Studies in Nonlinear Dynamics and Econometrics

Quarterly Journal
January 1998, Volume 2, Number 4
The MIT Press

Studies in Nonlinear Dynamics and Econometrics (ISSN 1081-1826) is a quarterly journal published electronically on the Internet by The MIT Press, Cambridge, Massachusetts, 02142. Subscriptions and address changes should be addressed to MIT Press Journals, Five Cambridge Center, Cambridge, MA 02142; tel.: (617) 253-2889; fax: (617) 577-1545; e-mail: journals-orders@mit.edu Subscription rates are: Individuals $40.00, Institutions $130.00. Canadians add 7% GST. Prices subject to change without notice.

Subscribers are licensed to use journal articles in a variety of ways, limited only as required to insure fair attribution to authors and the Journal, and to prohibit use in a competing commercial product. See the Journal's World Wide Web site for further details. Address inquiries to the Subsidiary Rights Manager, MIT Press Journals, Five Cambridge Center, Cambridge, MA 02142; tel.: (617) 253-2864; fax: (617) 258-5028; e-mail: journals-rights@mit.edu.

©1998 by the Massachusetts Institute of Technology

GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model

Eric Ghysels
Department of Economics
Pennsylvania State University
University Park, Pennsylvania, USA
eghysels@psu.edu

Joanna Jasiak
Department of Economics
York University
Toronto, Ontario, Canada
jasiakj@yorku.ca

Acknowledgment. Eric Ghysels would like to acknowledge the financial support of the Social Sciences and Humanities Research Council of Canada.

Keywords. tick-by-tick data, subordinated processes, duration models, volatility

Abstract. *We develop a class of ARCH models for series sampled at unequal time intervals set by trade or quote arrivals. Our approach combines insights from the temporal aggregation for GARCH models discussed by Drost and Nijman (1993) and Drost and Werker (1996), and the autoregressive conditional duration model of Engle and Russell (1996) proposed to model the spacing between consecutive financial transactions.*

The class of models introduced here will be called ACD-GARCH. It can be described as a random coefficient GARCH, or doubly stochastic GARCH, where the durations between transactions determine the parameter dynamics. The ACD-GARCH model becomes genuinely bivariate when past asset-return volatilities are allowed to affect transaction durations, and vice versa. Otherwise, the spacings between trades are considered exogenous to the volatility dynamics. This assumption is required in a two-step estimation procedure. The bivariate setup enables us to test for Granger causality between volatility and intratrade durations. Under general conditions, we propose several Generalized Method of Moments (GMM) estimation procedures, some having a Quasi Maximum Likelihood Estimation (QMLE) interpretation. As illustration, we present an empirical study of the IBM 1993 tick-by-tick data. We find some evidence that volatility of IBM stock prices Granger-causes intratrade durations. We also find that the persistence in GARCH drops dramatically once intratrade durations are taken into account.

1 Introduction

The autoregressive conditional heteroskedastic (ARCH) class of models introduced by Engle (1982) is widely used to capture the temporal dynamics of asset-return volatility. Several recent papers, including Bera and

© 1998 by the Massachusetts Institute of Technology

Higgins (1995), Bollerslev, Chou, and Kroner (1992), Bollerslev, Engle, and Nelson (1994), Diebold and Lopez (1995), and Palm (1996) survey the extensive number of applications to financial time series and the impressive theoretical developments that took place over the last decade. In their generic form, ARCH processes model the conditional variance as a measurable function of past returns. The volatility dynamics is explicitly modeled in discrete time, and while the sampling frequency is in general left unspecified, data are assumed to be equally spaced in time. This setup is adequate for the many applications involving financial time series sampled weekly, daily, or intradaily, at hourly or finer time intervals. There is a growing interest in the modeling of financial time series sampled at the transaction-based frequency. Such data became more widely available over the last years, owing to the implementation of the electronic trading systems on major financial markets.

In this paper we introduce a class of ARCH models for return series sampled at time intervals set by the trade arrivals. Any general formulation of GARCH for irregularly spaced financial data is extremely complex. We propose a class of processes that provides an approximation to the analytically and computationally intractable general case. Our approach combines insights from the temporal aggregation procedure for GARCH models proposed by Drost and Nijman (1993) and Drost and Werker (1996), and the autoregressive conditional duration model (ACD) of Engle and Russell (1996) for intratrade duration sequences. The class of models introduced here will be called ACD-GARCH, although strictly speaking they do not belong to the GARCH class of models. Indeed, the ACD-GARCH models are intrinsically bivariate, as they involve past returns and the time intervals between past transactions.[1] They can also be viewed as a time-deformed GARCH model, and therefore there is some legitimacy in calling it GARCH. An ACD-GARCH model is in fact a random-coefficient GARCH model, or doubly stochastic GARCH, where the duration between transactions determines the dynamics of the entire parameter vector. The parameter behavior is described by the temporal aggregation formulas of weak GARCH. In this sense, our specification differs from the recently proposed ISAR-ARCH model of Pai and Polasek (1995) where the unequal spacing in data is accommodated by ad hoc varying autoregressive coefficients in the conditional mean of returns and the coefficients on lagged-squared errors in the conditional variance equation.

Since the ACD-GARCH model is genuinely bivariate, past volatilities may determine the transaction durations, and vice-versa. In this context, causal relationships between volatility and duration arise as an interesting issue to investigate. In the univariate ACD model discussed by Engle and Russell (1996), the spacings between trades is not a function of the volatility process. We propose GMM as an estimation method applicable under general conditions. A sequential procedure can be adopted when we assume that the spacing of transactions in time is weakly exogenous in the sense of Engle, Hendry, and Richard (1983). In our empirical illustration we use tick-by-tick data on IBM stock extracted from the ISSM (Institute for the Study of Security Markets) 1993 data set. We estimate ACD-GARCH models and introduce tests for causality between IBM stock volatility and intratrade durations. We find such causality to be significant.

The paper is organized as follows. The general structure is presented in Section 2. Next, we discuss the estimation methods in Section 3. The empirical study IBM tick-by-tick data is presented in Section 4. Conclusions appear in Section 5.

2 The ACD-GARCH Class of Models

We consider asset prices recorded on a real-time basis. A simple example would be transaction prices of a particular stock where each transaction has a time stamp. We assume that observations are represented by the pairs $(y_{t_i}, t_i)_{i=0}^{N}$ where $y_{t_i} = \ln p_{t_i}$, the log of the transaction price recorded at time t_i. The data are unequally

[1] Hence, the conditional volatility is not exclusively a measurable function of past returns, a characteristic attributed to the definition of the class of ARCH models; see e.g., Bollerslev, Engle, and Nelson (1994).

spaced and tied to the point process of transaction events. In many cases, the arrival times are modeled as a Poisson process, possibly nonhomogeneous, with the intensity determined by a set of covariates.[2] Engle and Russell (1996) argued that the class of Poisson processes does not well fit the financial tick-by-tick data, as the occurrence of market transactions shows a strong temporal dependence. They suggested an autoregressive conditional duration model, henceforth denoted ACD, which is very similar to ARCH, to accommodate the persistence in intratrade durations. We will focus our attention almost exclusively on the sequential trading process, taking each tick into account. The framework we develop, however, is generic with respect to the transaction event. For instance, the ticks may be restricted to large block trades instead of all trades, or may be the times that elapse before a predetermined number of shares is traded (say 10,000 shares).

In the subsection, we present the structure of the models. Next we provide a time-deformation interpretation of ACD-GARCH models. The third subsection covers some special cases, and the final subsection discusses causality between trading and volatility.

2.1 The Model Structure

Following the approach of Engle and Russell, we call a *duration* the difference between subsequent transaction times $t_{i+1} - t_i$, $i = 0, \ldots, N$. The conditional expected intratrade time $\psi_{i+1} = E[(t_{i+1} - t_i) \mid (t_j - t_{j-1})_{j \le i}]$ can be represented as a measurable function of past durations. In analogy to GARCH(1,1):

$$\psi_{i+1} = \omega^d + \alpha^d (t_i - t_{i-1}) + \beta^d \psi_i, \tag{1}$$

where the parameter vector $\theta^d = (\omega^d, \alpha^d, \beta^d)$ is indexed by d to indicate that the parameters pertain to the duration dynamics. Dividing the durations by their conditional means yields a sequence of i.i.d. variables:

$$(t_{i+1} - t_i)/\psi_{i+1} \sim g(\theta^d), \quad i = 0, \ldots, N, \tag{2}$$

where g denotes the assumed exponential distribution function; see Engle and Russell (1996). Eventually, the conditional mean-duration statement, Equation 1, can be extended from ACD(1,1) to a more general setup:

$$\begin{aligned} &E[(t_{i+1} - t_i) \mid (t_j - t_{j-1})_{j \le i}, (y_{t_j} - y_{t_{j-1}})_{j \le i}, (x_{t_j})_{j \le i}] = \psi_{i+1} \\ &\equiv \psi[(t_j - t_{j-1})_{j=i}^{i-p+1}, (\psi_j)_{j=i}^{i-q+1}, (y_{t_j} - y_{t_{j-1}})_{j=i}^{i-r+1}, (x_{t_j})_{j=i}^{i-s+1}; \theta^d], \end{aligned} \tag{3}$$

where x_{t_i} represents a state-variable process which may include deterministic components or other variables such as the trading volume of the transactions occurring at or before t_i. Whenever past returns or especially past volatilities appear in the above equation, we allow other market-dependent variables to affect the trading process beyond the dynamics captured by the autoregressive structure in Equation 1. This suggests, of course, testing for causality—an issue that will be discussed later.

Consider now the asset-price process $\{y_{t_i}\}$, $i = 0, \ldots, N$. The return series is defined as a sequence of differences $\{(y_{t_i} - y_{t_{i-1}})\}$, indexed by the transaction numbers $i = 1, \ldots, N$. Note that the current return value is expected to be observed over the period ψ_{i+1} before the next trade arrives.

Ultimately, we would like to construct the expected volatility at the next tick (or large block trade or sale of the next 10,000 shares, depending on the definition of events); more precisely:

$$\bar{\sigma}_{i+1}^2 \equiv E[(y_{t_{i+1}} - y_{t_i})^2 \mid (t_j - t_{j-1})_{j \le i}, (y_{t_j} - y_{t_{j-1}})_{j \le i}, (x_{t_j})_{j \le i}],$$

where we assume that $(y_{i+1} - y_i)$ has conditional mean zero. This will be difficult to accomplish, however, if we aim for analytic expressions and stay within the class of ARCH models. Indeed, the ARCH class of models,

[2] This is the so-called proportional hazard model. See, e.g., Lancaster (1990) for a detailed discussion. Gouriéroux, Jasiak, and Le Fol (1996) propose nonparametric methods for hazard models and their associated duration and survival functions applied to French tick-by-tick data from the *Bourse de Paris* (Paris Stock Exchange).

as originally defined by Engle (1982), does not temporally aggregate. Since we deal with unequally spaced data, temporal aggregation is crucial. Hence, a first step is to weaken the definition of ARCH such that temporal aggregation is possible. This was developed by Drost and Nijman (1993). In particular, we assume that the first differences $\{(y_{t_i} - y_{t_{i-1}}), i = 1, \ldots, N\}$ have finite fourth moments and form a stationary sequence satisfying the weak GARCH(1,1) definition; see Drost and Nijman (1993). Especially we assume that $\{(y_{t_i} - y_{t_{i-1}}), i = 1, \ldots, N\}$ consists of selected observations drawn from an underlying data-generating process evolving on a time grid set by equal, very small time increments.[3] Hence, for the moment we will remain in a discrete time setting that will cause us difficulty. Some of this difficulty will be solved by adopting a continuous-time framework, to which we turn our attention later in this section. Drost and Nijman (1993) have shown that the class of weak GARCH(1,1) processes is closed under temporal aggregation. We will rely on this result to model the dynamics of the volatility process over unequal intratrade intervals.

Let us denote $\varepsilon_i = y_{t_i} - y_{t_{i-1}}$, and fix the next trade $t_{i+1} = t_i + \delta$; i.e., when time is measured in seconds, the intratrade duration will be δ. For any given δ, the weak GARCH(1,1) models provide the best linear projection:

$$\sigma^2_{i+1}(\delta) \equiv P_\delta[\varepsilon^2_{i+1} \mid 1, (\varepsilon_j)_{j \le i}, (\varepsilon^2_j)_{j \le i}], \tag{4}$$

where P_δ denotes the best linear projection. The temporal aggregation formula of Drost and Nijman (1993), which will be presented shortly, enables us to compute this linear projection for any given δ. There is an important technical issue emerging here that we cannot accommodate properly in this framework. Indeed, the temporal aggregation formula provides a mapping between equally spaced representations with different sampling frequencies, and does not, strictly speaking, apply to unequally spaced data. We will ignore this issue until we adopt a continuous-time framework later in this section. The continuous-time framework is more suitable to deal with this, but for the moment we will remain with the discrete-time setup. It is worth noting that P_δ is only a function of past returns, a feature which is proper to ARCH models. At this stage, however, we can introduce information about intratrade duration, because the ACD model provides us with a conditional duration distribution for δ. The ACD model allows us to compute:

$$\tilde{\sigma}^2_{i+1} = \int \sigma^2_{i+1}(\delta)\tilde{g}(\delta; \psi_{i+1}, \theta^d)\, d\delta, \tag{5}$$

which is the expected conditional volatility, based on the distribution function appearing in Equation 2 with conditional mean ψ_{i+1}. While the integration in Equation 5 is in principle feasible, it is analytically intractable and can only be accomplished numerically.[4] To sidestep this difficulty, we propose to use $\sigma^2_{i+1}(\psi_{i+1})$, which yields the structure for the class of ACD-GARCH models we introduce in this paper. Using the temporal aggregation formula of Drost and Nijman (1993) we can write the ACD-GARCH class of models more explicitly as:

$$\sigma^2_{i+1} = \omega_i + \alpha_i\ \varepsilon^2_i + \beta_i\ \sigma^2_i \tag{6}$$

where:

$$\omega_i = \psi_{i+1}\omega\frac{1 + (\alpha + \beta)^{\psi_{i+1}}}{1 - (\alpha + \beta)}, \tag{7}$$

$$\alpha_i = (\alpha + \beta)^{\psi_{i+1}} - \beta_i, \tag{8}$$

[3] In practice, we will consider as a very small time unit the accuracy limit of the records, which is usually 1 s.

[4] Particularly, keep in mind that this numerical integration has to be performed possibly at each transaction tick involving a different conditional duration distribution, which makes this option less appealing.

$$\beta_i/(1+\beta_i^2) = [a(\alpha,\kappa,\psi_{i+1})(\alpha+\beta)^{\psi_{i+1}} - b(\alpha,\beta,\psi_{i+1})] \\ /[a(\alpha,\kappa,\psi_{i+1})\{1+(\alpha+\beta)^{2\psi_{i+1}}\} - 2b(\alpha,\beta,\psi_{i+1})], \tag{9}$$

$$a(\alpha,\beta,\kappa,\psi_{i+1}) = \psi_{i+1}(1-\beta)^2 + 2\psi_{i+1}(\psi_{i+1}-1)(1-\alpha-\beta)^2 \\ \times \{1-(\alpha+\beta)^2+\alpha^2\}\ /\ [(\kappa-1)(1-(\alpha+\beta)^2)] \\ + 4c(\alpha,\beta,\psi_{i+1})/(1-(\alpha+\beta)^2), \tag{10}$$

$$b(\alpha,\beta,\psi_{i+1}) = \{\alpha(1-(\alpha+\beta)^2)+\alpha^2(\alpha+\beta)\}(1-(\alpha+\beta)^{2\psi_{i+1}}) \\ /(1-(\alpha+\beta)^2), \tag{11}$$

$$c(\alpha,\beta,\psi_{i+1}) = \{\psi_{i+1}(1-\alpha-\beta)-1+(\alpha+\beta)^{\psi_{i+1}}\} \\ \times \{\alpha(1-(\alpha+\beta)^2)-\alpha^2(\alpha+\beta)\}. \tag{12}$$

The constant parameters ω, α, and β in Equations 7–12 characterize the conditional volatility of the underlying process, which is defined at the smallest time interval. The parameter k in Equations 9 and 10 denotes the kurtosis. For the purpose of parameter identification, we will also use the relationship $\alpha_{\psi_{i+1}} + \beta_{\psi_{i+1}} = (\alpha+\beta)^{\psi_{i+1}}$.

When we define an error term $\eta_i = \varepsilon_i^2 - \sigma_i^2$ such that:

$$\varepsilon_i^2 = \omega_i + (\alpha_i+\beta_i)\varepsilon_{i-1}^2 + \eta_i + \beta_{i+1}\,\eta_{i-1}, \tag{13}$$

we obtain an alternative representation for the model. It reveals that there are two sources of error in the η_i process; one related to the innovation in the duration process $(t_{i+1}-t_i)-\psi_{i+1}$, and the other related to the usual GARCH volatility innovation. This mixture of errors suggests that the η_i process is most likely fat-tailed.

Finally, the ACD-GARCH can be written as a system of equations:

$$\psi_{i+1} = \omega^d + \alpha^d(t_i - t_{i-1}) + \beta^d\psi_i, \text{ and} \tag{14}$$

$$\sigma_i^2 = \omega_i + \alpha_i\varepsilon_{i-1}^2 + \beta_i\sigma_{i-1}^2. \tag{15}$$

2.2 Time Deformation

As noted in the Introduction, the ACD-GARCH model does not satisfy the weak GARCH or the ARCH definitions, in the sense that it does not represent the conditional variance as a measurable function defined on the filtration of past returns. In some way the ACD-GARCH model is, in fact, a stochastic volatility model. It was also noted that the GARCH aggregation formula applies in principle to equally spaced data with different sampling frequencies; this is a technical issue that we ignored when dealing with irregularly spaced series. In this section, we address some of these issues by noting that the class of ACD-GARCH models could also be viewed as a time-deformed GARCH diffusion.[5] Following Drost and Werker (1996), one can define a diffusion:

$$dy(t) = \sigma(t)dW_1(t), \text{ and} \tag{16}$$

$$d\sigma(t)^2 = \delta(\omega - \sigma(t)^2)dt + \sqrt{2\lambda\delta}\,\sigma(t)^2 dW_2(t), \tag{17}$$

[5]For more details on time deformation and its applications to financial data, see Ghysels and Jasiak (1996) and Ghysels, Gouriéroux, and Jasiak (1998a, 1998b).

where $W_1(t)$ and $W_2(t)$ are two independent Brownian motions with $\omega > 0$, $\delta > 0$, and $\lambda \epsilon (0, 1)$. Using the results of Drost and Werker (1996), one interprets the ACD-GARCH model as a time-deformed GARCH diffusion with time deformation ψ_{i+1}, yielding:

$$\omega_i = \psi_{i+1}\, \omega\{1 - \exp(-\psi_{i+1}\delta)\}, \tag{18}$$

$$\alpha_i = \exp(-\psi_{i+1}\delta) - \beta_i, \tag{19}$$

$$\begin{aligned} \beta_i = (\beta_i^2 + 1)(c_i)(\exp(-\psi_{i+1}\delta) - 1)/ \\ [c_i\{1 + \exp(-2\psi_{i+1}\delta)\} - 2], \end{aligned} \tag{20}$$

where

$$\begin{aligned} c_i \quad = \quad & [4\{\exp(-\psi_{i+1}\delta) - 1 + \psi_{i+1}\delta\} + 2\psi_{i+1}\delta\{1 + \psi_{i+1}\delta(1 - \lambda)/\lambda\}] \\ & /[1 - \exp(-2\psi_{i+1}\delta)]. \end{aligned} \tag{21}$$

Equations 18–21 yield the ACD-GARCH appearing in Equation 6 as an *exact* discretization of the diffusion (Equations 16 and 17) with the ACD conditional mean φ_{i+1} as the directing process. The time deformation is purely exogenous to volatility when we restrict our attention to pure ACD models. However, in the more general case where past volatility affects expected trade durations, we may have a partially endogenous directing process.[6] This time-deformation interpretation clarifies, as noted earlier, that we are in fact dealing with a class of stochastic volatility models.

2.3 Special Cases

The ACD-GARCH provides a convenient framework encompassing some well-known models. We discuss two special cases that arise when additional restrictions are imposed on the duration dynamics.

The first is the equally spaced traditional GARCH(1,1) model, which is obtained by setting $\psi_{i+1} \equiv \psi \ \forall i$. It applies to *unequally* spaced data with a constant conditional duration distribution. Such a situation would occur when the point process governing transactions is in fact a time-homogeneous Poisson process. We noted in the introduction that the event process is not exclusively tied to every transaction tick. For instance, the intratrade durations may reflect the occurrence of large block trades or other specific trades. If these arrive at a Poisson rate, we can simply use standard GARCH processes to predict the volatility of returns despite the fact that such trades are irregularly spaced.

Another special case worth mentioning is the situation where ψ_i follows a deterministic periodic pattern, i.e., $\psi_i = \psi_{i-S}$ for $\forall i$ and given S. Then, one obtains a periodic GARCH model, introduced by Bollerslev and Ghysels (1996), for unequally spaced data. Such a situation occurs if transactions arrive as a Poisson process that is time nonhomogeneous with a periodic pattern in the hazard rates.

2.4 Causality

An interesting issue to investigate in the context of ACD-GARCH is the causality from the volatility process to the intratrade duration sequence. More explicitly, the question is whether past returns or past volatilities should appear on the right-hand side of Equation 3.

There are various ways to test the exclusion restrictions associated with the absence of returns and/or volatility in the intratrade duration process. The most appealing, though not the most optimal, is to estimate

[6] This distinction has consequences with respect to estimation. If we were to estimate directly the diffusion instead of the discrete-time weak GARCH, we could in principle use the methods proposed by Conley et al. (1997) when the directing process is exogenous. Otherwise, one needs to adopt the methods discussed in Ghysels and Jasiak (1996).

the ACD model without any consideration for volatility, i.e., the univariate process of intratrade durations as in Engle and Russell (1996). The ACD-GARCH is estimated next using the return data, and tests are applied to the bivariate process of residuals generated by the two univariate models. Such a procedure is quite similar to recently proposed causality tests in conditional variances across various financial asset-price movements. Cheung and Ng (1996) developed such a procedure to test causality in conditional variances based on the early work by Haugh (1976) and McLeod and Li (1983), based on estimating the univariate ARCH models and using the cross-correlation function (CCF) of the squared standardized residuals. In theory, two issues can be investigated; namely, the traditional causality between the conditional mean of returns and the intratrade durations, as well as the causality between the volatility and durations, i.e., the conditional variance-mean causality.

The first problem can be addressed by applying the CCF-based Pierce-Haugh test to the standardized residuals, i.e., the i.i.d. sequence d_i and the return innovation process ε_i:

$$d_i = (t_{i+1} - t_1)/\psi_{i+1}, \text{ and} \tag{22}$$

$$e_i = (y_i - y_{i-1})/\sigma_i. \tag{23}$$

For independent sequences $\{d_i\}$ and $\{e_i\}$, the existence of their second moments implies that:

$$\begin{pmatrix} \sqrt{N}\hat{c}_{de}(k) \\ \sqrt{N}\hat{c}_{de}(k') \end{pmatrix} \rightarrow AN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right), k \neq k', \tag{24}$$

where $\hat{c}_{de}(k)$ is the sample cross-correlation at lag k,

$$\hat{c}_{de}(k) = c\hat{o}v_{de}(k)/\sqrt{c\hat{o}v_{dd}(0)c\hat{o}v_{ee}(0)}, \tag{25}$$

and $c\hat{o}v_{de}(k)$ is the k^{th} lag sample cross-covariance; see Hannan (1970), Cheung and Ng (1996):

$$c\hat{o}v_d(k) = \quad N^{-1}\sum(d_i - \bar{d})(e_{i-k}).$$

Causality at lag k can be tested by comparing $\sqrt{N}\hat{c}_{de}(k)$ to a critical value from the standard normal distribution. Alternatively, a chi-square test statistic based on:

$$S_k = N\sum_{i=j}^{k}\hat{c}_{de}(i)^2 \tag{26}$$

can be computed. It is chi-square distributed with $(k - j + 1)$ degrees of freedom, and can be used to test the null hypothesis of no causality from lag j to lag k.

The test of causality between the volatility process and the intratrade duration involves the conditional mean of durations and the conditional variance of returns. Hence the k^{th} lag sample cross-correlation of the squared standardized residuals e_i^2, denoted by E_i and d_i, are computed:

$$\hat{c}_{dE}(k) = c\hat{o}v_{dE}(k)/\sqrt{c\hat{o}v_{dd}(0)c\hat{o}v_{EE}(0)}, \tag{27}$$

where

$$c\hat{o}v_{dE}(k) = N^{-1}\sum(d_i - \bar{d})(E_{i-k} - \bar{E}).$$

For d_i and e_i^2 defined by Equations 22 and 23, we show in the Appendix that under suitable regularity conditions, $\sqrt{N}(\hat{c}_{dE}(k_1), \ldots, \hat{c}_{dE}(k_m))$ converges to $N(0, I_m)$ as $N \rightarrow \infty$, where $k_1, \ldots, k_m$ are m different integers.[7]

[7] For the above results to hold, the conditional mean of the returns has to be included in the model and consistently estimated. We need to remain cautious, however, as the eventual causality in mean violates the independence condition and affects the size of the mean-variance causality test [see Cheung and Ng (1996)].

It should be recalled that the residual-based causality tests are conservative tests, because they are constructed from univariate time-series processes. In the context of high-frequency data, this negative feature of such tests is probably not of great concern. More elaborate tests based on the genuine bivariate representation of the trade and volatility processes can be considered as well. They will be briefly covered in the next section. To conclude, we stress the importance of causality testing. Indeed, when trading and volatility are not mutually exogenous, we really have to deal with the joint process of d_i and e_i^2. Most microstructure models suggest such a bivariate setup; see, e.g., O'Hara (1995) for a survey of the literature.

3 Estimation

It is assumed that the econometrician has a high-frequency data set of financial time series consisting of observations $((y_{t_i}, t_i, x_{t_i})_{i=0}^{N}, Y_o, T_o, X_o)$ where Y_o, T_o, and X_o are sets of initial conditions of y_{t_i}, t_i, and x_{t_i}, which contain the p, r, s lagged presample observations (see Equation 3). As noted in Section 2, the x_{t_i} process may include deterministic components as well as other variables included in the ACD-GARCH model, with trading volume or past returns being particularly relevant examples.

It would be tempting to formulate a maximum-likelihood estimator (henceforth MLE) for ACD-GARCH models, since taken separately, both ARCH and ACD models are typically estimated using MLE. However, in estimating the two components simultaneously, we encounter a major obstacle to maintaining a fully specified distributional framework. Indeed, Drost and Nijman (1993) argued that the temporal aggregation results hold for a weaker-than-standard (strong) definition of ARCH, and this requires an estimation via a set of moment conditions or via quasi-MLE, or QMLE. The necessary moment conditions are provided by the projection equations that define weak GARCH models. Alternatively, the theoretical likelihood of GARCH could serve as a distributional form for the QMLE. Under certain circumstances, related to the causality issues discussed in Section 2.3, we could even consider maintaining an ML-estimation approach for the ACD model. Indeed, a two-step procedure can be considered, starting first with the estimation of the ACD parameters, followed by the estimation of the conditional variance model for the tick-by-tick data. In an MLE framework, such a two-step procedure would amount to a situation where the spacing of transactions in time is weakly exogenous, in the sense of Engle, Hendry, and Richard (1983). We first cover the general joint-estimation framework in Section 3.1, and discuss the two-step procedures in Section 3.2. In Section 3.1 we focus on the GMM estimation and show a particular GMM estimator having a QMLE interpretation.

3.1 Joint ACD and GARCH Estimation

We assume that the parameter vector θ has two components: θ^d, a subvector describing the parameters of the ACD model, and θ^r, a subvector that contains the parameters of the conditional mean and variance of tick-by-tick returns. This setup is the same as in Section 2.3. When we assume that x_{t_i} contains at least a constant, then the dimension of θ^d equals $p + q + r + s$, using the notation of Equation 4 with $s \geq 1$. The vector θ^r can also be divided into two components: μ^r, which governs the conditional mean, and may contain as components the constant drift or seasonal dummies; and the vector $\zeta \equiv (\omega, \alpha, \beta)$, which determines Equations 7–12. To set the scene for the discussion of the estimation procedure, we consider a generic GMM setup based on a set of moment conditions:

$$Ef_i(\theta) = 0, \tag{28}$$

where $f_i(\theta) \equiv (f_i^d(\theta)', f_i^r(\theta)')'$,

$$f_i^d(\theta) \equiv f^d((t_j - t_{j-1})_{j=i+1}^{i-p}, (\psi_j)_{j=i}^{i-q}, (y_{t_j} - y_{t_{j-1}})_{j=i}^{i-r}, (x_{t_j})_{j=i}^{i-s}; \theta), \text{ and} \tag{29}$$

$$f_i^r(\theta) \equiv f^r((t_j - t_{j-1})_{j=i}^{i-p+1}, (\psi)_{j=i}^{i-q+1}, (y_{t_j} - y_{t_{j-1}})_{j=i+1}^{i-r+1}, (x_{t_j})_{j=i}^{i-s+1}; \theta). \tag{30}$$

It should be noted that $t_{i+1} - t_i$ enters $f_i^d(\theta)$ while $y_{t_{i+1}} - y_{t_i}$ enters $f_i^r(\theta)$. Hence, the first set of moment conditions pertains to the durations between transactions, while the second set of moment conditions draws on projection equations of the conditional mean and variance of returns. Moreover, within this general framework the entire parameter vector θ enters both sets of moment conditions. Hence, as we discussed in Section 2.3, past volatility, σ_i, could possibly help to determine expected future durations such as ψ_{i+1}. This can be accommodated in Equation 3 by letting $x_{t_i} \equiv (1, \sigma_i)$.[8] In the sequel we consider this general framework. Later, in Section 3.2, we examine the more restrictive setting where $f_i^d(\theta^d)$ does not involve θ^r. While in principle a wide variety of choices for f_i^d and f_i^r exist, there is a rather natural choice that emerges from the discussion in Section 2. Although we cannot consider MLE for the joint duration/conditional volatility model, it would be logical to start with distributional properties of the normalized durations, i.e., $(t_i - t_{i-1})/\psi_i$ in Equation 2, to construct $f_i^d(\theta)$. It has the advantage that the two-step procedures examined in the next section are special cases of the procedures considered here. For the exponential version of the ACD model, the log-likelihood function can be written:

$$L^E(\theta) \equiv -\sum_{i=1}^{N} l_i^d(\theta) \equiv -\sum_{i=1}^{N} [\ln(\psi_i) + (t_i - t_{i-1})/\psi_i],$$

resulting in the score moment condition:

$$f_i^d(\theta) = \partial l_{i+1}^d(\theta)/\partial\theta^d = [(t_{i+1} - t_i)/\psi_{i+1}(\theta) - 1]/\psi_{i+1}(\theta)\, \partial\psi_{i+1}(\theta)/\partial\theta^d. \tag{31}$$

The score function of the ACD model, required in this approach, is its derivative with respect to θ^d. The dependence on the entire vector θ is revealed by making ψ_i an explicit function of θ. The computation of this moment condition is relatively straightforward. As indicated by Engle and Russell, it involves recursive computations of $\partial\psi_i/\partial\theta^d$, which are quite similar to those encountered in score functions of GARCH models. Engle and Russell also consider a Weibull-distributed ACD requiring an appropriate expression for $\partial l_{i+1}^d(\theta)/\partial\theta^d$.[9]

We turn our attention now to the second set of moment conditions, $f_i^r(\theta)$. There are two particular choices for $f_i^r(\theta)$ that would naturally emerge from the discussion in Section 2. The first choice is computationally simple, while the second is more involved but has the advantage of yielding a QMLE interpretation of the resulting GMM estimator. The moment conditions for the first approach are obtained from the projection equations resulting from Equations 4 and 6. In particular, consider the instrument set $Z_{mi} \equiv [1, (t_j - t_{j-1})_{j=i}^{i-p+1}, (y_{t_j} - y_{t_{j-1}})_{j=1}^{i-r+1}, (x_{t_j})_{j=i}^{i-r+1}]$. Then for ε_{i+1} defined in Equation 4 we have:

$$E\varepsilon_{i+1}Z_{mi}' = 0. \tag{32}$$

For the conditional variance, we can use:

$$E[\varepsilon_{i+1}^2 - \omega_{i+1} - (\alpha_{i+1} + \beta_{i+1})\varepsilon_i^2]Z_{Vi-1}' = 0, \tag{33}$$

where $Z_{Vi} \equiv [1, \varepsilon_i, \ldots, \varepsilon_{i-l+1}, \varepsilon_i^2, \ldots, \varepsilon_{i-k+1}^2]$. The combination of Equations 32 and 33 yields:

$$f_i^r(\theta) \equiv ((\varepsilon_{i+1}Z_{mi}')', ([\varepsilon_{i+1} - \omega_{i+1} - (\alpha_{i+1} + \beta_{i+1})\varepsilon_i^2]Z_{Vi-1}')')', \tag{34}$$

[8] Needless to say, such a situation could become numerically quite involved. Indeed, with $q = p = 1$, $r = 0$, and $x_{t_i} \equiv (1, \sigma_i)$ in Equation 3, it would require recursive calculations of ψ_{i+1} given σ_i, and σ_{i+1} given ψ_{i+1} and σ_i starting with σ_0.

[9] In fact, we could entertain the possibility of modeling normalized durations via SNP density [see Gallant and Tauchen (1989)], and use the SNP score in (3.4).

which amounts to $r + p + s + k + l + 2$ moment restrictions. The sample equivalent of Equation 28 forms the basis of the following GMM estimation procedure:

$$\hat{\theta}_N^G \equiv \underset{\theta}{Arg\min} \left(N^{-1} \sum_{i=1}^{N} f_i(\theta) \right)' \Omega_N^{-1} \left(N^{-1} \sum_{i=1}^{N} f_i(\theta) \right), \tag{35}$$

where Ω_N^{-1} is the optimal weighting matrix. Hansen (1982) showed that the matrix should be estimated from:

$$\Omega \equiv E \begin{bmatrix} f_i^d(\theta)' f_i^d(\theta) & f_i^d(\theta)' f_i^r(\theta) \\ f_i^r(\theta)' f_i^d(\theta) & f_i^r(\theta)' f_i^r(\theta) \end{bmatrix}. \tag{36}$$

Equation 36 emphasizes the fact that cross-products between $f_i^d(\theta)$ and $f_i^r(\theta)$ are involved in the computation of the optimal GMM estimator defined in Equation 35. A priori, we do not expect these cross-products to be zero. Indeed, it was shown in Equation 13 that the error process η_i is related to the innovation in the duration process error $(t_{i+1} - t_i) - \psi_{i+1}$, which makes the first element of $f_i^r(\theta)$ nonorthogonal to $f_i^d(\theta)$. It is important to emphasize this feature at this stage, as the two-step procedure presented in the next section yields an estimator inefficient relative to $\hat{\theta}_N^G$ because the off-diagonal elements in Equation 36 are set equal to zero.

We denoted the GMM estimator in Equation 35 by $\hat{\theta}_N^G$ to distinguish it from the second GMM estimator, denoted $\hat{\theta}_N^Q$, where the index Q refers to a QMLE interpretation of the estimator. We mentioned before that maintaining a fully specified distribution of normalized returns, i.e., $(y_{t_i} - y_{t_{i-1}})/\sigma_i \equiv Z_i$, is not compatible with the temporal aggregation condition, and therefore MLE of ACD-GARCH is ruled out. Yet, as stressed by Drost and Nijman, it is possible to provide a QMLE interpretation to the usual ARCH-type MLE procedures. This suggests that we construct the moment condition from a score function using a QMLE argument for θ^r. Indeed, the efficiency of $\hat{\theta}_N^G$ could be improved by using instruments other than Z_{mi} and Z_{Vi}. In fact, Meddahi and Renault (1995) examine QMLE estimation of weak GARCH processes and provide an optimal instrument GMM interpretation. We replace $f_i^r(\theta)$ specified in Equation 34 by:

$$f_i^r(\theta) = \partial l_{i+1}^r(\theta)/\partial \theta^r, \tag{37}$$

where the quasi-log likelihood function l_i^r for return $y_{t_i} - y_{t_{i-1}}$, assuming Z_i is i.i.d. $N(0, 1)$ for all i, is:

$$\begin{aligned} l_i^r(\theta) = (2\pi)^{-1/2} \exp\{-0.5[(y_{t_i} - y_{t_{i-1}})/\sigma_i(\theta)]^2\} \\ - \ln|J_i(\theta)| - 0.5 \ \ln(\sigma_i^2(\theta)), \end{aligned} \tag{38}$$

with $J_i(\theta) = \partial \sigma_i(\theta)/\partial \psi_{i+1} \mid \psi_{i+1}$. Since the score function in Equation 37 involves the derivative with respect to θ^r, we obtain an exactly identified system of equations when the scores $\partial l_i^d(\theta)/\partial \theta^d$ and $\partial l_i^r(\theta)/\partial \theta^r$ are stacked. This results in a GMM estimator that is independent of its weighting matrix and therefore yields a QMLE interpretation to the joint duration/volatility estimator $\hat{\theta}_N^Q$. The fact that we need to compute the score of Equation 38 may be computationally more intensive compared to the GMM estimator based on the projection equations defined in Equations 32 and 33. Hence there is some advantage in using $\hat{\theta}_N^G$ instead of $\hat{\theta}_N^Q$ in terms of computer time.

Several concluding comments can be made before turning our attention to the two-step procedure. Consider first the asymptotic distribution. Under suitable regularity conditions, which may or may not be easy to verify or establish [see Bollerslev, Engle, and Nelson (1994) for further discussion and references], the estimators have standard normal asymptotic distributions [see, e.g., Hansen (1982)]. It is worth noting that the MLE of Engle and Russell will be inconsistent, even under proper distributional assumptions when $\partial \psi_i(\theta)/\partial \theta^r \neq 0$, as is the case with $x_{t_i} \equiv (1, \sigma_i)$, since it would require joint estimation of transaction durations and volatilities. It was already noted that in principle, Equation 35 may involve over-identifying restrictions under the $\hat{\theta}_N^G$ specification when $\dim(\theta) < 2(1 + p + r + s) + k + l + q$. In such cases, one must choose an estimator for $\hat{\Omega}_N$, an issue covered in detail by Newey and West (1987) and Andrews and Monahan (1992), among others. Finally, a standard J-statistic could be applied as a model diagnostic.

Table 1
Hourly Averages of Durations[a]

Time	Mean
9:00–10:00	33.0949
10:00–11:00	34.3788
11:00–12:00	36.2974
12:00–13:00	39.3115
13:00–14:00	34.9770
14:00–15:00	33.4801
15:00–16:00	38.5714

[a]These intradaily seasonal hourly components were removed by using splines to compute the deterministic means conditioned on the time of day and dividing each observation by this value [see Engle and Russell (1996)]. No daily or weekly seasonal components were detected.

3.2 Two-Step Estimators

Engle, Hendry, and Richard (1983) drew attention to properties of likelihood functions which allow for separate estimation of θ^d and θ^r. As noted in the previous section, we do not really operate in the context of MLE to straightforwardly apply the concepts of weak and strong exogeneity introduced by Engle, Hendry, and Richard (1983). Yet it is only the weakening of the GARCH structure that prevents us from using MLE. Hence, in principle, we can focus on the properties of the duration process between transactions, and discuss its ML estimation separately. This raises the possibility of introducing two-step procedures.

If we assume that the spacing between transactions is weakly exogenous, in the sense of Engle, Hendry, and Richard, with respect to the return process, then the parameter vector θ^d can be estimated first and separately. This amounts to assuming that the score function $\partial l_{i+1}^d(\theta)/\partial\theta^d$ is independent of θ^r, or, put differently, $l_{i+1}^d(\theta)$ is a function of θ^d only. With this restriction, we can estimate the ACD-GARCH model in two steps. First, we estimate the ACD model following the procedure described in Engle and Russell (1996). Under the correctly specified distributional assumptions, this estimator will in fact be the most efficient by virtue of MLE. This estimator, denoted $\hat{\theta}_1^d$, can be used to construct an expected duration sequence $\psi_i(\hat{\theta}_1^d)$ using Equation 1 in the ACD(1,1) case, or a more general recursion formula for ACDX(p, q) models. The sequence of expected durations, $\psi_i(\hat{\theta}_1^d)$, can be used to estimate the weak GARCH model. As noted in the previous section, one may choose a different instrument setup, either using projection equation arguments as in Equation 34 or a score function with a Gaussian QMLE specification as in Equation 37. Other possibilities could be considered as well, of course. The former will be denoted $\hat{\theta}_{NG}^r$, while the latter will be denoted $\hat{\theta}_{NQ}^r$.

4 Empirical Illustration: IBM Stock Prices

In this section, we present an application of the ACD-GARCH to intradaily observations on IBM prices and intratrade durations recorded in November 1993. The sample originally consisted of 43,328 observations. Prior to estimation, the durations between the market closure and the next day's opening were deleted. We also removed observations corresponding to the market openings and trades occurring simultaneously at the same price, yielding zero returns at zero-length durations. The sample reduced to 13,038 observations. These data were adjusted for seasonal effects in the duration series and higher-order autocorrelation in the returns. The data exhibited intraday seasonality in the mean durations. Table 1 shows the average durations throughout the eight-hour trading day.

The resulting transformed series has mean 0.9999 and standard deviation 0.4469, i.e., the series features an under-dispersion compared to the exponentially distributed series with $\lambda = 1$. It is also characterized by a

Table 2
Maximum Likelihood Estimates of the ACD (1,1) Model[b]

Parameter	Estimate	Standard Error	p-value
ω^d	0.0028	0.0055	0.3041
α^d	0.0128	0.0064	0.0231
β^d	0.9843	0.0089	0.0000

[b]We note that the sum of the α^d and β^d parameters is close to unity, revealing the possible presence of a long memory pattern. [The issue of long memory in intratrade durations is discussed in Ghysels and Jasiak (1996), where supporting evidence for the long-memory hypothesis is, in fact, found.] The residuals of the ACD(1,1) feature a high degree of autocorrelation [Q(20) = 169.5615], and the null hypothesis of a white-noise process is rejected. To improve the fit, we increase the lag length and estimate an ACD(2,2) model, summarized in Table 3.

Table 3
Maximum Likelihood Estimates of the ACD (2,2) Model[c]

Parameter	Estimate	Standard Error	p-value
ω^d	0.0009	0.0026	0.3689
α_1^d	0.0760	0.0218	0.0002
α_2^d	−0.0710	0.0208	0.0003
β_1^d	1.3805	0.4241	0.0006
β_2^d	−0.3863	0.4187	0.1781

[c]This specification turns out to be more successful in accommodating the persistence in durations. The Portmanteau statistic associated with the residuals is Q(20) = 20.0516, and the white-noise hypothesis cannot be rejected at level 0.05.

significant degree of autocorrelation, as indicated by the value of the Portmanteau statistic Q(20) = 523.5944.

The original return series exhibited some persistence in the conditional mean, and was replaced by residuals of an AR(3) process with mean 0.0000, variance 0.0578, and kurtosis 5.4939. The Portmanteau statistics computed for the levels and squares of the transformed process are Q(20) = 14.3650 and 264.4364, respectively.

The first estimation method is the two-stage approach under the duration-exogeneity assumption. In the first stage, we estimated the ACD(1,1) model by maximum likelihood, based on an exponential density function with intensity parameter $\lambda = 1$. The parameter estimates are presented in Table 2.

The sequence of predicted duration $\{\psi_i\}$ was used in the second step of the estimation procedure to evaluate the time-varying parameters α_i and β_i, appearing in Equations 6–12. We relied on a normal approximation of the density of returns yielding QMLE estimates of the parameters ω and α of the underlying process reported in Table 4. The coefficient β was computed from the relation appearing in Equation 8. Given the large number of observations, we computed the average over all ψ in the sample. In principle, one could obtain standard errors, but as this is quite computationally involved we did not calculate them for β. The duration-dependent parameters α_i and β_i have means 0.1578 and 0.0004. They take values within the intervals (0., 0.2212) and (0., 0.0141), respectively. The expected intratrade durations range from 0.8082 to 2.2995 with a mean 1 by construction. The results in Table 4 are quite interesting, as they reveal that the persistence in volatility has completely disappeared once the intratrade duration persistence is driving the volatility process. Indeed, the well-known IGARCH phenomenon does not appear in the underlying process, as the sum of α and β is very small. Moreover, the random coefficients α_i and β_i also take values indicating

Table 4
Quasi-Maximum Likelihood Estimates of the ACD-GARCH(1,1)

Parameter	Estimate	StandardError	p-value
ω	0.0503	0.0003	0.0000
α	0.0729	0.0083	0.0000
β	0.0838	—	—

Table 5
Cross-Correlations in Levels and Squares of StandardizedResiduals[d]

Lag k	Levels r_{de}	Squares r_{dE}
1	−0.0025	0.0081
2	0.0003	0.0079
3	0.0087	0.0036
4	0.0064	*0.0209
5	−0.0042	0.0019
6	−0.0123	−0.0048
7	−0.0073	0.0035
8	0.0115	−0.0109
9	−0.0082	0.0019
10	0.0077	*0.0198
	Q(10) = 7.9255	Q(10) = 14.8523

[d] $\chi^2(10) = 18.3$. Asterisksindicate significance at the 5% level.

low persistence.[10] The setup of Equations 6 and 7 reveals perhaps why this is the case. Indeed, the drift ω_i depends directly on ψ_{i+1}, i.e., the expected durationwhich absorbs all the persistence. Since intratrade durationstake account of the persistence, it is interesting to investigatewhether they are exogenous in the sense of Granger causality.The series of rescaled ACD-GARCHand ACD(2,2) residuals were used to test causality,as discussed in Section 2.4. We investigatedthe impact of past returns (levels) on the durationsand of past volatilities(squares) on the durationsas well. We proceeded by computing the cross-correlation functions at variouslags k, reported in Table 5.

The first column in Table 5 contains the lags k corresponding to the number of ticks. The second one shows the values of the cross-correlation function between the levels of returns and durations (r_{de}), while the third contains the cross-correlations between past squared returns and durations (r_{dE}). The null hypothesis of independence cannot be rejected at any lag in the levels. We find, however, some evidence confirming the existence of Granger causalitybetween past volatilitiesand durationsat lags 4 and 10 at the 5% significance level. The joint hypothesis of independence at lags 1–10 cannot be rejected at 5% for both (r_{de}) and (r_{dE}).

5 Conclusion

In this paper we introduced ACD-GARCHmodels, which are a class of ARCHmodels for asset-return series sampled at time intervalsset by the trade arrivals.Our approach combines insights from the temporal aggregation procedure for GARCHmodels and ACDfor intratradedurationsequences. Since the ACD-GARCH models are intrinsicallybivariate,we also tested for Granger causal relationshipsbetween the volatilityand intratradedurations.Several interesting results emerged from our analysisof IBM data. First, we find that once the persistence in intratradedurationsis taken into account, the volatilitydynamicsshow very little temporal

[10] One can drawsome comparisonwith findingsin Lamoureux and Lastrapes(1990), which shows that a (persistent) volume series can explain the IGARCHphenomenon as well.

dependence. However, intratrade durations are not exogenous in a Granger causality sense. Indeed, the results based on tick-by-tick data on IBM stock suggest that volatility and trading are interdependent.

We also propose a GMM estimation method applicable under general conditions. A sequential procedure can also be adopted when we assume that the spacing of transactions in time is weakly exogenous. Admittedly, the methods we propose are still too computationally complex to be implemented on a real-time basis in a tick-by-tick trading scheme. As we noted throughout the paper, it is straightforward to consider events other than the next trading tick. One may, for instance, consider the durations between the sale of a total of 10,000 shares and predict price volatility that is associated with this event. Another example is that of large block trades, like the sale of 10,000 shares in one single trade. Both of these events are less frequent, and therefore make the use of ACD-GARCH easier in the context of real-time trading.

Appendix

In principle, the causality from durations to volatility and vice-versa can be investigated. Although in the framework of the ACD-GARCH model we restrict our attention to causality from volatility to intratrade durations, however, we will consider the general case. The demonstration follows arguments also used in Cheung and Ng (1996). We will use the following notation:

$$\frac{[(y_i - y_{i-1})]^2}{\sigma_i^2} = \frac{\Delta y_i^2}{\sigma_i^2} = E_i, \text{ and}$$

$$\frac{t_i - t_{i-1}}{\psi_i} = \frac{\Delta t_i}{\psi_i} = d_i,$$

where d_i has mean 1, and ε_i has mean 0 and variance 1. We adopt a general specification of the conditional mean of durations:

$$\psi_i = \varphi_0^d + \sum_{l=1}^{\infty} \varphi_l(\theta^d)\{\Delta t_{i-l} - \varphi_0^d\},$$

where $\varphi_l(\theta^d)$ are uniquely defined functions of θ^d, the subvector of θ that parameterizes the intraduration process. We assume that the above specification satisfies stationarity conditions. We first show that:

$$\sqrt{N}\frac{\partial\, cov_{dE}(k)}{\partial\theta_i}|_{\theta=\theta^0} = O(1), \ \forall\theta_i\epsilon\theta. \tag{A.1}$$

Consider $\theta_j\epsilon\theta^d$:

$$\left.\frac{\partial\, cov_{dE}(k)}{\partial\theta_j^d}\right|_{\theta=\theta^0} = N^{-1}\sum_i p_i q_{i-k}, \tag{A.2}$$

where

$$p_i = -\frac{d_i}{\psi_i}\sum_{l=1}^{\infty}\frac{\partial\varphi_l(\theta^d)}{\partial\theta_j^d}(\Delta t_{i-l} - \varphi_0^d), \text{ and} \tag{A.3}$$

$$q_{i-k} = E_{i-k} - 1. \tag{A.4}$$

The highest order of d_i in p_i is 2. Under the null hypothesis of independence, p_i and q_{i-k} are independent. Under the null the same result holds for $\theta_j \epsilon \theta^r$, the parameter vector of the return process rather than θ^d for the duration process, under the following specification of the conditional variance of returns:

$$\sigma_i^2 = \varphi_0^r + \sum_{l=1}^{\infty}\varphi_l(\theta^r, \bar{\delta}_i)\{\Delta y_{i-l}^2 - \varphi_0^r\}, \tag{A.5}$$

where $\varphi_l(\theta^r, \bar{\delta}_i)$ are functions of θ^r and a fixed value of the duration $\bar{\delta}$, and satisfy stationarity assumptions. We then have:

$$\left.\frac{\partial\, cov_{dE}(k)}{\partial\theta_j^r}\right|_{\theta=\theta^0} = N^{-1}\sum_i p_i q_{i-k}, \tag{A.6}$$

where $p_i = d_i - 1$ and:

$$q_{i-k} = -\frac{E_{i-k}}{\sigma_{i-k}}\sum_{l=1}^{\infty}\frac{\partial\varphi_l(\theta^r,\bar{\delta})}{\partial\theta_j^r}(\Delta y_{i-k-l}^2 - \varphi_0^r). \tag{A.7}$$

The highest order of $\varepsilon_{i-k}q_{i-k}$ is 4.

The stationarity and existence of the fourth moment of d_i and the fourth moment of E_i imply that:

$$\sqrt{N}\frac{\sum p_i q_{i-k}}{N} \underset{N\to\infty}{\longrightarrow} N(0,\sigma^2), \tag{A.8}$$

[Hannan(1970), Theorem 14, p. 232]. Hence for any θ_j epsilon θ:

$$\sqrt{N}\left.\frac{\partial\, cov_{dE}(k)}{\partial\theta_j}\right|_{\theta=\theta^0} = O(1). \tag{A.9}$$

We now expand $c\hat{o}v_{dE}(k)$ around the true parameter θ^0,

$$\begin{aligned}\sqrt{N}c\hat{o}v_{dE}(k) \;=\; & \sqrt{N}cov_{dE}(k) + \sqrt{N}(\hat{\theta}-\theta_0)\left.\frac{\partial\, cov_{dE}(k)}{\partial\theta}\right|_{\theta=\theta^0} \\ & + \sqrt{N}(\hat{\theta}-\theta^0)'\left.\frac{\partial^2 cov_{dE}(k)}{\partial\theta\partial\theta'}\right|_{\theta=\theta^*}(\hat{\theta}-\theta^0),\end{aligned} \tag{A.10}$$

where $\| \theta^0 - \theta^* \| \leq \| \theta^0 - \hat{\theta} \|$. Since by assumption $\hat{\theta}$ is consistent,

$$(\hat{\theta}-\theta^0) = O(1). \tag{A.11}$$

By this result, and if we assume that for all θ in an open convex neighborhood of (θ^0) and for all N, $\sqrt{N}\partial^2 c\hat{o}v_{dE}(k)/\partial\theta_i\partial\theta_j$ exists and is bounded in probability for $\theta_i, \theta_j \epsilon \theta$, where $\theta = (\theta^d, \theta^r)$ are the parameters of the duration and return processes, and $\theta^0 = (\theta^{d^0}, \theta^{r^0})$ is the true parameter vector, then we have that:

$$\sqrt{N}c\hat{o}v_{dE}(k) = \sqrt{N}cov_{dE}(k) + \sqrt{N}(\hat{\theta}-\theta^0)\left.\frac{\partial\, cov_{dE}(k)}{\partial\theta}\right|_{\theta=\theta^0} + o(1). \tag{A.12}$$

The second term is $o(1)$ by (1) and (8). Hence

$$\sqrt{N}c\hat{o}v_{dE}(k) = \sqrt{N}cov_{dE}(k) + o(1). \tag{A.13}$$

For $k \neq k'$ we have

$$\begin{aligned}Nc\hat{o}v_{dE}(k)c\hat{o}v_{dE}(k') \;&=\; (\sqrt{N}cov_{dE}(k) + o(1))(\sqrt{N}cov_{dE}(k') + o(1)) \\ &=\; N\, cov_{dE}(k)cov_{dE}(k') + o(1)\end{aligned} \tag{A.14}$$

By a similar argument,

$$c\hat{o}v_{dd}(\theta) = cov_{dd}(\theta) + o(1),$$

and

$$c\hat{o}v_{EE}(\theta) = cov_{EE}(\theta) + o(1).$$

Applying the Slutsky and Cramer-Rao theorems imply that $\hat{c}_{dE}(k)$ and $c_{dE}(k)$ have the same asymptotic distribution, and:

$$\sqrt{N}(\hat{c}d_E(k_1), \ldots, \hat{c}_{dE}(k_m)) \to N(0, I_m).$$

References

Andrews, D. W. K., and J. C. Monahan (1992). "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator." *Econometrica* , 60:953–966.

Bera, A. K., and M. L. Higgins (1995). "On ARCH models: Properties, estimation and testing." In L. Exley, D. A. R. George, C. J. Roberts, and S. Sawyer (eds.), *Surveys in Econometrics* . Oxford: Basil Blackwell.

Bollerslev, T., Y. C. Chou, and K. Kroner (1992). "ARCH modeling in finance: A selective review of the theory and empirical evidence." *Journal of Econometrics* , 52:201–224.

Bollerslev, T., R. Engle, and D. Nelson (1994). "ARCH models." In R. F. Engle and D. McFadden (eds.), *Handbook of Econometrics* , vol. IV. Amsterdam: North-Holland.

Bollerslev, T., and E. Ghysels (1996). "On periodic autoregression conditional heteroskedasticity." *Journal of Business and Economic Statistics*, 14:139–152.

Cheung, Y. W., and L. K. Ng (1996). "A causality-in-variance test and its applications to financial market prices." *Journal of Econometrics* , 72:33–48.

Conley, T., L. Hansen, E. Luttmer, and J. Scheinkman (1997). "Estimating subordinated diffusions from discrete time data." *Review of Financial Studies*, 10:525–577.

Diebold, F. X., and J. A. Lopez (1995). "Modeling volatility dynamics." In K. Hoover (ed.), *Macroeconomic Developments, Tensions and Prospects*. The Netherlands, Kluwer.

Drost, F. C., and T. E. Nijman (1993). "Temporal aggregation of GARCH processes." *Econometrica* , 61:909–927.

Drost, F. C., and B. J. M. Werker (1996). "Closing the GARCH gap: Continuous time GARCH modeling." *Journal of Econometrics* , 74:31–57.

Engle, R. F. (1982). "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation." *Econometrica* , 50:987–1007.

Engle, R. F., and J. R. Russell (1996). "Forecasting transaction rates: The autoregressive conditional duration model." Discussion Paper, Economics Department, University of California, San Diego.

Engle, R. F., D. Hendry, and J. F. Richard (1983). "Exogeneity." *Econometrica* , 51:277–304.

Gallant, A. R., and G. Tauchen (1989). "Semiparametric estimation of conditionally constrained heterogeneous processes: Asset pricing applications." *Econometrica* , 57:1091–1120.

Ghysels, E., C. Gouriéroux, and J. Jasiak (1998a). "Market time and asset price movements: Theory and estimation." In D. Hand and S. Jacka (eds.), *Statistics in Finance* . London, Edward Arnold.

Ghysels, E., C. Gouriéroux, and J. Jasiak (1998b). "High frequency financial time series data: Some stylized facts and models of stochastic volatility." In C. Dunis and B. Zhou (eds.), *Nonlinear Modeling of High Frequency Financial Time Series*. New York: John Wiley.

Ghysels, E., and J. Jasiak (1994). "Stochastic volatility and time deformation: An application of trading volume and leverage effects." Paper presented at the Western Finance Association Meeting, Santa Fe, New Mexico.

Ghysels, E., and J. Jasiak (1996). "Long term dependence in trading." Discussion Paper.

Gouriéroux, C., J. Jasiak, and G. LeFol (1996). "Intra-day market activity." Discussion Paper, CREST, Paris.

Hannan, E. J. (1970). "Multiple time series." New York: Wiley.

Hansen, L. P. (1982). "Large sample properties of generalized method of moments estimators." *Econometrica* , 50:1029–1054.

Haugh, L. D. (1976). "Checking the independence of two covariance stationary time series: A univariate residual cross-correlation approach." *Journal of the American Statistical Association* , 71:378–385.

Lamoureux, C., and W. Lastrapes (1990). "Heteroskedasticity in stock return data: Volume versus GARCH effect." *Journal of Finance* , 45:221–229.

Lancaster, T. (1990). "The econometric analysis of transition data." Cambridge: Cambridge University Press.

McLeod, A. I., and W. K. Li (1983). "Diagnostic checking ARMA time series models using squared-residual autocorrelations." *Journal of Time Series Analysis*, 4:269–274.

Meddahi, N., and E. Renault (1995). "Linear statistical inference for ARCH-type processes." Discussion Paper, GREMAQ, Toulouse, France.

Newey, W. K., and K. West (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica*, 55:703–708.

O'Hara, M. (1995). *Market Microstructure Theory*. Cambridge: Basil Blackwell Publ.

Pai, J. C., and W. Polasek (1995). "Irregularly spaced AR and ARCH (ISAR-ARCH) models." Working Paper, University of Basel.

Palm, F. C. (1996). "GARCH models of volatility." In G. S. Maddala and C. R. Rao (eds.), *Handbook of Statistics*, vol. 14. Amsterdam: North-Holland.

Advisory Panel

Jess Benhabib, New York University

William A. Brock, University of Wisconsin-Madison

Jean-Michel Grandmont, CEPREMAP-France

Jose Scheinkman, University of Chicago

Halbert White, University of California-San Diego

Editorial Board

Bruce Mizrach (editor), Rutgers University

Michele Boldrin, University of Carlos III

Tim Bollerslev, University of Virginia

Carl Chiarella, University of Technology-Sydney

W. Davis Dechert, University of Houston

Paul De Grauwe, KU Leuven

David A. Hsieh, Duke University

Kenneth F. Kroner, BZW Barclays Global Investors

Blake LeBaron, University of Wisconsin-Madison

Stefan Mittnik, University of Kiel

Luigi Montrucchio, University of Turin

Kazuo Nishimura, Kyoto University

James Ramsey, New York University

Pietro Reichlin, Rome University

Timo Terasvirta, Stockholm School of Economics

Ruey Tsay, University of Chicago

Stanley E. Zin, Carnegie-Mellon University

Editorial Policy

The *SNDE* is formed in recognition that advances in statistics and dynamical systems theory may increase our understanding of economic and financial markets. The journal will seek both theoretical and applied papers that characterize and motivate nonlinear phenomena. Researchers will be encouraged to assist replication of empirical results by providing copies of data and programs online. Algorithms and rapid communications will also be published.

ISSN 1081-1826