Evaluating models of autoregressive conditional duration

Mika Meitz* and Timo Teräsvirta†

Department of Economic Statistics, Stockholm School of Economics
P. O. Box 6501, SE–113 83 Stockholm, Sweden

SSE/EFI Working Paper Series in Economics and Finance No. 557

First version March 2004, revised December 2004

Abstract

This paper contains two novelties. First, a unified framework for testing and evaluating the adequacy of an estimated autoregressive conditional duration (ACD) model is presented. Second, two new classes of ACD models, the smooth transition ACD model and the time-varying ACD model, are introduced and their properties discussed.

A number of new misspecification tests for the ACD class of models are introduced. They are Lagrange multiplier and Lagrange multiplier type tests against general forms of additive and multiplicative misspecification of the conditional mean function. These forms include tests against higher-order models, tests of no remaining ACD in the standardized durations, as well as tests of linearity and parameter constancy. In addition to its generality, the advantage of this testing approach is its ease of application, since all the resulting asymptotic null distributions are standard. The finite sample properties of the tests are investigated by simulation. A general observation is that the tests are well-sized and have good power. Versions of the test statistics robust to deviations from distributional assumptions other than those being explicitly tested are also given.

The smooth transition and time-varying ACD models are introduced, their main properties are examined, and they serve as alternatives in the tests of linearity and parameter constancy. Finally, the tests are applied to ACD models of the IBM stock traded at the New York Stock Exchange.

JEL classification: C22; C41; C52

Key words: ACD model; Model misspecification test; Lagrange multiplier test; Smooth transition ACD model; Nonlinear time series; Parameter constancy

---

*e-mail: mika.meitz@hhs.se

†e-mail: timo.terasvirta@hhs.se

Acknowledgements. This research has been supported by the Jan Wallander's and Tom Hedelius' Foundation, Grant No. J03–41. We thank Torben Andersen, Luc Bauwens, Tim Bollerslev, Annastiina Silvennoinen, an associate editor, and two anonymous referees for helpful comments and suggestions. Participants of the conferences and workshops 'Econometric Time Series Analysis - Methods and Applications', Linz, September-October 2003, 'Econometrics and Computational Economics', Helsinki, November 2003, 'Econometrics of Stock Markets - Analysis and Prediction', Paris, April 2004, 'Economics and Econometrics of Market Microstructure', Constance, June 2004, and the European Meeting of the Econometric Society, Madrid, August 2004, also provided useful remarks. We also wish to thank Yongmiao Hong and Tae-Hwy Lee who generously shared with us their program code for the generalized spectral test. The responsibility for any errors and shortcomings in this paper remains ours.

1 Introduction

High-frequency financial time series have become widely available during the past decade or so. Records of all transactions and quoted prices, so-called "ultra-high-frequency" data; Engle (2000), are available from many stock exchanges. An inherent feature is that such data are irregularly spaced in time. There are several approaches to tackling this feature of the data. In this paper we follow the line of work originated by Engle and Russell (1998), where the durations between events (trades, quotes, price changes etc.) are the quantities being modeled. These authors proposed a class of models called the Autoregressive Conditional Duration, or ACD, models, where conditional expected durations are modeled in a fashion similar to the way conditional variances are modeled using ARCH and GARCH models of Engle (1982) and Bollerslev (1986).

Following the GARCH literature, a number of extensions to the original linear ACD model by Engle and Russell (1998) have been suggested. These include the logarithmic ACD model of Bauwens and Giot (2000), and the threshold ACD model of Zhang, Russell, and Tsay (2001). The distribution associated with the conditional durations has also been suggested to have several different shapes. Examples include the the exponential and Weibull distributions as in Engle and Russell (1998), and the Burr and generalized gamma distributions suggested by Grammig and Maurer (2000) and Lunde (1999), respectively.

Despite the surge of different models for financial durations the issue of model evaluation using misspecification tests has not yet received as much attention as it should deserve. In time series econometrics, estimated models of the conditional mean, and more recently, ones of conditional variance, are typically subjected to a variety of evaluation tests to determine the adequacy of the specification. For models of conditional duration, some misspecification tests have also been proposed in the literature. These can be divided into two categories: tests of misspecification in the distribution of the error term, and of misspecification in the functional form of the conditional mean duration. Fernandes and Grammig (in press) as well as Bauwens, Giot, Grammig, and Veredas (2004) have suggested tests of the first type, and Hautsch (2002) proposed tests of the second type. Li and Yu (2003) derived a portmanteau test that can be used to evaluate the adequacy of an estimated ACD model. Finally, Hong and Lee (2003) considered a general diagnostic test which can be used as a misspecification test for ACD models.

In this paper, we present a framework for evaluating models of conditional duration based on Lagrange multiplier misspecification tests of the functional form of the conditional mean duration. Our goal is to derive easily applicable tests that can reveal various types of misspecification. We present our results in a general form, from which misspecification tests against specific alternatives are derived in a straightforward fashion. Our tests include ones against higher-order models and remaining ACD effects in the standardized durations, as well as tests of linearity and parameter constancy. In the process of deriving linearity and parameter constancy tests, we propose two new ACD specifications, namely the smooth transition ACD (STACD) model, and the time-varying ACD (TVACD) model.

This paper has similarities with the one by Lundbergh and Teräsvirta (2002) who derived misspecification tests for GARCH models. The two papers share the same goal: to derive easily applicable evaluation tools based on Lagrange multiplier test statistics. The types of model misspecification considered in these papers are similar. Derivations of the test statistics differ in that the error distributions of the ACD and GARCH models are not the same. Furthermore, the present paper contains a discussion of nonlinear alternatives to the standard ACD model.

The rest of the paper is organized as follows. In Section 2 we briefly review previous work on misspecification testing in ACD models and present the general results that form the basis of our misspecification tests. In Section 3 we derive tests against higher-order models and remaining ACD in the standardized durations. Section 4 presents the smooth transition ACD model and deduces a test of linearity, and Section 5 presents the time-varying ACD model and the test of parameter constancy. Section 6 contains the results of a simulation experiment. In Section 7 we estimate and evaluate ACD models using data from the New York Stock Exchange. Finally, Section 8 concludes.

2 Testing ACD models against general additive and multiplicative alternatives

2.1 Previous work on misspecification testing of ACD models

Evaluation of estimated ACD models by misspecification tests has not been commonplace in empirical work. Often the only diagnostic test applied for the purpose has been the Ljung-Box Q–statistic applied to the standardized or squared standardized durations. In the latter case the test is commonly called the McLeod-Li test (McLeod and Li, 1983). Nevertheless, as already mentioned, there are some papers proposing misspecification tests for ACD models. Bauwens, Giot, Grammig, and Veredas (2004) as well as Fernandes and Grammig (in press) discussed the testing for distributional misspecification. The former authors evaluated duration models using density forecast evaluation methods of Diebold, Gunther, and Tay (1998). Their method relied on the fact that the sequence of probability integral transforms of the one-step-ahead forecasts of the conditional densities of durations will be distributed as independent and identically distributed uniform (0,1) random variables when the one-step-ahead forecasts of the conditional densities of the durations coincide with the true densities. This is the null hypothesis to be tested. It may be rejected either because the error distribution of the model is misspecified or because the conditional mean is misspecified. The latter alternative is due to the fact that the choice of the conditional mean function affects the one-step-ahead forecasts. Fernandes and Grammig (in press) tested the distribution of the error term by comparing parametric and nonparametric estimates of the density of the standardized durations. Their test explicitly assumes that the conditional mean is correctly specified but again a rejection may also be a consequence of a misspecified conditional mean. In order to obtain more information about the situation, complementing tests of the distributional assumption by tests of the conditional mean specification is quite important.

The question of testing the functional form of the conditional mean of an ACD model was addressed in Hautsch (2002). He mentioned Lagrange multiplier, conditional moment, and integrated conditional moment (ICM) tests as potential tools for detecting misspecification and focuses on the latter two methods. Conditional moment tests (see e.g. Newey (1985)) are based on the fact that correct specification implies the validity of certain moment conditions. These tests are, however, known to be heavily dependent on the choice of weighting of the moment conditions. They do not require a well-specified alternative and are thus rather general misspecification tests. But then, they are based on a finite number of moment restrictions and cannot therefore be consistent against all possible alternatives. The ICM test (see e.g. Bierens (1990) or de Jong (1996)) employs an infinite amount of moment conditions and is consistent against every deviation from the null hypothesis. A consequence of this property is that the test is not very powerful against any particular alternative, which may be considered a disadvantage. Another drawback of the ICM test is that application requires approximating the asymptotic null distribution of the test statistic by simulation. This makes the use of the test computationally burdensome in the ACD case where the time series in applications can be quite long.

Recently, Li and Yu (2003) derived a portmanteau test for ACD models.[1] Their test is based on the residual autocorrelations of an estimated ACD model in the spirit of the Ljung-Box test. More generally, Hong and Lee (2003) proposed the generalized spectrum based test of Hong (1999) as a general diagnostic test for ACD and many other models. When applied to the standardized durations resulting from the estimation of an ACD model, this test is consistent against any type of pairwise serial dependence left in the standardized durations. We provide a description of this test in Appendix C, and use it as benchmark test in our power simulations in Section 6.

The purpose of this paper is to present a unified framework for evaluating ACD models using Lagrange multiplier (LM) tests. Using LM tests makes misspecification testing easy without sacrificing power. Since the model is only estimated under the null hypothesis, the need for, say, nonlinear ACD models can be investigated without the often burdensome task of actually estimating such a model. We derive general results from which tests against specific alternatives are easily derived. Such alternatives include higher-order models, remaining ACD in the standardized durations, as well as nonlinearity and parameter nonconstancy.

2.2 General theory

Let t_i be the time at which the ith event (trade, quote, price change etc.) occurs and denote by $x_i = t_i - t_{i-1}$ the duration between two consecutive events. Let $\mathcal{F}_{i-1}$ be the information set consisting of all information up to and including time t_{i-1}. Following Engle and Russell (1998), the class of exponential autoregressive conditional duration (ACD) models is defined as follows:

$$\begin{aligned} x_i &= \psi_i \varepsilon_i & (1) \\ \psi_i &= \psi_i\left(x_{i-1}, \ldots, x_1; \theta_1\right) & (2) \\ \varepsilon_i &\sim \text{i.i.d. } \exp(1) & (3) \end{aligned}$$

where $\psi_i = \psi_i\left(x_{i-1}, \ldots, x_1; \theta_1\right)$ is the duration conditional on $\mathcal{F}_{i-1}$.

The types of misspecification of this structure can be divided into two broad categories: the conditional duration is either additively or multiplicatively misspecified. This means that the true process is governed either by

$$x_i = \left(\psi_i + \varphi_i\right) \varepsilon_i \qquad (4)$$

(additive misspecification) or

$$x_i = \psi_i \varphi_i \varepsilon_i \qquad (5)$$

(multiplicative misspecification). Both in (4) and (5) the additional component of the conditional duration,

$$\varphi_i = \varphi_i\left(x_{i-1}, \ldots, x_1; \theta_1, \theta_2\right) \qquad (6)$$

is assumed to be an $\mathcal{F}_{i-1}$–measurable function that depends on additional parameters θ_2.

Let

$$\begin{aligned} \mathbf{a}_i\left(\theta_1\right) &= \frac{1}{\psi_i\left(\theta_1\right)} \frac{\partial \psi_i\left(\theta_1\right)}{\partial \theta_1} \\ \mathbf{b}_i\left(\theta_1, \theta_2\right) &= \frac{1}{\psi_i\left(\theta_1\right)} \frac{\partial \varphi_i\left(\theta_1, \theta_2\right)}{\partial \theta_2} \\ c_i\left(\theta_1\right) &= \frac{x_i}{\psi_i\left(\theta_1\right)} - 1. \end{aligned}$$

[1]We thank a referee for bringing this paper into our attention.

Furthermore, let the superscript '0' denote values evaluated at the true (under the null hypothesis) parameter values, and the hat '^' values evaluated at the maximum value of the likelihood under the null hypothesis. The following theorem defines the test against a general additive alternative.

Theorem 1 *Consider the model (4), (2), (3), (6), and, in addition to assuming that standard regularity conditions (see Theorem* 1 *of Engle (2000)) apply, assume that under the null hypothesis $H_0 : \theta_2 = \theta_2^0$, the function φ_i satisfies $\varphi_i\left(x_{i-1},\ldots,x_1;\theta_1,\theta_2^0\right) \equiv 0$. Then, under the null hypothesis $H_0 : \theta_2 = \theta_2^0$, the LM statistic*

$$LM = \left\{\sum_{i=1}^{n} \hat{c}_i \hat{\mathbf{b}}_i'\right\} \left\{\sum_{i=1}^{n} \hat{\mathbf{b}}_i \hat{\mathbf{b}}_i' - \left(\sum_{i=1}^{n} \hat{\mathbf{b}}_i \hat{\mathbf{a}}_i'\right) \left(\sum_{i=1}^{n} \hat{\mathbf{a}}_i \hat{\mathbf{a}}_i'\right)^{-1} \left(\sum_{i=1}^{n} \hat{\mathbf{a}}_i \hat{\mathbf{b}}_i'\right)\right\}^{-1} \left\{\sum_{i=1}^{n} \hat{c}_i \hat{\mathbf{b}}_i\right\} \tag{7}$$

has an asymptotic χ^2 distribution with $\dim \theta_2$ *degrees of freedom.*

Proof. See Appendix A. ■

Under the null hypothesis $H_0 : \theta_2 = \theta_2^0$, model (4) reduces to (1)–(3). In practice the LM test in Theorem 1 is most easily carried out using an auxiliary least squares regression on particular transformed variables. This can be done as follows (see for example Engle (1984)):

Procedure 2

(i) *Obtain the quasi maximum likelihood estimate of θ_1 under the null hypothesis, and compute $\hat{\mathbf{a}}_i' = \frac{1}{\psi_i(\hat{\theta}_1)} \frac{\partial \psi_i(\hat{\theta}_1)}{\partial \theta_1'}$, $\hat{\mathbf{b}}_i' = \frac{1}{\psi_i(\hat{\theta}_1)} \frac{\partial \varphi_i(\hat{\theta}_1, \theta_2^0)}{\partial \theta_2'}$, $\hat{c}_i = \frac{x_i}{\psi_i(\hat{\theta}_1)} - 1$, $i = 1,\ldots,n$, and $SSR_0 = \sum_{i=1}^{n} \hat{c}_i^2$.*

(ii) *Regress $\hat{c}_i$ on $\hat{\mathbf{a}}_i'$ and $\hat{\mathbf{b}}_i'$, $i = 1,\ldots,n$, and compute SSR_1.*

Then, under the null hypothesis, the test statistic $LM = n\left(SSR_0 - SSR_1\right)/SSR_0$ has an asymptotic χ^2 distribution with $\dim \theta_2$ *degrees of freedom.*

There is considerable empirical evidence in the ACD literature against exponentially distributed errors. Engle and Russell (1998) already considered the Weibull distribution, Grammig and Maurer (2000) suggested the Burr distribution, and Lunde (1999) used the generalized gamma distribution. Therefore one might consider basing the QML estimation and the derivation of the asymptotic distributions of the test statistics on these distributions. There is, however, a drawback in this approach. As Gouriéroux, Monfort, and Trognon (1984) originally showed, the QML parameter estimators of a correctly specified conditional mean model are consistent if, and only if, the quasi-maximum likelihood is based on a distribution belonging to the linear exponential family, regardless of what the true density is. For a discussion of this, see for example White (1994, pp. 62–70). The exponential distribution does belong to the linear exponential family, while the Weibull, Burr, and generalized gamma distributions do not (except for special cases). Therefore, the QML approach based on the exponential distribution will produce consistent estimators regardless of the true error distribution, while QML based on these other distributions will not unless the distribution used is the true density.

Drost and Werker (2004) pointed out that the use of QML based on the ordinary gamma distribution produces consistent estimators. Furthermore, the score vectors of the gamma distribution (suitably normalized) and the exponential distribution are proportional to each other.

Hence the estimators for the parameters of the conditional mean based on these two error distributions are identical.

While the QML approach based on the exponential distribution yields consistent estimators, misspecification of the conditional distribution of the durations may still affect the properties of our LM test statistics. This is because applying the test statistics involves implicit assumptions about this conditional distribution. In particular, it is tacitly assumed that the conditional variance of the durations is correctly specified under the null hypothesis. As our interest lies in the specification of the conditional mean, we do not wish other properties of the conditional distribution to affect the properties of our test statistics. The results of Wooldridge (1991) are helpful here. Since we are using consistent estimators based on QML, his results allow us to derive 'robust' versions of the test statistics such that their asymptotic behaviour is unaffected by possible misspecification of the conditional distribution beyond the conditional mean. Applying these results leads to the procedure given below. They cannot be used, however, when the QML estimators are obtained assuming Weibull, Burr, or generalized gamma distributed errors, because a key requirement in Wooldridge (1991) is the consistency of the estimators.

Procedure 3

(i) *Obtain the quasi maximum likelihood estimate of* θ_1 *under the null hypothesis, and compute* $\hat{\mathbf{a}}_i' = \frac{1}{\psi_i(\hat{\theta}_1)} \frac{\partial \psi_i(\hat{\theta}_1)}{\partial \theta_1'}$, $\hat{\mathbf{b}}_i' = \frac{1}{\psi_i(\hat{\theta}_1)} \frac{\partial \varphi_i(\hat{\theta}_1, \theta_2^0)}{\partial \theta_2'}$ *and* $\hat{c}_i = \frac{x_i}{\psi_i(\hat{\theta}_1)} - 1$, $i = 1, \ldots, n$.

(ii) *Regress* $\hat{\mathbf{b}}_i'$ *on* $\hat{\mathbf{a}}_i'$, $i = 1, \ldots, n$, *and save the* $(\dim \theta_2 \times 1)$ *residual vectors* $\hat{\mathbf{r}}_i$.

(iii) *Regress* 1 *on* $\hat{c}_i \hat{\mathbf{r}}_i$, $i = 1, \ldots, n$, *and compute the sum of squared residuals, SSR, from this regression.*

Then, under the null hypothesis, the test statistic $nR^2 = n - SSR$ *has an asymptotic* χ^2 *distribution with* $\dim \theta_2$ *degrees of freedom.*

For the general multiplicative alternative the asymptotic distribution theory has the following form:

Theorem 4 *Consider the model (5), (2), (3), (6), and, in addition to assuming that standard regularity conditions (see Theorem 1 of Engle (2000)) apply, assume that under the null hypothesis* $H_0 : \theta_2 = \theta_2^0$, *the function* φ_i *satisfies* $\varphi_i\left(x_{i-1}, \ldots, x_1; \theta_1, \theta_2^0\right) \equiv 1$. *Then, under the null hypothesis* $H_0 : \theta_2 = \theta_2^0$, *the LM statistic*

$$LM = \left\{\sum_{i=1}^{n} \hat{\psi}_i \hat{c}_i \hat{\mathbf{b}}_i'\right\} \left\{\sum_{i=1}^{n} \hat{\psi}_i^2 \hat{\mathbf{b}}_i \hat{\mathbf{b}}_i' - \left(\sum_{i=1}^{n} \hat{\psi}_i \hat{\mathbf{b}}_i \hat{\mathbf{a}}_i'\right) \left(\sum_{i=1}^{n} \hat{\mathbf{a}}_i \hat{\mathbf{a}}_i'\right)^{-1} \left(\sum_{i=1}^{n} \hat{\psi}_i \hat{\mathbf{a}}_i \hat{\mathbf{b}}_i'\right)\right\}^{-1} \left\{\sum_{i=1}^{n} \hat{\psi}_i \hat{c}_i \hat{\mathbf{b}}_i\right\} \tag{8}$$

has an asymptotic χ^2 *distribution with* $\dim \theta_2$ *degrees of freedom.*

Under the null hypothesis $H_0 : \theta_2 = \theta_2^0$, model (5) reduces to (1)–(3). The procedures for the practical application of the test are almost identical to the ones given in Procedures 2 and 3. The only modification needed is to replace $\hat{\mathbf{b}}_i$ by $\hat{\psi}_i \hat{\mathbf{b}}_i$ throughout.

It may be mentioned that there is a close connection between the LM tests presented here and the conditional moment tests discussed in Hautsch (2002) in the sense that our tests can be interpreted as particular conditional moment tests. By choosing appropriate moment conditions and weighting functions when deriving a conditional moment test one obtains a test that is asymptotically equivalent to our LM test, the only difference being a different consistent estimator of the information matrix.

3 Applications of the theory

In this section we present three misspecification tests that are applications of our general theory. The first two are tests of the original ACD model of Engle and Russell (1998) and the LOGACD model of Bauwens and Giot (2000) against higher-order alternatives. The third one is a test for remaining ACD in the standardized durations. We shall present the precise alternative under consideration, state the distributional results corresponding to Theorem 1 or 4, and give explicit formulas for the quantities $\hat{\mathbf{a}}_i$ and $\hat{\mathbf{b}}_i$ needed for applying the testing procedures.

3.1 Testing ACD(m,q) against higher-order alternatives

Engle and Russell (1998) defined their original ACD(m,q) model by parameterizing the conditional duration (2) as

$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}. \tag{9}$$

A natural benchmark and starting point for modeling durations is a low-order ACD(m,q) model, but then, too low an order is an obvious source of misspecification. An estimated ACD(m,q) model is tested against higher-order alternatives in the same way as Bollerslev (1986) tested a GARCH(p,q) model against higher-order alternatives. Consequently, either

$$\begin{aligned}
x_i &= (\psi_i + \varphi_i)\,\varepsilon_i & (10)\\
\psi_i + \varphi_i &= \omega + \sum_{j=1}^{m+r} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \left(\psi_{i-j} + \varphi_{i-j}\right) & (11)\\
\varphi_i &= \sum_{j=m+1}^{m+r} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \varphi_{i-j} & (12)\\
\varepsilon_i &\sim \text{i.i.d. } \exp(1) & (13)
\end{aligned}$$

or

$$\begin{aligned}
x_i &= (\psi_i + \varphi_i)\,\varepsilon_i & (14)\\
\psi_i + \varphi_i &= \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q+r} \beta_j \left(\psi_{i-j} + \varphi_{i-j}\right) & (15)\\
\varphi_i &= \sum_{j=q+1}^{q+r} \beta_j \left(\psi_{i-j} + \varphi_{i-j}\right) + \sum_{j=1}^{q} \beta_j \varphi_{i-j} & (16)\\
\varepsilon_i &\sim \text{i.i.d. } \exp(1). & (17)
\end{aligned}$$

The null hypothesis equals $H_0 : \varphi_i \equiv 0$, i.e. $\alpha_{m+1} = \cdots = \alpha_{m+r} = 0$ in the former and $\beta_{q+1} = \cdots = \beta_{q+r} = 0$ in the latter case. Under the alternative the first model is an ACD(m+r,q) model and the second model is an ACD(m,$q+r$) model, while under the null both models collapse to an ACD(m,q) model. The ACD(m,q) model cannot be tested directly against an ACD($m+r$,$q+s$) model, r, $s > 0$, using standard techniques because of the identification problem already discussed in Bollerslev (1986).

These higher-order alternatives belong to the additive class of alternatives mentioned above. The following two corollaries of Theorem 1 define the test statistics. The tests are most easily carried out using the auxiliary regression procedures given in the previous section.

Corollary 5 *Consider the model (10)–(13) with* $\theta_1 = (\omega, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_q)'$ *and* $\theta_2 = (\alpha_{m+1}, \ldots, \alpha_{m+r})'$. *Under the null hypothesis* $H_0 : \theta_2 = \mathbf{0}$, *the statistic (7), where*

$$\begin{aligned} \hat{\mathbf{a}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\psi}_i}{\partial \theta_1} = \hat{\psi}_i^{-1}\left(1, x_{i-1}, \ldots, x_{i-m}, \hat{\psi}_{i-1}, \ldots, \hat{\psi}_{i-q}\right)' + \hat{\psi}_i^{-1}\sum_{j=1}^{q}\hat{\beta}_j\frac{\partial \hat{\psi}_{i-j}}{\partial \theta_1} \\ \hat{\mathbf{b}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\varphi}_i}{\partial \theta_2} = \hat{\psi}_i^{-1}\left(x_{i-m-1}, \ldots, x_{i-m-r}\right)' + \hat{\psi}_i^{-1}\sum_{j=1}^{q}\hat{\beta}_j\frac{\partial \hat{\varphi}_{i-j}}{\partial \theta_2} \end{aligned}$$

has an asymptotic χ^2 *distribution with* r *degrees of freedom.*

Corollary 6 *Consider the model (14)–(17) with* $\theta_1 = (\omega, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_q)'$ *and* $\theta_2 = (\beta_{q+1}, \ldots, \beta_{q+r})'$. *Under the null hypothesis* $H_0 : \theta_2 = \mathbf{0}$, *the statistic (7), where*

$$\begin{aligned} \hat{\mathbf{a}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\psi}_i}{\partial \theta_1} = \hat{\psi}_i^{-1}\left(1, x_{i-1}, \ldots, x_{i-m}, \hat{\psi}_{i-1}, \ldots, \hat{\psi}_{i-q}\right)' + \hat{\psi}_i^{-1}\sum_{j=1}^{q}\hat{\beta}_j\frac{\partial \hat{\psi}_{i-j}}{\partial \theta_1} \\ \hat{\mathbf{b}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\varphi}_i}{\partial \theta_2} = \hat{\psi}_i^{-1}\left(\hat{\psi}_{i-q-1}, \ldots, \hat{\psi}_{i-q-r}\right)' + \hat{\psi}_i^{-1}\sum_{j=1}^{q}\hat{\beta}_j\frac{\partial \hat{\varphi}_{i-j}}{\partial \theta_2} \end{aligned}$$

has an asymptotic χ^2 *distribution with* r *degrees of freedom.*

3.2 Testing LOGACD(m,q) against higher-order alternatives

Bauwens and Giot (2000) advocate the use of a logarithmic version of the ACD model instead of the linear one. In their LOGACD(m,q) model (2) is parameterized as

$$\ln \psi_i = \omega + \sum_{j=1}^{m}\alpha_j \ln x_{i-j} + \sum_{j=1}^{q}\beta_j \ln \psi_{i-j}.$$

Also in this case the starting point for modeling would be a low-order model, which is then evaluated. Testing the logarithmic model against higher-order alternatives is done in a similar fashion than in the linear case. We consider either

$$\begin{aligned} x_i &= \psi_i \varphi_i \varepsilon_i && (18) \\ \ln(\psi_i \varphi_i) &= \omega + \sum_{j=1}^{m+r}\alpha_j \ln x_{i-j} + \sum_{j=1}^{q}\beta_j \ln(\psi_{i-j}\varphi_{i-j}) && (19) \\ \ln \varphi_i &= \sum_{j=m+1}^{m+r}\alpha_j \ln x_{i-j} + \sum_{j=1}^{q}\beta_j \ln \varphi_{i-j} && (20) \\ \varepsilon_i &\sim \text{i.i.d. } \exp(1) && (21) \end{aligned}$$

or

$$\begin{aligned} x_i &= \psi_i \varphi_i \varepsilon_i && (22) \\ \ln(\psi_i \varphi_i) &= \omega + \sum_{j=1}^{m}\alpha_j \ln x_{i-j} + \sum_{j=1}^{q+r}\beta_j \ln(\psi_{i-j}\varphi_{i-j}) && (23) \\ \ln \varphi_i &= \sum_{j=q+1}^{q+r}\beta_j \ln(\psi_{i-j}\varphi_{i-j}) + \sum_{j=1}^{q}\beta_j \ln \varphi_{i-j} && (24) \\ \varepsilon_i &\sim \text{i.i.d. } \exp(1). && (25) \end{aligned}$$

Under the null hypothesis $H_0 : \varphi_i \equiv 1$, i.e. $\alpha_{m+1} = \cdots = \alpha_{m+r} = 0$ in the former and $\beta_{q+1} = \cdots = \beta_{q+r} = 0$ in the latter case, the models reduce to the LOGACD(m,q) model. Under the alternative the models are the LOGACD($m+r$,q) model and the LOGACD(m,$q+r$) model, respectively. The alternatives belong to the multiplicative class of alternatives mentioned above. Corollaries 7 and 8 define the tests.

Corollary 7 *Consider the model (18)–(21) with* $\theta_1 = (\omega, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_q)'$ *and* $\theta_2 = (\alpha_{m+1}, \ldots, \alpha_{m+r})'$*. Under the null hypothesis* $H_0 : \theta_2 = \mathbf{0}$*, the statistic (8), where*

$$\begin{aligned}
\hat{\mathbf{a}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\psi}_i}{\partial \theta_1} = \left(1, \ln x_{i-1}, \ldots, \ln x_{i-m}, \ln \hat{\psi}_{i-1}, \ldots, \ln \hat{\psi}_{i-q}\right)' + \sum_{j=1}^{q} \hat{\beta}_j \frac{\partial \ln \hat{\psi}_{i-j}}{\partial \theta_1} \\
\hat{\mathbf{b}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\varphi}_i}{\partial \theta_2} = \hat{\psi}_i^{-1} \left(\ln x_{i-m-1}, \ldots, \ln x_{i-m-r}\right)' + \hat{\psi}_i^{-1} \sum_{j=1}^{q} \hat{\beta}_j \frac{\partial \hat{\varphi}_{i-j}}{\partial \theta_2}
\end{aligned}$$

has an asymptotic χ^2 *distribution with* r *degrees of freedom.*

Corollary 8 *Consider the model (22)–(25) with* $\theta_1 = (\omega, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_q)'$ *and* $\theta_2 = (\beta_{q+1}, \ldots, \beta_{q+r})'$*. Under the null hypothesis* $H_0 : \theta_2 = \mathbf{0}$*, the statistic (8), where*

$$\begin{aligned}
\hat{\mathbf{a}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\psi}_i}{\partial \theta_1} = \left(1, \ln x_{i-1}, \ldots, \ln x_{i-m}, \ln \hat{\psi}_{i-1}, \ldots, \ln \hat{\psi}_{i-q}\right)' + \sum_{j=1}^{q} \hat{\beta}_j \frac{\partial \ln \hat{\psi}_{i-j}}{\partial \theta_1} \\
\hat{\mathbf{b}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\varphi}_i}{\partial \theta_2} = \hat{\psi}_i^{-1} \left(\ln \hat{\psi}_{i-q-1}, \ldots, \ln \hat{\psi}_{i-q-r}\right)' + \hat{\psi}_i^{-1} \sum_{j=1}^{q} \hat{\beta}_j \frac{\partial \ln \hat{\varphi}_{i-j}}{\partial \theta_2}
\end{aligned}$$

has an asymptotic χ^2 *distribution with* r *degrees of freedom.*

3.3 Testing the hypothesis of no remaining ACD

After estimating an ACD model one may also ask whether the estimated disturbances still contain some structure. One possibility is that all the ACD effects are not captured by the estimated model but that some are still present in the disturbances. In the ACD literature it is common to evaluate the properties of standardized durations resulting from the estimation of an ACD model using Ljung-Box or McLeod-Li tests (see Ljung and Box (1978) and McLeod and Li (1983), respectively). As was shown by Li and Mak (1994) in the context of GARCH models, this is somewhat misleading. The reason is that these test statistics do not have the usual asymptotic χ^2 distribution under the null hypothesis when they are applied to standardized residuals from an estimated GARCH model. Li and Mak (1994) proposed a corrected statistic and Lundbergh and Teräsvirta (2002) presented a Lagrange multiplier statistic asymptotically equivalent to it. A similar test statistic for the ACD(m,q) model is presented next.

To this end, let

$$x_i = \psi_i \varphi_i \varepsilon_i \tag{26}$$

$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} \tag{27}$$

$$\varphi_i = 1 + \sum_{j=1}^{m^*} \alpha_j^* \frac{x_{i-j}}{\psi_{i-j}} \tag{28}$$

$$\varepsilon_i \sim \text{i.i.d. } \exp(1). \tag{29}$$

The null hypothesis equals $H_0 : \varphi_i \equiv 1$, i.e., $\alpha_1^* = \ldots = \alpha_{m^*}^* = 0$. The test will be based on the following corollary to Theorem 4.

Corollary 9 *Consider the model (26)–(29) with $\theta_1 = (\omega, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_q)'$ and $\theta_2 = (\alpha_1^*, \ldots, \alpha_{m^*}^*)'$. Under the null hypothesis $H_0 : \theta_2 = \mathbf{0}$, the statistic (8), where*

$$\begin{aligned}
\hat{\mathbf{a}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\psi}_i}{\partial \theta_1} = \hat{\psi}_i^{-1}\left(1, x_{i-1}, \ldots, x_{i-m}, \hat{\psi}_{i-1}, \ldots, \hat{\psi}_{i-q}\right)' + \hat{\psi}_i^{-1}\sum_{j=1}^{q}\hat{\beta}_j\frac{\partial \hat{\psi}_{i-j}}{\partial \theta_1} \\
\hat{\mathbf{b}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\varphi}_i}{\partial \theta_2} = \hat{\psi}_i^{-1}(x_{i-1}\hat{\psi}_{i-1}^{-1}, \ldots, x_{i-m^*}\hat{\psi}_{i-m^*}^{-1})'
\end{aligned}$$

has an asymptotic χ^2 distribution with m^* degrees of freedom.

As x_{i-1}/ψ_{i-1} are standardized durations, the test is one of the standardized durations being iid against the alternative that they follow an ACD process. When $m = q = 0$ it collapses to a test of no ACD effects in the original series. Tests of no remaining ACD effects after estimating other ACD type models are obtained by redefining $\hat{\mathbf{a}}_i$ in Corollary 9.

Recently, Li and Yu (2003) derived a portmanteau test of testing the null hypothesis that the exponentially distributed errors are independent. It turns out that their test is asymptotically equivalent to the statistic given in Corollary 9. For a proof, see Appendix B.

4 Smooth transition ACD models

Engle and Russell (1998) report that their linear ACD model generates expected durations that are on the average too long after the shortest and the longest durations. This suggests that a nonlinear specification for the conditional duration would be more appropriate than the standard linear one. Alternatives in the literature include the LOGACD model of Bauwens and Giot (2000), the Box-Cox and Exponential ACD models of Dufour and Engle (2000), and the threshold ACD model of Zhang, Russell, and Tsay (2001). A smooth transition version of the ACD model also appears to be a possibility and is considered here.

The inspiration for smooth transition ACD models comes from the GARCH literature. Smooth transition GARCH models are treated in Hagerud (1996), González-Rivera (1998) and Anderson, Nam, and Vahid (1999); see also Lundbergh and Teräsvirta (2002). In the present work, the smooth transition ACD(m,q) model is defined as follows:

$$\begin{aligned}
\psi_i &= \omega + \sum_{j=1}^{m}\alpha_j x_{i-j} + \sum_{j=1}^{m}\left(\omega_j^* + \alpha_j^* x_{i-j}\right) G\left(x_{i-j};\gamma,\mathbf{c}\right) + \sum_{j=1}^{q}\beta_j\psi_{i-j} \\
&= \omega + \sum_{j=1}^{m}\omega_j^* G\left(x_{i-j};\gamma,\mathbf{c}\right) + \sum_{j=1}^{m}\left(\alpha_j + \alpha_j^* G\left(x_{i-j};\gamma,\mathbf{c}\right)\right) x_{i-j} + \sum_{j=1}^{q}\beta_j\psi_{i-j}
\end{aligned}$$

where $G\left(x_{i-j};\gamma,\mathbf{c}\right)$ is a suitably chosen bounded and non-negative transition function. A natural candidate for the transition function could at first sight be the logistic function. A disadvantage of this transition function is, however, that the logistic function is defined on the whole real axis, whereas in the present case the potential transition variable x_{i-j} only takes positive values. Another candidate for the transition function would be a cumulative distribution function of a random variable with a positive support. The shortcoming of this alternative is that in this case the transition function would inevitably be a non-decreasing function, which will not produce nonlinearities of the type we are interested in.

As the logarithmic transformation is a common and often convenient way of transforming positive-valued objects to ones defined on the whole real axis, we retain the logistic function but use $\ln x_{i-j}$ as the transition variable. This leads to the following smooth transition ACD(m,q) (STACD) specification

$$\begin{aligned}\psi_i &= \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{m} \left(\omega_j^* + \alpha_j^* x_{i-j}\right) G\left(\ln x_{i-j}; \gamma, \mathbf{c}\right) + \sum_{j=1}^{q} \beta_j \psi_{i-j} \\ &= \omega + \sum_{j=1}^{m} \omega_j^* G\left(\ln x_{i-j}; \gamma, \mathbf{c}\right) + \sum_{j=1}^{m} \left(\alpha_j + \alpha_j^* G\left(\ln x_{i-j}; \gamma, \mathbf{c}\right)\right) x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} \quad (30)\end{aligned}$$

where

$$G\left(\ln x_{i-j}; \gamma, \mathbf{c}\right) = \left(1 + \exp\left\{-\gamma \prod_{k=1}^{K} \left(\ln x_{i-j} - c_k\right)\right\}\right)^{-1}, \qquad c_1 \leq \cdots \leq c_K, \gamma > 0 \quad (31)$$

and where the order $K \in \mathbb{Z}_+$ determines the general shape of the transition function. We also propose the smooth transition LOGACD(m,q) specification

$$\begin{aligned}\ln \psi_i &= \omega + \sum_{j=1}^{m} \alpha_j \ln x_{i-j} + \sum_{j=1}^{m} \left(\omega_j^* + \alpha_j^* \ln x_{i-j}\right) G\left(\ln x_{i-j}; \gamma, \mathbf{c}\right) + \sum_{j=1}^{q} \beta_j \ln \psi_{i-j} \\ &= \omega + \sum_{j=1}^{m} \omega_j^* G\left(\ln x_{i-j}; \gamma, \mathbf{c}\right) + \sum_{j=1}^{m} \left(\alpha_j + \alpha_j^* G\left(\ln x_{i-j}; \gamma, \mathbf{c}\right)\right) \ln x_{i-j} + \sum_{j=1}^{q} \beta_j \ln \psi_{i-j}.\end{aligned}$$

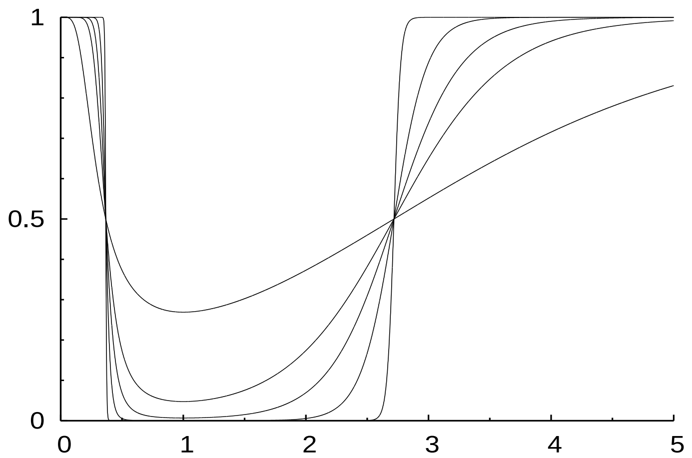
As $K = 1$ the transition function is an increasing function of the lagged duration. In order to capture the effects of very short and long durations we concentrate on the choice $K = 2$, which allows these extreme durations to have an impact different from the one of the more average durations. For illustration, Figure 1 contains the transition function (31) for $K = 2$ and a set of values for γ, c_1 and c_2. Assuming that the durations are transformed to take diurnal variation into account (see Section 7.1 for an explanation on how this is done) the average duration equals one, and shorter and longer than expected durations (at that time of the day) will be represented by durations less than and greater than one, respectively. For example, with the second steepest transition function in Figure 1, durations between (approximately) 1/2 and 2 would belong to the "normal" regime, whereas durations less than 1/3 and greater than 3 would belong to the "extreme" regime (here a transformed duration of x represents a duration of x times the expected duration at that time of the day).

The smooth transition ACD model is closely related to the threshold ACD model of Zhang, Russell, and Tsay (2001). We restrict our comparison to the model they refer to as the TACD(1,1) model. It is defined as follows:

$$\begin{aligned}x_i &= \psi_i \varepsilon_i^{(j)} \\ \psi_i &= \omega^{(j)} + \alpha^{(j)} x_{i-1} + \beta^{(j)} \psi_{i-1}\end{aligned}$$

whenever $x_{i-1} \in [r_{j-1}, r_j)$, $j = 1, 2, \ldots, J$, where J is the number of different regimes and $0 = r_0 < r_1 < \cdots < r_J = \infty$ are the threshold values. The parameter values $\omega^{(j)}$, $\alpha^{(j)}$ and $\beta^{(j)}$ as well as the distribution of $\varepsilon_i^{(j)}$ are allowed to vary depending on the regime. The two-regime

Figure 1: The logistic transition function $G(\ln x; \gamma, \mathbf{c})$ as a function of x for $K = 2$, $c_1 = -1$, $c_2 = 1$ and for $\gamma = 1$ (the smoothest), 3, 5, 10 and 50 (the steepest).



TACD(1,1) model with the restrictions $\beta^{(1)} = \beta^{(2)}$ and $\varepsilon_i^{(1)} \sim \varepsilon_i^{(2)} \sim \exp(1)$ is achieved as the limiting case of the STACD(1,1) model with $K = 1$ as $\gamma \to \infty$. Similarly, the three-regime TACD(1,1) model with $\beta^{(1)} = \beta^{(2)} = \beta^{(3)}$, $\omega^{(1)} = \omega^{(3)}$, $\alpha^{(1)} = \alpha^{(3)}$ and $\varepsilon_i^{(j)} \sim \exp(1)$, $j = 1, 2, 3$, is the limiting case of the STACD(1,1) model with $K = 2$ as $\gamma \to \infty$.

4.1 Testing ACD(m,q) against smooth transition ACD(m,q)

We now consider testing the ACD model against its smooth transition counterpart. It is seen that model (30) is only identified under the alternative. For example, when $\gamma = 0$ (this is one form of the null hypothesis), parameters ω_j^* and α_j^*, $J = 1, \ldots, m$, as well as $\mathbf{c}$, are not identified. Testing when some parameters are identified only under the alternative is discussed for example in Hansen (1996). He studies the (non-standard) asymptotic distribution theory for such tests, and develops a procedure to approximate these distibutions by simulation. As our goal is to derive easily applicable misspecification tests, we do not follow Hansen's approach, but instead use the method suggested in Luukkonen, Saikkonen, and Teräsvirta (1988). In their approach the identification problem is solved by approximating the transition function with its first-order Taylor expansion around $\gamma = 0$. This will lead to an approximate alternative, which is free of nuisance parameters under the null.

To this end, define

$$\begin{aligned} x_i &= \psi_i \varepsilon_i \\ \psi_i &= \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} \\ &\quad + \sum_{j=1}^{m} \left(\omega_j^* + \alpha_j^* x_{i-j}\right) \bar{G}\left(\ln x_{i-j}; \gamma, \mathbf{c}\right) \qquad (32) \\ \varepsilon_i &\sim \text{i.i.d. } \exp(1) \end{aligned}$$

where $\bar{G}\left(\ln x_{i-j}; \gamma, \mathbf{c}\right) = G\left(\ln x_{i-j}; \gamma, \mathbf{c}\right) - \frac{1}{2}$ (subtracting $\frac{1}{2}$ simplifies the derivation below but does not affect the conclusions, because we can replace $\bar{G}$ by G with a simple reparameterization).

Using Taylor's theorem one obtains as

$$\begin{aligned}
\bar{G}\left(\ln x_{i-j};\gamma,\mathbf{c}\right) &= \bar{G}\left(\ln x_{i-j};0,\mathbf{c}\right)+\frac{\partial\bar{G}\left(\ln x_{i-j};0,\mathbf{c}\right)}{\partial\gamma}\left(\gamma-0\right)+\bar{G}\left(\ln x_{i-j};\tilde{\gamma},\mathbf{c}\right)\\
&= \frac{1}{4}\gamma\prod_{k=1}^{K}\left(\ln x_{i-j}-c_k\right)+\bar{G}\left(\ln x_{i-j};\tilde{\gamma},\mathbf{c}\right)\\
&= \sum_{l=0}^{K}\gamma\tilde{c}_l\left(\ln x_{i-j}\right)^l+\bar{G}\left(\ln x_{i-j};\tilde{\gamma},\mathbf{c}\right) \qquad (33)
\end{aligned}$$

where $\tilde{\gamma}\in[0,\gamma]$. Applying (33) to (32) yields

$$\begin{aligned}
\psi_i &= \omega+\sum_{j=1}^{m}\alpha_j x_{i-j}+\sum_{j=1}^{q}\beta_j\psi_{i-j}\\
&\quad+\sum_{j=1}^{m}\left[\left(\omega_j^*+\alpha_j^*x_{i-j}\right)\sum_{l=0}^{K}\gamma\tilde{c}_l\left(\ln x_{i-j}\right)^l\right]+\sum_{j=1}^{m}\left(\omega_j^*+\alpha_j^*x_{i-j}\right)\bar{G}\left(\ln x_{i-j};\tilde{\gamma},\mathbf{c}\right)\\
&= \omega+\sum_{j=1}^{m}\alpha_j x_{i-j}+\sum_{j=1}^{q}\beta_j\psi_{i-j}+\sum_{j=1}^{m}\left(\gamma\omega_j^*\tilde{c}_0+\gamma\alpha_j^*\tilde{c}_0x_{i-j}\right)\\
&\quad+\sum_{j=1}^{m}\sum_{l=1}^{K}\left[\gamma\omega_j^*\tilde{c}_l\left(\ln x_{i-j}\right)^l+\gamma\alpha_j^*\tilde{c}_lx_{i-j}\left(\ln x_{i-j}\right)^l\right]+\sum_{j=1}^{m}\left(\omega_j^*+\alpha_j^*x_{i-j}\right)\bar{G}\left(\ln x_{i-j};\tilde{\gamma},\mathbf{c}\right)\\
&= \left[\omega+\sum_{j=1}^{m}\gamma\omega_j^*\tilde{c}_0\right]+\sum_{j=1}^{m}\left[\alpha_j+\gamma\alpha_j^*\tilde{c}_0\right]x_{i-j}+\sum_{j=1}^{q}\beta_j\psi_{i-j}\\
&\quad+\sum_{j=1}^{m}\sum_{l=1}^{K}\left[\gamma\omega_j^*\tilde{c}_l\left(\ln x_{i-j}\right)^l+\gamma\alpha_j^*\tilde{c}_lx_{i-j}\left(\ln x_{i-j}\right)^l\right]\\
&\quad+\sum_{j=1}^{m}\left(\omega_j^*+\alpha_j^*x_{i-j}\right)\bar{G}\left(\ln x_{i-j};\tilde{\gamma},\mathbf{c}\right). \qquad (34)
\end{aligned}$$

This form does not lead to an operational test statistic as $\tilde{\gamma}$ is unknown. If we instead use Taylor's theorem to approximate $\bar{G}\left(\ln x_{i-j};\gamma,\mathbf{c}\right)$, we can drop the remainder term $\bar{G}\left(\ln x_{i-j};\tilde{\gamma},\mathbf{c}\right)$ from the last expression in (34). Doing this and renaming parameters yields the following approximation to the conditional mean of the alternative:

$$\psi_i\approx w+\sum_{j=1}^{m}a_jx_{i-j}+\sum_{j=1}^{q}b_j\psi_{i-j}+\sum_{j=1}^{m}\sum_{l=1}^{K}\left(d_{jl}\left(\ln x_{i-j}\right)^l+e_{jl}x_{i-j}\left(\ln x_{i-j}\right)^l\right).$$

Using this approximation we have transformed the original testing problem into testing the

ACD(m,q) model against the approximate alternative

$$x_i = (\psi_i + \varphi_i)\,\varepsilon_i \tag{35}$$

$$\psi_i + \varphi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j (\psi_{i-j} + \varphi_{i-j}) + \sum_{j=1}^{m} \sum_{l=1}^{K} \left(d_{jl} (\ln x_{i-j})^l + e_{jl} x_{i-j} (\ln x_{i-j})^l \right) \tag{36}$$

$$\varphi_i = \sum_{j=1}^{q} \beta_j \varphi_{i-j} + \sum_{j=1}^{m} \sum_{l=1}^{K} \left(d_{jl} (\ln x_{i-j})^l + e_{jl} x_{i-j} (\ln x_{i-j})^l \right) \tag{37}$$

$$\varepsilon_i \sim \text{i.i.d. } \exp(1). \tag{38}$$

Model (35)–(38) reduces to the null model when $d_{jl} = e_{jl} = 0$ for $j = 1, \ldots, m$ and $l = 1, \ldots, K$, and there are no unindentified parameters under the null. This enables us to use Theorem 1 to derive the test, which is given in the following corollary.

Corollary 10 *Consider the model (35)–(38) and denote* $\theta_1 = (\omega, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_q)'$ *and* $\theta_2 = ((vecD)', (vecE)')'$, *where* $D = [d_{jl}]$, $E = [e_{jl}]$, $j = 1, \ldots, m$, $l = 1, \ldots, K$, *are* $(m \times K)$ *matrices, and the vec-operator stacks the columns of the matrix. Furthermore, denote* $X_{i,1} = \left[(\ln x_{i-j})^l \right]$ *and* $X_{i,2} = \left[x_{i-j} (\ln x_{i-j})^l \right]$, $j = 1, \ldots, m$, $l = 1, \ldots, K$. *Under the null hypothesis* $H_0 : \theta_2 = \mathbf{0}$ *the statistic (7), where*

$$\hat{\mathbf{a}}_i = \frac{1}{\hat{\psi}_i} \frac{\partial \hat{\psi}_i}{\partial \theta_1} = \hat{\psi}_i^{-1} \left(1, x_{i-1}, \ldots, x_{i-m}, \hat{\psi}_{i-1}, \ldots, \hat{\psi}_{i-q} \right)' + \hat{\psi}_i^{-1} \sum_{j=1}^{q} \hat{\beta}_j \frac{\partial \hat{\psi}_{i-j}}{\partial \theta_1}$$

$$\hat{\mathbf{b}}_i = \frac{1}{\hat{\psi}_i} \frac{\partial \hat{\varphi}_i}{\partial \theta_2} = \hat{\psi}_i^{-1} ((vecX_{i,1})', (vecX_{i,2})')' + \hat{\psi}_i^{-1} \sum_{j=1}^{q} \hat{\beta}_j \frac{\partial \hat{\varphi}_{i-j}}{\partial \theta_2}$$

has an asymptotic χ^2 *distribution with* $2mK$ *degrees of freedom.*

5 Time-varying ACD models

In standard econometric analysis of low-frequency time series the observation period easily spans over several years or decades of data. In such a situation it may not be realistic to expect the parameters of any model to remain constant over the whole period. For this reason testing parameter constancy is important. When using ultra-high-frequency data a period of a few days or weeks often yields a sufficient amount of observations. Thus it does not seem inappropriate to assume that the parameters actually remain constant over the observation period. On the other hand, certain events affecting the economic or institutional environment could cause the structure of the trading process to change. In such a situation, fitting an ACD model with constant parameters to the observed durations may yield unsatisfactory results. One remedy to the problem is to split the sample into several periods and estimate separate models for each of them. Identifying the number and location of the break-points becomes, however, a demanding task, but see Zhang, Russell, and Tsay (2001) for an example. An alternative to abrupt changes in the parameters would be a model where the parameters are allowed to change smoothly over time. This can be achieved for example using the logistic transition functions (31) with time as the transition variable.

Even when the model builder does not want to fit ACD models with time-varying parameters to data, they can be used as tools for detecting misspecification. For example, if an ACD model is tested and rejected against a model with time-varying parameters, this might be seen as evidence that the structure of the duration series changes during the period in question and that a more careful analysis is required. On the other hand, if an ACD model is rejected against a model with time-of-day-varying parameters, this could be an indication of the test that the approach used for removing the diurnal pattern is not satisfactory.

Two definitions of time are considered for ACD models with time-varying parameters. The first one is the total trading time (in seconds) from the beginning of the sample to the end, and is called the (*total*) *time*. In our empirical application in Section 7, each one of the samples consists of one week of data, and the time thus runs from the beginning of the first trading day of the week till the end of the last day of the week. Hence the term *intraweek time* is also used. The second definition, called the *intraday time*, is time measured in seconds from the beginning of the trading day. Both measures are for convenience and numerical stability standardized to obtain values between 0 and 1. As each financial event considered has a precise time-stamp attached to it, these two time definitions are readily available.

This leads to the following time-varying ACD (TVACD) specification

$$\begin{aligned}
\psi_i &= \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} + \left(\omega^* + \sum_{j=1}^{m} \alpha_j^* x_{i-j} + \sum_{j=1}^{q} \beta_j^* \psi_{i-j} \right) \bar{G}\left(t_{i-1}; \gamma, \mathbf{c}\right) \\
&= \left(\omega + \omega^* \bar{G}\left(t_{i-1}; \gamma, \mathbf{c}\right)\right) + \sum_{j=1}^{m} \left(\alpha_j + \alpha_j^* \bar{G}\left(t_{i-1}; \gamma, \mathbf{c}\right)\right) x_{i-j} + \sum_{j=1}^{q} \left(\beta_j + \beta_j^* \bar{G}\left(t_{i-1}; \gamma, \mathbf{c}\right)\right) \psi_{i-j}
\end{aligned}$$

where $\bar{G}\left(t_{i-1}; \gamma, \mathbf{c}\right) = G\left(t_{i-1}; \gamma, \mathbf{c}\right) - \frac{1}{2}$ and G is the transition function given in (31) except that the transition variable used now is t_{i-1}, which can correspond to either one of the two time definitions. If we consider parameter constancy or structural breaks in the process, then the total time is the one to be used. If the issue is how well the diurnal pattern has been removed then the intraday time is the appropriate measure. Analogous definitions are available for the LOGACD model.

5.1 Testing parameter constancy

We now consider testing an estimated ACD model against these time-varying alternatives. The identification problem already discussed is present in the current situation as well. Therefore it is necessary to test for the presence of time-varying parameters before estimating any TVACD model. The identification problem is again solved using a Taylor series approximation of the transition function. Arguments similar to the ones in Section 4.1 lead to the following approximation to the TVACD(m,q) model:

$$x_i = (\psi_i + \varphi_i)\,\varepsilon_i \tag{39}$$

$$\begin{aligned}\psi_i + \varphi_i &= \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \left(\psi_{i-j} + \varphi_{i-j}\right) \\ &\quad + \sum_{l=1}^{K} d_l t_{i-1}^{l} + \sum_{j=1}^{m}\sum_{l=1}^{K} e_{jl} x_{i-j} t_{i-1}^{l} + \sum_{j=1}^{q}\sum_{l=1}^{K} f_{jl} \psi_{i-j} t_{i-1}^{l}\end{aligned} \tag{40}$$

$$\varphi_i = \sum_{j=1}^{q} \beta_j \varphi_{i-j} + \sum_{l=1}^{K} d_l t_{i-1}^{l} + \sum_{j=1}^{m}\sum_{l=1}^{K} e_{jl} x_{i-j} t_{i-1}^{l} + \sum_{j=1}^{q}\sum_{l=1}^{K} f_{jl} \psi_{i-j} t_{i-1}^{l} \tag{41}$$

$$\varepsilon_i \sim \text{i.i.d. } \exp(1). \tag{42}$$

The model reduces to the null model when $d_l = 0$ ($l = 1, \ldots, K$), $e_{jl} = 0$ ($j = 1, \ldots, m$ and $l = 1, \ldots, K$) and $f_{jl} = 0$ ($j = 1, \ldots, q$ and $l = 1, \ldots, K$). This enables us to use Theorem 1 to derive the test, which is given in the following corollary. The total time can be replaced with the intraday time without affecting the validity of the result.

Corollary 11 *Consider the model (39)–(42) and denote* $\theta_1 = (\omega, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_q)'$ *and* $\theta_2 = (d_1, \ldots, d_K, (vecE)', (vecF)')'$, *where* $E = [e_{jl}]$ *(*$j = 1, \ldots, m$, $l = 1, \ldots, K$*) is a* $(m \times K)$ *matrix and* $F = [f_{jl}]$ *(*$j = 1, \ldots, q$, $l = 1, \ldots, K$*) is a* $(q \times K)$ *matrix, and the vec-operator stacks the columns of the matrix. Furthermore, denote* $X_{i,1} = \left[x_{i-j} t_{i-1}^{l}\right]$ *(*$j = 1, \ldots, m$, $l = 1, \ldots, K$*) and* $X_{i,2} = \left[\hat{\psi}_{i-j} t_{i-1}^{l}\right]$ *(*$j = 1, \ldots, q$, $l = 1, \ldots, K$*). Under the null hypothesis* $H_0 : \theta_2 = \mathbf{0}$ *the statistic (7), where*

$$\begin{aligned}\hat{\mathbf{a}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\psi}_i}{\partial \theta_1} = \hat{\psi}_i^{-1}\left(1, x_{i-1}, \ldots, x_{i-m}, \hat{\psi}_{i-1}, \ldots, \hat{\psi}_{i-q}\right)' + \hat{\psi}_i^{-1} \sum_{j=1}^{q} \hat{\beta}_j \frac{\partial \hat{\psi}_{i-j}}{\partial \theta_1} \\ \hat{\mathbf{b}}_i &= \frac{1}{\hat{\psi}_i}\frac{\partial \hat{\varphi}_i}{\partial \theta_2} = \hat{\psi}_i^{-1}(t_{i-1}, \ldots, t_{i-1}^{K}, (vecX_{i,1})', (vecX_{i,2})')' + \hat{\psi}_i^{-1} \sum_{j=1}^{q} \hat{\beta}_j \frac{\partial \hat{\varphi}_{i-j}}{\partial \theta_2}\end{aligned}$$

has an asymptotic χ^2 *distribution with* $(1 + m + q)\,K$ *degrees of freedom.*

6 Simulation experiment

6.1 Size simulations

We investigate the finite sample properties of the test statistics by simulation and begin with size simulations. The data generating process we use has the form

$$\begin{aligned}x_i &= \psi_i \varepsilon_i \\ \psi_i &= 0.15 + 0.10 x_{i-1} + 0.80 \psi_{i-1} \\ \varepsilon_i &\sim \text{i.i.d. } \exp(1).\end{aligned} \tag{43}$$

The parameters in (43) have been chosen such that the model is representative for the estimated ACD(1,1) models reported in the literature. We use sample sizes $n = 1000$, 5000, and 10000. The smallest size is extremely small in the context of ACD models, and the largest one is still considerably smaller than the sample sizes in our empirical example. To avoid initialization effects we discard 1000 observations from the beginning of each generated series. The number of replications is 10000. In each replication, an ACD(1,1) model is estimated and then evaluated using five different tests. These are the tests against ACD(2,1) and ACD(1,2) models, test of no remaining ACD (of order one) in the standardized durations, and linearity tests against smooth transition ACD of orders one and two. For each test both the ordinary version using the auxiliary regression and the robust version are computed.

Results of the experiment are presented graphically in Figure 2. For each test we calculate the actual rejection frequencies for the nominal significance levels 0.1%, 0.2%, ..., 5.0%. The graphs show the discrepancies in size, i.e. the difference between the actual and the nominal size. In each subgraph we present the results for one of the tests using all the three sample sizes.

As can be seen, all the tests are rather well-sized for $n = 5000$ and 10000. The distortions are not very severe for the smallest sample size either. It can be concluded that the asymptotic null distributions of the test statistics are reasonably good approximations to the unknown finite-sample distributions for $n \geq 5000$. Such sample sizes are standard in the analysis of ultra-high-frequency data.

In order to complete the experiment, we also investigated the effect of having more persistence in the data generating process. The design of the experiment was the same as before except that the parameter values in (43) were changed to $\psi_i = 0.05 + 0.09x_{i-1} + 0.90\psi_{i-1}$. The results were similar to those reported in Figure 2 and are not presented here (they are available at *http://swopec.hhs.se/hastef/abs/hastef0557.htm*). The only notable change was a slight increase in the empirical size of the test against the ACD(1,2) alternative.

To explore the effects of misspecification of the error distribution (and hence of the conditional distribution of the durations) on the finite-sample properties of our test statistics, we repeated the simulations with different error distributions. They were the Weibull distribution with shape parameter values 0.8 and 0.9, and the generalized gamma distribution with shape parameter values 8 and 0.3 as well as 10 and 0.5; see for example Lunde (1999) for definitions of density functions of these distributions. The distributions were scaled to have an expected value of one. With the chosen parameter values the Weibull distribution has a monotonically decreasing hazard function, whereas for the generalized gamma distribution the hazard function is inverted U–shaped. As the results from these simulations were similar to the ones already reported, we do not show them here. Our general conclusion is that both the non-robust and the robust versions of the tests remain well-sized for sample sizes over 5000. It appears that if the conditional distribution of the durations is only mildly misspecified and if the sample size is sufficiently large, even the non-robust versions of the tests have satisfactory size properties.

6.2 Power simulations

To evaluate the performance of our test statistics we perform a power comparison to a benchmark test. As the test to compare with we choose the generalized spectral density based test introduced in Hong (1999). This test has been proposed as a general diagnostic test for a wide class of time series models, including ACD models, by Hong and Lee (2003). In the context of ACD models, it is a test of the standardized durations being iid against an unspecified alternative. As such it is an omnibus test against any kind of pairwise dependence structure in the standardized durations, and the simulation study in Hong and Lee (2003) suggests that the test has good power against a wide variety of alternatives. Users of this test have to make some parameter choices, and we both describe the test and discuss our choices in Appendix C.

We begin by repeating the size simulation experiment of the previous subsection for the generalized spectral test. The design of the experiment is exactly the same as earlier, except

Figure 2: Results from size simulations of the LM tests. In the figures the size discrepancy (i.e. the actual size less the nominal size) is plotted against the nominal size. Both of them are measured in percentage points. Performed tests are the tests against ACD(2,1) and ACD(1,2) models, tests of no remaining ACD (of order one) in the standardized durations, and tests of no smooth transition ACD of orders one and two. Both the ordinary and robust versions of the tests are used. The three lines in each subfigure correspond to sample sizes 1000 (+), 5000 (×) and 10000 (∗).

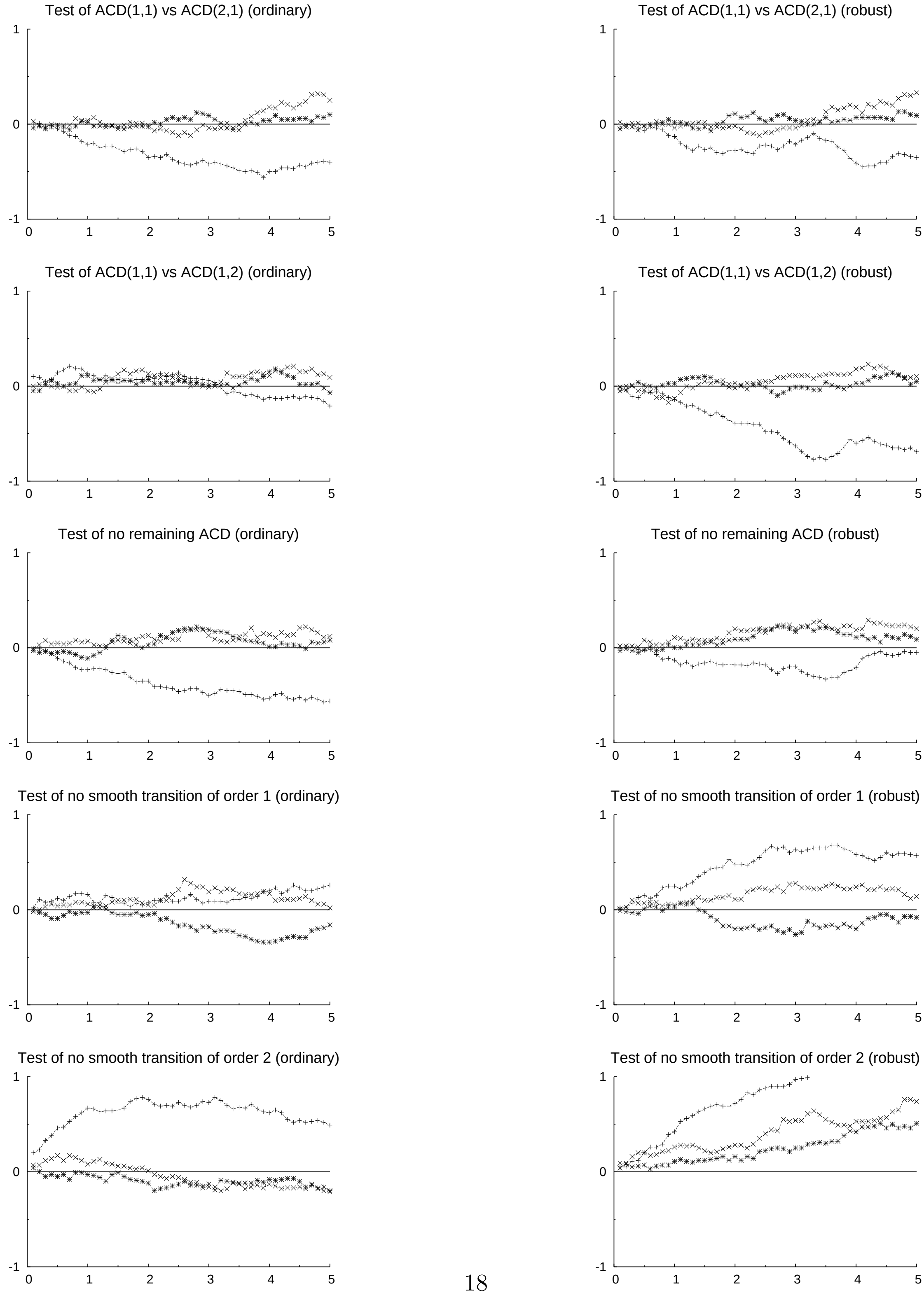
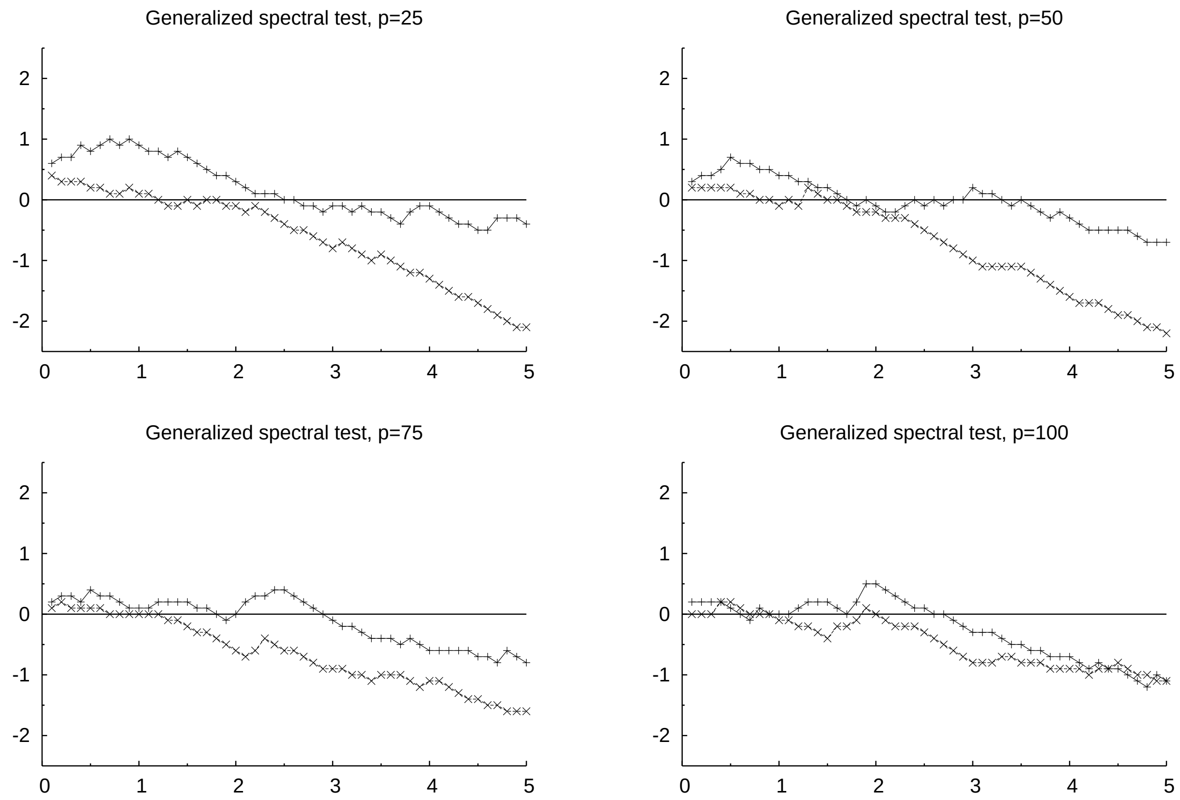


Figure 3: Results from size simulations of the generalized spectral tests. In the figures the size discrepancy (i.e. the actual size less the nominal size) is plotted against the nominal size. Both of them are measured in percentage points. Performed tests are the generalized spectral tests for preliminary bandwidths $\bar{p} = 25$, 50, 75, and 100. The two lines in each subfigure correspond to sample sizes 1000 (+) and 5000 ($\times$).



that we consider only sample sizes of 1000 and 5000, and perform only 1000 replications. This parsimony is due to the fact that the generalized spectral test is computationally much more burdensome than our tests. As explained in Appendix C, performing the test also involves the choice of a preliminary bandwidth, $\bar{p}$. We use the values 25, 50, 75, and 100. The size discrepancies of the test are presented in Figure 3 (note the different scale on the y-axes compared to Figure 2). The tests seem to be slightly undersized for all choices of the preliminary bandwidth when the nominal significance level approaches 5%. Note, however, that the size distortion diminishes with increasing $\bar{p}$.

In the power simulations the restricted model to be estimated is always an ACD(1,1) model. We consider two alternative data generating processes. The first one of these is an ACD(2,1) specification given by

$$x_i = \psi_i \varepsilon_i \tag{44}$$
$$\psi_i = 0.15 + 0.10x_{i-1} + 0.05x_{i-2} + 0.80\psi_{i-1} \tag{45}$$
$$\varepsilon_i \sim \text{i.i.d. } \exp(1). \tag{46}$$

Because of relatively large sample sizes, the coefficient of x_{i-2} is chosen to be close to zero. This is an alternative against which our test against higher-order models is expected to have very good power. As the second alternative model we consider a three-regime threshold ACD(1,1) model given by

$$x_i = \psi_i \varepsilon_i \tag{47}$$

$$\psi_i = \begin{cases} 0.05 + 0.20x_{i-1} + 0.85\psi_{i-1} & \text{for } 0 < x_{i-1} < 0.25 \\ 0.10 + 0.05x_{i-1} + 0.90\psi_{i-1} & \text{for } 0.25 \leq x_{i-1} < 1.5 \\ 0.20 + 0.03x_{i-1} + 0.80\psi_{i-1} & \text{for } 1.5 \leq x_{i-1} < \infty \end{cases} \tag{48}$$

$$\varepsilon_i \sim \text{i.i.d. } \exp(1). \tag{49}$$

The parameter values in this model are chosen such that it resembles the threshold ACD models estimated in Zhang, Russell, and Tsay (2001). This model is not a special case of our STACD model. The first regime (for the smallest values of x_{i-1}) is an explosive one. The other two regimes are stable, the middle one being more persistent than the third one.

Our sample sizes are 1000 and 5000, and we perform 1000 replications. As before, we discard 1000 observations from the beginning of each series. An ACD(1,1) model is fitted to each series. The diagnostic tests performed are the tests against ACD(2,1) and ACD(1,2) models, the test of no remaining ACD (of order one) in standardized durations, and tests of no smooth transition ACD of orders one and two. We also apply the generalized spectral test using preliminary bandwidths 25, 50, 75, and 100. We compute both the ordinary and the robustified versions of the LM tests.

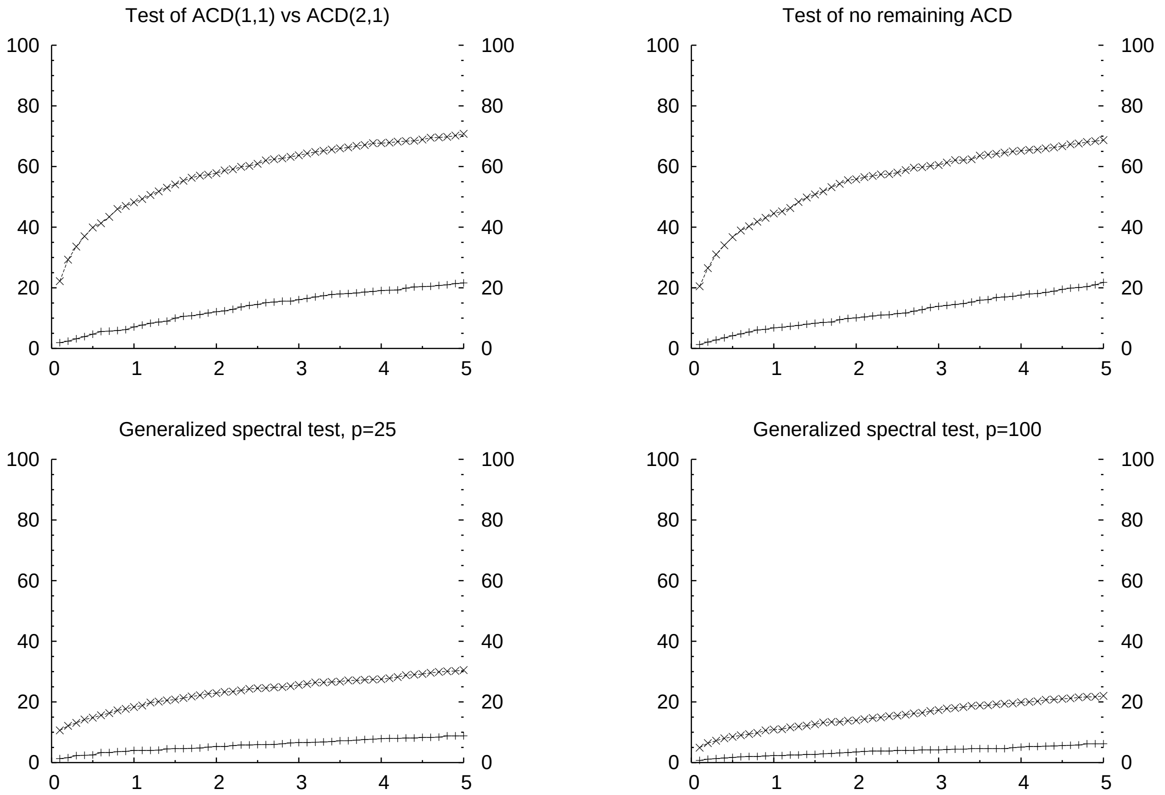
Results of the power simulations are presented in Figures 4a and 4b. In order to conserve space we do not present all of the results, but they are available upon request. These figures show the rejection frequencies of the tests for both sample sizes at the nominal significance levels of 0.1%, 0.2%, ..., 5.0%. Since our size simulations indicated that all the tests in question are well-sized, the power results are not size-adjusted.

Figure 4a reports the powers of the tests when the alternative data generating process is the ACD(2,1) model given in equations (44)–(46). The upper panel presents the power levels for the non-robustified versions of the test against an ACD(2,1) model and the test of no remaining ACD in the standardized durations. Both of these tests have rather good power for $n = 5000$, whereas the power is still low when $n = 1000$. Without presenting the results we note that the test against an ACD(1,2) model has almost equally good power, whereas the tests against smooth transition ACD have no power at all against this alternative. Furthermore, the robustified versions of the tests have power very close to the non-robust ones.

In the lower panel we show the power of the generalized spectral test using preliminary bandwidths 25 and 100. Both of these tests have moderate power, the one with $\bar{p} = 25$ being somewhat more powerful. The powers of the tests with $\bar{p} = 50$ and 75 (not shown) are very similar to the ones shown. Our tests against higher-order ACD thus have clearly higher power than the benchmark. This is natural because these tests are designed to have power against this particular alternative. If they had been only slightly more powerful or even less powerful than the generalized spectral density based tests, that would have warned us that the small-sample properties of our test would leave much to desire.

Power results for the threshold ACD(1,1) model of equations (47)–(49) can be found in Figure 4b. The non-robustified versions of the linearity tests against smooth transition of orders one and two both have very good power for $n = 5000$, and low power for $n = 1000$ (upper panel). Without presenting the results we note that the tests against higher-order models and of no remaining ACD have no power at all, and that the power of the robustified versions of the tests again have power very close to the non-robust ones. The generalized spectral test using preliminary bandwidths 25 and 100 has rather low power against this alternative (the powers for the tests with $\bar{p} = 50$ and 75 are very similar).

Figure 4a: Results from power simulations of the tests using the ACD(2,1) model given in equations (44)–(46) as the alternative data generating process. In the figures the power (rejection frequency) is plotted against the nominal size. Both of them are measured in percentage points. Performed tests are the test against a higher-order ACD(2,1) model, test of no remaining ACD, and the generalized spectral tests for preliminary bandwidths $\bar{p} = 25$ and 100. The two lines in each subfigure correspond to sample sizes 1000 (+) and 5000 (×).



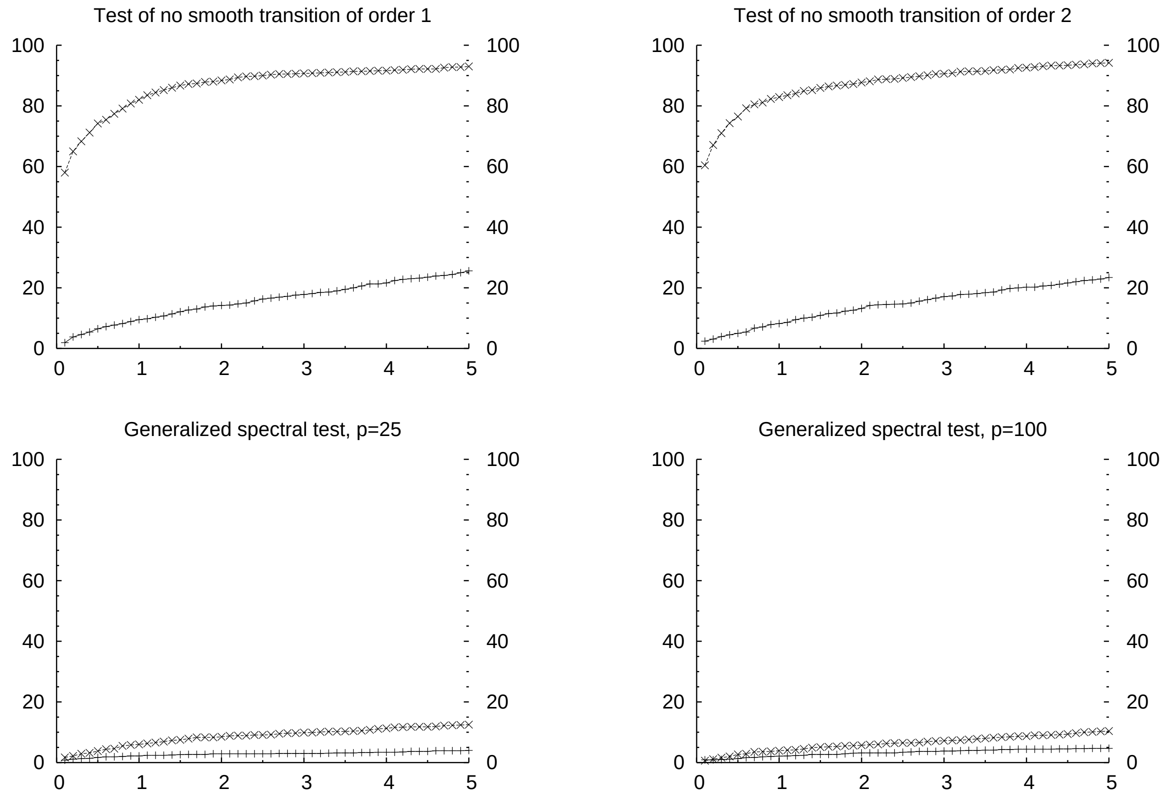
6.3 Performance of the Ljung-Box and McLeod-Li tests

As mentioned in Section 3.3, it is common in the ACD literature to evaluate the properties of estimated standardized durations using Ljung-Box or McLeod-Li tests. This practice can, however, result in misleading conclusions, because these test statistics do not have the usual asymptotic χ^2 distribution under the null hypothesis when they are applied to standardized durations from an estimated ACD model. In order to take a closer look at this possibility, we shall investigate the properties of these test statistics when they are applied to standardized durations from an estimated ACD model.

We perform exactly the same simulation experiments as in the previous subsection, except that we also use the sample size 10000. The tests applied to the standardized durations from the estimated ACD(1,1) models are the Ljung-Box and McLeod-Li tests with lag lengths 1, 5, 10, 15, and 20. The rejection frequencies in both the size and power simulations are based on the (incorrect) asymptotic χ^2 distribution with degrees of freedom equal to the lag length used.

The size discrepancies of the tests with lag length 15 are presented in Figure 5a (results with the other lag lengths are similar and thus omitted). The Ljung-Box test seems to be somewhat undersized. On the other hand, the McLeod-Li test is oversized: quite strongly for $n = 1000$

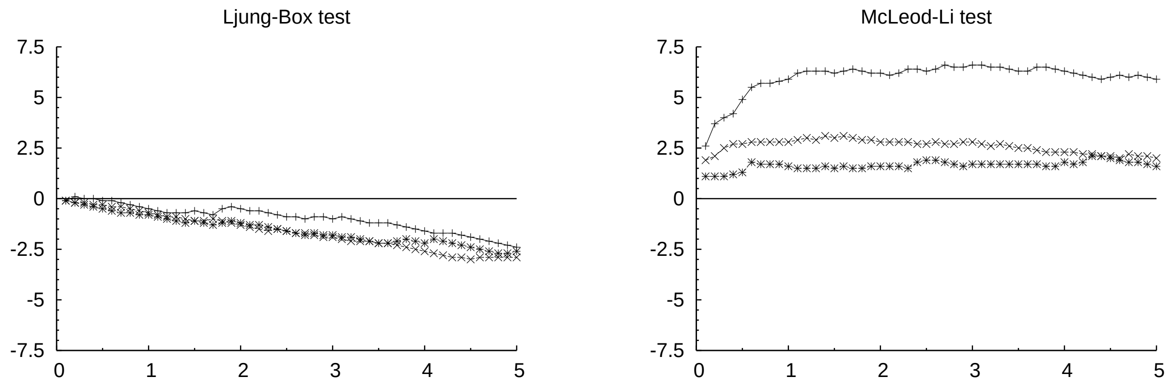
Figure 4b: Results from power simulations of the tests using the threshold ACD(1,1) model given in equations (47)–(49) as the alternative data generating process. In the figures the power (rejection frequency) is plotted against the nominal size. Both of them are measured in percentage points. Performed tests are the tests of no smooth transition ACD of orders 1 and 2, and the generalized spectral tests for preliminary bandwidths $\bar{p} = 25$ and 100. The two lines in each subfigure correspond to sample sizes 1000 (+) and 5000 (×).



and less so for the two larger sample sizes.

In Figure 5b we present the power of the tests against the ACD(2,1) (the two upper figures) and threshold ACD(1,1) (the two lower figures) alternatives given in equations (44)–(46) and (47)–(49), respectively. Again, we only present the results with lag length 15. It can be seen that the Ljung-Box test has moderate power against the ACD(2,1) alternative. A comparison with the results in Figure 4a indicates that the test is less powerful than the tests considered there. This is not surprising because the size simulations showed that the Ljung-Box test is conservative. The McLeod-Li test has almost no power at all against this alternative, which may not be unexpected either. This is because one cannot expect to discover a misspecified lag length using a test based on squared standardized durations. Finally, neither of the tests has any notable power against the threshold ACD model.

Figure 5a: Results from size simulations of the Ljung-Box and McLeod-Li tests. In the figures the size discrepancy (i.e. the actual size less the nominal size) is plotted against the nominal size. Both of them are measured in percentage points. Performed tests are the Ljung-Box (left) and McLeod-Li (right) tests with lag length 15. The three lines in each subfigure correspond to sample sizes 1000 (+), 5000 (×) and 10000 (∗).



7 Application to Trades and Quotes series

7.1 Description of the data

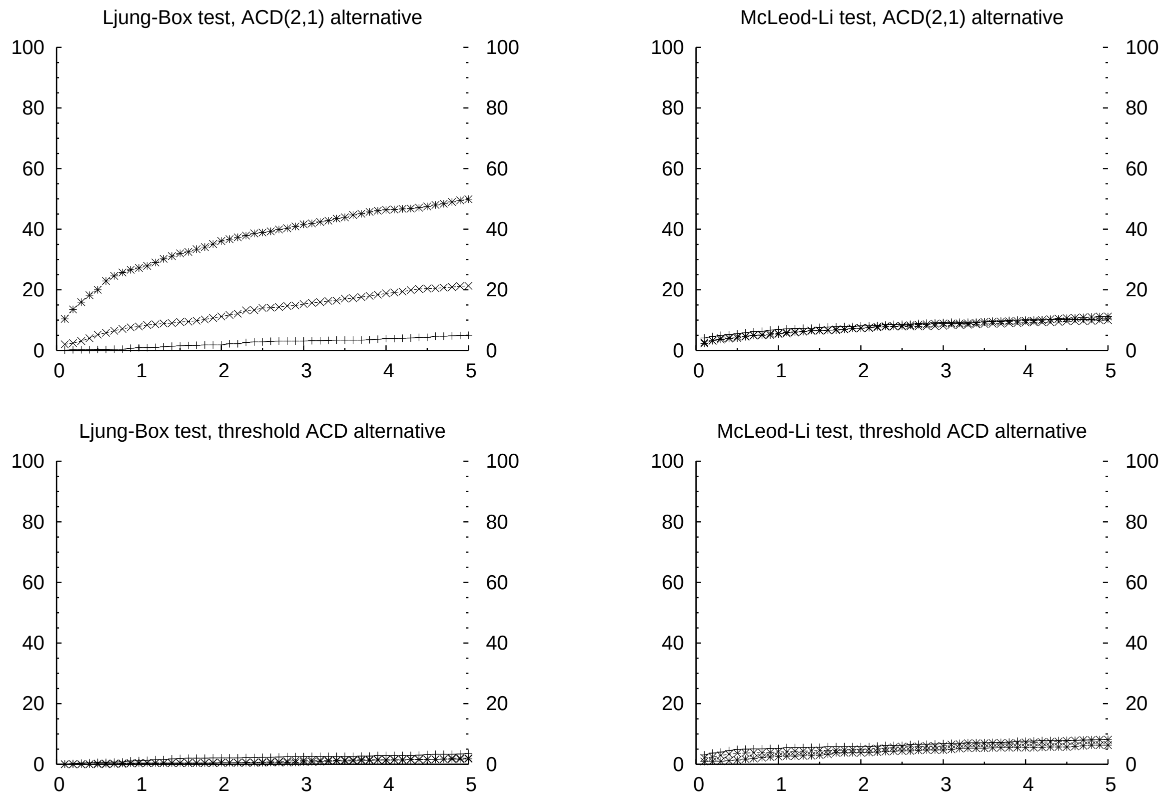
In this section we apply our evaluation tests to ACD models fitted to Trades and Quotes (TAQ) data available from the New York Stock Exchange (NYSE). We use a period of six months from the beginning of July 2002 to the end of December 2002 and concentrate on intertrade durations between transactions of IBM shares.

Before using the data we remove the trades which are uncorrected and irregular. This can be done using the correction indicator attached to each trade. We also remove all trades occuring before 9:30 am. and after 4:00 pm. The NYSE was entirely closed on July 4 (Independence Day), September 2 (Labor Day), November 28 (Thanksgiving Day) and December 25 (Christmas Day), and partly closed on July 5, September 11, November 29 and December 24. For this reason all these days have been removed from the sample. Furthermore, we only consider unique trading times and hence simultaneously recorded trades are regarded as a single trade. Finally, the trades are treated consecutively from day to day, ignoring the overnight duration.

Owing to the enormous amount of trades, nearly half a million during the period in question prior to removals, we consider every week in the sample separately and perform the estimation and tests only on complete five-day weeks. This leaves us with 21 weeks with approximately 15000–20000 trades in each. The exact dates, numbers of trades and some summary statistics of durations can be found in Table 1. We note that the consecutive treating of the trades implies, in particular, that over each of the 21 weeks, estimation is not re-initialized at the beginning of every day. This practice differs from the approach of Engle and Russell (1998) but is applied by Bauwens and Giot (2000).

As is well documented in the literature, there is a strong diurnal component in the duration series. The durations tend to be shorter around the beginning and the end of a trading day, when traders open and close their positions, respectively, and longer around lunchtime. This results in an inverted U–shape pattern in the moving average of durations over the day. A common practice in the literature is to first "diurnally adjust" the series by approximating the average durations using a cubic spline and then removing this diurnal component from the

Figure 5b: Results from power simulations of the tests using the ACD(2,1) model given in equations (44)–(46) and the threshold ACD(1,1) model given in equations (47)–(49) as the alternative data generating process (two upper and two lower figures, respectively). In the figures the power (rejection frequency) is plotted against the nominal size. Both of them are measured in percentage points. Performed tests are the Ljung-Box (left) and McLeod-Li (right) tests with lag length 15. The three lines in each subfigure correspond to sample sizes 1000 (+), 5000 (×) and 10000 (∗).

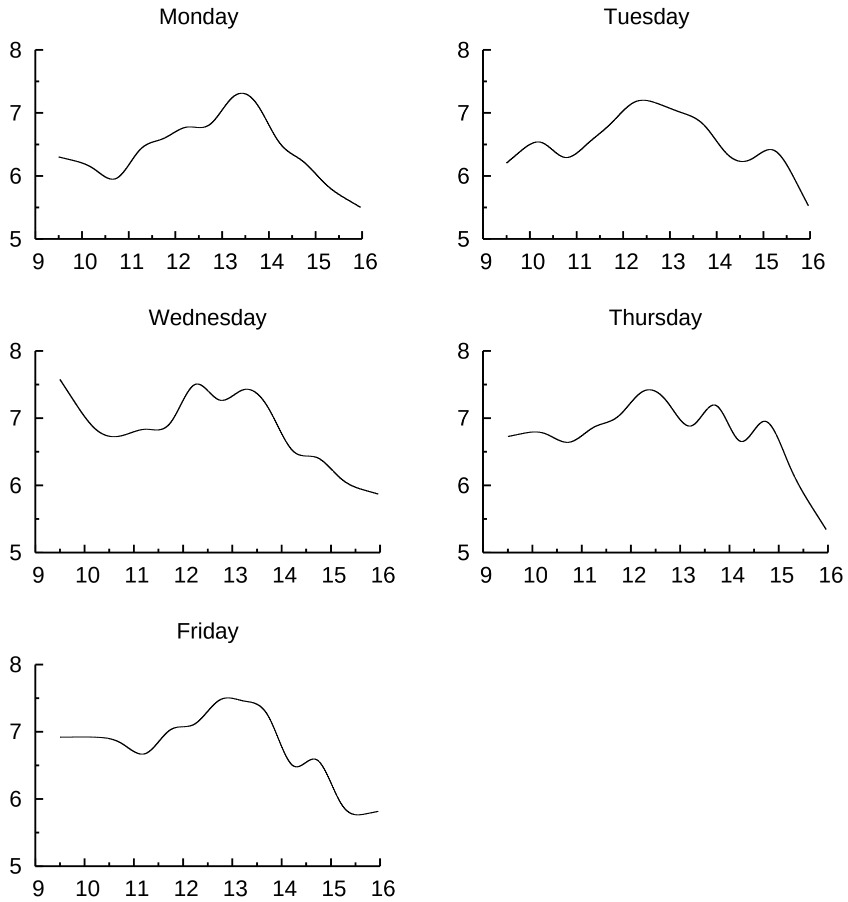


durations. Such a procedure was used for example by Bauwens and Giot (2000). We do this by first averaging the durations over 30–minute periods (9:30–10:00, 10:00–10:30, etc.), setting the average durations of the midpoints of these intervals (9:45, 10:15, etc.) to the resulting averaged values, and then fitting a cubic spline over the course of the day using these midpoints as fixed. The original durations series is then diurnally adjusted by dividing it with the estimated diurnal component. This is done separately for each day of the sample, since the time-of-day component varies depending on the day of the week.

We illustrate the diurnal components in Figure 6 that contains the estimated diurnal components averaged over the 21 weeks, for each day of the week separately.[2] It can be seen that the durations are longest around the midday, typically somewhat shorter before midday, and shortest near the closing. It should be pointed out, however, that the graphs in Figure 6 are merely for illustrating a general pattern, and that the diurnal patterns differ quite strongly from day to day and from week to week. This, however, is a casual observation not based on any statistical inference. Nevertheless, the individual diurnal pattern for any particular day obtained

[2]The averaging is performed to produce a single graph for each day of the week and is made for illustrational purposes only.

Figure 6: Graphs of the averaged estimated diurnal components. The diurnal component is estimated separately for each day of the sample. The resulting diurnal components are then averaged over the 21 weeks, for each day of the week separately. Therefore the subfigures represent the (averaged) mean duration (in seconds) at a particular time of the day versus the time of the day (in hours).



by the cubic spline technique can be very different from the average one for the corresponding day of the week in Figure 6. This raises the question of how a diurnal component of this type should be interpreted: is it that even other events than the time of the day affect its structure? Should the diurnal pattern in fact be estimated jointly with the parameters? This problem is left for further research.

7.2 Estimation and evaluation of ACD(1,1) models

We fit an ACD(1,1) model to each of the 21 diurnally adjusted duration series assuming exponentially distributed errors. The estimation as well as all the other computations are carried out using Ox version 3.30 (see Doornik (2002)). Maximum likelihood estimation of the ACD(1,1) model is performed using the sequential quadratic programming algorithm of Lawrence and Tits (2001) with analytical first derivatives. The parameter estimates that can be found in Table 2 are significant in all of the 21 cases.

We now subject our estimated models to a battery of evaluation tests. We perform tests against ACD(2,1) and ACD(1,2) models, remaining ACD in the standardized durations, STACD(1,1) models of order one and two, and TVACD(1,1) models of order one and two. In the TVACD case, both the intraday time and the intraweek time are used. The p–values of the tests are given in Tables 3 and 4. The values less than 0.01 are shown in boldface.

As to the order of the model, the ACD(1,1) model is rejected in favour of the ACD(2,1) and/or the ACD(1,2) model in seven cases out of 21. As may be expected, in these cases the hypothesis of no remaining ACD in the standardized durations is also (typically) rejected.

The ACD(1,1) model is almost always rejected against the STACD(1,1) model of orders 1 and 2. Perhaps surprisingly, given the findings of Engle and Russell (1998) mentioned in Section 4, rejections using tests based on assuming $K = 2$ (nonmonotonic change) are not systematically stronger than the ones obtained by assuming $K = 1$. More research is needed to find out what kind of nonlinear ACD model would fit the data best. Nevertheless, the results indicate that the linear ACD model does not capture the dynamics of the duration process in a satisfactory fashion and that a nonlinear model should be considered. In addition to the family of STACD models, the TACD model of Zhang, Russell, and Tsay (2001) could be a viable alternative.

Of the tests against time-varying ACD models, we first consider the ones based on intraday time. Almost all of the 21 models pass the test against the first-order TVACD model, but about one third of them are rejected against the second-order one. This rejection may be interpreted as showing that the removal of the diurnal component has not been successful, since there still is an identifiable parameter change in the process within the trading day. If this is a valid interpretation, there seems to be room for improvement in methods for diurnal adjustment of the durations. It is also possible that an erroneous linearity assumption causes these rejections.

When the time is measured as intraweek time all the 21 models pass the tests against time-varying ACD models of orders 1 and 2. It can be concluded that the structure of the duration process does not seem to change within the week in any of the cases. We have also attempted to make the diurnal adjustment using the same time-of-day-curve for all the days of the week. The results from this evaluation test (not shown) are different from the previous ones: the ACD(1,1) model is rejected or nearly rejected at the 1% level in favour of a TVACD model of order 1 and/or 2 in about two thirds of the cases. This suggests that the diurnal pattern is not the same for all days of the week. A tentative conclusion is that conditioning the diurnal adjustment on the day of the week may be of importance.

8 Conclusions

In this paper we present a general framework for evaluating ACD models using Lagrange multiplier or Lagrange multiplier type tests. We derive several misspecification tests of the functional form of the conditional mean of an ACD model. The alternatives considered are parametric, and hence, in case of rejection, they may suggest a direction in which to extend the model. Clearly, the test battery may also be viewed as a set of misspecification indicators that convey information about the fit instead of prompting a particular action to extend the model.

Our tests are simple to use, since the model only has to be estimated under the null hypothesis, and computation of any of the test statistics simply requires one or two additional ordinary linear regressions. Versions of the test statistics robust to deviations from distributional assumptions other than those being explicitly tested are also presented but it appears that robustifying the tests is not as important in practice as it is in some other time series applications. All the tests are found to have good size properties.

Results of the application to Trades and Quotes data clearly point out the need for nonlinear ACD models. This issue will be taken up in future work. Diurnal adjustment of durations appears to be another topic worth further consideration.

Week	Dates	Number of observations	Original durations			Transformed durations		
			min	mean	max	min	mean	max
1	Jul 8–12	18025	1	6.4814	217	0.0654	0.9992	26.8210
2	Jul 15–19	18179	1	6.4278	118	0.0763	0.9982	14.9018
3	Jul 22–26	18240	1	6.4078	92	0.1024	0.9994	11.7863
4	Jul 29 – Aug 2	17088	1	6.8295	69	0.0996	1.0004	9.8779
5	Aug 5–9	18585	1	6.2917	89	0.1001	1.0015	11.6193
6	Aug 12–16	17544	1	6.6656	126	0.1043	1.0007	17.0263
7	Aug 19–23	16900	1	6.9173	79	0.0972	1.0009	14.0726
8	Aug 26–30	15172	1	7.7050	118	0.0727	1.0008	15.3554
9	Sep 16–20	17388	1	6.7126	76	0.0949	0.9985	9.8126
10	Sep 23–27	16273	1	7.1859	117	0.1000	0.9994	14.3419
11	Sep 30 – Oct 4	17732	1	6.5917	238	0.0831	1.0001	22.4224
12	Oct 7–11	17077	1	6.8220	183	0.0993	0.9987	24.1145
13	Oct 14–18	17129	1	6.8193	90	0.0930	0.9990	12.1304
14	Oct 21–25	18500	1	6.3204	106	0.1090	1.0000	14.9705
15	Oct 28 – Nov 1	18702	1	6.2453	80	0.1207	0.9997	10.9473
16	Nov 4–8	19464	1	6.0036	83	0.1113	0.9985	13.6595
17	Nov 11–15	17383	1	6.7255	151	0.0888	0.9998	20.0673
18	Nov 18–22	20317	1	5.7546	93	0.1164	1.0005	17.5540
19	Dec 2–6	20811	1	5.6180	142	0.1253	1.0003	21.4300
20	Dec 9–13	19248	1	6.0743	81	0.1135	1.0004	13.1855
21	Dec 16–20	18818	1	6.2150	82	0.0951	0.9999	12.1802

Table 1: Statistics of the durations. “Transformed durations” refers to the diurnally adjusted durations.

Week	Omega		Alpha		Beta		Value of the
	estimate	(stdev)	estimate	(stdev)	estimate	(stdev)	log-likelihood
1	0.2526	(0.0304)	0.1681	(0.0130)	0.5811	(0.0396)	-17573.1365
2	0.1652	(0.0162)	0.1119	(0.0076)	0.7230	(0.0207)	-17888.7339
3	0.1333	(0.0143)	0.1112	(0.0073)	0.7557	(0.0194)	-17945.3494
4	0.1573	(0.0183)	0.0743	(0.0064)	0.7686	(0.0220)	-16992.9188
5	0.1800	(0.0210)	0.0767	(0.0065)	0.7436	(0.0248)	-18495.7041
6	0.1396	(0.0166)	0.0774	(0.0064)	0.7832	(0.0205)	-17415.2212
7	0.1415	(0.0153)	0.0807	(0.0064)	0.7780	(0.0187)	-16783.2353
8	0.1662	(0.0220)	0.0679	(0.0067)	0.7661	(0.0254)	-15098.1953
9	0.1327	(0.0229)	0.0741	(0.0077)	0.7931	(0.0288)	-17222.3036
10	0.1235	(0.0152)	0.0889	(0.0071)	0.7878	(0.0197)	-16059.7855
11	0.1060	(0.0137)	0.0858	(0.0072)	0.8085	(0.0187)	-17509.7025
12	0.1610	(0.0154)	0.1026	(0.0073)	0.7367	(0.0193)	-16847.6630
13	0.1337	(0.0147)	0.0823	(0.0065)	0.7841	(0.0182)	-16957.2471
14	0.1476	(0.0161)	0.0974	(0.0073)	0.7553	(0.0209)	-18287.0412
15	0.1634	(0.0211)	0.0929	(0.0076)	0.7439	(0.0264)	-18534.1774
16	0.1528	(0.0221)	0.0741	(0.0068)	0.7730	(0.0269)	-19295.7451
17	0.1898	(0.0246)	0.0977	(0.0086)	0.7130	(0.0299)	-17162.8802
18	0.1918	(0.0245)	0.0851	(0.0075)	0.7234	(0.0295)	-20199.7797
19	0.2095	(0.0229)	0.0865	(0.0073)	0.7042	(0.0271)	-20673.7455
20	0.1289	(0.0251)	0.0640	(0.0074)	0.8071	(0.0311)	-19128.6345
21	0.1612	(0.0278)	0.0781	(0.0080)	0.7608	(0.0339)	-18669.6529

Table 2: Results from the maximum likelihood estimation of the ACD(1,1) models with exp(1) errors.

Week	ACD(1,1) vs ACD(2,1) (ordinary)	ACD(1,1) vs ACD(2,1) (robust)	ACD(1,1) vs ACD(1,2) (ordinary)	ACD(1,1) vs ACD(1,2) (robust)	No remaining ACD (ordinary)	No remaining ACD (robust)	No STACD K=1 (ordinary)	No STACD K=1 (robust)	No STACD K=2 (ordinary)	No STACD K=2 (robust)
1	0.0749	0.0932	0.0928	0.1136	0.2982	0.2949	$\mathbf{1 \times 10^{-5}}$	$\mathbf{6 \times 10^{-8}}$	$\mathbf{3 \times 10^{-6}}$	$\mathbf{2 \times 10^{-9}}$
2	0.6507	0.6474	0.6047	0.5996	0.1590	0.1332	$\mathbf{3 \times 10^{-6}}$	$\mathbf{3 \times 10^{-6}}$	$\mathbf{8 \times 10^{-7}}$	$\mathbf{4 \times 10^{-7}}$
3	0.1920	0.1997	0.2304	0.2374	0.2152	0.2049	**0.0065**	**0.0040**	0.0276	0.0158
4	**0.0037**	**0.0050**	**0.0013**	**0.0026**	**0.0057**	**0.0075**	$\mathbf{8 \times 10^{-5}}$	$\mathbf{4 \times 10^{-5}}$	**0.0007**	**0.0003**
5	0.0108	0.0127	**0.0099**	0.0112	0.0117	0.0109	$\mathbf{5 \times 10^{-5}}$	$\mathbf{9 \times 10^{-6}}$	**0.0004**	$\mathbf{3 \times 10^{-5}}$
6	0.1177	0.1125	0.1647	0.1544	0.1675	0.1525	**0.0002**	**0.0001**	**0.0020**	**0.0007**
7	**0.0055**	**0.0046**	**0.0018**	**0.0016**	**0.0086**	**0.0065**	**0.0033**	**0.0030**	0.0124	0.0126
8	0.0138	0.0139	0.0189	0.0167	0.0286	0.0303	0.1217	0.1202	0.0895	0.1004
9	0.3133	0.3003	0.3014	0.2834	0.3392	0.3198	**0.0009**	**0.0005**	**0.0070**	**0.0039**
10	0.4766	0.4804	0.4709	0.4758	0.2667	0.2699	$\mathbf{1 \times 10^{-5}}$	$\mathbf{2 \times 10^{-6}}$	$\mathbf{4 \times 10^{-6}}$	$\mathbf{3 \times 10^{-6}}$
11	0.0765	0.0965	0.0949	0.1152	0.0578	0.0956	**0.0002**	**0.0012**	**0.0005**	**0.0009**
12	**0.0028**	**0.0028**	**0.0019**	**0.0018**	**0.0013**	**0.0007**	**0.0009**	**0.0002**	**0.0038**	**0.0012**
13	**0.0003**	**0.0004**	**0.0003**	**0.0003**	$\mathbf{8 \times 10^{-5}}$	**0.0001**	$\mathbf{2 \times 10^{-6}}$	$\mathbf{3 \times 10^{-7}}$	$\mathbf{3 \times 10^{-5}}$	$\mathbf{5 \times 10^{-6}}$
14	**0.0060**	**0.0066**	**0.0038**	**0.0046**	**0.0023**	**0.0032**	**0.0001**	$\mathbf{2 \times 10^{-5}}$	**0.0008**	$\mathbf{7 \times 10^{-5}}$
15	0.1585	0.1481	0.0797	0.0772	0.1814	0.1760	$\mathbf{6 \times 10^{-7}}$	$\mathbf{4 \times 10^{-8}}$	$\mathbf{7 \times 10^{-6}}$	$\mathbf{6 \times 10^{-7}}$
16	0.1036	0.0962	0.1012	0.0926	0.0687	0.0577	**0.0024**	**0.0018**	**0.0020**	**0.0013**
17	0.6923	0.6942	0.7378	0.7384	0.8032	0.7945	**0.0007**	**0.0002**	**0.0019**	**0.0004**
18	**0.0084**	0.0120	**0.0055**	**0.0084**	**0.0096**	0.0146	$\mathbf{3 \times 10^{-5}}$	$\mathbf{2 \times 10^{-5}}$	$\mathbf{6 \times 10^{-6}}$	$\mathbf{7 \times 10^{-6}}$
19	0.0407	0.0397	0.0231	0.0235	0.0379	0.0479	$\mathbf{6 \times 10^{-5}}$	**0.0002**	$\mathbf{4 \times 10^{-5}}$	$\mathbf{5 \times 10^{-5}}$
20	0.9678	0.9679	0.8619	0.8607	0.7678	0.7661	$\mathbf{4 \times 10^{-5}}$	$\mathbf{5 \times 10^{-6}}$	**0.0002**	$\mathbf{5 \times 10^{-5}}$
21	0.8595	0.8587	0.6393	0.6392	0.8243	0.8280	**0.0041**	**0.0033**	**0.0033**	**0.0026**

Table 3: p–values of the tests of ACD(1,1) models. The tests are the ones against ACD(2,1) and ACD(1,2) models, remaining ACD in the standardized durations, and STACD(1,1) model of orders 1 and 2. Two versions of each test are performed: the *ordinary* refers to the test done using the auxiliary regression, and the *robust* is the robustified version. The p–values less than 0.01 are shown in boldface.

Week	Tests using intraday time				Tests using intraweek time			
	No TVACD K=1		No TVACD K=2		No TVACD K=1		No TVACD K=2	
	(ordinary)	(robust)	(ordinary)	(robust)	(ordinary)	(robust)	(ordinary)	(robust)
1	0.9077	0.8900	**0.0037**	**0.0009**	0.0883	0.1180	0.3286	0.3951
2	0.2452	0.2713	0.5399	0.5758	0.7698	0.7435	0.8239	0.8043
3	0.9014	0.8994	0.8765	0.8805	0.1810	0.1168	0.0496	0.0398
4	0.7490	0.7258	0.4092	0.3564	0.9691	0.9670	0.9811	0.9735
5	0.7167	0.7058	0.5430	0.5800	0.3774	0.3463	0.7864	0.7600
6	0.4052	0.3715	**0.0059**	**0.0054**	0.2823	0.3279	0.4972	0.6410
7	0.6626	0.6475	0.1771	0.1148	0.5990	0.5840	0.9243	0.9096
8	0.9419	0.9414	0.3068	0.3283	0.3701	0.3699	0.5281	0.5192
9	0.9435	0.9381	0.3428	0.2993	0.2480	0.2510	0.1611	0.1409
10	0.5326	0.5702	0.0296	0.0103	0.0169	0.0165	0.0877	0.0825
11	0.8669	0.8753	0.0574	0.0610	0.2726	0.2357	0.6489	0.5912
12	0.8500	0.8437	0.8431	0.8195	0.2056	0.1657	0.4842	0.4175
13	0.9243	0.9092	0.0328	0.0133	0.9671	0.9650	0.9878	0.9851
14	0.5777	0.5600	**0.0002**	**0.0001**	0.5353	0.5337	0.4324	0.4131
15	0.0578	0.0576	**0.0052**	**0.0070**	0.1779	0.2016	0.2514	0.2346
16	0.4290	0.4047	**0.0077**	**0.0033**	0.4569	0.3988	0.7721	0.7155
17	0.0325	0.0169	0.0563	0.0448	0.4896	0.5390	0.4291	0.5608
18	0.2992	0.2149	0.6958	0.5817	0.7600	0.7932	0.8832	0.9088
19	**0.0002**	$\mathbf{4 \times 10^{-5}}$	$\mathbf{5 \times 10^{-7}}$	$\mathbf{1 \times 10^{-7}}$	0.4018	0.3797	0.6769	0.6237
20	0.2875	0.2413	0.2091	0.1262	0.5438	0.5316	0.5325	0.4821
21	**0.0083**	**0.0025**	$\mathbf{8 \times 10^{-5}}$	$\mathbf{1 \times 10^{-5}}$	0.2991	0.3125	0.3797	0.4107

Table 4: p–values of the tests of ACD(1,1) models (continued). The tests are the ones against TVACD(1,1) models of orders 1 and 2. Time is defined either as intraday time or as intraweek time. Two versions of each test are performed: the *ordinary* refers to the test done using the auxiliary regression, and the *robust* is the robustified version. The p–values less than 0.01 are shown in boldface.

References

ANDERSON, H. M., K. NAM, AND F. VAHID (1999): "Asymmetric nonlinear smooth transition GARCH models," in *Nonlinear Time Series Analysis of Economic and Financial Data*, ed. by P. Rothman, pp. 191–207. Kluwer Academic Publishers, Boston.

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, 117–149.

BAUWENS, L., P. GIOT, J. GRAMMIG, AND D. VEREDAS (2004): "A comparison of financial duration models via density forecasts," *International Journal of Forecasting*, 20, 589–609.

BIERENS, H. J. (1990): "A consistent conditional moment test of functional form," *Econometrica*, 58, 1443–1458.

BOLLERSLEV, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

DE JONG, R. M. (1996): "The Bierens test under data dependence," *Journal of Econometrics*, 72, 1–32.

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.

DOORNIK, J. A. (2002): *Object-Oriented Matrix Programming Using Ox*. Timberlake Consultants Press, 3rd edn., see also www.nuff.ox.ac.uk/Users/Doornik.

DROST, F. C., AND B. J. M. WERKER (2004): "Semiparametric duration models," *Journal of Business and Economic Statistics*, 22, 40–50.

DUFOUR, A., AND R. F. ENGLE (2000): "The ACD model: predictability of the time between consecutive trades," Discussion Paper, ISMA Centre, University of Reading.

ENGLE, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50, 987–1008.

——— (1984): "Wald, likelihood ratio, and Lagrange multiplier tests in econometrics," in *Handbook of Econometrics*, ed. by Z. Griliches, and M. D. Intriligator, vol. II, chap. 13. Elsevier Science Publishers.

——— (2000): "The econometrics of ultra-high-frequency data," *Econometrica*, 68, 1–22.

ENGLE, R. F., AND J. R. RUSSELL (1998): "Autoregressive conditional duration: a new model for irregularly spaced transaction data," *Econometrica*, 66, 1127–1162.

FERNANDES, M., AND J. GRAMMIG (in press): "Nonparametric specification tests for conditional duration models," *Journal of Econometrics*.

GONZÁLEZ-RIVERA, G. (1998): "Smooth-transition GARCH models," *Studies in Nonlinear Dynamics and Econometrics*, 3, 61–78.

GOURIÉROUX, C., A. MONFORT, AND A. TROGNON (1984): "Pseudo maximum likelihood methods: theory," *Econometrica*, 52, 681–700.

GRAMMIG, J., AND K.-O. MAURER (2000): "Non-monotonic hazard functions and the autoregressive conditional duration model," *Econometrics Journal*, 3, 16–38.

HAGERUD, G. E. (1996): "A smooth transition ARCH model for asset returns," Working Paper Series in Economics and Finance No. 162, Stockholm School of Economics.

HANSEN, B. E. (1996): "Inference when a nuisance parameter is not identified under the null hypothesis," *Econometrica*, 64, 413–430.

HAUTSCH, N. (2002): "Testing the conditional mean function of autoregressive conditional duration models," Discussion paper, University of Konstanz.

HONG, Y. (1999): "Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach," *Journal of the American Statistical Association*, 94, 1201–1220.

HONG, Y., AND T.-H. LEE (2003): "Diagnostic checking for the adequacy of nonlinear time series models," *Econometric Theory*, 19, 1065–1121.

LAWRENCE, C. T., AND A. L. TITS (2001): "A computationally efficient feasible sequential quadratic programming algorithm," *SIAM Journal on Optimization*, 11, 1092–1118.

LI, W. K., AND T. K. MAK (1994): "On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity," *Journal of Time Series Analysis*, 15, 627–636.

LI, W. K., AND P. L. H. YU (2003): "On the residual autocorrelation of the autoregressive conditional duration model," *Economics Letters*, 79, 169–175.

LJUNG, G. M., AND G. E. P. BOX (1978): "On a measure of lack of fit in time series models," *Biometrika*, 65, 297–303.

LUNDBERGH, S., AND T. TERÄSVIRTA (2002): "Evaluating GARCH models," *Journal of Econometrics*, 110, 417–435.

LUNDE, A. (1999): "A generalized gamma autoregressive conditional duration model," Working paper, Aalborg University.

LUUKKONEN, R., P. SAIKKONEN, AND T. TERÄSVIRTA (1988): "Testing linearity against smooth transition autoregressive models," *Biometrika*, 75, 491–499.

MCLEOD, A. I., AND W. K. LI (1983): "Diagnostic checking ARMA time series models using squared-residual autocorrelations," *Journal of Time Series Analysis*, 4, 269–273.

NEWEY, W. K. (1985): "Maximum likelihood specification testing and conditional moment tests," *Econometrica*, 53, 1047–1070.

WHITE, H. (1994): *Estimation, Inference and Specification Analysis*. Cambridge University Press, New York.

WOOLDRIDGE, J. M. (1991): "Specification testing and quasi-maximum-likelihood estimation," *Journal of Econometrics*, 48, 29–55.

ZHANG, M. Y., J. R. RUSSELL, AND R. S. TSAY (2001): "A nonlinear autoregressive conditional duration model with applications to financial transaction data," *Journal of Econometrics*, 104, 179–207.

Appendix A

In this appendix we prove Theorem 1. The proof of Theorem 4 is almost identical and is omitted.

Proof of Theorem 1. Let $\theta = (\theta_1', \theta_2')'$ be the parameter vector and define the conditional quasi log-likelihood function for observation x_i as

$$l_i(\theta) = -\frac{x_i}{\psi_i + \varphi_i} - \ln(\psi_i + \varphi_i). \tag{50}$$

The partial derivatives of (50) with respect to θ_1 and θ_2 are

$$\begin{aligned} \frac{\partial l_i(\theta)}{\partial \theta_1} &= \tfrac{1}{\psi_i+\varphi_i}\left(\tfrac{\partial \psi_i}{\partial \theta_1} + \tfrac{\partial \varphi_i}{\partial \theta_1}\right)\left(\tfrac{x_i}{\psi_i+\varphi_i} - 1\right) \\ \frac{\partial l_i(\theta)}{\partial \theta_2} &= \tfrac{1}{\psi_i+\varphi_i}\tfrac{\partial \varphi_i}{\partial \theta_2}\left(\tfrac{x_i}{\psi_i+\varphi_i} - 1\right). \end{aligned}$$

Letting $\theta^0 = (\theta_1^{0\prime}, \theta_2^{0\prime})'$ be the true (under the null hypothesis) parameter vector, the score for observation x_i evaluated at the true parameter values equals

$$\begin{aligned} \frac{\partial l_i(\theta^0)}{\partial \theta} &= \begin{bmatrix} \frac{1}{\psi_i(\theta_1^0)+\varphi_i(\theta_1^0,\theta_2^0)}\left(\frac{\partial \psi_i(\theta_1^0)}{\partial \theta_1} + \frac{\partial \varphi_i(\theta_1^0,\theta_2^0)}{\partial \theta_1}\right)\left(\frac{x_i}{\psi_i(\theta_1^0)+\varphi_i(\theta_1^0,\theta_2^0)} - 1\right) \\ \frac{1}{\psi_i(\theta_1^0)+\varphi_i(\theta_1^0,\theta_2^0)}\frac{\partial \varphi_i(\theta_1^0,\theta_2^0)}{\partial \theta_2}\left(\frac{x_i}{\psi_i(\theta_1^0)+\varphi_i(\theta_1^0,\theta_2^0)} - 1\right) \end{bmatrix} \\ &= \left(\frac{x_i}{\psi_i(\theta_1^0)} - 1\right)\begin{bmatrix} \frac{1}{\psi_i(\theta_1^0)}\frac{\partial \psi_i(\theta_1^0)}{\partial \theta_1} \\ \frac{1}{\psi_i(\theta_1^0)}\frac{\partial \varphi_i(\theta_1^0,\theta_2^0)}{\partial \theta_2} \end{bmatrix} = c_i(\theta_1^0)\begin{bmatrix} \mathbf{a}_i(\theta_1^0) \\ \mathbf{b}_i(\theta_1^0,\theta_2^0) \end{bmatrix} = c_i^0 \begin{bmatrix} \mathbf{a}_i^0 \\ \mathbf{b}_i^0 \end{bmatrix}. \end{aligned}$$

Furthermore, let $\hat{\theta} = (\hat{\theta}_1', \theta_2^{0\prime})'$ be the vector of maximum likelihood estimates, estimated under the null. Then the score evaluated at the ML estimates is (note that the upper block of the score is now just a vector of zeros)

$$\frac{\partial l(\hat{\theta})}{\partial \theta} = \sum_{i=1}^{n} \frac{\partial l_i(\hat{\theta})}{\partial \theta} = \begin{bmatrix} \mathbf{0} \\ \sum_{i=1}^{n} c_i(\hat{\theta}_1)\,\mathbf{b}_i(\hat{\theta}_1, \theta_2^0) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \sum_{i=1}^{n} \hat{c}_i \hat{\mathbf{b}}_i \end{bmatrix}.$$

Under condition (v) of Theorem 1 in Engle (2000)

$$n^{-1/2}\frac{\partial l(\theta^0)}{\partial \theta} = n^{-1/2}\sum_{i=1}^{n} \frac{\partial l_i(\theta^0)}{\partial \theta} \xrightarrow{d} N\left(0, n^{-1}\sum_{i=1}^{n} E\left[\frac{\partial l_i(\theta^0)}{\partial \theta}\frac{\partial l_i(\theta^0)}{\partial \theta'}\right]\right).$$

The expectation in this expression can be written as

$$\begin{aligned} E\left[\frac{\partial l_i(\theta^0)}{\partial \theta}\frac{\partial l_i(\theta^0)}{\partial \theta'}\right] &= E\left[\left(c_i^0\right)^2 \begin{bmatrix} \mathbf{a}_i^0\mathbf{a}_i^{0\prime} & \mathbf{a}_i^0\mathbf{b}_i^{0\prime} \\ \mathbf{b}_i^0\mathbf{a}_i^{0\prime} & \mathbf{b}_i^0\mathbf{b}_i^{0\prime} \end{bmatrix}\right] \\ &= E\left(c_i^0\right)^2 \cdot E\left\{\begin{bmatrix} \mathbf{a}_i^0\mathbf{a}_i^{0\prime} & \mathbf{a}_i^0\mathbf{b}_i^{0\prime} \\ \mathbf{b}_i^0\mathbf{a}_i^{0\prime} & \mathbf{b}_i^0\mathbf{b}_i^{0\prime} \end{bmatrix}\right\} \end{aligned}$$

since $c_i^0 = \frac{x_i}{\psi_i(\theta_1^0)} - 1 = \varepsilon_i - 1$ is independent of the other terms, which are measurable with respect to $\mathcal{F}_{i-1}$. Furthermore, $E\left(c_i^0\right)^2 = E\left(\frac{x_i}{\psi_i(\theta_1^0)} - 1\right)^2 = Var(\varepsilon_i) = 1$, so that

$$n^{-1/2}\frac{\partial l(\theta^0)}{\partial \theta} \xrightarrow{d} N\left(0, n^{-1}\sum_{i=1}^{n} E\left\{\begin{bmatrix} \mathbf{a}_i^0\mathbf{a}_i^{0\prime} & \mathbf{a}_i^0\mathbf{b}_i^{0\prime} \\ \mathbf{b}_i^0\mathbf{a}_i^{0\prime} & \mathbf{b}_i^0\mathbf{b}_i^{0\prime} \end{bmatrix}\right\}\right).$$

This implies that the quadratic form

$$\left\{n^{-1/2}\frac{\partial l(\theta^0)}{\partial\theta'}\right\}\left\{n^{-1}\sum_{i=1}^{n}E\left\{\left[\begin{array}{cc}\mathbf{a}_i^0\mathbf{a}_i^{0\prime} & \mathbf{a}_i^0\mathbf{b}_i^{0\prime}\\ \mathbf{b}_i^0\mathbf{a}_i^{0\prime} & \mathbf{b}_i^0\mathbf{b}_i^{0\prime}\end{array}\right]\right\}\right\}^{-1}\left\{n^{-1/2}\frac{\partial l(\theta^0)}{\partial\theta}\right\}$$

has an asymptotic χ^2 distribution with $\dim\theta_2$ degrees of freedom.

Since $\hat{\theta}$ is a consistent estimator of θ^0 (Theorem 1, Engle (2000)) and

$$n^{-1}\sum_{i=1}^{n}\left[\begin{array}{cc}\mathbf{a}_i\left(\hat{\theta}_1\right)\mathbf{a}_i'\left(\hat{\theta}_1\right) & \mathbf{a}_i\left(\hat{\theta}_1\right)\mathbf{b}_i'\left(\hat{\theta}_1,\theta_2^0\right)\\ \mathbf{b}_i\left(\hat{\theta}_1,\theta_2^0\right)\mathbf{a}_i'\left(\hat{\theta}_1\right) & \mathbf{b}_i\left(\hat{\theta}_1,\theta_2^0\right)\mathbf{b}_i'\left(\hat{\theta}_1,\theta_2^0\right)\end{array}\right]=n^{-1}\sum_{i=1}^{n}\left[\begin{array}{cc}\hat{\mathbf{a}}_i\hat{\mathbf{a}}_i' & \hat{\mathbf{a}}_i\hat{\mathbf{b}}_i'\\ \hat{\mathbf{b}}_i\hat{\mathbf{a}}_i' & \hat{\mathbf{b}}_i\hat{\mathbf{b}}_i'\end{array}\right]$$

is a consistent estimator of

$$n^{-1}\sum_{i=1}^{n}E\left\{\left[\begin{array}{cc}\mathbf{a}_i^0\mathbf{a}_i^{0\prime} & \mathbf{a}_i^0\mathbf{b}_i^{0\prime}\\ \mathbf{b}_i^0\mathbf{a}_i^{0\prime} & \mathbf{b}_i^0\mathbf{b}_i^{0\prime}\end{array}\right]\right\}$$

(since the expressions are functions of $\hat{\theta}$ and θ^0, respectively), the LM statistic

$$\begin{aligned}&\left\{n^{-1/2}\frac{\partial l(\hat{\theta})}{\partial\theta'}\right\}\left\{n^{-1}\sum_{i=1}^{n}\left[\begin{array}{cc}\hat{\mathbf{a}}_i\hat{\mathbf{a}}_i' & \hat{\mathbf{a}}_i\hat{\mathbf{b}}_i'\\ \hat{\mathbf{b}}_i\hat{\mathbf{a}}_i' & \hat{\mathbf{b}}_i\hat{\mathbf{b}}_i'\end{array}\right]\right\}^{-1}\left\{n^{-1/2}\frac{\partial l(\hat{\theta})}{\partial\theta}\right\}\\
=&\left[\begin{array}{c}\mathbf{0}\\ \sum_{i=1}^n\hat{c}_i\hat{\mathbf{b}}_i\end{array}\right]'\left\{\left[\begin{array}{cc}\sum_{i=1}^n\hat{\mathbf{a}}_i\hat{\mathbf{a}}_i' & \sum_{i=1}^n\hat{\mathbf{a}}_i\hat{\mathbf{b}}_i'\\ \sum_{i=1}^n\hat{\mathbf{b}}_i\hat{\mathbf{a}}_i' & \sum_{i=1}^n\hat{\mathbf{b}}_i\hat{\mathbf{b}}_i'\end{array}\right]\right\}^{-1}\left[\begin{array}{c}\mathbf{0}\\ \sum_{i=1}^n\hat{c}_i\hat{\mathbf{b}}_i\end{array}\right]\\
=&\left\{\sum_{i=1}^{n}\hat{c}_i\hat{\mathbf{b}}_i'\right\}\left\{\sum_{i=1}^{n}\hat{\mathbf{b}}_i\hat{\mathbf{b}}_i'-\left(\sum_{i=1}^{n}\hat{\mathbf{b}}_i\hat{\mathbf{a}}_i'\right)\left(\sum_{i=1}^{n}\hat{\mathbf{a}}_i\hat{\mathbf{a}}_i'\right)^{-1}\left(\sum_{i=1}^{n}\hat{\mathbf{a}}_i\hat{\mathbf{b}}_i'\right)\right\}^{-1}\left\{\sum_{i=1}^{n}\hat{c}_i\hat{\mathbf{b}}_i\right\}\end{aligned}$$

also has an asymptotic χ^2 distribution with $\dim\theta_2$ degrees of freedom under the null hypothesis.
■

Appendix B

Here we show that the test statistic of Li and Yu (2003) is asymptotically equivalent to the one given in Corollary 9. Li and Yu (2003) only consider the case where the estimated model is an ACD(m,0) model, and for ease of exposition they restrict the derivations to the case $m = 1$. In this case $\hat{\mathbf{a}}_i$ in Corollary 9 reduces to $\hat{\mathbf{a}}_i = \hat{\psi}_i^{-1}(1, x_{i-1})'$. As usual, $\hat{\mathbf{b}}_i = \hat{\psi}_i^{-1}\left(x_{i-1}\hat{\psi}_{i-1}^{-1}, \ldots, x_{i-m^*}\hat{\psi}_{i-m^*}^{-1}\right)'$ and $\hat{c}_i = x_i\hat{\psi}_i^{-1} - 1$. The statistic of Li and Yu (2003) is

$$Q = n\hat{\mathbf{r}}'\left(I_{m^*} - \hat{X}\hat{G}^{-1}\hat{X}'\right)^{-1}\hat{\mathbf{r}} \tag{51}$$

where I_{m^*} is the $m^* \times m^*$ identity matrix, and $\hat{\mathbf{r}}$ is the following $m^* \times 1$ vector:

$$\begin{aligned} \hat{\mathbf{r}} &= n^{-1}\sum \begin{bmatrix} \left(\frac{x_i}{\hat{\psi}_i} - 1\right)\left(\frac{x_{i-1}}{\hat{\psi}_{i-1}} - 1\right) \\ \vdots \\ \left(\frac{x_i}{\hat{\psi}_i} - 1\right)\left(\frac{x_{i-m^*}}{\hat{\psi}_{i-m^*}} - 1\right) \end{bmatrix} \\ &= n^{-1}\sum \hat{c}_i\hat{\psi}_i\hat{\mathbf{b}}_i - n^{-1}\sum \hat{c}_i. \end{aligned} \tag{52}$$

Furthermore, the $m^* \times 2$ matrix $\hat{X}$ is given by

$$\begin{aligned} \hat{X} &= n^{-1}\sum \begin{bmatrix} \frac{x_i}{\hat{\psi}_i^2}\left(\frac{x_{i-1}}{\hat{\psi}_{i-1}} - 1\right) & \frac{x_i x_{i-1}}{\hat{\psi}_i^2}\left(\frac{x_{i-1}}{\hat{\psi}_{i-1}} - 1\right) \\ \vdots & \vdots \\ \frac{x_i}{\hat{\psi}_i^2}\left(\frac{x_{i-m^*}}{\hat{\psi}_{i-m^*}} - 1\right) & \frac{x_i x_{i-1}}{\hat{\psi}_i^2}\left(\frac{x_{i-m^*}}{\hat{\psi}_{i-m^*}} - 1\right) \end{bmatrix} \\ &= n^{-1}\sum \frac{x_i}{\hat{\psi}_i}\hat{\psi}_i\hat{\mathbf{b}}_i\hat{\mathbf{a}}_i' - n^{-1}\sum \frac{x_i}{\hat{\psi}_i}\mathbf{1}_{m^*}\hat{\mathbf{a}}_i' \end{aligned} \tag{53}$$

where $\mathbf{1}_{m^*}$ is an $m^* \times 1$ vector of ones, and

$$\begin{aligned} \hat{G} &= n^{-2}\sum \begin{bmatrix} 2\frac{x_i}{\hat{\psi}_i^3} - \frac{1}{\hat{\psi}_i^2} & 2\frac{x_i x_{i-1}}{\hat{\psi}_i^3} - \frac{x_{i-1}}{\hat{\psi}_i^2} \\ 2\frac{x_i x_{i-1}}{\hat{\psi}_i^3} - \frac{x_{i-1}}{\hat{\psi}_i^2} & 2\frac{x_i x_{i-1}^2}{\hat{\psi}_i^3} - \frac{x_{i-1}^2}{\hat{\psi}_i^2} \end{bmatrix} \\ &= n^{-2}\sum\left(2\frac{x_i}{\hat{\psi}_i} - 1\right)\hat{\mathbf{a}}_i\hat{\mathbf{a}}_i'. \end{aligned} \tag{54}$$

Now, (52), (53), and (54) converge in probability to the same quantities as $n^{-1}\sum \hat{c}_i\hat{\psi}_i\hat{\mathbf{b}}_i$, $n^{-1}\sum \hat{\psi}_i\hat{\mathbf{b}}_i\hat{\mathbf{a}}_i'$, and $n^{-2}\sum \hat{\mathbf{a}}_i\hat{\mathbf{a}}_i'$, respectively, as $n \to \infty$. Furthermore, $n^{-1}\sum \hat{\psi}_i^2\hat{\mathbf{b}}_i\hat{\mathbf{b}}_i'$ converges in probability to I_{m^*}. It follows that the statistic (51) and the one given in Corollary 9 are asymptotically equivalent.

Appendix C

In this appendix we provide a description of the generalized spectral test introduced in Hong (1999) and suggested as a misspecification test for ACD models in Hong and Lee (2003). We present the test in the context of using it as a misspecification test for the estimated standardized durations of an ACD model. Denoting the standardized duration series by ε_i, consider the covariance between the empirical characteristic functions of ε_i and ε_{i-j},

$$\sigma_j(u,v) = cov\,(e^{\imath u\varepsilon_i}, e^{\imath v\varepsilon_{i-j}}) \tag{55}$$

where $\imath = \sqrt{-1}$ and $j = 0, \pm1, \ldots$. Hong (1999) calls the Fourier transform of (55) (it exists under some regularity conditions)

$$f(\omega,u,v) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\sigma_j(u,v)e^{-\imath j\omega}$$

the generalized spectral density function of $\{\varepsilon_i\}$. This is estimated using the following kernel estimator:

$$\hat{f}(\omega,u,v) = \frac{1}{2\pi}\sum_{j=1-n}^{n-1}(1-|j|/n)^{1/2}k(j/p)\hat{\sigma}_j(u,v)e^{-\imath j\omega} \tag{56}$$

where $k(\cdot)$ is a kernel function, p is a bandwidth, n is the sample size, and $\hat{\sigma}_j(\cdot,\cdot)$ is a consistent estimator for the covariance of the two empirical characteristic functions. The generalized spectral test is defined as the (suitably weighted and standardized) L_2 norm of the difference of the estimate of f and the estimate of f under the null hypothesis of serial independence of the standardized duration series. For an exact formulation of the test statistic we refer to Hong (1999) or Hong and Lee (2003).

Using the generalized spectral test involves, among other things, the choice of the bandwidth p and the kernel function $k(\cdot)$. For the bandwidth p, Hong (1999) discusses a data-driven method for choosing an optimal bandwidth (in the sense of an integrated mean squared error criterion for the estimator $\hat{f}$ in equation (56)) given a preliminary bandwidth $\bar{p}$. The choice of $\bar{p}$ remains somewhat arbitrary, but the simulation studies of Hong (1999) and Hong and Lee (2003) suggest that the test statistic is quite robust to the choice of $\bar{p}$. We choose the values 25, 50, 75, and 100 for the preliminary bandwidth.

For the kernel function, Hong (1999) shows that the Daniell kernel is optimal in the sense that it maximizes the asymptotic power of the test over a class of kernel functions. The fact that the Daniell kernel has an unbounded support implies that when using the estimator (56), the covariance $\hat{\sigma}_j(u,v)$ has to be computed for all j between 1 and $n-1$. For ultra-high-frequency data, where sample sizes are quite large, this becomes computationally demanding. For this reason we choose to use the Parzen kernel that has a bounded support. This considerably reduces the time needed for computing the value of the statistic. According to the simulation results of Hong (1999) this should only have a minor effect on the power of the test.