JOURNAL OF
Econometrics

ELSEVIER Journal of Econometrics 106 (2002) 369–400
www.elsevier.com/locate/econbase

Modeling the interdependence of volatility and inter-transaction duration processes

Joachim Grammig[a,b,*], Marc Wellner[b]

[a] CORE, 34 Voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium
[b] University of Frankfurt, Frankfurt, Germany

Received 21 September 1999; revised 13 June 2001; accepted 13 July 2001

Abstract

This paper develops an approach for modeling the interdependence of intra-day volatility and trade duration processes, and extends the recursive specifications that have recently been proposed in the literature. We propose a suitable GMM estimation strategy that includes straightforward estimation of the autoregressive conditional duration model. A Monte Carlo study examines the performance of the estimation method. The empirical work investigates the impact of volatility on transaction intensity in the secondary equity market after a large initial public offering. We find that lagged volatility significantly reduces transaction intensity, which is consistent with predictions from microstructure theory. © 2002 Elsevier Science S.A. All rights reserved.

JEL classification: C32; C41; C51; G14

Keywords: Inter-trade duration and volatility; Financial market microstructure; Ultra-high-frequency data

1. Introduction

The accessibility of financial time series on the transaction level, labeled ultra-high-frequency data by Engle (2000), offered new perspectives for empirical finance and stimulated econometric model development. An essential benefit of ultra-high-frequency data is that it provides the appropriate basis

* Corresponding author. CORE, Voie du Roman Pays, 34, B-1348 Louvain-La-Neuve, Belgium. Tel.: +49-171-607-1385.
E-mail address: grammig@core.ucl.ac.be (J. Grammig).

PII: S0304-4076(01)00105-1

for empirical tests of market microstructure theories. An intriguing issue is to model the interaction of volatility and transaction intensity in an econometric framework. Since French and Roll (1986) have found evidence that volatility is caused by private information that affects prices when superiorly informed investors trade, a bivariate model for volatility and inter-trade duration processes offers means of studying the impact of asymmetric information effects on trading intensity.

In this paper, we extend the recently proposed GARCH models for transactions data to an interdependent specification in which the transaction intensity impacts on the volatility process, and vice versa. The model is applied to secondary market trading after a large IPO in which, as the corporate finance literature claims, asymmetrically informed trading is most likely taking place. GMM is proposed as a suitable estimation strategy, and recently proposed moment selection criteria are utilized. A Monte Carlo study is conducted to assess the performance of alternative moment conditions that are derived in this paper.

Both microstructure literature and empirical models have offered partially contradictory results regarding the relationship between trading intensity and information-based trading.[1] Easley and O'Hara (1992) and Admati and Pfleiderer (1988) have offered quite antipodal conclusions. In the Easley/O'Hara model informed agents trade only during days with news events that influence the asset price. If the news is good (bad), informed traders buy (sell). If there is no information event, only the uninformed agents trade with the market maker. The conclusion that transaction intensity is high if informed trading takes place rests on a crucial assumption: Although the market maker learns from the history of trades about the possible presence of informed traders, and adjusts bid and ask prices accordingly, the uninformed traders do not exhibit any strategic behavior. Instead, it is assumed that they trade with a constant intensity that is unaffected by any indicators of informed trading. This seems to be a restrictive assumption in the light of recent empirical evidence provided by Dufour and Engle (2000), who find that when the uninformed assume the presence of informed agents they will reduce their trading activity.

Admati and Pfleiderer (1988) consider a more sophisticated behavior pattern on the part of uninformed traders. In their model the uninformed agents concentrate trading activity in periods where there is no indication for informed trading, while they avoid to trade when there is evidence for informed trading. As Engle (2000) notes, the Admati/Pfleiderer model implies, in a nutshell, *slow trading means informed trading and high volatility*.[2]

[1] See O'Hara (1995) and Madhavan (2000) for surveys of the market microstructure literature.

[2] See Diamond and Verrecchia (1987) and Easley and O'Hara (1991) for further explanations regarding the relationship between informed trading and transaction intensity.

Like the theoretical literature, recently developed empirical models did not offer unambiguous results regarding the relationship between volatility and transaction intensity. Econometric frameworks that are appropriate for quote-driven markets (like the NYSE) have been proposed by Engle and Lunde (1999) and Russell (1999). Since Engle and Russell (1998) have shown that quote durations and volatility are inversely related measures, these models enable one study of the interactions of trade intensity and volatility (the latter measured via the quote duration process). An important result of Engle and Lunde's empirical application is that when quotes have not been revised for a long time, i.e. volatility was low, transaction intensities increase (which is, conceiving volatility as an indicator for informed trading, consistent with the Admati/Pfleiderer model). Russell's (1999) application is an analysis of market and limit order submission processes. He found that lagged volatility, measured as transaction price changes, exerts only little impact on market and limit order submission intensity.

In order-driven markets, volatility has to be measured in terms of transaction prices instead of quotes.[3] Suitable econometric frameworks for order-driven markets have been introduced by Russell and Engle (1999), who have developed the autoregressive conditional multinomial (ACM) model, and Engle (2000) and Ghysels and Jasiak (1998), who have proposed a family of ultra-high-frequency (UHF)-GARCH models. There are two main differences between the ACM and the UHF-GARCH framework. First, the ACM can be conceived as a dynamic competing risks model in which it is recognized that a non-trading spell can end in alternative states, determined by discrete price jumps. In contrast, UHF-GARCH models deal with transaction price changes as continuous random variables and volatility is dependent nonlinearly on the transaction intensity. Second, the UHF-GARCH framework offers a much more parsimonious parameterization: Due to the state changes that have to be tracked, an ACM application may easily involve estimation of more than 100 parameters, whereas Engle's (2000) most extensively parameterized UHF-GARCH includes only 12 unknown parameters.

In their ACM application, Russell and Engle (1999) allowed alternative volatility measures to impact on trading intensity and found that expected transaction intensity tended to be higher when conditional volatility *per transaction* is higher (which is in line with predictions from the Easley/O'Hara model).

[3] No official market makers act in an order-driven market. Instead, two instruments are available to trade an asset. Using a limit order one writes a free option to buy (sell) a given number of shares at a maximum (minimum) price. Non-executed limit buy and sell orders constitute the limit order book that provides the source of liquidity. Using a market order one can buy or sell a given amount of shares at the best possible price which is determined by the current state of the limit order book.

In the UHF-GARCH framework, the impact of volatility on trading intensity has not yet been considered. This recursive specification greatly reduces complexity, since the trade intensity and volatility equations can be estimated separately. However, the recursive model structure also limits the testing and quantification of the asymmetric information effects discussed above.

The contribution of the present paper is the formulation of an interdependent model for volatility and transaction intensity, including the development and performance testing of a GMM estimation strategy. We seek to extend the power of the recursive UHF-GARCH framework, and provide a specification that makes it possible to address empirically the interesting issue of how volatility influences transaction intensity.

The remainder of this paper is organized as follows. Section 2 develops the econometric framework. The estimation strategy is outlined in Section 3, where we also discuss GMM moment selection criteria, and in Section 4 we report the results of a Monte Carlo experiment. Section 5 describes the data and presents the results of the empirical application, including a comparison with the forecasting performance of a recursive UHF-GARCH model. Section 6 contains the concluding remarks, which are mainly an overall summary of the results and an outlook for research perspectives.

2. Econometric framework

Following Engle (2000), the sequence of trading events is conceived as a self-exciting marked point process. According to Cox and Lewis (1966), Hawkes (1971) and Rubin (1972) a point process is defined a sequence of event arrival times $\{t_0, t_1, \ldots, t_n, \ldots\}$ with $t_0 < t_1 < \cdots < t_n \ldots$, and an associated function $N(t)$ counting the number of events that have occurred by time t. The point process is described as self-exciting when the past evolution impacts on the probability of future events. Marks are random variables that contain the information that is associated with these events. In the present paper we are interested in jointly modeling the trade arrival times, t_i and the transaction price change, denoted r_i, the mark of interest that is associated with the trade event.

The basic reference for modeling intertemporally correlated event arrival times is the autoregressive conditional duration (ACD) model introduced by Engle and Russell (1998). The ACD specification is motivated by the observation that the waiting times between events during the intra-day trading process often exhibit a significant serial dependence. This may be due to information events occurring in clusters. Some of these events take place with certainty, for example the opening and closing of other exchanges or lunch breaks. Hence, a part of the duration persistence is attributable to intra-day seasonality or, more precisely, diurnality. Engle and Russell (1998) therefore propose that financial duration processes be broken down into a deterministic time-of-day effect (also referred to as a diurnal factor), $\Phi(t_{i-1})$, and a random variable x_i. With X_i as the duration between two trade events occurring at t_i and t_{i-1}, we have $X_i = x_i \Phi(t_{i-1})$. Engle and Russell (1995) propose approximating $\Phi(t_i)$ by a regression of X_i on a spline function using polynomials of time as explanatory variables and each full hour as a node. Dividing each duration in the sample by the corresponding spline function value, one obtains a sequence of diurnally adjusted durations, $x_i = X_i \hat{\Phi}(t_{i-1})^{-1}$.

In the ACD framework, the conditional expected inter-trade duration, $\psi_i = \mathrm{E}(x_i \mid \mathscr{F}_i)$, where $\mathscr{F}_i$ denotes the conditioning information set generated by the durations preceding x_i, is dependent on past (expected) durations. An ACD(1,1), augmented with a set of pre-determined variables z_i, is given by

$$\psi_{i+1} = \omega_d + \alpha_d x_i + \beta_d \psi_i + \zeta z_i. \tag{1}$$

Engle and Russell (1998) assume that the standardized durations

$$\varepsilon_i = \frac{x_i}{f(\psi_i)}, \tag{2}$$

are independent and identically distributed, and that their density satisfies

$$g\left(\frac{x_i}{f(\psi_i)} \mid \mathscr{F}_i; \theta_g\right) = g\left(\frac{x_i}{f(\psi_i)}; \theta_g\right). \tag{3}$$

$f(\cdot) > 0$ is a function with positive support that follows from the choice of $g(\cdot)$. [4]

As in Ghysels and Jasiak's (1998) recursive ACD-GARCH specification, we model the price change sequence as a GARCH process with permanently revised sampling frequency. The rationale of our econometric specification, however, is different. Our point of departure is the hypothetical situation in which the inter-trade durations are exactly equal to the value predicted by the diurnal component $\Phi(t_i)$, i.e. all diurnally adjusted inter-trade durations are equal to one. Expressed in transaction time, this implies that the data can be conceived of as being sampled at a constant frequency. Assume that in this hypothetical situation the squared price changes follow a weak GARCH(1,1),

$$P(r_{i+1}^2 \mid r_i, r_{i-1} \ldots) = h_{i+1} = \omega_h + \alpha_h r_i^2 + \beta_h h_i, \tag{4}$$

where $P(r_{i+1}^2 \mid r_i, r_{i-1} \ldots)$ denotes the best linear predictor in terms of $(1, r_{i-1}, r_{i-2}, \ldots, r_{i-1}^2, r_{i-2}^2, \ldots)$, i.e. $\mathrm{E}((r_i^2 - P(r_{i+1}^2 \mid r_i, r_{i-1} \ldots)) r_{i-n}^l) = 0$ for $i \geqslant 1$, $l = 0, 1, 2$ and $n = 1, 2, \ldots$ (see Drost and Nijman (1993) for the definitions of strong, semi-strong, and weak GARCH).

Because the inter-trade durations are equal to the values predicted by the deterministic diurnal component, we will refer to Eq. (4) as "normal

[4] Alternative distributions have been employed by Engle and Russell (1998), Grammig and Maurer (2000) and Lunde (1999).

duration GARCH process". Although unobservable in reality, this latent process is pivotal for our further considerations. The crucial question is, how the properties of the normal duration GARCH process are altered when the inter-trade durations differ from the diurnal component $\Phi(t_i)$. To address this issue, we maintain the assumption that until time t_i the inter-trade durations had been equal to the diurnal factor. If at time t_i the next inter-trade duration is expected to be higher than predicted by the diurnal factor, i.e. $\psi_{i+1} > 1$, this can be perceived as an expected aggregation of the normal duration GARCH to a lower sampling frequency. On the other hand, an expected inter-trade duration that is smaller than predicted by $\Phi(t_i)$, i.e. $\psi_{i+1} < 1$, implies that the expected sampling frequency is higher than for the normal duration GARCH process. After the next trades at $t_{i+1}, t_{i+2}, \ldots$ expected inter-trade durations will be revised again. The normal duration GARCH process is therefore subject to a permanent change of expected sampling frequencies. This suggests employing Drost and Nijman's (1993) temporal aggregation formulae to obtain a GARCH model with time-varying parameters that are functions of the expected inter-trade duration

$$h_{i+1} = \omega_{\mathrm{DN}}(\psi_{i+1}, \theta_h) + \alpha_{\mathrm{DN}}(\psi_{i+1}, \theta_h) r_i^2 + \beta_{\mathrm{DN}}(\psi_{i+1}, \theta_h) h_i, \tag{5}$$

where $\theta_h = (\omega_h, \alpha_h, \beta_h, k_h)'$, and k_h denotes the kurtosis of the price change distribution that is associated with the normal duration GARCH process. k_h is not treated as a parameter to be estimated. Instead, we assume the kurtosis of the normal distribution. $\omega_{\mathrm{DN}}(\psi_{i+1}, \theta_h)$, $\alpha_{\mathrm{DN}}(\psi_{i+1}, \theta_h)$ and $\beta_{\mathrm{DN}}(\psi_{i+1}, \theta_h)$ denote the elements of Drost and Nijman's (1993) temporal aggregation formulae for symmetric weak GARCH processes. Appendix A.2 contains the exact formulae as well as the theorem that justifies their application using arbitrary real aggregation parameters. This theorem is a crucial result for our approach, since in Eq. (5) the aggregation parameter, ψ_{i+1}, is a positive real number.

The effect of a changing sampling frequency on the GARCH parameters implied by the temporal aggregation formulae is depicted in Fig. 1. It illustrates that the temporal aggregation formulae generalize Diebold's (1988) intuitively appealing result that conditional heteroskedasticity disappears if the sampling interval increases to infinity. It is important to note that the Drost/Nijman formulae are valid for weak GARCH(1,1) processes, whereas strong and semi-strong GARCH processes do not temporally aggregate. This has important consequences for the estimation strategy, since ML estimation is not feasible.

In Section 1 we have argued for allowing the volatility process to influence the inter-trade duration process. Accordingly, we add two volatility measures to the ACD model's conditional expected duration equation. The first is h_i (expected volatility) and the second is the ratio of the squared price change

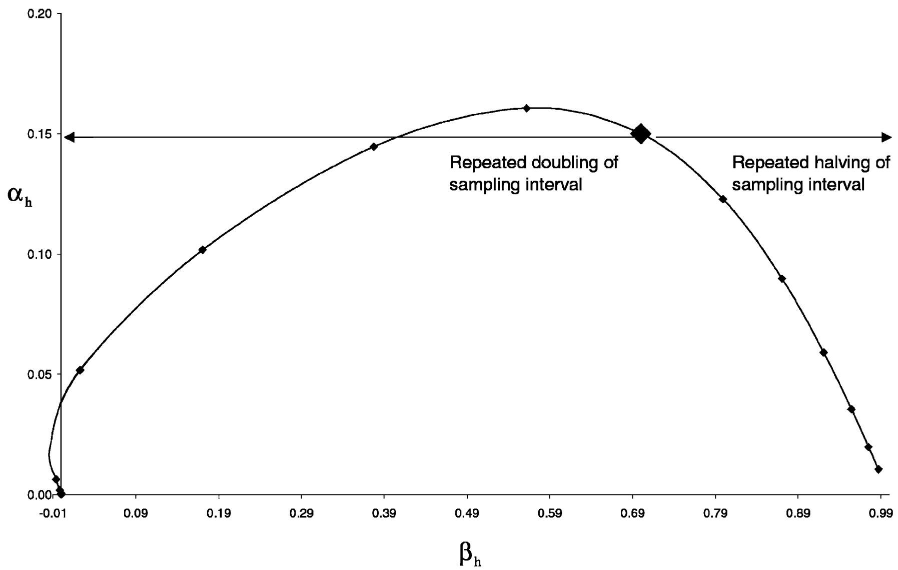

Fig. 1. Temporal aggregation of a symmetric weak GARCH(1,1) process for flow variables. $h_i = \omega_h + \alpha_h r_{i-1}^2 + \beta_h h_{i-1}$ where $\alpha_h = 0.2, \omega_h = 4, \beta_h = 0.7$, kurtosis: $k_h = 10$.

and h_i (the volatility shock)

$$\psi_{i+1} = \omega_d + \alpha_d x_i + \beta_d \psi_i + \zeta_1 h_i + \zeta_2 \frac{r_i^2}{h_i}. \tag{6}$$

Eqs. (5) and (6) constitute the interdependent duration–volatility (IDV) model. The parameter vector to be estimated is $\theta = (\omega_h, \alpha_h, \beta_h, \omega_d, \alpha_d, \beta_d, \zeta_1, \zeta_2)'$. The IDV specification implies that parameter estimation of the GARCH and ACD parameters cannot be separated, since the conditional volatility and inter-trade duration processes evolve simultaneously: Computation of h_{i+1} requires the availability of the conditional expected duration ψ_{i+1}, and in order to compute ψ_{i+1}, h_i is needed. This precludes applying the two-step estimation strategy for UHF-GARCH models pursued by Engle (2000) and Ghysels and Jasiak (1998).[5]

Conceiving volatility as an indicator for informed trading (French and Roll, 1986), significantly negative estimates of ζ_1 and ζ_2 indicate that evidence for informed trading induces a higher trading intensity. Such a result would be predicted by the Easley and O'Hara (1992) model. In contrast, one could conclude from significantly positive estimates of ζ_1 and ζ_2 that evidence for

[5] This procedure involves the estimation of an ACD model and the computation of a sequence of estimated conditional expected durations, $\{\hat{\psi}_i\}_1^T$ in the first step. In the second step, the ψ_i-sequence is used for the estimation of the GARCH parameters.

informed trading leads to a lower trading intensity. This result would be consistent with the Admati/Pfleiderer model, suggesting that the liquidity traders refrain from trading when informed trading takes place.

3. Estimation issues

3.1. Moment conditions and simultaneous GMM estimation

We have argued above that the temporal aggregation formulae hold only for a weak GARCH. Although a weak GARCH does not lead to a feasible ML estimation, orthogonality conditions that can be utilized for GMM estimation of the IDV model are readily available. In the following we will first present distribution-free orthogonality conditions that can be employed for GMM estimation of the ACD model, and then proceed to outline the moment conditions needed to perform simultaneous estimation of the interdependent ACD-GARCH model. [6]

Utilizing the definitions of strong, semi-strong and weak ACD processes in Appendix A.1, we are able to derive a set of orthogonality conditions which provide the basis for the GMM estimation of the ACD model. First, we can utilize the serial independence (strong ACD) or zero autocorrelation (semi strong ACD) assumption for the standardized durations $x_i\psi_i^{-1}$ to obtain the following orthogonality conditions:

$$E(f_i^e(\theta, v_i)) = E\left(\frac{x_i}{\psi_i} - 1\right) = 0, \tag{7}$$

$$E(f_i^c(\theta, v_i)) = E\left(\left(\frac{x_i}{\psi_i} - 1\right)\left(\frac{x_{i-j}}{\psi_{i-j}} - 1\right)\right) = 0, \quad \text{for } j = 1, \ldots, L, \tag{8}$$

where v_i denotes a vector of variables observable at time t_i. Additional moment conditions follow from the weak ACD definition. Utilizing the assumption that the implicit error in forecasting the duration at t_{i+1};

$$u_{i+1} = x_{i+1} - \psi_{i+1} \tag{9}$$

is uncorrelated with the lagged durations we have

$$E(f_i^u(\theta, v_i)) = E(u_{i+1} z_i^u) = 0, \tag{10}$$

[6] Dropping a distributional assumption for the ACD model does not seem to be a major drawback. Bauwens et al. (2000) show that all strong ACD models proposed in the literature assume duration distributions that do not seem suitable for modeling trade duration processes.

where

$$z_i^u = (1, x_i, x_{i-1}, \ldots, \tilde{z}_i^{u'})'. \tag{11}$$

$\tilde{z}_i^u$ denotes a vector of additional instruments assumed to be uncorrelated with u_{i+1}. Orthogonality conditions for the volatility equation can be derived from Drost and Nijman's (1993) weak GARCH definition which is characterized by the assumption that the implicit error in forecasting the squared price change

$$w_{i+1} = r_{i+1}^2 - h_{i+1} \tag{12}$$

is uncorrelated with lagged (squared) price changes (see Bates and White, 1988; Rich et al., 1991; Ghysels and Jasiak, 1998)

$$\mathrm{E}(f_i^w(\theta, v_i)) = \mathrm{E}(w_{i+1} z_i^w) = 0, \tag{13}$$

where

$$z_i^w = (1, r_i, r_{i-1}, \ldots, r_i^2, r_{i-1}^2, \ldots, \tilde{z}_i^{w'})'. \tag{14}$$

$\tilde{z}_i^w$ denotes a vector of additional instruments assumed to be uncorrelated with w_{i+1}. Assuming weak market efficiency (see Campbell et al., 1997, Chapter 1) one can utilize additional orthogonality conditions, r_{i+1}, too

$$\mathrm{E}(f_i^m(\theta, v_i)) = \mathrm{E}(r_{i+1} z_i^m) = 0, \tag{15}$$

where

$$z_i^m = (1, r_i, r_{i-1}, \ldots, \tilde{z}_i^{m'})'. \tag{16}$$

$\tilde{z}_i^m$ denotes additional instruments assumed to be uncorrelated with r_{i+1}. The moment conditions in (15) become especially important if the model contains, besides Eqs. (5) and (6), an equation for the price change, too.

For simultaneous estimation we combine the ACD and GARCH moment conditions

$$\mathrm{E}(f_i^w(\theta, v_i)) = \mathrm{E}(w_{i+1} z_i^w) = 0, \tag{17}$$

$$\mathrm{E}(f_i^m(\theta, v_i)) = \mathrm{E}(r_{i+1} z_i^m) = 0, \tag{18}$$

$$\mathrm{E}(f_i^e(\theta, v_i)) = \mathrm{E}\left(\frac{x_i}{\psi_i} - 1\right) = 0, \tag{19}$$

$$\mathrm{E}(f_i^c(\theta, v_i)) = \mathrm{E}\left(\left(\frac{x_i}{\psi_i} - 1\right)\left(\frac{x_{i-j}}{\psi_{i-j}} - 1\right)\right) = 0, \quad j = 1, \ldots, L, \tag{20}$$

$$\mathrm{E}(f_i^u(\theta, v_i)) = \mathrm{E}(u_{i+1} z_i^u) = 0. \tag{21}$$

To account for the interdependence of the IDV model, the vector of instruments z_i^u is augmented with lagged squared price changes

$$z_i^u = (1, x_i, x_{i-1}, \ldots, r_i^2, r_{i-1}^2, \ldots). \tag{22}$$

In the following we restrict our attention to the instrument vectors

$$z_i^m = (1, r_i, r_{i-1}, \ldots)', \tag{23}$$

$$z_i^w = (1, r_i, r_{i-1}, \ldots, r_i^2, r_{i-1}^2, \ldots)'. \tag{24}$$

Arrange the GMM disturbances in a vector

$$f(\theta, v_i) = (f_i^m(\theta, v_i)', f_i^w(\theta, v_i)', f_i^e(\theta, v_i), f_i^c(\theta, v_i)', f_i^u(\theta, v_i)')' \tag{25}$$

and define the sample means of $f(\theta, v_i)$ as

$$g(\theta; S_T) = T^{-1} \sum_{i=1}^{T} f(\theta, v_i), \tag{26}$$

where $S_T = (v'_T, v'_{T-1}, \ldots, v'_1)$ contains the observations of a sample size T. The GMM estimate $\hat{\theta}$ is then obtained by minimizing

$$Q(\theta; S_T) = g(\theta; S_T)' W_T g(\theta; S_T) \tag{27}$$

with respect to θ. W_T is a positive-semidefinite weighting matrix which satisfies $\lim_{T\to\infty} W_T = W_0$. When computing an estimate for the minimum asymptotic variance weighting matrix $W_T = \Omega^{-1}$, where

$$\Omega = \lim_{T\to\infty} T\mathrm{E}(g(\theta; S_T) g(\theta; S_T)') \tag{28}$$

we take into account that the vector process $\{f(\theta, v_i)\}_{i=-\infty}^{\infty}$ is serially correlated, by using an autocorrelation consistent estimator such as the one proposed by Newey and West (1987)[7]

$$\hat{\Omega} = \hat{\Gamma}_{0,T} + \sum_{v=1}^{M} \left(1 - \frac{v}{M+1}\right)(\hat{\Gamma}_{v,T} + \hat{\Gamma}'_{v,T}),$$

$$\hat{\Gamma}_{v,t} = T^{-1} \sum_{i=v+1}^{T} (f(\hat{\theta}, v_i))(f(\hat{\theta}, v_{i-v}))'. \tag{29}$$

After obtaining initial estimates of the parameters θ using $W_T = I$ we perform two more optimizations using an updated optimal weighting matrix $\hat{\Omega}$.

[7] Alternative estimators have been proposed by Gallant (1987), Andrews (1991), Andrews and Monahan (1992), and West (1997).

3.2. Moment selection criteria

In time series analysis the number of moment conditions becomes arbitrarily large (for example, if x_{i-1} is a valid instrument in Eq. (22), so are $x_{i-2}, x_{i-3}, \ldots$). It can be shown that in a linear regression context the asymptotic benefits of using efficiently all available lags, compared to a small number of moments, are sometimes large (see West et al., 1998). One has to be careful, however: based on a Monte Carlo experiment, Andersen and Sørenson (1996) demonstrate that in small samples the inclusion of an excessive number of moments results in more pronounced biases and larger root mean squared errors (see also Koenker and Machado, 1999). One way of approaching the moment selection problem would be to apply Gallant and Tauchen's (1996) efficient method of moments (EMM), another would be to extend optimal instruments approaches (see Hayashi and Sims, 1983; Hansen, 1985; West et al., 1998). However, extending optimal IV and adapting EMM for estimation of the IDV is beyond the scope of the present paper. We will therefore pursue the conventional GMM approach outlined before, and employ the moment selection criteria (MSC) proposed by Andrews (1999) as a guide to selecting lagged instruments and autocovariances in Eqs. (20), (22)–(24).

Andrews' basic idea is as follows. Assume we have a vector of moment conditions $g(\theta; S_T); \Theta \to \mathbb{R}^r$. Let $c \in \mathbb{R}^r$ denote a moment selection vector (MSV) whose jth element is one if the jth moment condition is included, and zero otherwise, and let $|c|$ denote the number of moments selected by c, i.e. $|c| = \sum_{i=1}^{r} c_j$. Let C denote parameter space for moment selection vector. The J-statistic based on the vector of moment conditions selected is

$$J_T(c) = T g_c(\theta; S_T)' W_T(c) g_c(\theta; S_T). \tag{30}$$

The subscript c indicates the dependence of the J-statistic on the selected moment selection vector. When all moment conditions selected by c are correct, then $J_T(c)$ has an asymptotic χ^2 distribution with $|c| - \min(p_\theta, |c|)$ degrees-of-freedom, where p_θ denotes the dimension of the parameter vector θ.[8]

The MSC estimator $\hat{c}_{\mathrm{MSC}}$ is the value that minimizes $MSC_T(c)$ over C where

$$MSC_T(c) = J_T(c) - h(|c|)\kappa_T. \tag{31}$$

The function $h(\cdot)$ and the constants $\{\kappa_T\colon T \geqslant 1\}$ are specified by the researcher. $h(\cdot)$ is assumed to be strictly increasing, and $\kappa_T \to \infty$ and $\kappa_T = \mathrm{o}(T)$. This implies that $h(|c|)\kappa_T$ is a "bonus term" that rewards selection vectors

[8] For the conditions under which this result holds, see Hansen (1982) for the case of moment conditions that are smooth in θ and Andrews (1997) for the case of moment conditions that may be nondifferentiable and/or discontinuous.

that utilize more moment conditions. This term is necessary to offset the increase in $J_T(c)$ that inevitably occurs when moment conditions are added, even if they are correct. Andrews (1999) introduces the following analogue of the familiar Schwartz Bayesian criterion (BIC):

$$\text{GMM-BIC}: MSC_{\text{BIC},T}(c) = J_T(c) - (|c| - p_\theta)\ln T; \tag{32}$$

and proves that the minimization of the GMM-BIC yields consistent moment selection estimators, i.e. the procedure selects the true moment selection vector c^0 with probability one as $T \to \infty$.[9]

4. Monte Carlo study

Since we assume a weak GARCH, the simulation of the IDV model is not feasible. We therefore focus on the ACD part to investigate the performance of the alternative moment conditions and the selection criteria that have been proposed above. We are especially interested in examining the relationship between number of moments and estimation performance using samples of a size typically available for high-frequency studies. The settings of our experiments varied with (a) the true data generating process (DGP), (b) the type, (c) number of moment conditions, and (d) the number of sample observations used in estimation. We consider two (strong) ACD-DGP that correspond to Engle and Russell's (1998) basic models, namely an exponential and a Weibull-ACD(1,1). In calibrating our DGP we tried to choose parameters that resembled those observed in trade duration data. By selecting the parameters in the ψ-equation to be equal to $\omega = 0.2$, $\alpha = 0.1$ and $\beta = 0.7$ (suppressing the d-subscript for convenience), we ensured that $\mathrm{E}(x_i) = 1$ (which results in real-world application as a consequence of the diurnal adjustment procedure outlined in Section 2. The γ-parameter of the Weibull distribution is 0.7. We consider three types of moment selection vectors (MSV). First, the pure instrumental variable-type (IV) of Eq. (10). Second, the pure autocovariance restrictions (AC) of Eq. (8), and third, mixed versions (denoted AC/IV). For each type we examine three alternative specifications with increasing number of lagged instruments or autocovariances, respectively. All nine MSVs utilize the moment restriction $\mathrm{E}(\frac{x_j}{\psi_j}) = 1$ and $v = 10$ for the Newey–West covariance matrix in Eq. (29).

Table 1 provides the details of this experimental setup. Two sample sizes, $T = 5000$ and $10\,000$, are investigated. In all, therefore, we have 36 experi-

[9] The associate editor pointed out that Andrews' MSC focus on the validity of the moment conditions, rather than their redundancy. Hall and Peixe (2001) have recently proposed a method, based on a canonical correlations information criterion, for selecting instruments on the basis of their relevance. They show that the method is consistent in the sense that it selects all relevant instruments from a candidate set of instruments which are orthogonal.

Table 1
Moment selection vectors (MSV) and number of moment conditions $|c|$ employed in the Monte Carlo study[a]

MSV	lags x_{i-j} in Eq. (10)	lags (x_{i-j}/ψ_{i-j}) in Eq. (8)	$\|c\|$
IV_1	3	0	5
IV_2	12	0	14
IV_3	24	0	26
AC_1	0	3	4
AC_2	0	12	13
AC_3	0	24	25
AC/IV_1	3	3	8
AC/IV_2	12	12	26
AC/IV_3	24	24	50

[a]All MSV also use $E(x_j/\psi_j - 1) = 0$. If the number of lags used for Eq. (10) is nonzero, then a constant is also used as an instrument.

ments in our design constructed from 2 DGPs, 9 MSVs, and 2 sample sizes. For each of the DGPs and sample sizes we generated 5000 random samples. The estimation is performed on the simulated data as outlined in the previous section. The programming language is GAUSS, the maximization algorithm Newton–Raphson.

Tables 2 and 3 display the sample mean and standard deviation (SD) of the parameter estimates as well as the root mean squared errors (RMSE). For each MSV we report the average GMM-BIC, and the frequencies of rejection at 5% significance level to illustrate how the finite-sample distribution of the J-statistic conforms to the asymptotic distribution (to conserve space, these results are reported only for $T = 5000$).

The kernel density estimates of the sampling distribution for the estimates ω, α, β are displayed in Fig. 2. These graphs are only depicted for the Weibull-DGP, since those for the exponential look similar and tell the same story. The left panels of Fig. 2 show the results for $T = 5000$. The right panels depict the results for the larger sample size, where we restrict our attention to the specification that employs the largest number of moments for each MSV type ($IV_3, AC_3, AC/IV_3$) and, for comparison, the densities that correspond to the most parsimonious specification, AC_1. As a benchmark for estimation precision we also display, for $T = 10\,000$, the kernel densities of the sampling distributions that result from a ML estimation of the correct model. All MSV converged without problems, but the difference in computation time varied significantly with the number of moments selected. On a Pentium III/650 MHz the estimation of specification AC_1 ($|c| = 4$), based on a $T = 10\,000$ sample, took about 1 min, whereas AC/IV_3 (($|c| = 50$)) used about 15 min CPU time. The main findings of the Monte Carlo experiment can be summarized as follows.

Table 2
Monte Carlo results exponential DGP[a]

MSV		$T=5000$ Mean	SD	RMSE	Rej. 5%	GMM BIC		$T=10\,000$ Mean	SD	RMSE
AC_1	ω	0.1994	0.1098	0.1098	5.4%	−7	ω	0.1964	0.0821	0.0822
	α	0.1041	0.0153	0.0158			α	0.1020	0.0110	0.0111
	β	0.6969	0.1159	0.1160			β	0.7019	0.0872	0.0872
AC_2	ω	0.2058	0.0441	0.0445	4.9%	−75	ω	0.2029	0.0300	0.0302
	α	0.1000	0.0140	0.0140			α	0.1001	0.0099	0.0099
	β	0.6938	0.0528	0.0532			β	0.6968	0.0362	0.0363
AC_3	ω	0.2077	0.0428	0.0435	4.8%	−165	ω	0.2039	0.0291	0.0294
	α	0.1003	0.0141	0.0141			α	0.1002	0.0099	0.0099
	β	0.6916	0.0513	0.0520			β	0.6957	0.0351	0.0354
IV_1	ω	0.2261	0.0871	0.0910	4.4%	−15	ω	0.2109	0.0561	0.0571
	α	0.1000	0.0172	0.0172			α	0.0997	0.0124	0.0124
	β	0.6728	0.0997	0.1034			β	0.6889	0.0655	0.0665
IV_2	ω	0.2120	0.0487	0.0501	3.4%	−83	ω	0.2055	0.0333	0.0338
	α	0.0962	0.0147	0.0152			α	0.0979	0.0105	0.0108
	β	0.6887	0.0586	0.0596			β	0.6950	0.0403	0.0406
IV_3	ω	0.2145	0.0460	0.0483	3.0%	−174	ω	0.2080	0.0323	0.0333
	α	0.0952	0.0144	0.0152			α	0.0972	0.0104	0.0108
	β	0.6855	0.0554	0.0573			β	0.6922	0.0390	0.0398
AC/IV_1	ω	0.2474	0.0987	0.1095	4.5%	−38	ω	0.2226	0.0707	0.0742
	α	0.0969	0.0155	0.0158			α	0.0979	0.0111	0.0113
	β	0.6535	0.1081	0.1177			β	0.6782	0.0779	0.0809
AC/IV_2	ω	0.2170	0.0416	0.0449	3.1%	−174	ω	0.2124	0.0295	0.0320
	α	0.0929	0.0130	0.0148			α	0.0949	0.0097	0.0109
	β	0.6884	0.0491	0.0505			β	0.6913	0.0351	0.0362
AC/IV_3	ω	0.2126	0.0406	0.0425	2.3%	−355	ω	0.2104	0.0291	0.0309
	α	0.0902	0.0122	0.0157			α	0.0929	0.0093	0.0117
	β	0.6954	0.0474	0.0476			β	0.6953	0.0345	0.0348

[a]The mean and SD column report the mean and the standard deviation of the sampling distribution of the estimates (5000 Monte Carlo replications). RMSE is the root mean squared error with respect to the true parameters $\omega=0.2$, $\alpha=0.1$, $\beta=0.7$. Rej. 5% reports the frequency of rejections at 5% significance level of the J-statistic.

(a) *Number of moments and potential bias*: The evidence regarding the number of moments is unambiguous. The selection of more moment conditions generates a clear improvement in estimation precision, relative to the use of few moments. For both DGPs and sample sizes, the RMSEs are

Table 3
Monte Carlo results, DGP: Weibull-ACD[a]

		$T=5000$						$T=10\,000$		
MSV		Mean	SD	RMSE	Rej. 5%	GMM BIC		Mean	SD	RMSE
AC_1	ω	0.2042	0.1279	0.1279	4.4%	−8	ω	0.1973	0.0983	0.0983
	α	0.1038	0.0197	0.0201			α	0.1022	0.0142	0.0143
	β	0.6930	0.1397	0.1398			β	0.7015	0.1072	0.1072
AC_2	ω	0.2142	0.0532	0.0551	6.3%	−75	ω	0.2082	0.0360	0.0369
	α	0.0992	0.0187	0.0187			α	0.0995	0.0132	0.0132
	β	0.6839	0.0653	0.0673			β	0.6908	0.0443	0.0452
AC_3	ω	0.2165	0.0507	0.0533	9.9%	−163	ω	0.2098	0.0344	0.0358
	α	0.0997	0.0188	0.0188			α	0.0999	0.0132	0.0132
	β	0.6809	0.0624	0.0652			β	0.6886	0.0425	0.0440
IV_1	ω	0.2273	0.1168	0.1199	2.2%	−15	ω	0.2132	0.0867	0.0877
	α	0.0935	0.0249	0.0257			α	0.0954	0.0197	0.0202
	β	0.6760	0.1370	0.1391			β	0.6897	0.1034	0.1039
IV_2	ω	0.2346	0.0631	0.0719	2.0%	−83	ω	0.2251	0.0474	0.0537
	α	0.0891	0.0196	0.0224			α	0.0929	0.0152	0.0168
	β	0.6655	0.0773	0.0847			β	0.6752	0.0581	0.0631
IV_3	ω	0.2311	0.0541	0.0624	2.4%	−174	ω	0.2270	0.0413	0.0493
	α	0.0874	0.0175	0.0216			α	0.0911	0.0142	0.0168
	β	0.6662	0.0670	0.0751			β	0.6718	0.0508	0.0581
AC/IV_1	ω	0.2611	0.1158	0.1309	4.3%	−37	ω	0.2300	0.0882	0.0931
	α	0.0910	0.0203	0.0222			α	0.0934	0.0152	0.0166
	β	0.6411	0.1300	0.1427			β	0.6724	0.0995	0.1032
AC/IV_2	ω	0.2227	0.0482	0.0533	4.3%	−173	ω	0.2187	0.0348	0.0395
	α	0.0881	0.0165	0.0203			α	0.0905	0.0124	0.0156
	β	0.6835	0.0582	0.0604			β	0.6860	0.0422	0.0444
AC/IV_3	ω	0.2173	0.0458	0.0489	4.5%	−353	ω	0.2163	0.0333	0.0371
	α	0.0856	0.0154	0.0210			α	0.0885	0.0118	0.0165
	β	0.6914	0.0548	0.0555			β	0.6901	0.0402	0.0414

[a]The mean and SD column report the mean and the standard deviation of the sampling distribution of the estimates (5000 Monte Carlo replications). RMSE is the root mean squared error with respect to the true parameters $\omega=0.2$, $\alpha=0.1$, $\beta=0.7$. Rej. 5% reports the frequency of rejections at 5% significance level of the J-statistic.

substantially reduced when we employ more lags and/or autocovariances to generate additional moment conditions. This can be seen, for example, by comparing the moment selection vectors AC_1 and AC_2. Further increasing the number of moment conditions (compare, for example, AC_2 to AC_3 where the number of autocovariances used as moments is doubled) leads to a further,

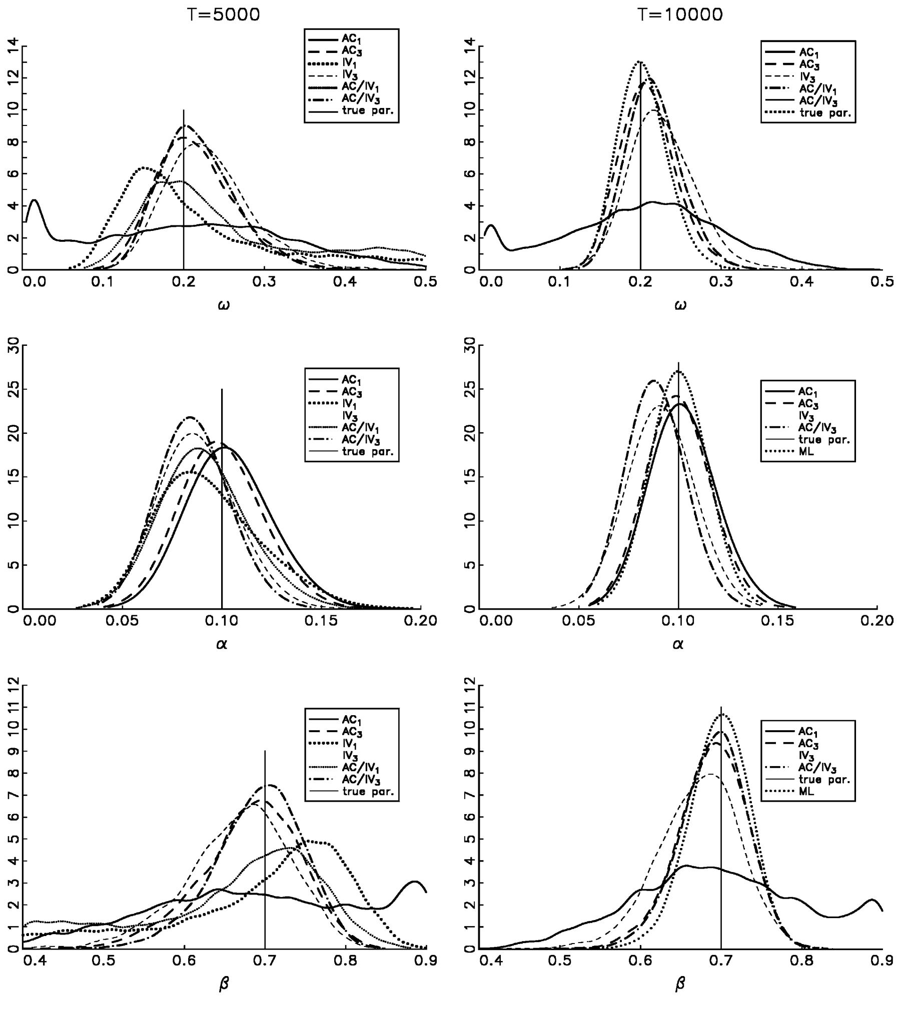

Fig. 2. Kernel density estimates of the parameter sampling distributions. DGP: Weibull-ACD. Gaussian kernel, bandwidth = 0.01.

but smaller improvement. The same pattern can be observed for the pure IV and pure autocovariance MSV and the mixed types. Fig. 2 shows graphically how estimation precision is increased by selecting more moment conditions. This effect is particularly visible when comparing AC_1 and AC_3. We conclude that the problem, found in small samples, of inducing a bias when selecting an increasing number of moments is not present when high-frequency data sample sizes are available, at least not with the dimension of the moment selection vectors that we have considered (and that are still small compared to the sample sizes)

(b) *Preferred type of moment condition*: In terms of RMSE, the moment selection vectors that employ pure AC moments perform better than the pure IV specifications. Furthermore, the kernel densities for the pure IV moment selection vectors depict a bias for α. This bias is clearly reduced in the larger sample. For $T = 5000$ employing both IV-type and autocovariances leads to an improvement of the RMSE, although the mixed strategy inherits the bias for α that is associated with pure IV. Again, the bias is reduced for $T = 10\,000$. Nevertheless, the performance of the IV specifications improve when AC-type moment conditions are added. This turns out to have important implications for our empirical application.

(c) *Moment selection by GMM-BIC*: The GMM-BIC favors MSVs that utilize a larger number of moments. For both the Weibull- and the exponential-DGP the moment selection vector with the smallest average GMM-BIC is AC/IV_3.

(d) *Reliability of J-statistic*: In their Monte Carlo study investigating GMM estimation of stochastic volatility models, Andersen and Sørenson (1996) reported that J-statistic p-values are inflated with increasing number of moments. This effect cannot be detected in our study. Instead, it seems that GMM inference based on the J-statistic is quite reliable.

(e) *Loss of precision compared to maximum likelihood*: The loss of precision of the GMM specifications that employ MSVs with a sufficiently large number of moments is small compared to maximum likelihood. In the case of the exponential DGP the loss is almost negligible (the exponential DGP graphs are not displayed).

5. Empirical application

5.1. Data: The Deutsche Telekom IPO

For the application of the IDV model we use transactions data from the first 5 trading weeks after the November 1996 Deutsche Telekom IPO, the first step of the largest ever privatization project in Germany. With Deutsche and Dresdner Bank as well as Goldman Sachs as underwriters 713 million shares, about 25% of the total shares of the formerly 100% state owned telecommunications monopolist, were offered to private investors. 23 million of the offered shares were distributed among the employees. The remaining shares were bought by German private (174 million shares) and institutional investors (254 million shares). International investors were located in the USA (14% of shares), the UK (8% of shares), other European countries (6% of shares) and Asia (5% of shares). The total issue proceeds amounted to 13.3

billion USD. Transactions data are obtained from the electronic trading system IBIS—(short for Integriertes Boersenhandels- und Informations-System—Integrated Stock Exchange Trading and Information System), an anonymous, electronic order-driven market.

Analyzing the Telekom IPO/IBIS data promises to be interesting for two main reasons. First, Hau (1999) has documented considerable asymmetric information effects in this order-driven market.[10] More precisely, he found that the local proximity of the trader to the corporate headquarters of the traded stock results in superior trading profits. Given the international distribution of the Telekom stock outlined above, the proximity issue definitely matters. Second, the corporate finance literature emphasizes that in the case of an IPO the market participants are clearly asymmetrically informed in that the underwriters are assumed to possess superior knowledge about the asset value (see Madhavan, 2000 for a survey).

When working with financial microstructure data the idiosyncrasies of market design and the data generating process inevitably have to be dealt with. IBIS is no exception, being designed as a "hit and take" system according to the taxonomy of Domowitz (1992). Therefore, buy and sell orders were not automatically matched by the system. Rather, traders willing to buy shares had to explicitly accept sell orders submitted by others and vice versa. A trader wishing to buy or sell a large number of shares had to accept several standing orders. From an economic point of view, this may be considered as one trade. In our data set, however, one transaction is recorded for each of the accepted orders. Since these transactions do not have identical time stamps, we cannot identify the component transactions to a large trade with certainty. We therefore used the following algorithm to bundle the split-transactions. All transactions that occurred within at most one second from the previous transaction within a sequence of transactions at either non-increasing or non-decreasing prices where considered as one large trade. This algorithm thus treats a sequence of buyer-initiated transactions that occurred with a delay of no more than one second between two consecutive transactions as one large trade, and similarly for seller-initiated transactions.

The motivation of Russell and Engle's (1999) ACM model was the observation that at NYSE, transaction price changes are discrete. In the raw IBIS data price discreteness is present to a much smaller extend, since in IBIS, unlike at NYSE, there is no minimum tick rule. After consolidating split transactions the remainder of discreteness in the data is lost. The resulting sample contains 12 057 trade events occurring between November 18, 1996 and December 20, 1996 during the regular IBIS trading hours (8:30 a.m.–5:00 p.m. CET). The first trade of each day was excluded. We also removed the observation immediately after a system breakdown on December 13 between

[10] Hau (1999) used data from the IBIS system-update, XETRA.

Table 4
Sample descriptives[a]

	X_i	x_i	R_i	r_i
Mean	61.393	1.009	−0.011	−0.002
SD	110.073	1.322	2.277	2.220
Skewness	5.115	3.225	0.057	0.103
Kurtosis	48.671	17.751	5.341	4.067
Max	2356.100	18.866	20.000	18.033
Min	0.130	0.002	−23.000	−16.086
Median	20.660	0.527	0.000	0.013
1% quantile	1.150	0.016	−6.000	5.820
99% quantile	517.090	6.151	6.000	6.024
$Q(6)$	2450.530	248.790	424.51	11.32
	(0.000)	(0.000)	(0.000)	(0.079)

[a]X_i, raw inter-transaction duration in seconds; x_i, diurnally adjusted duration; R_i, transaction price change in DEM/100; r_i, bid-ask bounce and diurnally adjusted transaction price change (in DEM/100). $Q(6)$ reports the value of the Ljung–Box Q-statistic with six degrees of freedom for each series. Under the null hypothesis of white noise the statistic is $\chi^2(6)$. The associated p-values are reported in parentheses. Sample size $T = 12\,057$.

09:00 a.m. and 10:30 a.m. In order to account for duration diurnality, we followed Engle and Russell (1998) and estimated cubic spline functions using half hours as nodes. The diurnal adjustment of inter-trade durations was performed as explained in Section 2. As in Bauwens et al. (2000) separate splines were fitted for the weekdays, US holidays and expiry dates at the Frankfurt futures and options exchange. In order to obtain a price change sequence which is free of the bid-ask bounce that affects financial transaction prices (see Campbell et al., 1997, Chapter 3), we follow Ghysels and Jasiak (1998) and use the residuals of an ARMA(2,0) model estimated on the price change data. Intra-day volatility seasonality was accounted for by applying the method proposed by Andersen and Bollerslev (1997).

Fig. 3 depicts the resulting filtered price change time series, the sample autocorrelation functions (ACF) for the series of interest, and kernel density estimates based on the filtered duration and price change data. Descriptives of the raw price change and duration data, as well as the filtered series, are reported in Table 4. The mean raw inter-transaction duration is about 1 min, and the diurnally adjusted duration data exhibit the familiar overdispersion (Engle and Russell, 1998). The ACF charts and the Q-statistics indicate no significant autocorrelations in the filtered price change series, whereas both squared price changes and durations exhibit significant serial dependence.

5.2. Estimation results

Before turning to the IDV estimation results we want to assess whether the transaction intensity influences significantly the volatility process. For this

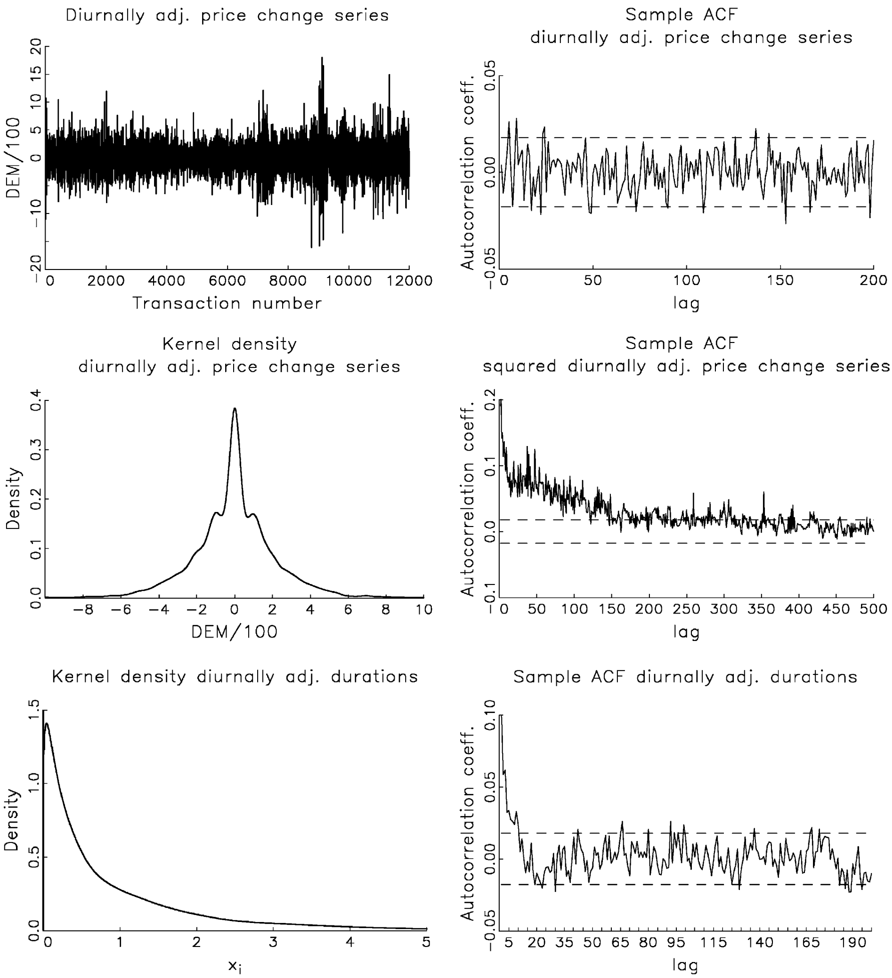

Fig. 3. The Telekom IPO dataset. *Note*: A Gaussian kernel with bandwidth as proposed by Silverman (1986) was used for the price change density. For the duration density the Gamma kernel with bandwidth $b=((0.9sT)^{-0.2})^2$ (s is the standard deviation of the duration sequence) was employed (Chen, 2000). The horizontal lines in the ACF charts represent the limits of the 95% confidence interval ($\pm 1.96/\sqrt{T}$).

purpose, the following test was carried out.[11] Consider the following mean value expansion of Eq. (5) about an arbitrary value for ψ (and ignoring θ_h)

$$h_{i+1} = \omega_1 + \omega_2\psi_{i+1} + \alpha_1 r_i^2 + \alpha_2\psi_{i+1} r_i^2 + \beta_1 h_i + \beta_2 \psi_i h_i \tag{33}$$

[11] We are grateful to the associate editor for suggesting this test idea.

Table 5
Alternative moment selection vectors for the IDV model[a]

MSV	lags r_{i-j} in (23)	lags r_{i-j} in (24)	lags r_{i-j}^2 in (22) and (24)	lags x_{i-j} in (22)	lags(x_{i-j}/ψ_{i-j}) in (20)	$\lvert c\rvert$	p-val.	GMM BIC
M_1	0	2	2	2	2	13	25.0	−40
M_2	0	3	3	3	3	18	0.0	−59
M_3	0	6	6	6	6	33	0.0	−172
M_4	0	12	12	12	12	63	0.8	−433
M_5	0	12	12	12	24	75	2.0	−537
M_6	0	12	12	24	24	87	2.5	−637
M_7	0	12	24	24	24	111	20.2	−853
M_8	0	24	24	24	24	123	23.8	−955
M_9	0	30	30	30	30	153	9.2	−1194
M_{10}	2	2	2	2	2	16	0.3	−51
M_{11}	3	3	3	3	3	27	0.1	−133
M_{12}	6	6	6	6	6	39	0.0	−227
M_{13}	12	12	12	12	12	75	0.8	−532
M_{14}	12	12	12	12	24	87	1.7	−634
M_{15}	12	12	12	24	24	99	3.2	−738
M_{16}	12	12	24	24	24	123	21.0	−954
M_{17}	12	24	24	24	24	135	25.8	−1057
M_{18}	24	24	24	24	24	147	33.9	−1161
M_{19}	30	30	30	30	30	183	50.2	−1470

[a]The $|c|$ column reports the number of moment conditions employed and the p-value. Column p-value of the J-statistic in %. All specifications employ the moment condition $\mathrm{E}(x_i/\psi_i - 1) = 0$. When lagged instruments are used in Eqs. (22)–(24) then a constant is also used as an instrument.

and assume that in a (strong) Gaussian GARCH the conditional variance evolves according to Eq. (33). A conventional GARCH(1,1) then follows from restricting $\omega_2 = \alpha_2 = \beta_2 = 0$. Hence, a straightforward likelihood ratio test can be conducted to test the significance of the impact of ψ on conditional volatility. The ψ-sequence required for the estimation of the unrestricted model is obtained by ML estimation of a Weibull-ACD(1,1). The estimation produces

$$\begin{aligned} h_{i+1} &= \underset{(0.173)}{0.169} - \underset{(0.158)}{0.017}\psi_{i+1} + \underset{(0.0327)}{0.067} r_i^2 + \underset{(0.026)}{0.026}\psi_{i+1} r_i^2 + \underset{(0.063)}{0.999} h_i - \underset{(0.058)}{0.123}\psi_i h_i \\ \mathscr{L} &= -25658; \quad \mathscr{L}^* = -25694. \end{aligned}$$

Robust standard errors are given in parentheses. $\mathscr{L}$ and $\mathscr{L}^*$ denote the log-likelihood of the unrestricted and restricted model, respectively. With a $\mathscr{LR}$ p-value < 0.0001 the restricted model is clearly rejected.

Andrews' (1999) GMM-BIC was used to choose the moment selection vector for the IDV model. All MSVs employed the restriction $\mathrm{E}(x_i/\psi_i - 1) = 0$ and

Table 6
GMM parameter estimates[a]

MSV	$\hat{\omega}_h$	$\hat{\alpha}_h$	$\hat{\beta}_h$	$\hat{\omega}_d$	$\hat{\alpha}_d$	$\hat{\beta}_d$	$\hat{\zeta}_1$	$\hat{\zeta}_2$	p-val.	GMM BIC	F-stat.
M_{19}	0.2845	0.1117	0.8203	0.2963	0.0636	0.5786	0.0154	0.0127	0.5024	−1470	37.33
	(0.0274)	(0.0082)	(0.0122)	(0.0494)	(0.0072)	(0.0603)	(0.0025)	(0.0035)			[0.000]
M_9	0.2791	0.1195	0.8150	0.3753	0.0578	0.5186	0.0156	0.0058	0.0915	−1194	27.37
	(0.0279)	(0.0092)	(0.0131)	(0.0601)	(0.0073)	(0.0694)	(0.0024)	(0.0036)			[0.000]
M_{18}	0.2989	0.1088	0.8192	0.2968	0.0678	0.5895	0.0103	0.0089	0.3389	−1161	27.80
	(0.0313)	(0.0088)	(0.0136)	(0.0473)	(0.0075)	(0.0573)	(0.0018)	(0.0033)			[0.000]
M_{17}	0.3165	0.1111	0.8146	0.3143	0.0677	0.5727	0.0098	0.0090	0.2576	−1057	25.44
	(0.0333)	(0.0091)	(0.0140)	(0.0492)	(0.0077)	(0.0593)	(0.0018)	(0.0034)			[0.000]
M_8	0.3512	0.1191	0.8012	0.3198	0.0668	0.5689	0.0087	0.0095	0.2377	−955	22.28
	(0.0387)	(0.0105)	(0.0160)	(0.0497)	(0.0076)	(0.0599)	(0.0017)	(0.0036)			[0.000]
M_{16}	0.3453	0.1226	0.7987	0.3074	0.0679	0.5812	0.0095	0.0072	0.2102	−954	23.10
	(0.0367)	(0.0104)	(0.0156)	(0.0485)	(0.0077)	(0.0582)	(0.0017)	(0.0034)			[0.000]
M_7	0.3841	0.1338	0.7819	0.3128	0.0675	0.5785	0.0085	0.0064	0.2020	−853	19.73
	(0.0422)	(0.0120)	(0.0176)	(0.0485)	(0.0077)	(0.0581)	(0.0016)	(0.0035)			[0.000]
M_1	0.0471	0.1058	0.8885	0.4918	0.0979	0.3651	0.0091	0.0057	0.2498	−40	12.34
	(0.0225)	(0.0219)	(0.0396)	(0.0968)	(0.0128)	(0.1031)	(0.0018)	(0.0105)			[0.000]

[a]MSV accepted at 5% significance level of J-statistic ordered by GMM-BIC. Standard errors in parentheses. The F-stat. column reports the $F(2, 12055)$ statistic and (in brackets) the associated p-value for a test of the null hypothesis that both $\widehat{\zeta_1}$ and $\widehat{\zeta_2}$ are equal to zero. The critical value at 5% significance is 2.99.

for the weighting matrix in Eq. (29) $v = 24$ was chosen. Table 5 reports the types and the number of moment conditions, the GMM-BIC, and the p-value of the J-statistic that is associated with each moment selection vector. GMM estimation was performed as described in Section 3. Starting values were obtained by a estimating a Gaussian GARCH and a Weibull-ACD and setting initial values of ζ_1 and ζ_2 equal to zero, or previous estimates employing similar moment conditions, respectively.

The moment selection vectors M_1-M_9 utilize, for the ACD part of the model, both IV- and AC-type moment conditions with increasing number of lagged instruments and autocovariances. For the volatility equation, a varying number of lagged (squared) price changes were used as instruments. The number of lagged squared price changes in Eqs. (23) and (24) is always the same. Moment selection vectors M_{10}-M_{19} additionally utilize the moment conditions in Eq. (15) that were derived from the weak market efficiency assumption. One could argue that this introduces certain redundancies since the IDV model does not contain a specific equation for r_i, and the moment conditions in Eq. (15) do not depend on the model parameters. On the other hand, by construction of the optimal weighting matrix (see Eq. (29)), utilizing these moment conditions can lead to an increase in efficiency. This effect seems to dominate in the present empirical application, since the specifications that utilize the moment conditions (15) produce slightly smaller parameter standard errors (see Table 6). For a more formal investigation of this issue one could employ the moment selection criteria focusing on relevance of moment conditions by Hall and Peixe (2001). The GMM-BIC points to moment selection vectors that employ more moment conditions. Weak market efficiency moment restrictions as well as additional AC-type moment conditions, that augment the IV moments for the ACD equation, are accepted by the GMM-BIC. The latter is consistent with the findings of Section 4 where it was shown that augmenting IV-type moment selection vectors with AC-type moments increased estimation precision. Generally, the moment selection vectors with superior GMM-BIC values are also those that achieve higher J-statistic p-values. [12] We have also tried to estimate specifications that do not employ the IV-type moment conditions of Eq. (21). However, the GMM algorithm collapsed for all of these specifications due to complex numbers produced by the GARCH temporal aggregation formulae during the minimization of the objective function (27). We conclude that it is crucial to employ lagged squared price changes as instruments for the moment condition $\mathrm{E}(u_{i+1}z_i^u)=0$ to facilitate estimation of the IDV model. For the other specifications no numerical problems occurred, but the required computer resources increased with the number of moments included. Moment selection vector M_{19} ($|c|=183$) used 7 h 20 min CPU time for parameter estimation and computation of standard errors, whereas M_1 required only 53 min (Pentium III/650 MHz).

Parameter estimates and standard errors of those moment selection vectors that yield p-values greater than 0.05 are reported in Table 6. Parameter estimates are similar across MSVs, but the standard errors associated with the MSV that employ more moment conditions tend to be smaller. The estimate $\hat{\alpha}_h$ is larger compared to what is typically found in GARCH applications estimated on equally spaced data, and $\hat{\beta}_h$ is smaller. Their sum is significantly smaller than one, such that the normal duration GARCH process is clearly not an integrated-GARCH. Furthermore, compared to univariate ACD applications, the estimate $\hat{\beta}_d$ is smaller. [13]

As outlined in Section 2, the normal duration GARCH is a latent process. After each transaction the GARCH parameters change, depending on the expected duration until the next trade. For an illustration, the estimated parameters of specification M_{19} are used to depict the sequence of time varying GARCH parameters, together with the conditional volatility and expected duration series, in Fig. 4.

[12] The associate editor noted that MSV M_1, but not M_{10}, is accepted by the J-statistic, indicating problems associated with the market efficiency moment conditions. However, both M_1 and M_{10} employ only a small number of moment conditions. In the light of the Monte Carlo and the GMM-BIC results we advocate the use of a larger set of moment conditions for the estimation of the IDV model.

[13] ML estimation of a Weibull-ACD(1,1) on the Telekom duration data yields an estimate $\hat{\beta}_d=0.6817$ with standard error 0.0324.

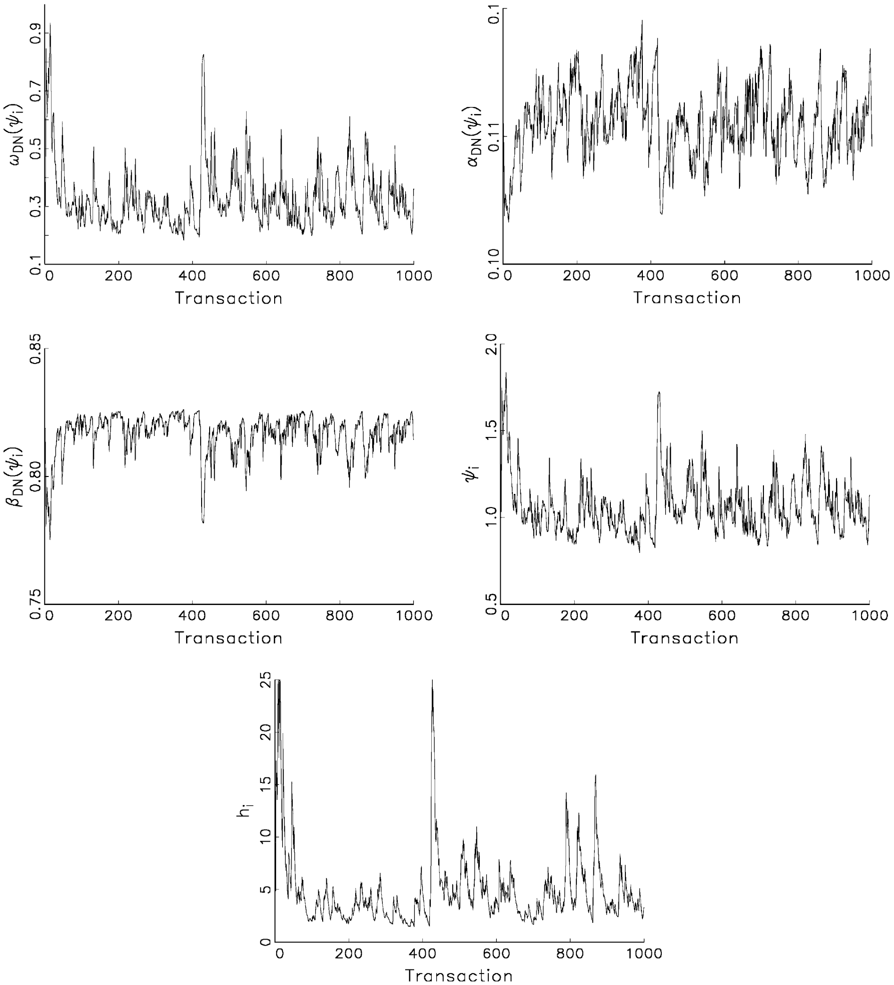

Fig. 4. First 1000 observations: time varying GARCH parameters, conditional expected duration, and conditional volatility.

To benchmark the IDV model we have also estimated the recursive UHF-GARCH specification proposed by Engle (2000), and compared the root mean squared errors of the implied one-step-ahead volatility forecast. The point of departure for the UHF-GARCH is the volatility per unit of time, conditional on the inter-trade duration, $V_i(r_{i+1}/\sqrt{x_{i+1}}\,|\,x_i) \equiv \sigma^2_{i+1}$ for which Engle proposes the dynamic specification

$$\sigma^2_{i+1} = \omega_u + \alpha_u \tilde{r}^2_i + \beta_u \sigma^2_i + \gamma_1 x_i^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \xi_i + \gamma_4 \psi_i^{-1}, \tag{34}$$

where $\tilde{r}_i = r_i^2/\sqrt{x_i}$. ξ_i is a long run volatility component computed by exponentially smoothing r^2/x with a proposed smoothing parameter equal to 0.005 (Engle, 2000). The conditional volatility per transaction forecast (i.e. the equivalent of Eq. (5) in the IDV) is then given by $V_i(r_{i+1} \mid x_{i+1}) \equiv \tilde{h}_{i+1} = x_{i+1}\sigma_{i+1}^2$. The IDV model presented in this paper and the UHF-GARCH differ in that the latter conditions the volatility forecast on the inter-trade duration x_{i+1} which is, however, unknown at t_i. We have estimated Eq. (34) using for r_i the diurnally adjusted, and bid-ask bounce corrected price change sequence that was also used for IDV estimation. The required ψ-sequence was obtained by ML estimation of a Weibull-ACD(1,1). Once this is available, the estimation of the UHF-GARCH can be performed using standard econometric software. Assuming Gaussian innovations we obtain

$$\begin{aligned} \sigma_{i+1}^2 = {} & \underset{(0.299)}{0.581} + \underset{(0.011)}{0.064}\tilde{r}_i^2 + \underset{(0.005)}{0.005}\sigma_i^2 + \underset{(0.142)}{3.390}x_i^{-1} - \underset{(0.013)}{0.042}\frac{x_i}{\psi_i} \\ & -\underset{(0.159)}{0.718}\xi_i + \underset{(0.329)}{0.217}\psi_i^{-1}. \end{aligned}$$

Robust standard errors are given in parentheses. The parameter estimates are comparable to Engle's (2000) result, especially the phenomenon that the persistence of the GARCH model is made up from duration and other variables. The exception is the long volatility parameter that changes its sign. Comparing the RMSE of the IDV volatility forecast, $(T^{-1}\sum(r_{i+1}^2 - h_{i+1})^2)^{0.5} = 11.6$, and the UHF-GARCH volatility forecast, $(T^{-1}\sum(r_{i+1}^2 - \tilde{h}_{i+1})^2)^{0.5} = 14.3$, we see that the IDV-RMSE is 19% smaller, although the inter-trade duration x_{i+1} is not used for the volatility forecast.

To assess how well the model explains the variance of the duration process we employ the following pseudo-R^2 measure

$$GoF = \frac{\mathrm{SD}(x_i) - \mathrm{SD}(x_i/\hat{\psi}_i)}{\mathrm{SD}(x_i)}. \tag{35}$$

This goodness-of-fit measure relates the standard deviation (SD) of the "ACD-residual" $x_i/\hat{\psi}_i$ to the standard deviation of a residual that corresponds to a model that assumes iid unit exponential durations, i.e. a Poisson process for the arrival of trades. The Poisson model simply issues the conditional expected duration forecast $\psi_i = 1$. Positive GoF-values therefore indicate improvements over this simple forecast. Computing the goodness of fit measure based on the ψ-sequence based on a ML estimation of a Weibull-ACD

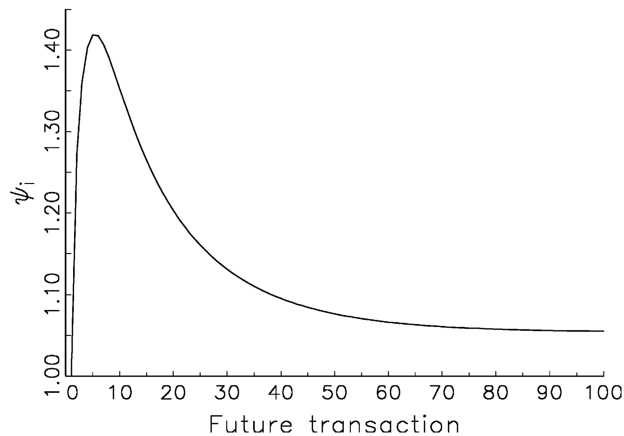

Fig. 5. The response of the conditional expected duration to a return shock of ± 10 $DEM/100$ at t_0. *Note*: The initial conditional volatility h_{-1} was set equal to 4.9 (the unconditional return variance) and the initial conditional expected duration ψ_{-1} and the lagged duration x_{-1} were set equal to one.

yields a GoF-statistic equal to 0.036, whereas the IDV model produces a GoF-statistic equal to 0.119. This indicates the improved explanatory power of the interdependent specification.

Table 6 shows that all moment selection vectors produce estimates of ζ_1 and ζ_2 that are greater than zero. For the GMM-BIC selected specification M_{19}, the t-values associated with the parameter estimates $\hat{\zeta}_1$ and $\hat{\zeta}_2$ are 6.2 and 3.6. Their correlation is -0.349. When testing the joint significance of the volatility indicators we obtain an $F(2, 12\,055)$-statistics that range from 12.34 (M_1) to 37.33 (M_{19}) (see the last column of Table 6). These numbers are clearly above the critical value at 5% significance (2.99). The magnitude of the ζ-parameters shows that their effect on the conditional expected duration is significant also from the economic point of view. This is illustrated in Fig. 5 which depicts the response of the conditional expected duration to a return shock of ± 10 $DEM/100$ at time t_0. The initial conditional volatility h_{-1} was set equal to 4.9 (the unconditional return variance) and the initial conditional expected duration ψ_{-1} and the lagged duration x_{-1} were set equal to one. Based on the parameter estimates of specification M_{19}, the IDV was then solved forward with future values of x_i and r_i^2 being replaced by the predictors ψ_i and h_i. Fig. 5 shows that in response to the volatility shock the conditional expected duration increases up to 1.42. To measure this effect in seconds one has to multiply ψ_i by the time-of-day function $\Phi(\cdot)$. By construction (spline regression), the mean of the time-of-day function is very close to the mean of the raw duration series which is about 1 min (see

Table 4). Evaluating at $\Phi(\cdot) = 60$ s the volatility shock therefore induces a maximum increase in the expected duration of about 25 s, an effect that can be considered economically important in the context of intra-day trading processes.

We conclude that in the secondary market after the Telekom IPO, both expected volatility and volatility shocks lengthened expected inter-trade durations (i.e. slowed down trade intensity). We have argued above that this finding, which is in line with a result recently reported by Engle and Lunde (1998), is consistent with the predictions from the Admati/Pfleiderer model. In an order-driven system like IBIS a reduction of transaction intensity can be caused by uninformed traders ceasing to submit market orders and a reduced submission and increased cancellation of their limit orders when the presence of informed traders is assumed. The resulting lack of liquidity can make informed market order trading unprofitable.

6. Conclusion and outlook

This paper has presented a new approach for modeling the interdependence of intra-day volatility and inter-trade duration processes. The IDV model extends the power of the recursive UHF-GARCH framework, and makes it possible to address interesting issues of financial microstructure. A suitable GMM estimation strategy for the IDV model was proposed that also facilitates straightforward estimation of the autoregressive conditional duration model.

In a Monte Carlo experiment it was demonstrated that the problem of inducing a bias when selecting an increasing number of moments, that was found in small samples, is not present when high-frequency data sample sizes are available. The empirical application has investigated the impact of volatility on transaction intensity in the secondary equity market after a large initial public offering (IPO). It was found that lagged volatility has a significantly negative impact on transaction intensity which is consistent with predictions from the Admati/Pfleiderer microstructure model.

Interesting further research projects are conceivable. From the methodological point of view, further investigating the opportunity to extend optimal instruments approaches, Hall and Peixe's (2001) alternative moment selection criterion, and EMM for the estimation of the model presented in this paper seems a promising task. Furthermore, the IDV model can be employed to address a variety of issues in empirical finance. For example, one might consider analyzing the influence of different research efforts by stock market analysts: Does the result that volatility shocks decrease transaction intensity durations hold, in general, for blue chip securities which are subject to intensive research, and would the result be different for less frequently traded, less intensively analyzed stocks, where informational signals from trading may be given more weight?

Acknowledgements

We thank the audiences at the North American Econometric Society 2000 Winter Meeting, the 9th $(EC)^2$ Meeting, the 54th Econometric Society European Meeting, and the seminar participants at Frankfurt, Louvain-la-Neuve and Mannheim who contributed to the discussion of this paper. We benefited from comments by T. Amemiya, D. Bates, L. Bauwens, R. Gencay, P. Giot, N. Hautsch, R. Hujer, M. Melvin, M. Paolella, and the Editor, Associate Editor and two referees. Financial support is gratefully acknowledged from the German Economic Association and the Landeszentralbank Hessen. We thank Roland Berger & Partners for providing access to their computer resources.

Appendix A.

A.1. Autoregressive conditional duration processes

The following definitions of autoregressive conditional duration processes correspond to Drost and Nijman's (1993) GARCH definitions. Let $\{x_i,\ i \in Z\}$ be a sequence of stationary durations normalized such that $\mathrm{E}(x_i) = 1$. Define operators $A(L) = 1 + \sum_j^q \alpha_{d,j} L^j$ and $B(L) = 1 - \sum_j^p \beta_{d,j} L^j$ and let the sequence $\{\psi_i,\ i \in Z\}$ be defined as the stationary solution of

$$B(L)\psi_i = \omega_d + (A(L) - 1)x_i.$$

We assume that $A(L)$ and $B(L) + 1 - A(L)$ have roots outside the unit circle and hence are invertible.

Definition 1 (Strong ACD). The sequence $\{x_i,\ i \in Z\}$ is defined to be generated by a strong ACD(p,q) process, if ω_d, $A(L)$ and $B(L)$ can be chosen such that

$$\frac{x_i}{f(\psi_i)} \sim \text{i.i.d. } g\left(\frac{x_i}{f(\psi_i)}; \theta_g\right),$$

where $g(\cdot)$ specifies a p.d.f. and $f(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$.

Definition 2 (Semi-strong ACD). The sequence $\{x_i,\ i \in Z\}$ is defined to be generated by a semi-strong ACD(p,q) process, if ω_d, $A(L)$ and $B(L)$ can be chosen such that

$$\mathrm{E}(x_i \mid x_{i-1}, x_{i-2}, \ldots) = \psi_i \quad \text{implying,}$$

$$\mathrm{E}\left(\frac{x_i}{\psi_i}\right) - 1 = 0 \quad \text{and}$$

$$\mathrm{E}\left(\left(\frac{x_i}{\psi_i} - 1\right)\left(\frac{x_{i-j}}{\psi_{i-j}} - 1\right)\right) = 0 \quad \text{for } j = 1, 2, \ldots .$$

Definition 3 (Weak ACD). The sequence $\{x_i,\ i \in Z\}$ is defined to be generated by a weak ACD(p,q) process, if ω_d, $A(L)$ and $B(L)$ can be chosen such that

$$P(x_i \mid x_{i-1}, x_{i-2}, \ldots) = \psi_i,$$

where $P(x_i \mid x_{i-1}, x_{i-2}, \ldots) = \psi_i$ denotes the best linear predictor of x_i in terms of $1, x_{i-1}, x_{i-2}, \ldots$ i.e.,

$$\mathrm{E}((x_i - P(x_i \mid x_{i-1}, x_{i-2}, \ldots))x_{i-n}^{\ell}) = 0$$
$$\text{for } i \geqslant 1 \ \ell = 0, 1 \quad \text{and} \quad n = 1, 2, \ldots .$$

A.2. Discrete time GARCH aggregation

Theorem A.1 (Drost and Werker, 1996, pp. 47–48). *Let $h > 0$ and suppose $(y_{(h)t}, t \in hN)$ is a weak GARCH process with parameter $\theta_h = (\omega_h, \alpha_h, \beta_h, k_h)$, where k_h is the kurtosis of $h_{(h)t}$. Then, for each integer $m \geqslant 1$ the process $(y_{(mh)t}^{(m)} = \sum_{i=1}^{m-1} y_{(h)t+ih}, t \in mhN)$ is symmetric weak GARCH with parameter $\theta_{mh} = (\omega_{mh}, \alpha_{mh}, \beta_{mh}, k_{mh})$ (with $|\beta_{mh}| < 1$) where*

$$\omega_{mh} = \omega_{\mathrm{DN}}(m, \theta_h) = m\omega_h \frac{1 - (\alpha_h + \beta_h)^m}{1 - (\alpha_h + \beta_h)},$$

$$\alpha_{mh} = \alpha_{\mathrm{DN}}(m, \theta_h) = (\alpha_h + \beta_h)^m - \beta_h,$$

$\beta_{mh} = \beta_{\mathrm{DN}}(m, \theta_h)$ *is the real solution of*

$$\frac{\beta_{mh}}{1 + \beta_{mh}^2} = \frac{a(\alpha_h, \beta_h, k_h, m)(\alpha_h + \beta_h)^m - b(\alpha_h, \beta_h, m)}{a(\alpha_h, \beta_h, k_h, m)(1 + (\alpha_h + \beta_h)^{2m}) - 2b(\alpha_h, \beta_h, m)},$$

$$b(\theta_h, m) = (\alpha_h(1 - (\alpha_h + \beta_h)^2) + \alpha_h^2(\alpha_h + \beta_h)) \frac{1 - (\alpha_h + \beta_h)^{2m}}{1 - (\alpha_h + \beta_h)^2},$$

$$a(\alpha_h, \beta_h, k_h, m) = m(1 - \beta_h)^2 + 2m(m - 1)$$
$$\times \frac{(1 - \alpha_h - \beta_h)^2(1 - (\alpha_h + \beta_h)^2 + \alpha_h^2)}{(k_h - 1)(1 - (\alpha_h + \beta_h)^2)} + 4$$

$$\times \frac{(m(1-\alpha_h-\beta_h)-1+(\alpha_h+\beta_h)^m)(\alpha_h(1-(\alpha_h+\beta_h)^2)+\alpha_h^2(\alpha_h+\beta_h))}{1-(\alpha_h+\beta_h)^2},$$

$$k_{mh} = 3 + \frac{k_h - 3}{m} + 6(k_h - 1) \times$$

$$\frac{(m(1-\alpha_h-\beta_h)-1+(\alpha_h+\beta_h)^m)(\alpha_h(1-(\alpha_h+\beta_h)^2)+\alpha_h^2(\alpha_h+\beta_h))}{m^2(1-\alpha_h-\beta_h)(1-(\alpha_h+\beta_h)^2+\alpha_h^2)}.$$

Let θ_{DN} be the transfer function corresponding to Theorem $A.1$ that transforms high-frequency parameters into low-frequency ones, i.e. $\theta_{\mathrm{DN}}(m, \theta_h)$. Theorem $A.1$ implies $\theta_{\mathrm{DN}}(\theta_{\mathrm{DN}}(m, \theta_h), n) = \theta_{\mathrm{DN}}(mn, \theta_h)$. Drost and Werker (1996) argue that the latter equality holds if the integers m and n are replaced by arbitrary real numbers. If a weak GARCH process with parameter θ_h is known to be the aggregate over m periods of some other higher-frequency GARCH process, then the parameter of the latter high-frequency process is given by $\theta_{h,n} = \theta_{\mathrm{DN}}(\theta_h, 1/m)$. If one assumes that the observed process at, say, frequency g is infinitely divisible, i.e. if one assumes that for each integer m there exists an underlying high-frequency GARCH process such that the observed process is the sum over m periods of the high-frequency process, then the transfer function q determines the parameters by $\theta_h = \theta_{\mathrm{DN}}(\theta_g, h/g)$.

References

Admati, A., Pfleiderer, P., 1988. A theory of intraday patterns: volume and price variability. Review of Financial Studies 1, 3–40.

Andersen, T.G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115–158.

Andersen, T.G., Sørenson, B.E., 1996. GMM estimation of a stochastic volatility model: a Monte Carlo study. Journal of Business and Economic Statistics 14, 328–352.

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.

Andrews, D.W.K., 1997. A stopping rule for the computation of generalized method of moments estimators. Econometrica 65, 913–931.

Andrews, D.W.K., 1999. Consistent moment selection procedures for generalized method of moments estimation. Econometrica 67, 543–564.

Andrews, D.W.K., Monahan, C., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. Econometrica 60, 953–966.

Bates, C., White, H., 1988. Efficient instrumental variables estimation of systems of implicit heterogeneous nonlinear dynamic equations with nonspherical errors. In: Barnett, W.A., Berndt, E.R., White, H. (Eds.), Dynamic Econometric Modelling. Cambridge University Press, Cambridge.

Bauwens, L., Giot, P., Grammig, J., Veredas, D., 2000. A comparison of financial duration models via density forecasts. CORE Discussion Paper 0060. Université Catholique de Louvain.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton.

Chen, S.X., 2000. Gamma kernel estimators for density functions. Annals of the Institute of Statistical Mathematics 52, 471–480.

Cox, D.R., Lewis, P.A.W., 1966. The Statistical Analysis of Series of Events. Wiley, London, New York.

Diamond, D.W., Verrecchia, R.E., 1987. Constraints on short-selling and asset price adjustments to private information. Journal of Financial Economics 18, 227–311.

Diebold, F.X., 1988. Empirical Modeling of Exchange Rates. Springer, New York.

Domowitz, I., 1992. Automating the price discovery process: some international comparisons and regulatory implications. Journal of Financial Sevices Research 6, 305–326.

Drost, F.C., Nijman, T.E., 1993. Temporal aggregation of GARCH processes. Econometrica 61, 907–927.

Drost, F.C., Werker, B.J.M., 1996. Closing the GARCH gap: continuous time GARCH modeling. Journal of Econometrics 74, 31–57.

Dufour, A., Engle, R.F., 2000. Time and the impact of a trade. Journal of Finance 55, 2467–2498.

Easley, D., O'Hara, M., 1991. Order form and information in securities markets. Journal of Finance 46, 905–927.

Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment. Journal of Finance 47, 577–605.

Engle, R.F., 2000. The econometrics of ultra high frequency data. Econometrica 68 (1), 1–22.

Engle, R.F., Lunde, A., 1999. Trades and quotes: a bivariate point process. Working paper, Universities San Diego and Aalborg.

Engle, R.F., Russell, J.R., 1995. Autoregressive conditional duration: a new model for irregularly spaced time series data. Discussion Paper 9427, University of California.

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

French, K.R., Roll, R., 1986. Stock return variances. The arrival of information and the reaction of traders. Journal of Financial Economics 17, 5–26.

Gallant, A.R., 1987. Nonlinear Statistical Models. Wiley, New York.

Gallant, A.R., Tauchen, G., 1996. Which moments to match? Econometric Theory 12, 657–681.

Ghysels, E., Jasiak, J., 1998. GARCH for irregularly spaced financial data: the ACD-GARCH model. Studies in Nonlinear Economics and Econometrics 2, 133–149.

Grammig, J., Maurer, K.-O., 2000. Non-monotonic hazard functions and the autoregressive conditional duration model. The Econometrics Journal 3, 16–38.

Hall, A., Peixe, F., 2001. A consistent method for the selection of relevant instruments. Working Paper, North Carolina State University.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hansen, L.P., 1985. A method for calculating bounds on the asymptotic variance–covariance matrix of generalized method of moments estimators. Journal of Econometrics 30, 203–228.

Hau, H., 1999. Information and geography: evidence from the German stock market. Discussion Paper 2297, CEPR.

Hawkes, A.G., 1971. Spectra of some self-exciting and mutually exciting point processes. Biometrika 58, 83–90.

Hayashi, F., Sims, C.A., 1983. Nearly efficient estimation of time series models with predetermined, but not exogenous, instruments. Econometrica 51, 783–798.

Koenker, R., Machado, J.A.F., 1999. GMM inference when the number of moment conditions is large. Journal of Econometrics, pp. 327–344.

Lunde, A., 1999. A generalized gamma autoregressive conditional duration model. Aalborg University.

Madhavan, A., 2000. Market microstructure: a survey. Journal of Financial Markets 3, 205–258.

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

O'Hara, M., 1995. Market Microstructure Theory. Basil Blackwell, Oxford.

Rich, R.W., Raymond, J., Butler, J.S., 1991. Generalized instrumental variables estimation of autoregressive conditional heteroskedastic models. Economics Letters 35, 179–185.

Rubin, I., 1972. Regular point processes and their detection. IEEE Transactions on Information Theory ITT-18, 547–557.

Russell, J., 1999. Econometric modeling of irregularly-spaced multivariate transaction data. Working Paper, University of Chicago, Graduate School of Business.

Russell, J.R., Engle, R.F., 1999. Econometric analysis of discrete-valued, irregularly-spaced financial transactions data. Working Paper, University of Chicago and University of San Diego.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman & Hall, London.

West, K.D., 1997. Another heteroskedasticity and autocorrelation consistent covariance matrix estimator. Journal of Econometrics 76, 171–191.

West, K.D., Wong, K., Anatolyev, S., 1998. Feasible optimal instrumental variables estimation of linear models with moving average disturbances. SSRI Working Paper 9901, University of Wisconsin.