour and Engle (2000). At the 5% level, rejection ensues for the Box-Cox ACD model, while rejecting the logarithmic ACD Type II specification is somewhat arguable given that the p-value is very close to 0.05. The figures in Table 4 are much rosier: There is indeed no clear rejection, though we find a borderline result for the asymmetric ACD model at the 5% significance level. The D-test results also indicate that the asymmetric logarithmic ACD specification is the most successful model, achieving a quite large p-value. Figure 3 illustrates this pattern by plotting the kernel and parametric density estimates of the log-residual for the two groups of models in the first and second column, respectively. While there are striking discrepancies in the first column, the nonparametric density estimates nicely oscillate around the parametric density estimates of the log-residuals in the second column.

Next, we inspect whether the standardized durations are serial independent using the BDS test, which has excellent power against several forms of dependence (Brock, Dechert, Scheinkman, and LeBaron, 1996). The fact that the BDS statistic is nuisance parameter free for additive models (de Lima, 1996) is quite convenient given that we are dealing with estimated residuals rather than true errors. Since a simple log-transformation renders additive any duration process in the AACD

family, we apply the testing procedure to log-residuals with tuning parameter set to the sample standard deviation and dimension m varying from 2 to 4 (see Brock, Dechert, Scheinkman, and LeBaron, 1996).³ The BDS statistics are far above the critical values provided by the standard normal limiting distribution, rejecting the null of serial independence in all instances. This means that we must enhance the lag structure of the duration processes inasmuch as there are no omitted variables.⁴

Table 5 displays the out-of-sample results, which are not very favorable. The serial independence assumption is clearly rejected for every model in Table 3 with exception to the logarithmic ACD Type I model that gives rise to borderline results at the 1% level of significance. The results for our novel specifications are slightly better in that borderline results at the 1% level becomes the rule. More importantly, the D-test results suggest that the Burr density estimates fit very poorly the out-of-sample residuals. Figure 4 indeed reveals that the parametric density estimates implied by the Burr assumption are too negatively skewed in comparison to the nonparametric kernel density estimates of the out-of-sample log-residuals. This poor out-of-sample performance is actually not unexpected. Bauwens, Giot, Grammig, and Veredas (2000) find similar results for a wide array of conditional duration processes applied to price durations. In particular, they estimate the linear and log versions of the ACD model, the stochastic volatility duration model (Ghysels, Gouriéroux, and Jasiak, 1997), the stochastic conditional duration model (Bauwens and Veredas, 1999) and the threshold ACD model (Zhang, Russell, and Tsay, 1999).

To complement the previous analyses, we evaluate density forecasts both in-sample and out-of-sample using the graphical techniques put forward by Diebold, Gunther, and Tay (1998). The procedure leans upon the fact that the sequence of probability integral transforms $z_i = F_{\hat{\theta}}(\hat{\epsilon}_i)$ of iid residuals is iid $U(0, 1)$ under the right specification of the distribution of the residuals. Figures 5 and 6 plot respectively the histograms for the probability integral transforms of the in-sample and out-of-sample residuals with confidence intervals based on the null of uniformity. As is apparent, all models have trouble in explaining (very) short durations both in-sample and out-of-sample. It seems that, for every model, the frequency of short durations ($0.05 < z < 0.2$) is usually too high in detriment to the frequency of very short durations ($0 < z < 0.05$). The out-of-sample analysis further identifies that very long durations ($0.9 < z < 1$) are underrepresented in most instances.

³ We have also experimented with two other values for the tuning parameter, namely 0.5 and 1.5 times the sample standard deviation, in order to check how sensitive are these results. It turns out, however, that there are virtually no changes in the p-values.

⁴ We have attempted to control for market microstructure effects by accounting for adverse selection costs in the conditional expected duration of an asymmetric logarithmic ACD process. As in Bauwens and Giot (2000) and Bauwens and Veredas (1999), we consider three measures of informed trading activity: number of trades, average volume per trade and the average bid-ask spread (all variables are lagged so as to ensure that they are predetermined). The results are not very encouraging, however. Even though the corresponding coefficients are significantly negative as expected (Easley and O'Hara, 1992), there are no material changes in the results of the BDS test, while the p-values of the D-test are far below 0.001 both in-sample and out-of-sample.

4 Conclusion

This paper introduces a family of augmented ACD models that encompasses most specifications in the literature. The nesting leans upon a Box-Cox transformation to the conditional duration process and an asymmetric shocks impact curve. The motivation for the latter stems from Engle and Russell's (1998) empirical findings, evincing that the linear ACD model tends to overpredict after either very long or very short durations. We derive sufficient conditions for the existence of higher-order moments, strict stationarity, geometric ergodicity and β -mixing property with exponential decay in this class of ACD models.

Our empirical results on IBM price durations show that the restrictions imposed by the existing models in the literature are incompatible with the data, warranting the extra flexibility granted by the augmented ACD models. Actually, inspecting the parameter estimates of the different specifications we conclude that imposing concavity in the shocks impact curve is the main issue and therefore the Box-Cox transformation and the asymmetric response to shocks work to some extent as substitutes. The power ACD and asymmetric logarithmic ACD models produce the best in-sample fit, though out-of-sample residual diagnostic tests reject all formulations. Moreover, the BDS test uncovers additional dependence structure in the residuals, so that considering higher-order processes might prove useful.

References

- BAUWENS, L., AND P. GIOT (2000): “The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks,” *Annales d’Economie et de Statistique*, 60, 117–150.
- BAUWENS, L., P. GIOT, J. GRAMMIG, AND D. VEREDAS (2000): “A comparison of financial duration models via density forecasts,” CORE Discussion Paper 0060, Université Catholique de Louvain.
- BAUWENS, L., AND D. VEREDAS (1999): “The stochastic conditional duration model: A latent factor model for the analysis of financial durations,” CORE Discussion Paper 9958, Université Catholique de Louvain.
- BOX, G. E. P., AND D. R. COX (1964): “An analysis of transformations,” *Journal of the Royal Statistical Society B*, 26, 211–243.
- BROCK, W. A., W. D. DECHERT, J. A. SCHEINKMAN, AND B. LEBARON (1996): “A test for independence based on the correlation dimension,” *Econometric Reviews*, 15, 197–235.
- CARRASCO, M., AND X. CHEN (2000): “Mixing and moment properties of various GARCH and stochastic volatility models,” *forthcoming in Econometric Theory*.
- DE LIMA, P. (1996): “Nuisance parameter free properties of correlation integral based statistics,” *Econometric Reviews*, 15, 237–259.
- DIEBOLD, F. X., T. A. GUNTHER, AND A. S. TAY (1998): “Evaluating density forecasts with applications to financial risk management,” *International Economic Review*, 39, 863–883.
- DUFOUR, A., AND R. F. ENGLE (2000): “The ACD model: Predictability of the time between consecutive trades,” University of Reading and University of California at San Diego.
- EASLEY, D., AND M. O’HARA (1992): “Time and the process of security price adjustment,” *Journal of Finance*, 47, 577–605.

- ENGLE, R. F., AND J. R. RUSSELL (1997): "Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model," *Journal of Empirical Finance*, 4, 187–212.
- (1998): "Autoregressive conditional duration: A new model for irregularly-spaced transaction data," *Econometrica*, 66, 1127–1162.
- FERNANDES, M., AND J. GRAMMIG (2000): "Non-parametric specification tests for conditional duration models," ECO Working Paper 2000/4, European University Institute.
- GHYSELS, E., C. GOURIÉROUX, AND J. JASIAK (1997): "Stochastic volatility duration models," CREST Working Paper 9746, INSEE.
- GIOT, P. (2000): "Time transformations, intraday data and volatility models," *Journal of Computational Finance*, 4, 31–62.
- HE, C., AND T. TERÄSVIRTA (1999): "Properties of moments of a family of GARCH processes," *Journal of Econometrics*, 92, 173–192.
- HE, C., T. TERÄSVIRTA, AND H. MALMSTEN (1999): "Fourth moment structure of a family of first-order exponential GARCH models," Department of Economics and Statistics, Stockholm School of Economics.
- HENTSCHEL, L. (1995): "All in the family: Nesting symmetric and asymmetric GARCH models," *Journal of Financial Economics*, 39, 71–104.
- MOKKADEM, A. (1990): "Propriétés de mélange des modèles autoregressifs polynomiaux," *Annales de l'Institut Henri Poincaré*, 26, 219–260.
- NELSON, D. B. (1990): "Stationary and persistence in the GARCH(1,1) model," *Econometric Theory*, 6, 318–334.
- PHAM, D. T. (1986): "The mixing property of bilinear and generalised random coefficient autoregressive models," *Stochastic Processes and Their Applications*, 23, 291–300.
- PRINGENT, J.-L., O. RENAULT, AND O. SCAILLET (1999): "An autoregressive conditional binomial option pricing model," Université de Cergy-Pontoise, CREST, and Université Catholique de Louvain.

- SILVERMAN, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
- WOLAK, F. A. (1991): “The local nature of hypothesis tests involving inequality constraints in nonlinear models,” *Econometrica*, 59, 981–995.
- ZHANG, M. Y., J. R. RUSSELL, AND R. S. TSAY (1999): “A nonlinear autoregressive conditional duration model with applications to financial transaction data,” *forthcoming in Journal of Econometrics*.

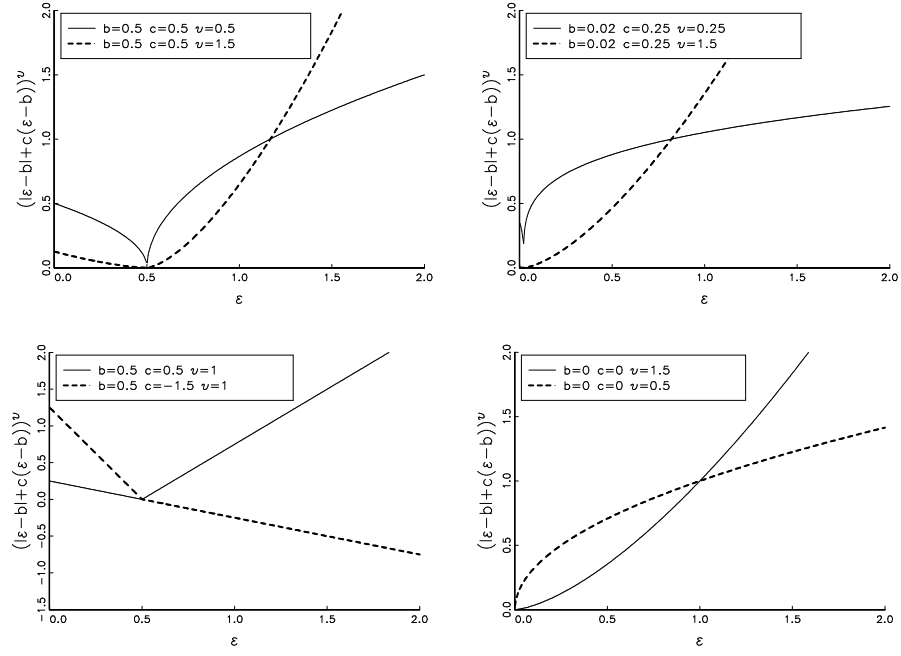


Figure 1: Feasible shocks impact curves for the augmented ACD model

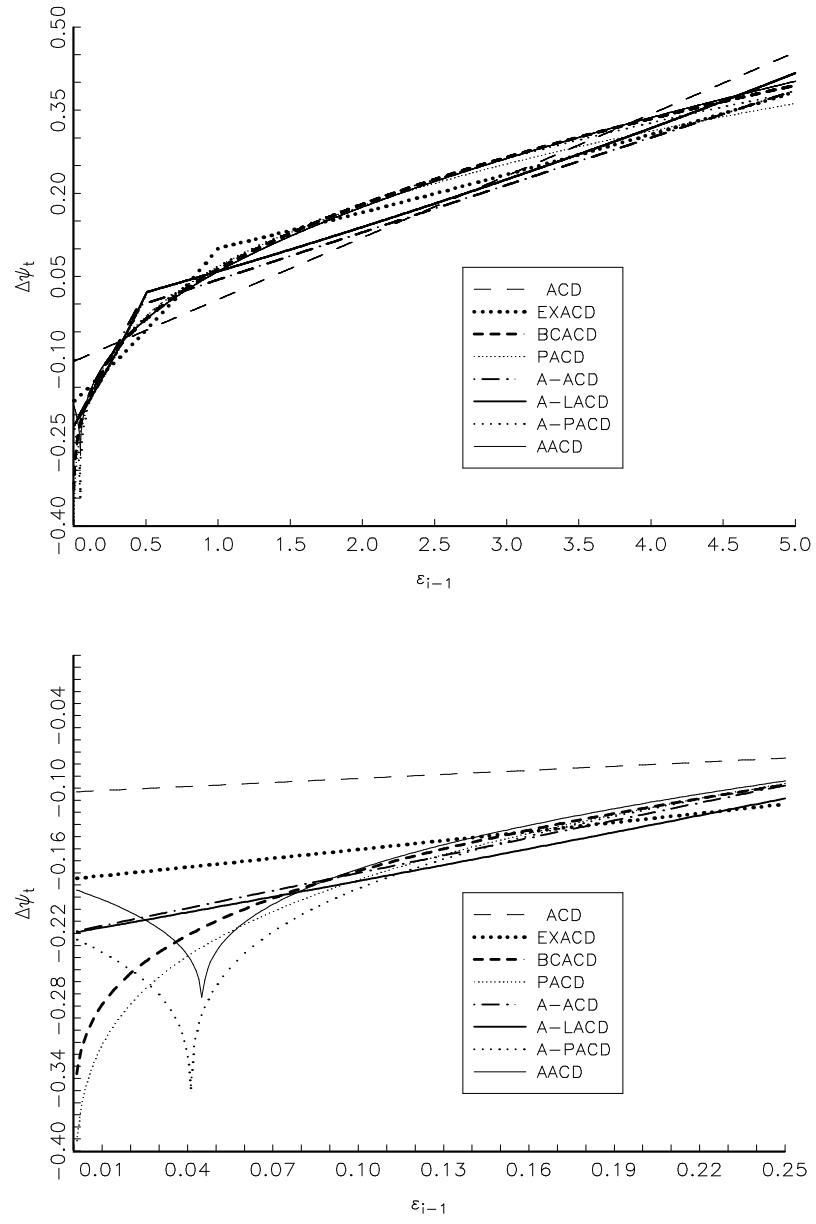


Figure 2: Empirical shocks impact curves

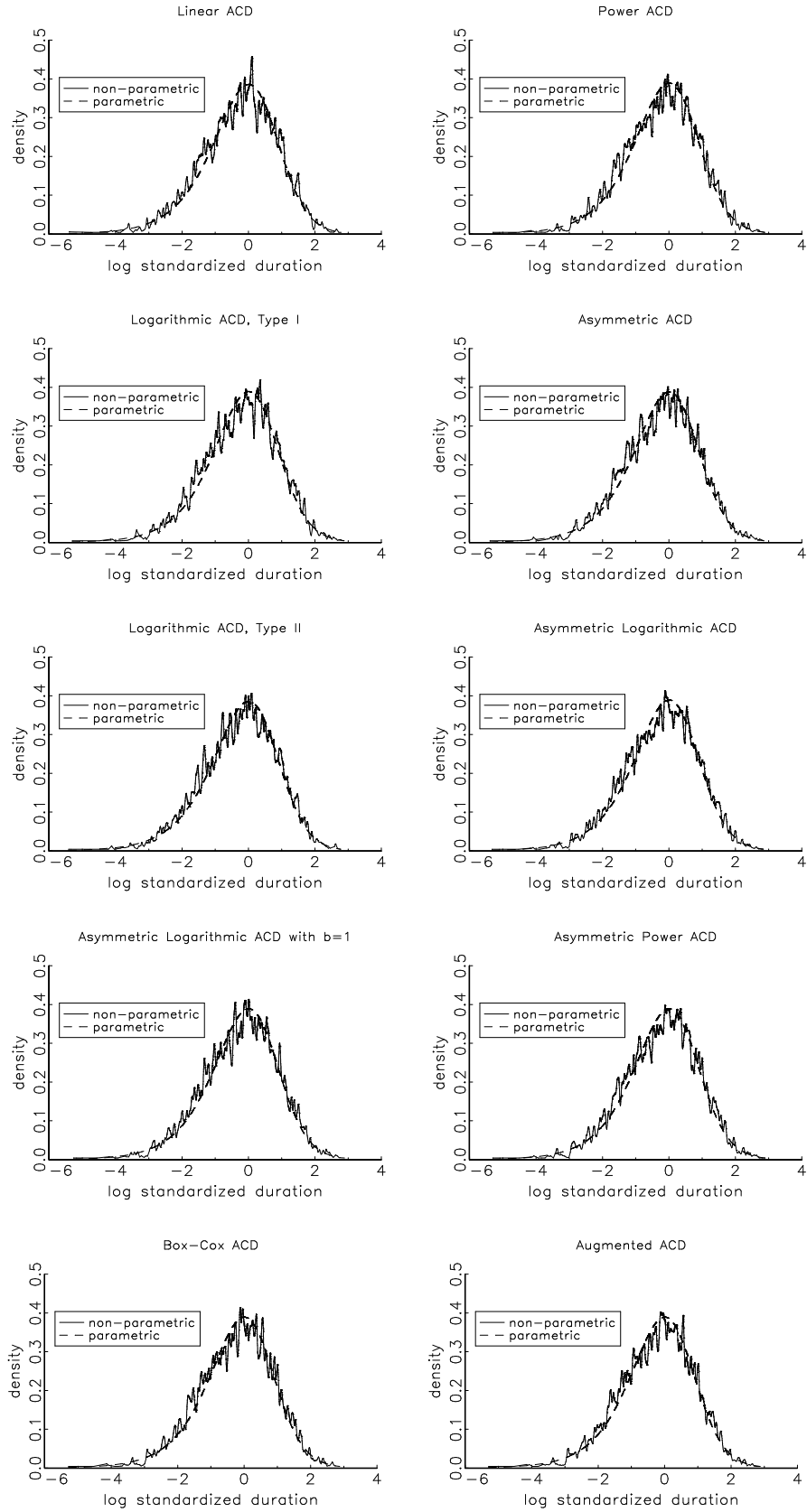


Figure 3: Density estimates for the log-residuals, in-sample

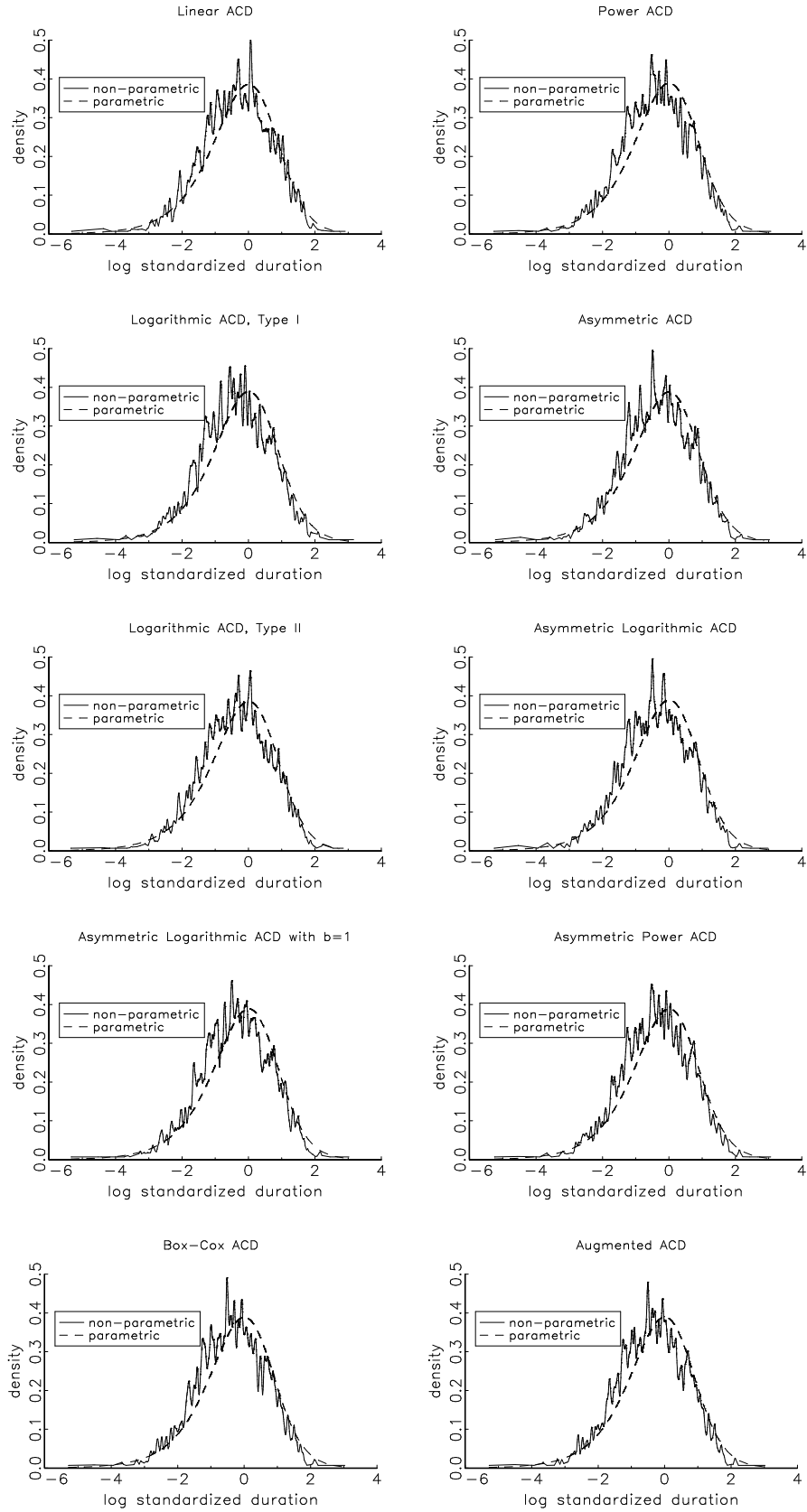


Figure 4: Density estimates for the log-residuals, out-of-sample

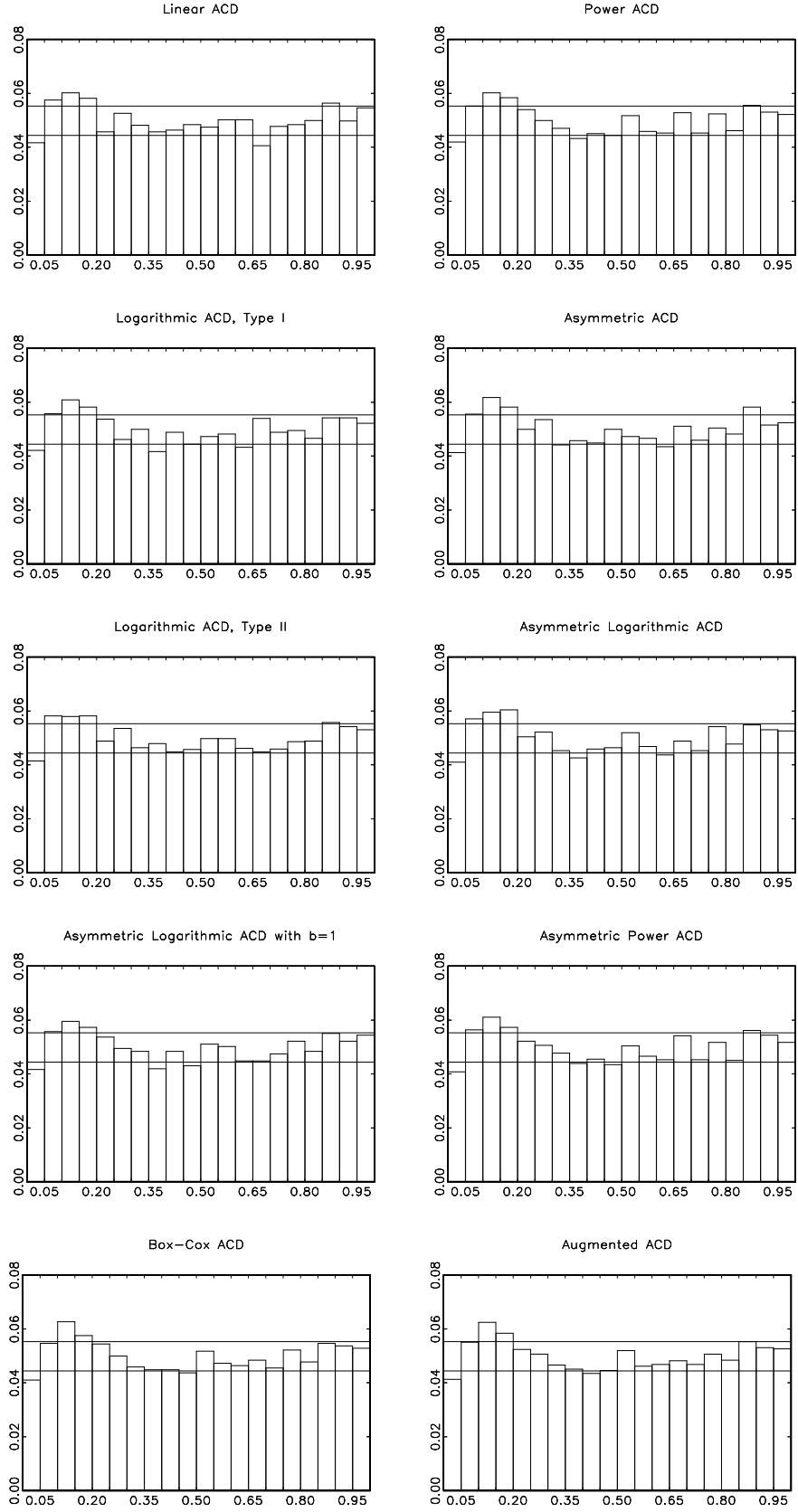


Figure 5: Histograms for the probability integral transforms, in-sample

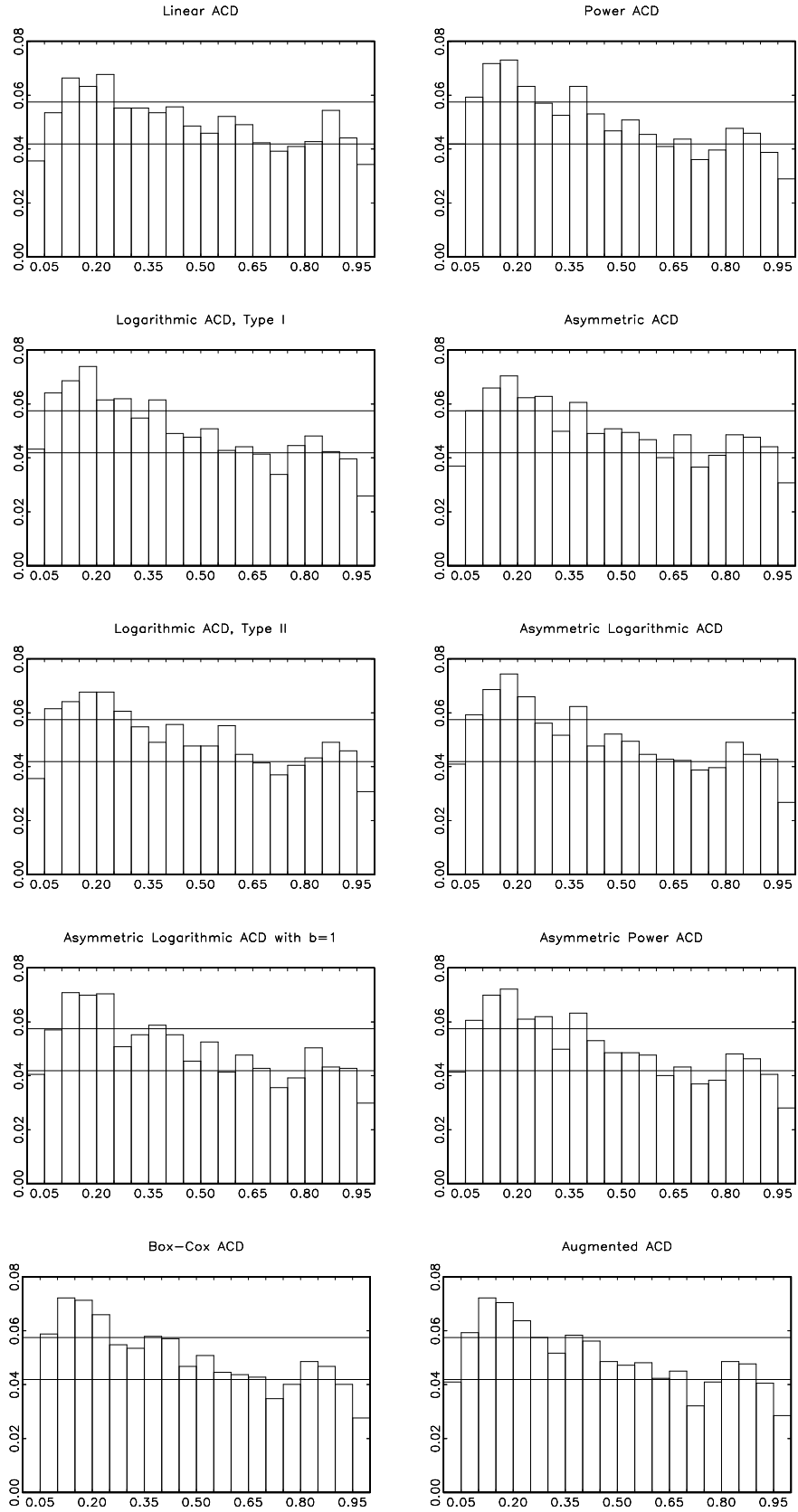


Figure 6: Histograms for the probability integral transforms, out-of-sample

Table 1
Typology of ACD models

Augmented ACD
$\psi_i^\lambda = \omega + \alpha \psi_{i-1}^\lambda \left[\epsilon_{i-1} - b + c(\epsilon_{i-1} - b) \right]^v + \beta \psi_{i-1}^\lambda$
Asymmetric Power ACD ($\lambda = v$)
$\psi_i^\lambda = \omega + \alpha \psi_{i-1}^\lambda \left[\epsilon_{i-1} - b + c(\epsilon_{i-1} - b) \right]^\lambda + \beta \psi_{i-1}^\lambda$
Asymmetric Logarithmic ACD ($\lambda \rightarrow 0$ and $v = 1$)
$\log \psi_i = \omega + \alpha \left[\epsilon_{i-1} - b + c(\epsilon_{i-1} - b) \right] + \beta \log \psi_{i-1}$
Asymmetric ACD ($\lambda = v = 1$)
$\psi_i = \omega + \alpha \psi_{i-1} \left[\epsilon_{i-1} - b + c(\epsilon_{i-1} - b) \right] + \beta \psi_{i-1}$
Power ACD ($\lambda = v$ and $b = c = 0$)
$\psi_i^\lambda = \omega + \alpha x_{i-1}^\lambda + \beta \psi_{i-1}^\lambda$
Box-Cox ACD ($\lambda \rightarrow 0$ and $b = c = 0$)
$\log \psi_i = \omega + \alpha \epsilon_{i-1}^v + \beta \log \psi_{i-1}$
Logarithmic ACD Type I ($\lambda, v \rightarrow 0$ and $b = c = 0$)
$\log \psi_i = \omega + \alpha \log x_{i-1} + \beta \log \psi_{i-1}$
Logarithmic ACD Type II ($\lambda \rightarrow 0, v = 1$ and $b = c = 0$)
$\log \psi_i = \omega + \alpha \epsilon_{i-1} + \beta \log \psi_{i-1}$
Linear ACD ($\lambda = v = 1$ and $b = c = 0$)
$\psi_i = \omega + \alpha x_{i-1} + \beta \psi_{i-1}$

Table 2
Descriptive statistics for IBM durations

	price		trade	
	plain	adjusted	plain	adjusted
sample size	6,728	6,728	60,454	60,454
mean	214.92	1.015	24.146	1.003
median	100	0.534	12	0.543
maximum	7,170	29.122	674	22.954
minimum	1	0.003	1	0.023
overdispersion	1.701	1.427	1.433	1.308
<i>n</i> -th order sample autocorrelation				
<i>n</i> = 1	0.280	0.221	0.212	0.165
<i>n</i> = 2	0.254	0.218	0.194	0.151
<i>n</i> = 3	0.262	0.204	0.183	0.147
<i>n</i> = 4	0.197	0.165	0.172	0.134
<i>n</i> = 5	0.185	0.145	0.159	0.119
<i>n</i> = 6	0.160	0.158	0.155	0.112
<i>n</i> = 7	0.148	0.139	0.147	0.107
<i>n</i> = 8	0.162	0.154	0.144	0.106
<i>n</i> = 9	0.127	0.133	0.151	0.109
<i>n</i> = 10	0.124	0.123	0.134	0.097
<i>n</i> = 11	0.115	0.122	0.140	0.100
<i>n</i> = 12	0.128	0.140	0.138	0.098
<i>n</i> = 16	0.099	0.122	0.137	0.095
<i>n</i> = 20	0.058	0.107	0.132	0.092
<i>n</i> = 24	0.061	0.117	0.129	0.086
<i>n</i> = 28	0.040	0.098	0.131	0.090
<i>n</i> = 32	0.030	0.092	0.116	0.081
<i>n</i> = 36	0.045	0.103	0.119	0.077

Table 3
Estimation results for the AACD family of models
IBM price durations (\$0.125 mid-price change)

parameter	ACD	LACD I	LACD II	EXACD	BCACD
ω	0.0171 (0.0038)	0.0774 (0.0069)	-0.0865 (0.0063)	-0.0964 (0.0131)	-0.5230 (0.1708)
α	0.1116 (0.0088)	0.1250 (0.0083)	0.0912 (0.0067)	-0.1157 (0.0152)	0.5843 (0.1768)
β	0.8799 (0.0089)	0.8327 (0.0127)	0.9759 (0.0042)	0.9614 (0.0064)	0.9616 (0.0067)
ν					0.2371 (0.0772)
c				-1.4927 (0.1288)	
κ	1.2616 (0.0318)	1.3036 (0.0331)	1.2592 (0.0318)	1.2892 (0.0327)	1.2954 (0.0330)
γ	0.4186 (0.0471)	0.4808 (0.0487)	0.4137 (0.0469)	0.4519 (0.0478)	0.4635 (0.0486)
$\log \mathcal{L}$	-4,952.4	-4,924.8	-4,950.5	-4,924.5	-4,920.5
BIC	2.2130	2.2007	2.2121	2.2013	2.1996
D-test	0.0029	0.0025	0.0488	0.0140	0.0266
BDS test					
$m = 2$	6.1662	3.6081	6.6112	4.2313	3.9808
$m = 3$	6.4405	4.0644	6.8937	4.5542	4.3355
$m = 4$	6.2835	4.0351	6.7120	4.4861	4.2609

Figures in parentheses correspond to standard errors based on the outer-product-of-the-gradient (OPG) estimator of the information matrix. $\log \mathcal{L}$ reports the value of the log-likelihood function, whereas BIC denotes the Bayesian information criterion. D-test displays the p-values of the nonparametric test proposed by Fernandes and Grammig (2000) applied to log-residuals using a Gaussian kernel. The BDS test figures correspond to the statistics for the log-residuals with tuning parameter set to the their sample standard deviation and dimension m .

Table 4
Estimation results for the AACD family of models
IBM price durations (\$0.125 mid-price change)

parameter	PACD	A-ACD	A-LACD	A-PACD	AACD
ω	0.0378 (0.0067)	0.0208 (0.0049)	0.0217 (0.0147)	0.0378 (0.0061)	0.0361 (0.0064)
α	0.1352 (0.0110)	-0.1990 (0.0452)	-0.2294 (0.0448)	0.1270 (0.0268)	0.00001 (0.1022)
β	0.8386 (0.0123)	0.9760 (0.0168)	0.9639 (0.0062)	0.8468 (0.0116)	0.9639 (0.0891)
λ	0.1751 (0.0848)			0.2573 (0.0681)	0.00003 (0.2621)
v					0.3244 (0.0747)
b		0.4456 (0.0741)	0.5066 (0.0680)	0.0411 (0.0019)	0.0451 (0.0015)
c		-1.4294 (0.1099)	-1.3172 (0.0724)	0.2326 (0.9446)	0.1117 (1.1718)
κ	1.2976 (0.0331)	1.2882 (0.0328)	1.2926 (0.0329)	1.2979 (0.0330)	1.2935 (0.0329)
γ	0.4688 (0.0487)	0.4556 (0.0484)	0.4589 (0.0485)	0.4685 (0.0482)	0.4588 (0.0483)
$\log \mathcal{L}$	-4,922.7	-4,930.1	-4,921.1	-4,921.0	-4,918.4
BIC	2.2006	2.2047	2.2007	2.2014	2.2011
D-test	0.0955	0.0494	0.4039	0.1353	0.1370
BDS test					
$m = 2$	3.8848	4.4200	4.1079	3.7222	4.0475
$m = 3$	4.2910	4.6834	4.3608	4.0894	4.3333
$m = 4$	4.2176	4.5590	4.2372	4.0154	4.2255

Figures in parentheses correspond to standard errors based on the outer-product-of-the-gradient (OPG) estimator of the information matrix. $\log \mathcal{L}$ reports the value of the log-likelihood function, whereas BIC denotes the Bayesian information criterion. D-test displays the p-values of the nonparametric test proposed by Fernandes and Grammig (2000) applied to log-residuals using a Gaussian kernel. The BDS test figures correspond to the statistics for the log-residuals with tuning parameter set to the their sample standard deviation and dimension m .

Table 5
Out-of-sample results for the AACD family of models
IBM price durations (\$0.125 mid-price change)

specification	D-test	BDS test		
		$m = 2$	$m = 3$	$m = 4$
ACD	0.0000	3.0788 (0.0021)	3.3353 (0.0009)	3.0781 (0.0021)
LACD I	0.0000	2.2648 (0.0235)	2.6152 (0.0089)	2.4650 (0.0137)
LACD II	0.0000	3.3325 (0.0009)	3.6466 (0.0003)	3.4494 (0.0006)
EXACD	0.0000	2.8164 (0.0049)	2.9517 (0.0032)	2.6325 (0.0085)
BCACD	0.0000	2.4313 (0.0151)	2.6859 (0.0072)	2.5238 (0.0116)
PACD	0.0000	2.3467 (0.0189)	2.6211 (0.0088)	2.4694 (0.0135)
A-ACD	0.0000	2.3588 (0.0183)	2.5710 (0.0101)	2.3313 (0.0197)
A-LACD	0.0000	2.5018 (0.0124)	2.6021 (0.0093)	2.3118 (0.0208)
A-PACD	0.0000	2.2576 (0.0240)	2.5811 (0.0098)	2.4364 (0.0148)
AACD	0.0000	2.3966 (0.0165)	2.6707 (0.0076)	2.4975 (0.0125)

D-test displays the p-values of the nonparametric test proposed by Fernandes and Grammig (2000) applied to out-of-sample log-residuals. The BDS test figures correspond to the statistics (and their p-values) for the log-residuals with tuning parameter set to the their sample standard deviation and dimension m .