The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks

Luc BAUWENS, Pierre GIOT *

ABSTRACT. – This paper introduces the logarithmic autoregressive conditional duration (Log-ACD) model and compares it with the ACD model of ENGLE and RUSSELL [1998]. The logarithmic version allows to introduce in the model additional variables without sign restrictions on their coefficients. We apply the Log-ACD model to price durations relative to the bid-ask quote process of three securities listed on the New York Stock Exchange, and we investigate the influence of some characteristics of the trade process (trading intensity, average volume per trade and average spread) on the bid-ask quote process.

Le modèle logarithmique auto-régressif de durée conditionnelle (Log-ACD) : une application aux processus des prix offerts et demandés de trois titres de la Bourse de New-York

RÉSUMÉ. – Ce papier introduit le modèle logarithmique auto-régressif de durée conditionnelle (Log ACD) d'ENGLE et RUSSELL [1998]. La version logarithmique permet d'introduire dans le modèle des variables supplémentaires sans restreindre le signe de leurs coefficients. Nous appliquons le modèle Log-ACD à des durées de prix construites à partir du processus d'annonces de prix offerts et demandés de trois titres cotés à la Bourse de New-York, et nous étudions l'influence de certaines caractéristiques du processus des échanges, comme l'intensité des échanges, le volume moyen par échange et le *spread*, sur le processus des annonces de prix.

---

* L. BAUWENS: CORE, Université Catholique de Louvain;
P. GIOT: CORE, Université Catholique de Louvain.

While remaining responsible for any errors in this paper, the authors would like to thank R. ANDERSON, J. JASIAK, G. LE FOL, A. ROELL, O. SCAILLET, N. SHEPHARD, S. WEI and an anonymous referee for useful remarks and suggestions, and Mr A. RUTTIENS from CERA Investment Bank for valuable discussions.

Support of the European Commission Human Capital and Mobility Program through the network *"Econometric inference using simulation methods"* is gratefully acknowledged.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsability is assumed by the authors.

1 Introduction

Over the past years, there has been a lot of research in financial market microstructure. This research has focused both on theoretical models and on empirical work. With the advent of high frequency databases and the wealth of information made available by these,[1] new fields of research are open. To deal with this kind of data, new econometric models had to be developed.

The so-called high frequency data econometric models originally appeared as *"fixed interval"* models: these belonged to the stochastic volatility and GARCH types of models, with data regularly sampled at a very high frequency. See, for example, HAFNER [1996], EDDELBUTEL and MCCURDY [1996] or the empirical work conducted by OLSEN *and Associates*, such as GUILLAUME, PICTET and DACOROGNA [1995], and GUILLAUME, DACOROGNA, DAVÉ, MULLER, OLSEN and PICTET [1997]. However, one main drawback of these models is that they do not take into account the irregular spacing of the data.[2]

Recently, ENGLE and RUSSELL [1997, 1998] proposed an econometric model for the durations between two successive market events, such as a buy or a sell of a security. ENGLE and RUSSELL find a clustering effect in the durations: short (respectively, long) durations tend to be followed by short (respectively, long) durations. This effect resembles the one found in the volatility of many financial series. ENGLE and RUSSELL call their model the ACD model, where ACD stands for *"Autoregressive Conditional Duration"*, and apply it to the foreign exchange market and to the IBM stock. In the same spirit, ENGLE [2000] combines such a duration model with a GARCH model for the returns, which provides a new way to model irregularly spaced data.

The ACD model raises several questions and calls for several extensions, such as:

(i) Are there alternative econometric models which can account for the structure exhibited by the durations? BAUWENS and VEREDAS [1999] put forward the stochastic conditional duration (SCD) model, which is the counterpart of the stochastic volatility model (in the same way as the ACD model is the counterpart of the GARCH model). As an alternative to the Weibull distribution, GRAMMIG and MAURER [1999] introduce an ACD model based on the Burr distribution. JASIAK [1998] considers the fractionally integrated ACD model which allows for long range dependence in the durations. GHYSELS, GOURIÉROUX, and JASIAK [1997] propose the stochastic volatility duration (SVD) model. This model

1. Nowadays, most exchanges (NYSE, NASDAQ, Paris Bourse, Chicago Mercantile Exchange, ...) make available complete databases of tick-by-tick data, which, depending on the exchange, give information on the trade process (time of the trade, price, volume) and the bid-ask quote process (time of the quotes, bid and ask quotes, depths) or the state of the order book.

2. As an alternative to regular sampling, OLSEN *and Associates* use time transformation techniques to transform the irregularly spaced data into fixed interval data.

introduces dependence in the durations through their conditional volatility, in addition to the dependence through their conditional expectation that is also present in the ACD and SCD models.

(ii) How can this model be used to model the way in which market variables (price, volume …) or features (price volatility) behave along a trading day? ENGLE'S paper [2000] is a first step to modelling irregularly spaced data. Other contributions are by GHYSELS and JASIAK [1998], BAUWENS and GIOT [1998], and RUSSELL and ENGLE [1998].

(iii) How can this model be used to test market microstructure hypotheses? This will typically require to include in the model other variables than lagged durations, such as volume, spread, volatility of the returns, in order to link the econometric model with testable hypotheses from the market microstructure literature.

In this paper, we focus on the first and third questions. We introduce a logarithmic version of the ACD model, called the Log-ACD model. In this version, the autoregressive equation is specified on the logarithm of the conditional expectation of the durations. This way of modelling the conditional behavior of the durations is more flexible than with the ACD model. Indeed we no longer have to impose non-negativity constraints on the coefficients of the autoregressive equation, as is needed in the ACD model to ensure positive expected durations. Although analytical results are not available for the moments and autocorrelations of the Log-ACD model, numerical Monte Carlo simulations and estimations of the model indicate that the Log-ACD model provides a very good alternative to the ACD model.

We use the logarithmic ACD model on the bid-ask quote process of three securities actively traded on the NYSE (BOEING, DISNEY, and IBM), in order to investigate the way specialists revise their beliefs on the bid and ask prices they fix. As expected, the bid-ask quote durations exhibit a highly autoregressive structure. Then, we link this quote revision process to three characteristics of the trade process: the trading intensity, the average volume per trade, and the average spread.[3] The empirical evidence is clearly in favor of the information models, such as the model of EASLEY and O'HARA [1992].

The rest of the paper is organized in the following way. In Section 2, we review some recent issues in market microstructure, both on the theoretical and empirical sides, focusing on the bid-ask quote revision process. In Section 3, we briefly review the ACD model of ENGLE and RUSSELL [1998] and we introduce the Log-ACD model. In Section 4, we apply the Log-ACD model to the bid-ask quote process for three NYSE stocks and examine the way specialists revise their bid-ask quote. We also link the specification of the Log-ACD model to characteristics of the trade process. Section 5 concludes.

3. Strictly speaking, the spread is a characteristic of the bid-ask quote process. However, it is also related to the trade process, as we use the existing spread when the past trades were made (see Section 4).

2 Market Microstructure Issues: Bid-Ask Prices and Market-Maker's Behavior

Over the past twenty years, there has been a considerable amount of research in market microstructure which has focused both on theoretical and on empirical models. In this section, we give a very brief summary of the models related to the behavior of market-makers and how they determine their bid and ask prices. O'Hara [1995], Biais, Foucault and Hillion [1997] and Goodhart and O'Hara [1997] are excellent surveys of the existing theoretical and empirical models developed in this field.

2.1 Theoretical Models

The first theoretical model explaining the market-maker's behavior was proposed by Garman [1976]. He considers a single monopolistic market-maker who is confronted with a succession of buy and sell orders, which are assumed to be independent stochastic processes. To avoid failure (bankruptcy and failure to provide for liquidity), this market-maker sets different buy (bid price) and sell prices (ask price). Garman's model was improved by Amihud and Mendelson [1980]. These early models are characterized by the fact that the market-maker is assumed to be a monopolist, whose bid and ask prices reflect his market power.

Stoll [1978] introduces the notion of a market-maker supplier of intermediary services. In his model, the market-maker is a market participant who buys and sells shares to other market participants. In doing so, he no longer has the optimal amount of wanted securities and thus faces an inventory risk. To safeguard himself against this risk, the market-marker buys and sells shares at different prices. Stoll's model was improved by Ho and Stoll [1981] who extend Stoll's model to a multi-period framework.

A major breakthrough was made in 1985 by Glosten and Milgrom with the introduction of information based models. In these models, traders and market-makers do not have the same information regarding the value of the security they are trading. Typically two kinds of traders are trading with the market-maker: informed and uninformed traders. Uninformed traders do not have superior information regarding the financial asset they are trading. They mainly trade for liquidity reasons. Informed traders, however, have superior information on the asset they are trading: they sell if they know bad news and they buy if they know good news. The market-maker, who is confronted with both types of traders, does not know if he deals with an informed or an uninformed trader. To protect himself from a possible incurring loss, the market-maker fixes different buy and sell prices: his buy price (the bid) is smaller than his sell price (the ask). Thus, the market-maker fixes his prices conditionally on the type of trade. In a multi-period framework, this gives rise to a Bayesian updating behavior.

Recently, EASLEY and O'HARA [1992] extended the GLOSTEN and MILGROM's model by focusing on the role of time in price adjustment. Indeed, in the GLOSTEN and MILGROM's model, time does not matter: it is exogenous to the price process. EASLEY and O'HARA argue that the duration between two trades conveys information. In their model, a no-trade outcome (a long duration) means that no new information has been released. Thus, the probability of dealing with an informed trader is small (relative to the case where the duration would be small). Consequently, with a low probability of dealing with an informed trader, the market-maker decreases his bid-ask spread. Their model has several important consequences:

– time is no longer exogenous to the price process: *"empirical investigations using transaction data will be biased because examining only transaction prices ignores the information content contained in the nontrading intervals"* (O'HARA [1995]);

– the sequence of prices matters and is informative;

– volume brings valuable information to the market-maker.[4]

Another consequence of their model is that the release of news (information event) should lead to an increase in the trading intensity and this should imply more frequent revisions of the bid-ask prices posted by the market-makers: *"... quotes converge to their strong form efficient values at exponential rate. Rates of convergence are increasing in the fraction of trades from the informed."* (EASLEY and O'HARA [1992]).

2.2 Empirical Research

The amount of research in empirical market microstructure is very large (although relatively few empirical studies are available using high frequency data). We do not attempt to give a detailed description of this field of research; an excellent introduction to this field can be found in CAMPBELL, LO and MACKINLAY [1997] and a recent survey is available in GOODHART and O'HARA [1997]. With the advent of high frequency database and the availability of all trades and quotes for a given security over a period of time, it is now possible to study more closely the price formation process and the way market variables such as the volume, the bid-ask prices and the spread behave along the trading day.

Recently, there has been a growing increase in empirical investigations of market microstructure models using high frequency econometric models. HAUSMAN, LO and MACKINLAY [1992] use an ordered probit model to examine such issues as the effect of a sequence of prices changes, the effect of trade size and the effect of price discreteness. BROCK and KLEIDON [1992] introduce a model of intraday bids and asks and examine the effect of stock market open and closure on these prices and on the resulting spread. Their

4. Thanks to the information models, volume (and especially unexpected volume) and the much critized technical analysis are found to have important implications on the behavior of the market-makers. See, for example, EASLEY and O'HARA [1992] or BLUME, EASLEY and O'HARA [1994].

analysis focuses on the NYSE and on the behaviour of the specialist. As pointed out by CHAN, CHRISTIE and SCHULTZ [1995] who conduct the same analysis for stocks listed on the NASDAQ market, there are key differences between a market with a monopolistic specialist (such as NYSE) and a market with competing market-makers (such as NASDAQ). This is also highlighted by GWILYM, BUCKLE and THOMAS [1997] who investigate the same problem for the FTSE-100 Stock Index Options (which are traded at LIFFE, a market with competing market-makers).

In a careful study on stocks traded on the NYSE, and using databases keeping track of the inventory holdings of the specialists, MADHAVAN and SOFIANOS [1998] examine the importance and relative share of dealer trading by NYSE specialists. They conclude that the proportion of the trades involving the specialist is relatively weak and depends very much on the stock. HASBROUCK [1988, 1991] investigates the interaction between security trades and bid-ask quote revisions for stocks listed on the NYSE and highlights the information content of trades. Information models are also dealt with in EASLEY, KIEFER and O'HARA [1997], as they extend the EASLEY and O'HARA [1992] paper and estimate the model with intraday data for NYSE stocks.

As pointed out in the introduction, ENGLE and RUSSELL [1998] introduced the ACD model to take into account the irregular spacing of the data; they also tested for the significance of additional market activity variables in their model. ENGLE and LANGE [1997] propose a new measure of liquidity, which is related to the depth of the market; in their paper, they combine an ACD model with a regression equation involving market microstructure variables in order to better explain the variation in liquidity over a trading day.

Regarding the microstructure of currency trading, BOLLERSLEV and DOMOWITZ [1993] examine the trading patterns and prices in the Interbank Foreign Exchange Market; they introduce GARCH models for price volatility and spread, that feature market activity variables as additional explicative variables. A survey of the recent developments in this field is given in GUILLAUME, DACOROGNA, DAVÉ, MULLER, OLSEN and PICTET [1997].While most studies focus on stocks traded on the NYSE and on currency trading, some recent publications also take a close look at the Paris Bourse; examples are BIAIS, HILLION and SPATT [1995] and BISIÈRE and KAMIONKA [2000].

3 The ACD and Logarithmic ACD Models

In this section, we start by a brief review of the ACD model introduced by ENGLE and RUSSELL [1998]. This model shares some features of the GARCH model. Instead of modelling an autoregressive process on the variance of the returns (as in the GARCH model), the ACD model bears on the autoregressive structure exhibited by the durations. This model is particularly well suited to the analysis of irregularly spaced data, such as stock market data, where the time elapsed between two trades conveys information.

In the second part of this section, we introduce a logarithmic version of the ACD model. This model is close to the original ACD model, but is more flexible as there are no sign constraints on the parameters of the autoregressive equation. In the third part, we compare the ACD and Log-ACD models through Monte Carlo simulations

In this paper, as we focus on the bid-ask quote process, a duration (x_i) is the time elapsed between two consecutive bid-ask quotes, which are released at times t_{i-1} and t_i, *i.e.* $x_i = t_i - t_{i-1}$.

3.1 The ACD Model and its Properties

3.1.1 Structure of the Model

Let x_i be the duration between two quotes (a quote being a collection of data relative to a buy or a sell of a security on a stock exchange). The assumption introduced by ENGLE and RUSSELL [1998] is that the time dependence in the durations can be subsumed in their conditional expectations $\Psi_i = E(x_i|I_{i-1})$, in such a way that x_i/Ψ_i is independent and identically distributed. I_{i-1} denotes the information set available at time t_{i-1} (*i.e.* at the beginning of duration x_i), supposed to contain at least $\tilde{x}_{i-1}$ and $\tilde{\psi}_{i-1}$, where $\tilde{x}_{i-1}$ denotes x_{i-1} and its past values, and likewise for $\tilde{\psi}_{i-1}$.

The ACD model specifies the observed duration as a mixing process:

$$x_i = \Phi_i \epsilon_i, \tag{1}$$

where the ϵ_i are IID and follow a Weibull (1, γ) probability distribution, while the Φ_i are proportional to the conditional expectation of x_i as explained below.

A second equation specifies an autoregressive model for the (expected) conditional durations:[5]

$$\Psi_i = \omega + \alpha x_{i-1} + \beta \Psi_{i-1} \tag{2}$$

with the following constraints on the coefficients: $\omega > 0$, $\beta \geqslant 0$, $\alpha \geqslant 0$ and $\alpha + \beta < 1$. The last constraint ensures the existence of the unconditional mean of the durations, the others ensure the positivity of the conditional durations.

The condition $\Psi_i = E(x_i|I_{i-1})$ provides us with a third equation, linking equations (1) and (2):

$$\Psi_i = \Gamma(1 + \frac{1}{\gamma})\Phi_i, \tag{3}$$

where $\Gamma(.)$ is the gamma function. If $\gamma = 1$, the Weibull distribution becomes an exponential one. In this case, $\Phi_i = \Psi_i$.

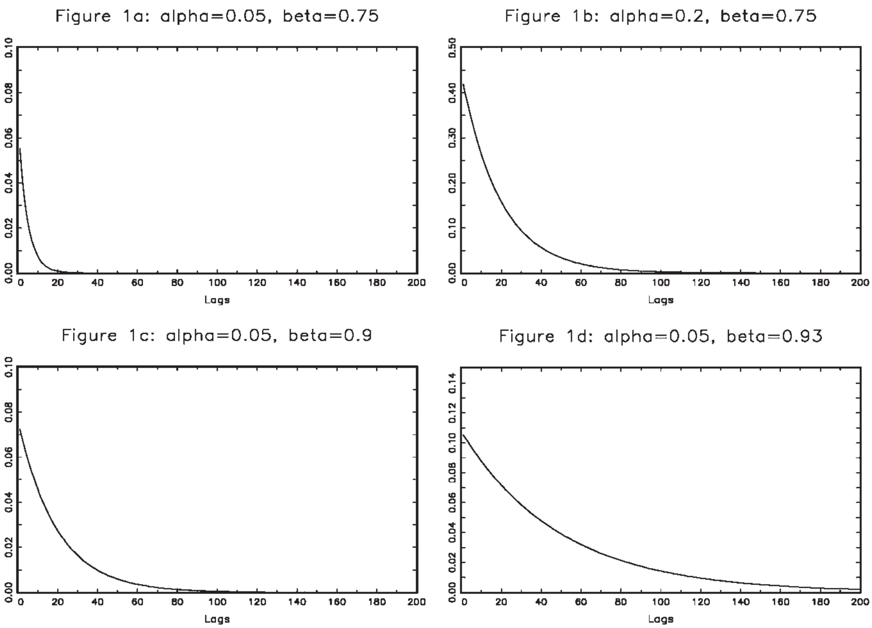
5. This model is the ACD (1,1). More lags of x_i and Ψ_i can be added. As in this paper we use only one lag, we use the short notation ACD.

TABLE 1
Overdispersion implied by the ACD Model

Coefficients	Overdispersion Ratio
$\beta = 0.75$, $\alpha = 0.05$	1.01
$\beta = 0.75$, $\alpha = 0.10$	1.04
$\beta = 0.75$, $\alpha = 0.20$	1.55
$\beta = 0.90$, $\alpha = 0.05$	1.03
$\beta = 0.90$, $\alpha = 0.08$	1.18
$\beta = 0.93$, $\alpha = 0.05$	1.07

The overdispersion ratio is the ratio standard deviation/mean, computed using (4) and (5). In all cases, $\omega = 1 - \alpha - \beta$ and $\gamma = 1$.

FIGURE 1
Autocorrelation Functions of the ACD Model



The autoregressive structure on the conditional expectation of the durations implies that small durations are more likely to be followed by small durations (and likewise for long durations). Thus, the model accounts for a clustering effect on the durations. The Weibull distribution is of course more flexible than the exponential one. If $\gamma < 1$, the model exhibits a decreasing hazard function: long durations will be less likely. If $\gamma > 1$, long durations will be more likely.

3.1.2 Statistical Properties of the ACD Model

By definition, the conditional expectation of x_i is equal to Ψ_i. Thus equation (2) gives us a way to forecast expected durations, based on the information set at the previous period; see ENGLE and RUSSELL [1997] for a detailed discussion. The unconditional expectation (μ) and variance (σ^2) of x_i are given by

$$\mu = E(x_i) = \frac{\omega}{1-\alpha-\beta} \tag{4}$$

and

$$\sigma^2 = \mu^2 \kappa \frac{1-2\alpha\beta-\beta^2}{1-(\alpha+\beta)^2-\alpha^2\kappa} \tag{5}$$

(provided that the denominators are positive), where

$$\kappa = \frac{\Gamma(1+2/\gamma)}{\Gamma(1+1/\gamma)^2} - 1. \tag{6}$$

From equation (5), we can see that σ is greater than μ, whenever α is greater than 0: the model can account for overdispersion.[6] For a proof of these results, see ENGLE and RUSSELL [1998]. In the appendix, we show how to compute the autocorrelation function (ACF) of the durations by a recursive formula.

3.1.3 Numerical Illustrations

In the case of the ACD model, the first and second unconditional moments, and the autocovariances can be computed analytically as shown above. It is nevertheless useful to give numerical results about these moments and autocovariances for several sets of parameters.

For several sets of parameters and using (4), (5) and the results given in the appendix, Table 1 gives the degree of overdispersion (measured by the ratio standard deviation/mean) while Figure 1 plots the autocorrelation function for six sets of parameters.[7]

As it is the case for the ARCH and GARCH class of models, a slowly decaying autocorrelation function requires β to be close to one. A large degree of overdispersion requires a *"large"* value of α.

3.2 The Logarithmic ACD Model

Although the ACD model has been applied to high frequency tick-by-tick data (IBM stock trades on NYSE by ENGLE and RUSSELL [1998], and foreign exchange prices by ENGLE and RUSSELL [1997]), the positivity constraints on

6 By definition, a series of observations is overdispersed when its standard deviation is greater than its mean.

7. We use an ACD (1,1) model with the unconditional mean set equal to one, *i.e.* $\omega = 1 - \alpha - \beta$.

the coefficients of the model may be quite restrictive. In particular, if we want to include additional explanatory variables in the autoregressive equation (2), we must ensure that the right-hand side of (2) remains strictly positive. As explained in Section 4, this can be a problem when additional variables suggested by market microstructure theories are included linearly in equation (2).[8] Our motivation for the Log-ACD model is thus to put forward a duration model, capitalizing on the ACD model, but with more flexibility.

3.2.1 Structure of the Model

As in the ACD model, let x_i be the duration between two quotes. The logarithmic version of the ACD model changes the mixing process (1) of the ACD model into the following equation:

$$x_i = e^{\phi_i} \varepsilon_i, \tag{7}$$

where the ε_i are IID and follow a Weibull$(1,\gamma)$ distribution, while ϕ_i is proportional to the logarithm of the conditional expectation of x_i as explained below.

Let ψ_i be the logarithm of the conditional expectation of x_i, so that $\psi_i = \ln E(x_i | I_{i-1})$. A second equation specifies an autoregressive model for the logarithm of the conditional durations:

$$\psi_i = \omega + \alpha g(x_{i-1}, \epsilon_{i-1}) + \beta \psi_{i-1}. \tag{8}$$

For positivity of e^{ψ_i} and thus of x_i, there are no restrictions on the sign of the parameters ω, α and β.

The condition $\psi_i = \ln E(x_i | I_{i-1})$ or $e^{\psi_i} = E(x_i | I_{i-1})$ provides us with a third equation, linking equations (7) and (8):

$$e^{\phi_i} \Gamma(1 + 1/\gamma) = e^{\psi_i}. \tag{9}$$

We can propose several choices of $g(x_{i-1}, \epsilon_{i-1})$, in particular:

(i) $g(x_{i-1}, \epsilon_{i-1}) = \ln x_{i-1}$. This gives rise to a Log-ACD model, where the logarithm of the conditional expectation depends on its past lagged value and on the lagged logarithm of the durations:

$$\psi_i = \omega + \alpha \ln x_{i-1} + \beta \psi_{i-1}. \tag{10}$$

This model is analogous to the Log-GARCH model (GEWEKE [1986]). For covariance stationarity of $\ln x_i$, $|\alpha + \beta|$ must be smaller than one. Using (7) and (9), equation (10) can also be written as

$$\psi_i = \omega' + \alpha \ln \varepsilon_{i-1} + (\alpha + \beta) \psi_{i-1} \tag{11}$$

where $\omega' = \omega - \alpha \ln[\Gamma(1 + 1/\gamma)]$. We call this model the Log-ACD$_1$ model.

8. The additional variables could be added in a non-linear way, such that the expected duration is surely positive. However, this raises the issue of which functional form to use. The logarithmic ACD model allows to enter the variables linearly even they have negative coefficients.

(ii) $g(x_{i-1}, \epsilon_{i-1}) = \epsilon_{i-1}$. With this choice, equation (8) can be written as

$$\psi_i = \omega + \alpha\varepsilon_{i-1} + \beta\psi_{i-1} = \omega + \alpha \frac{x_{i-1}\Gamma(1+1/\gamma)}{e^{\psi_{i-1}}} + \beta\psi_{i-1}. \tag{12}$$

In this specification, the logarithm of the conditional expectation depends on its past lagged value and on the lagged *"excess duration"*. This model is close to the exponential GARCH model[9] of NELSON [1991]. For covariance stationarity of ψ_i, $|\beta|$ must be smaller than one. Hereafter, this version of the model is called the Log-ACD$_2$ model.

By definition of the Weibull density, the density function of x_i can be written as

$$f(x_i) = \frac{\gamma}{x_i}\left(\frac{x_i\Gamma(1+1/\gamma)}{e^{\psi_i}}\right)^{\gamma} e^{-\left(\frac{x_i\Gamma(1+1/\gamma)}{e^{\psi_i}}\right)^{\gamma}}. \tag{13}$$

Using (13) we can write the log-likelihood function of the observations x_i, $i = 1 \ldots N$ as

$$\sum_1^N \ln(\gamma) - \ln(x_i) + \gamma\ln[x_i\Gamma(1+1/\gamma)] - \gamma\psi_i - \left(\frac{x_i\Gamma(1+1/\gamma)}{e^{\psi_i}}\right)^{\gamma}, \tag{14}$$

with ψ_i defined by (10) or (12). The estimation by the maximum likelihood method is then straightforward (as an initial condition, x_0 and e^{ψ_0} can be set equal to the unconditional mean of the x_i). The choice of the function g only affects the log-likelihood function through ψ_i.

3.2.2 Statistical Properties of the Log-ACD Model

By definition, the conditional expectation of x_i is equal to e^{ψ_i}. Thus equation (10) or (12) gives us a way to forecast expected durations, based on the information set at the previous period.

Unfortunately (as for the exponential GARCH model), no analytical expressions seem to be available for the unconditional moments. Indeed, if we try to compute the expectation of x_i using equation (10) for ψ_i, we arrive at

$$\mu = E(x_i) = E[(E(x_i|I_{i-1})] = E(e^{\psi_i}) = E(e^{\omega+\alpha\ln x_{i-1}+\beta\psi_{i-1}}), \tag{15}$$

which cannot be analytically computed. It seems also impossible to derive an analytical expression for the autocorrelation function. The same difficulties arise if we take equation (12) for ψ_i.

---

9. We refrain from calling our model the *exponential* ACD model since this term is used by ENGLE and RUSSELL [1998] for the ACD model of Section 3.1, when ε_i has an *exponential* distribution.

TABLE 2
Overdispersion of the Log-ACD $_1$

Coefficients	Overdispersion Ratio
$\beta = 0.75$, $\alpha = 0.05$	1.01
$\beta = 0.75$, $\alpha = 0.10$	1.05
$\beta = 0.75$, $\alpha = 0.20$	1.58
$\beta = 0.90$, $\alpha = 0.05$	1.04
$\beta = 0.90$, $\alpha = 0.08$	1.25
$\beta = 0.93$, $\alpha = 0.05$	1.10

Average overdispersion ratio of the Log-ACD $_1$ *based on Monte-Carlo simulations (200 samples of size 10,000) for the given sets of parameters. Log-ACD* $_1$ *corresponds to (7) and (10). In all cases,* γ *is equal to 1 while* ω *is chosen so that the empirical mean is close to 1.*

TABLE 3
Overdispersion of the Log-ACD $_2$

Coefficients	Overdispersion Ratio
$\beta = 0.90$, $\alpha = 0.05$	1.02
$\beta = 0.90$, $\alpha = 0.10$	1.06
$\beta = 0.90$, $\alpha = 0.20$	1.15
$\beta = 0.98$, $\alpha = 0.05$	1.07
$\beta = 0.98$, $\alpha = 0.08$	1.29
$\beta = 0.98$, $\alpha = 0.15$	1.71

Average overdispersion ratio of the Log-ACD $_2$ *based on Monte-Carlo simulations (200 samples of size 10 000) for the given sets of parameters. Log-ACD* $_2$ *corresponds to (7) and (12). In all cases,* γ *is equal to 1 while* ω *is chosen so that the empirical mean is close to 1.*

3.2.3 Numerical Simulations

In the case of the two versions of the Log-ACD model, the unconditional moments and the autocovariances cannot be computed analytically. To gain insight about the two models, it is thus interesting to conduct numerical simulations with several sets of parameters. In these numerical simulations, we focus on the overdispersion ratio and the general shape of the autocorrelation function associated with the models.

For several sets of parameters, Table 2 gives the numerical values taken by the overdispersion ratio (the unconditional mean is approximately set equal to one) for the Log-ACD$_1$ model, while Table 3 gives these results for the Log-ACD$_2$ model. The given results are average results for the overdispersion ratio based on Monte Carlo simulations for the given sets of parameters (200 samples of size 10,000). In both Log-ACD models, the parameter α has an important impact on the overdispersion: a *"large"* value of α is needed for the model to imply a large degree of overdispersion.

Figure 2 (for the Log-ACD$_1$) and Figure 3 (for the Log-ACD$_2$) give the empirical autocorrelation functions computed using Monte Carlo simulations (the given autocorrelation functions are averages, over 200 samples, of the autocorrelation functions associated with samples of size 10,000); in each case, four plots are given, one for each set of parameters. In the case of the second version of the Log-ACD model, it can be seen that a slowly decaying autocorrelation function is achieved with a value of β very close to one.

3.3 A Comparison of the Models

In order to compare the two models, we conduct Monte Carlo simulations. For the original ACD model and for the two specifications of the Log-ACD model, we generated 200 sets of 6,728 durations and we compare the empirical moments, autocorrelation functions, overdispersion ratios, and ranges of these durations. Using a numerical simulation, we also compare the forecasted durations (as given by the ACD and Log-ACD models) following an unexpected shock in a duration.

Examples of the autocorrelation functions of the ACD and the two specifications of the Log-ACD model are given in Figure 4. For comparison and anticipating Section 4, we also show the autocorrelation function of the IBM durations (Figure 4d), one of the stocks on which the empirical part of this paper focuses on. As the coefficients ω, α, β and γ do not have the same meaning in the three specifications, we chose values for these coefficients equal to their estimated values for the IBM stock.[10] The average results for the mean, standard deviation, minimum value, maximum value, overdispersion ratio, and first-order autocorrelation coefficient are given in Table 4. The last column of this table gives the results for the IBM stock durations.

The main conclusions of the comparisons are that:

- the ACD model and the two specifications of the Log-ACD model can account equally well for a slowly (but geometrically) decaying ACF, starting at a relatively low first autocorrelation (notice that the three ACF start too high in Figure 4a and 4b (compared to the ACF of the IBM data in Figure 4d);
- the ACF of the Log-ACD$_2$ model is closest to the ACF of the IBM data; of all three specifications, this one has the best fit as far as the ACF is concerned;[11]
- all three specifications exhibit overdispersion;
- for the selected parameters, the ACD model has the largest degree of overdispersion and the largest first autocorrelation coefficient, but it

10. See Table 10 for the estimates of the ACD and Log-ACD$_2$ models. For the Log-ACD$_1$, the estimates are $\widehat{\omega} = 0.040$, $\widehat{\alpha} = 0.076$, $\widehat{\beta} = 0.917$ and $\widehat{\gamma} = 0.994$.

11. Although the ACF for the three specifications are given for one sample of each model, this is a recurrent feature with different samples.

FIGURE 2
Autocorrelation Functions of the Log-ACD $_1$ ***Model***

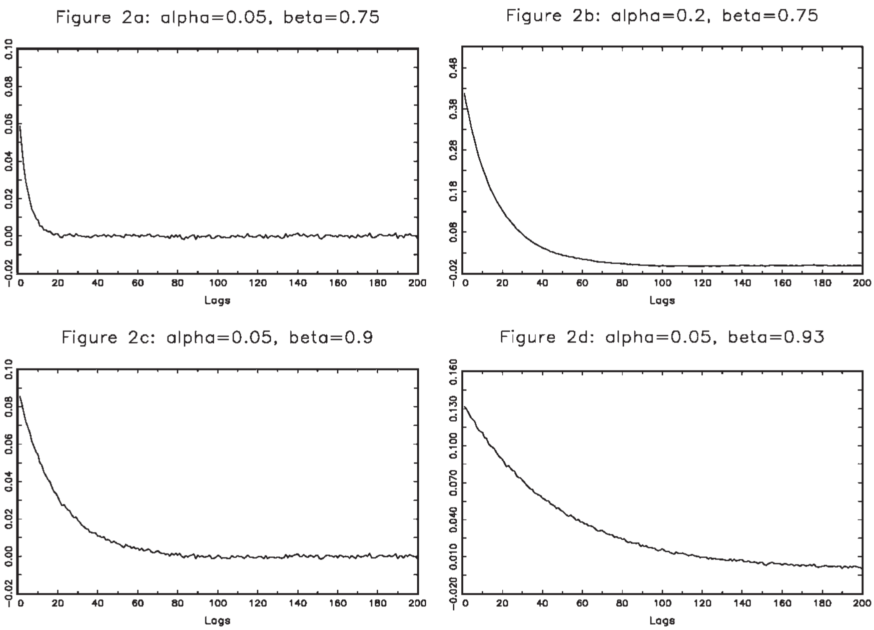
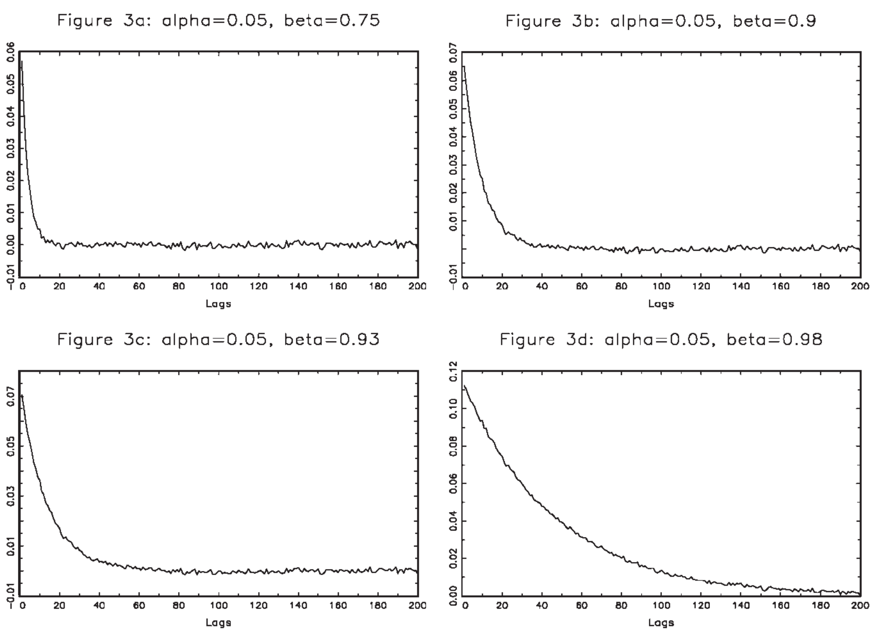


FIGURE 3
Autocorrelation Functions of the Log-ACD $_2$ ***Model***



has a mean much higher than the required mean (1.43 instead of 1.02), and it overestimates the overdispersion ratio (1.68 instead of 1.43). The Log-ACD$_2$ model has a mean that is very close to one, but has an overdispersion ratio below that of the empirical data (1.28 instead of 1.43). For both specifications of the Log-ACD model, the standard deviation is much closer to the empirical standard deviation of the IBM data than for the ACD model. It seems that the ACD model produces too many high durations.

Regarding the numerical simulations and the plots of the autocorrelation functions of the three models, the second version of the Log-ACD model appears to fit the data best. Indeed, its ACF is closest to the ACF of the data, its mean very closely matches the mean of the data, and its standard deviation and overdispersion ratio are not too far from their empirical counterparts. Regarding the general shape exhibited by the ACF of the data and while all three models feature a slowly decreasing ACF, it could be argued that the IBM data feature *"long-memory"*, which is not captured in the estimated models.[12] This issue is discussed more fully in Section 4.

Finally, we perform a numerical simulation to assess the impact of an unexpected shock in an observed duration on the subsequent forecasted durations. Using the estimated ACD and second version of the Log-ACD model, we compute the forecasted durations using as input the durations set to one except at the shocking time where the duration is set to one-tenth. Because both models are at the equilibrium at the start of the simulation, this is equivalent to shocking the models with a duration equal to one-tenth of the forecasted value.[13] The experiment is then repeated with the duration shock set to ten times its expected value. Figures 5a and 5b give the forecasted durations (as ratios of the long-run solutions) by the two models, with the shock occuring at the 100th observation.

As could be expected, a very small (large) observed duration at $t = 100$ leads to smaller (larger) forecasted durations in the subsequent periods. The forecasted durations gradually tend to their equilibrium values, indicating that the models are stationary for the given parameter values.[14] At $t = 140$, both systems are back to their long-run equilibrium values. The differences between the ACD and Log-ACD models are quite small. As indicated by Figure 5a, the Log-ACD model underreacts with respect to the ACD model when the duration is smaller than expected, while it overreacts when it is larger. However, even in the extreme cases considered here (ten times the expected value), the difference is rather small. Other simulations conducted with a shock equal to five times the expected value indicate that the difference is almost non-existent.

12. As reported in the next section, this seems to be a specific feature of the IBM durations.
13. In the market microstructure setting introduced in Section 2, this is as if the market-maker observed a duration equal to one-tenth of the expected duration. The subsequent forecasted durations can then be interpreted as the market response to the surprise in the length of the duration.
14. The ACD is weakly stationary if $1 - (\alpha + \beta)^2 - \alpha^2\kappa > 0$. For the Log-ACD, stationarity conditions are not known since the moments are not available. Heuristically, a necessary condition is that $|\beta| < 1$.

TABLE 4

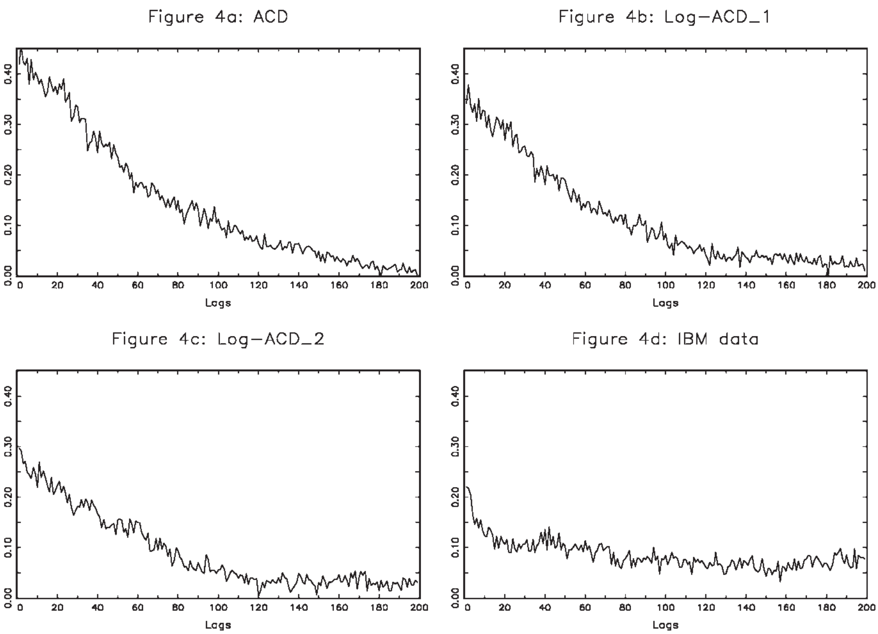
ACD and Log-ACD: A Comparison

	ACD	Log-ACD$_1$	Log-ACD$_2$	IBM
Mean	1.43	0.75	1.02	1.02
Standard deviation	2.55	1.27	1.30	1.45
Overdispersion	1.68	1.68	1.28	1.43
Minimum value	1.21e-4	5.46e-5	1.16e-4	0.003
Maximum value	47.81	22.75	18.97	29.12
First autocorrelation	0.42	0.34	0.30	0.22

Average results (except in the last column) based on Monte-Carlo simulations (200 samples of size 6,728). Log-ACD$_1$ corresponds to (7) and (10), and Log-ACD$_2$ to (7) and (12). The parameters (ω,α,β and γ) are set equal to their estimated value for the IBM data. The last column gives the statistics for the IBM data.

FIGURE 4

Autocorrelation Functions of 3 Simulated Data Sets Corresponding to ACD, Log-ACD$_1$ and Log-ACD$_2$ Specifications, and for the IBM Data.



4 An Application to the Bid-Ask Quote Process

In this section, we use the ACD and Log-ACD models[15] to examine the way market-makers revise their beliefs about their buy and sell prices. Market microstructure theory has already dealt extensively with this subject, as was outlined in Section 2. Keeping in mind these theoretical models, we also extend the Log-ACD model with additional variables in the autoregressive equation, so as to test some assumptions about the behavior of the market-makers.

An important feature of our approach is that we link the bid-ask quote dynamics to characteristics of the trade process. The bid-ask quote revision process is closely linked to the concept of liquidity[16] as the frequency of quote revisions is an important characteristic of the state of the market. We consider three variables related to the trade process: the trading intensity, the average volume per trade, and the average spread when the past trades were made. Thus, our model links the bid-ask quote revision process to characteristics of the trades that are actually carried out.

After a description of the database and the data transformations, we introduce the three additional explicative variables related to the trade process which are used in the specification of the Log-ACD model. Next, we report estimates for the ACD and Log-ACD models for the three stocks (BOEING, DISNEY and IBM) considered in this paper.

4.1 Data and Transformations

We worked on three securities actively traded on the NYSE: BOEING, DISNEY and IBM. The data was extracted from the Trade and Quote (TAQ) database released by the NYSE. Actually, this database consists of two parts: the first reports all trades, while the second lists the bid and ask prices posted by the specialists (on the NYSE). All records listed in the TAQ database are not valid (all trades have a correction and a condition indicator, giving information on the validity of the trade). Prior to using the data, we thus selected the *"regular"* trade quotes. Trades and bid-ask quotes recorded before 9:30 am and after 16:00 pm were also deleted. Notice that we treat the data consecutively from day to day.[17]

15. We use the second specification of the Log-ACD model introduced in Section 3. Although both specifications yield similar results, the second specification better captures the structure of the autoregressive process of the bid-ask quote process, as explained in Section 4.3. Therefore in this section we use the acronym Log-ACD for Log-ACD$_2$.

16. Kyle [1985] introduced three notions of liquidity: tightness (bid-ask spread), depth (amount of one sided volume that can be absorbed by the market without causing a revision of the bid-ask prices) and resiliency (speed of return to the equilibrium).

17. If the last event of day j is at 15:59:40, and the first of the next day is at 9:30:02, the duration between these two events is not used, so that the first duration for day $j+1$ will be the one between 9:30:02 and the next event.

TABLE 5

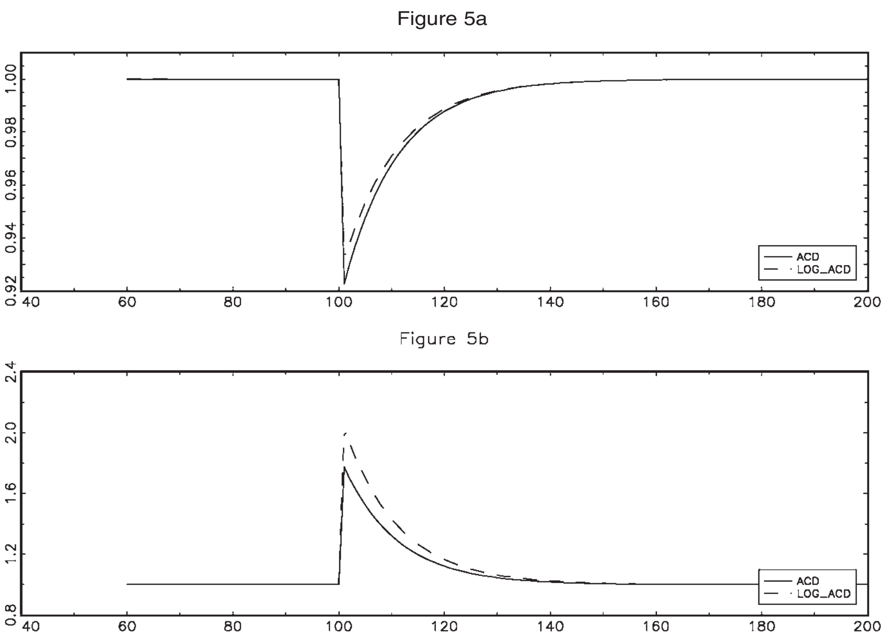
Information on a Data

	BOING	DISNEY	IBM
Number of trades	24.143	33.146	61.063
Number of b/a quotes	2.620	2.160	6.728
Mean of x_i and X_i	1 - 533.53	1 - 646.98	1.01 - 215.92
Standard deviation of x_i and X_i	1.34 - 867.83	1.21 - 966.34	1.45 - 365.46
Minimum value x_i and X_i	0.0048 - 3	0.0054 - 3	0.0033 - 1
Maximum value x_i and X_i	18.94 - 13059	14.45 - 9739	29.12 - 7170

Data extracted from the September, October and November 1996 TAQ CD-ROM. The given number of bid-ask quotes is the number obtained after filtering the data (the original number of bid-ask quotes is equal to 17,150 for BOEING, 37,325 for DISNEY and 34,321 for IBM). x_i is a time-of-day adjusted duration, see (16), and is measured in seconds. The mean of x_i is almost equal to 1, after the removal of the time-of-day effect. X_i are the non-adjusted durations.

FIGURE 5

Impact of a Smaller (figure 5 a) and Larger (figure 5b) than expected Duration on Subsequent Forecasted Durations by the ACD and Log-ACD$_2$ Models



In the IBM data, the number of bid-ask quotes (34,321) is approximately half the number of trades (61,063). For BOEING and DISNEY, the number of bid-ask quotes is 17,150 and 37,325, while the number of trades is equal to 24,143 and 33,146, respectively. Information about the data is given in Table 5.

For all three stocks, we computed and used price durations (see ENGLE and RUSSELL [1998], or GIOT [1999]). Price durations are defined by filtering the bid-ask quote durations and retaining those leading to a significant cumulated change in the mid-price of the specialist's quote. A significant change in the mid price of the specialist is defined as a change leading to at least a $0.125 cumulated change in the mid price. Thus, we did not take into account the numerous $0.0625 changes in the mid price, which are due to a $0.125 price change of the bid or the ask. Of course, two successive $0.0625 changes in the same direction yield a cumulative $0.125 change (and thus lead to a retained duration). The filtering can be justified by the presumption that the $0.0625 changes are transitory, *i.e.* are mainly due to the short term component of the bid-ask quote updating process. Indeed, BIAIS, HILLION and SPATT [1995] provide evidence that information effects in the order process lead to similar (successive) changes in quotes on both sides of the market (*i.e.* information events quickly lead to movements of the bid and ask quotes in the same direction).

As explained by ENGLE and RUSSELL [1998], it is necessary to transform the raw durations prior to using the ACD model. Indeed these durations can be thought of as consisting of two parts: a stochastic component (which is explained by the ACD model) and a deterministic part, which they call a *"time-of-day"* effect. This time-of-day effect arises from the systematic variations of the quote arrivals over the trading day. This deterministic effect should be extracted from the raw durations. We follow ENGLE and RUSSELL in defining this deterministic effect as a multiplicative component, *i.e.*

$$X_i = x_i \, \phi(t_i), \tag{16}$$

where X_i is the raw duration, $\phi(t_i)$ is the time-of-day effect and x_i denotes the time-of-day (tod) adjusted duration. The deterministic tod effect is defined as the expected duration conditioned on time-of-day and on the day of the week (so that, for example, the time-of-day effect of Monday can be different from the time-of-day of Tuesday), where the expectation is computed by averaging the durations over thirty minutes intervals for each day of the week. Cubic splines are then used on the thirty minutes intervals to smooth the time-of-day function.[18]

The time-of-day functions for the five days of the week are given in Figure 6 (for IBM). The well known inverted U shape documented for the trade durations (the trading activity is much higher in the morning and in the end of the afternoon than around lunch time) is also exhibited by the price durations. The corresponding functions of the other stocks display the same shape as for IBM.

18. An alternative approach is to estimate jointly (by maximum likelihood) the parameters of the ACD model and the parameters of the time-of-day functions. ENGLE and RUSSELL [1998] report that the two approaches yield almost the same results, which is why we use the simplest one.

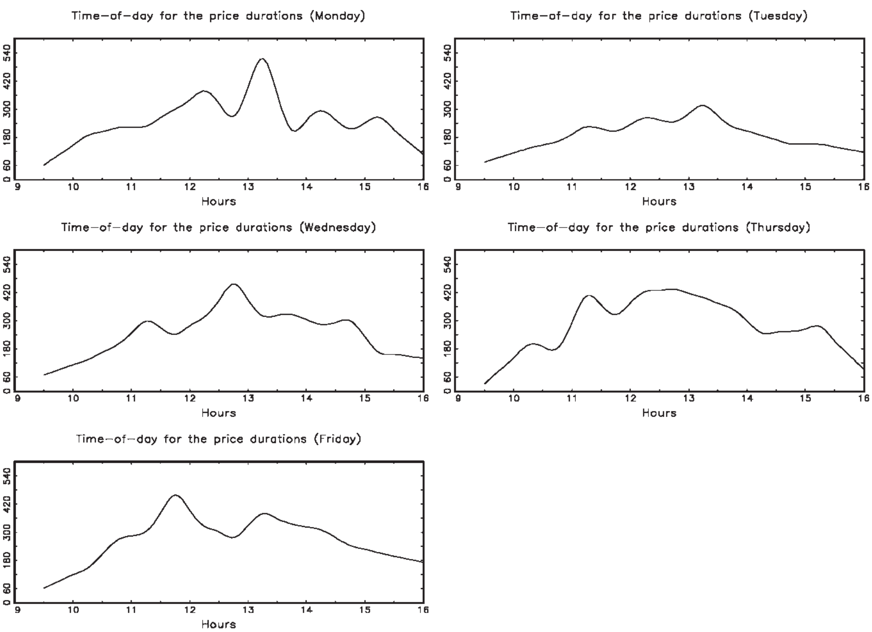
4.2 Additional Explicative Variables Related to the Trade Process

The specification of the conditional expectation of duration x_i given by equation (12) does not use possible relevant information that may be included in the information set available to market participants at the start of duration x_i. Indeed, as given by (12), the model is *"self-contained"*: durations depend on past durations through an autoregressive process. In order to test market microstructure hypotheses, we need to add explicative variables in the autoregressive equation. We focus on three variables related to characteristics of the trade process: the trading intensity, the average volume per trade, and the existing spread when the past trades were made. The choice of these variables is suggested by the information models developed in the recent market microstructure literature as explained hereafter.

4.2.1 Trading Intensity

Over each price duration, the trading intensity is defined as the number of trades recorded during the price duration divided by the length of the duration (X_i). Thus, for example, a large number of trades over a short duration leads to a high trading intensity. As briefly explained in Section 2, the model of Easley and O'Hara [1992] implies that an increase in the trading intensity (due for example to the release of news) should lead to more frequent revi-

Figure 6
Time-of-Day and Day-of-Week Functions for IBM Price Durations



sions of the quotes. By adding this variable, the specification of ψ_i in the Log-ACD model is extended to

$$\psi_i = \omega + \alpha\epsilon_{i-1} + \beta\psi_{i-1} + \eta_1\, tint_{i-1}, \tag{17}$$

where $tint_i$ is the trading intensity for duration X_i.

If an increase in the trading intensity leads to more frequent quotes revisions, η_1 should be negative. As the trading intensity $tint_i$ depends on the number of trades and the duration X_i, it also contains a time-of-day effect. This time-of-day effect is extracted from the trading intensity using the same method as the one used for X_i.

4.2.2 Average Volume per Trade

The important role of volume is highlighted in several recent papers: EASLEY and O'HARA [1992], EASLEY, KIEFER and O'HARA [1997], BLUME, EASLEY and O'HARA [1994]. Generally speaking, volume is found to exhibit an informational content that is not contained in the price process. For example, in EASLEY and O'HARA [1992], the excess volume (with respect to what is usually observed, or normal volume) is indicative of possible arrival of informed traders. We introduce volume in our model by defining the average volume per trade over duration X_i. This variable is given by the average, over duration X_i, of the volume of the trades made during this duration. With this variable, the Log-ACD model becomes

$$\psi_i = \omega + \alpha\epsilon_{i-1} + \beta\psi_{i-1} + \eta_2\, avol_{i-1}, \tag{18}$$

where $avol_i$ is the average volume per trade for duration X_i.

If an increase in the average volume per trade is indicative of possible informed trading, it should lead to more frequent quotes revisions, and η_2 should be negative. Because this variable is related to X_i and the number of trades, which both depend on the time-of-day, we use the variable $avol_i$ from which the time-of-day effect has been extracted, so that this variable measures the average volume relative to its normal value (*i.e.* excess volume).

4.2.3 Spread

As indicated in Section 2, one of the results of the EASLEY and O'HARA [1992] model is that a high spread is indicative of possible informed trading, and should be linked to short durations. To investigate this effect, we define the average spread over duration X_i as the average spread corresponding to the trades made during duration X_i. With this variable, our model is

$$\psi_i = \omega + \alpha\epsilon_{i-1} + \beta\psi_{i-1} + \eta_3\, sp_{i-1}, \tag{19}$$

where sp_i is the average spread for the trades made during duration X_i. According to the EASLEY and O'HARA [1992] model, a negative coefficient for η_3 is expected.

TABLE 6
ML Result for the ACD and Log-ACD Model (BOEING)

Coefficient	ACD	Log-ACD
ω	0.030 (0.009)	-0.104 (0.013)
α	0.113 (0.018)	0.095 (0.012)
β	0.859 (0.024)	0.953 (0.012)
γ	0.907 (0.013)	0.905 (0.012)
$Q(10)$	322.3	322.3
$Q(10)^*$	17.28	13.61

Asymptotic standard errors are given in parentheses (the number of observations is 2,620). $Q(10)$ denotes the Ljung-Box Q-statistic of order 10 on the x_i. $Q(10)^*$ gives the corresponding Q-statistic on the residuals e_i, defined in (20) or (21).

TABLE 7
ML Result for the Log-ACD Model (BOEING) (with the Additional Explicative Variables)

Coefficient	with $tint_{i-1}$	with $avol_{i-1}$	with sp_{i-1}	with all variables
ω	0.030 (0.016)	-0.092 (0.015)	-0.063 (0.043)	-0.141 (0.047)
α	0.058 (0.012)	0.096 (0.012)	0.096 (0.012)	0.080 (0.012)
β	0.914 (0.017)	0.953 (0.012)	0.918 (0.018)	0.884 (0.020)
γ	0.911 (0.013)	0.905 (0.013)	0.910 (0.013)	0.915 (0.013)
η_1	-0.069 (0.013)	–	–	0.071 (0.013)
η_2	–	-0.012 (0.009)	–	-0.003 (0.011)
η_3	–	–	-0.895 (0.233)	-0.867 (0.230)
$Q(10)$	322.3	322.3	322.3	322.3
$Q(10)^*$	16.25	13.1	11.1	14.7

See footnote of Table 6.

TABLE 8
ML Result for the ACD and Log-ACD Model (DISNEY)

Coefficient	ACD	Log-ACD
ω	0.013 (0.005)	-0.063 (0.008)
α	0.068 (0.010)	0.061 (0.008)
β	0.920 (0.012)	0.980 (0.006)
γ	0.980 (0.016)	0.982 (0.016)
$Q(10)$	137.31	137.31
$Q(10)^*$	8.04	7.80

Asymptotic standard errors are given in parentheses (the number of observations is 2,160). $Q(10)$ denotes the Ljung-Box Q-statistic of order 10 on the x_i. $Q(10)^*$ gives the corresponding Q-statistic on the residuals e_i, defined in (20) or (21).

4.3 Results with the ACD and Log-ACD Models

The ACD model, the Log-ACD model and the specifications of the Log-ACD model with the lagged trading intensity, the lagged average volume per trade and the lagged average spread were estimated using a maximum likelihood procedure (CML library in GAUSS). The results for BOEING, DISNEY, and IBM are given in Tables 6-7, Tables 8-9, and Tables 10-11, respectively. In analogy with the literature on GARCH models and as suggested by previous work on ACD models (such as ENGLE and RUSSELL [1998], for example), the performance of the ACD model in capturing the autocorrelation structure of the data can be evaluated by examining the residuals

$$e_i = \frac{x_i}{\widehat{\Psi}_i}, \tag{20}$$

where $\widehat{\Psi}_i$ is given by (2) evaluated at the MLE. The ACD model successfully captures the autocorrelation of the durations if the residuals look like white noise. This can be tested with Ljung-Box Q-statistics and can be visually confirmed by plotting the ACF of the residuals. Furthermore, according to (1), the residuals $e_i \Gamma(1 + \frac{1}{\widehat{\gamma}})$ should follow a Weibull$(1, \widehat{\gamma})$ distribution. This can be checked by using some characteristics of the residuals as explained below.

Regarding the Log-ACD model, its performance can be assessed by examining the residuals e_i, defined in this case as

$$e_i = \frac{x_i}{e^{\widehat{\psi}_i}}, \tag{21}$$

where $\widehat{\psi}_i$ is given by (12), (17), (18) or (19) evaluated at the MLE.

The Q(10)-statistics for the tod-adjusted durations and for the residuals are given in Tables 6 to 11. For each stock, plots of the ACF of the tod-adjusted durations and residuals are given in Figures 7a and 7b (BOEING), 8a and 8b (DISNEY) and 9a and 9b (IBM). Following the procedure outlined in ENGLE and RUSSELL [1998], we compute a nonparametric estimate of the hazard of the residuals $e_i \Gamma(1 + \frac{1}{\widehat{\gamma}})$. These are given in Figures 7c, 8c and 9c. Finally, we also estimate (using a gamma kernel) the density of these residuals, which are given in Figures 7d, 8d and 9d.

Several comments can be made from the numerical results in Tables 6-11 and the plots in Figures 7-9.

1) The tod-adjusted durations exhibit a strong autocorrelation structure, indicating that the time-of-day effect (complemented by a day-of-week effect) does not account fully for the dependence of the durations.

2) Both the ACD and Log-ACD models are successful in removing the autocorrelation in the data: for two stocks, BOEING and DISNEY, the residuals are not significantly autocorrelated at order 10, while for IBM the Ljung-Box Q-statistic of order 10 has been reduced from 1932.6 to

TABLE 9

ML Result for the Log-ACD Model (DISNEY) (with the Additional Explicative Variables)

Coefficient	with $tint_{i-1}$	with $avol_{i-1}$	with sp_{i-1}	with all variables
ω	0.010 (0.014)	-0.033 (0.011)	-0.013 (0.024)	-0.050 (0.028)
α	0.047 (0.008)	0.056 (0.008)	0.052 (0.008)	0.042 (0.008)
β	0.950 (0.010)	0.972 (0.007)	0.974 (0.006)	0.948 (0.010)
γ	0.987 (0.016)	0.985 (0.016)	0.984 (0.016)	0.989 (0.016)
η_1	-0.051 (0.011)	–	–	0.040 (0.012)
η_2	–	-0.025 (0.007)	–	-0.016 (0.008)
η_3	–	–	-0.415 (0.134)	-0.255 (0.155)
$Q(10)$	137.31	137.31	137.31	137.31
$Q(10)^*$	6.75	8.07	8.89	7.25

See footnote of Table 8.

TABLE 10

ML Result for the ACD and Log-ACD Model (IBM)

Coefficient	ACD	Log-ACD
ω	0.007 (0.002)	-0.079 (0.007)
α	0.088 (0.008)	0.077 (0.007)
β	0.907 (0.009)	0.988 (0.003)
γ	1.000 (0.009)	0.999 (0.009)
$Q(10)$	1932.64	1932.64
$Q(10)^*$	34.06	33.75

Asymptotic standard errors are given in parentheses (the number of observations is 6,728*).* $Q(10)$ *denotes the Ljung-Box Q-statistic of order* 10 *on the* x_i*.* $Q(10)^*$ *gives the corresponding Q-statistic on the residuals* e_i*, defined in* (20) *or* (21).

TABLE 11

ML Result for the Log-ACD Model (IBM) (with the Additional Explicative Variables)

Coefficient	with $tint_{i-1}$	with $avol_{i-1}$	with sp_{i-1}	with all variables
ω	0.039 (0.008)	-0.058 (0.007)	-0.048 (0.013)	-0.010 (0.015)
α	0.071 (0.007)	0.072 (0.006)	0.078 (0.007)	0.068 (0.007)
β	0.969 (0.005)	0.987 (0.003)	0.983 (0.004)	0.967 (0.005)
γ	1.003 (0.009)	1.001 (0.009)	1 (0.009)	1.004 (0.009)
η_1	-0.036 (0.006)	–	–	0.034 (0.006)
η_2	–	-0.016 (0.004)	–	-0.013 (0.004)
η_3	–	–	-0.169 (0.069)	-0.082 (0.071)
$Q(10)$	1932.64	1932.64	1932.64	1932.64
$Q(10)^*$	29.98	37.62	32.58	32.15

See footnote of Table 10.

about 30 (which has a p-value of 0.08 %). The plots given in Figures 7-9 provide the same message. For all stocks, the ACF of the tod adjusted durations are slowly decreasing to zero, indicating a strong autocorrelation structure. However, the ACF of the residuals show that these are much less autocorrelated. Regarding the IBM stock, an additional comment has to be made. As hinted in Section 3, the ACF of the tod adjusted durations for the IBM stock decreases to zero more slowly than for the other two stocks. This indicates that a long memory process could be in play. This issue has been recently studied by JASIAK [1998] with the fractionally integrated ACD model. She works with *trade* durations which display very large Q-statistics (much larger than the Q-statistics of the price durations considered here) and finds a long memory effect. While acknowledging the fact that FIACD models could perhaps enhance the specification (for the IBM stock), we do not use this technique here as only one stock (out of three) is concerned by this effect. Although the results are not reported here, the ACD and Log-ACD models were applied to other stocks traded on the NYSE, and no long memory effect was exhibited by the price durations: when working with price durations, the IBM stock seems to be the exception rather than the rule.

3) The nonparametric estimates of the hazards of the residuals indicate that these hazards are fairly constant around one, with a slightly decreasing hazard for the BOEING and DISNEY stocks. This is in agreement with the estimated γ coefficients for the three stocks, which are close to one. Monte Carlo simulations were also used to simulate data with values of γ larger and smaller than one. The nonparametric method was then applied to the simulated data. In both cases, the estimated hazard function was clearly downward sloping for the data with γ smaller than one, and upward sloping when γ was higher than one. This stresses the fact that the estimated hazard functions are in agreement with the estimated values for the γ parameter.

4) If $e_i\Gamma(1+1/\widehat{\gamma})$ follows a Weibull$(1,\widehat{\gamma})$ distribution, $\left(e_i\Gamma(1+\frac{1}{\widehat{\gamma}})\right)^{\widehat{\gamma}}$ follows an exponential distribution, which implies that the no-overdispersion hypothesis for this series should be accepted. As indicated in ENGLE and RUSSELL [1998], this can be tested with the statistic $Z=\sqrt{N}(\sigma^2-1)/\sqrt{8}$, where N is the number of observations, σ^2 is the variance of the series and Z follows a normal distribution when N is large. For the three stocks, the statistic is equal to 4.09, 3.89 and 8.81, respectively. At the five percent level, the critical value is equal to 1.96, which implies that the null hypothesis of no-overdispersion is rejected for all stocks. However, the computed statistics are not much larger than the critical value (except for IBM), especially given the fact that N is very large. Moreover, if $e_i\Gamma(1+1/\widehat{\gamma})$ follows a Weibull$(1,\widehat{\gamma})$ distribution, then $-\ln(S)$, the negative logarithm of the survivor function of the residuals, should be equal to $\left(e_i\Gamma(1+\frac{1}{\widehat{\gamma}})\right)^{\widehat{\gamma}}$. As

FIGURE 7

ACF of the Tod-Adjusted Durations for BOEING (figure 7a), Residuals of the Log-ACD Model (figure 7b), Nonparametric Hazard (figure 7c), and Density Function (figure 7d) of the Residuals of the Log-ACD Model

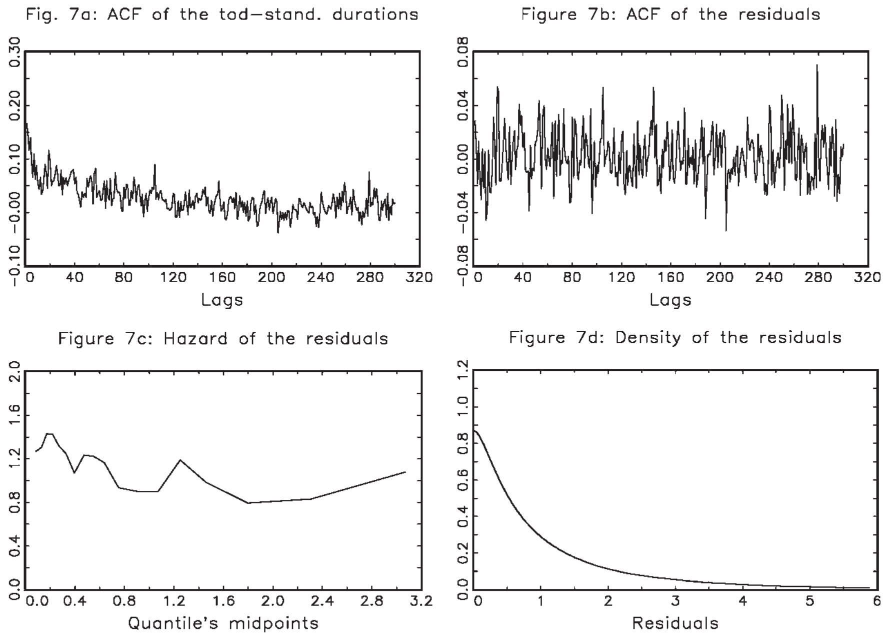


FIGURE 8

ACF of the Tod-Adjusted Durations for DISNEY (figure 8a), Residuals of the Log-ACD Model (figure 8b), Nonparametric Hazard (figure 8c), and Density Function (figure 8d) of the Residuals of the Log-ACD Model

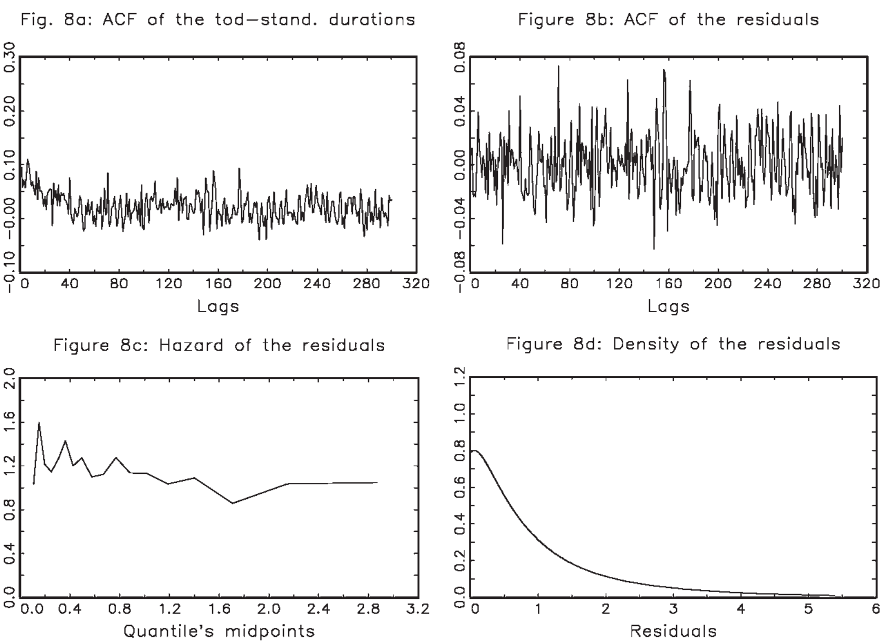


FIGURE 9

ACF of the Tod-Adjusted Durations for IBM (figure 9a), Residuals of the Log-ACD Model (figure 9b), Nonparametric Hazard (figure 9c), and Density Function (figure 9d) of the Residuals of the Log-ACD Model

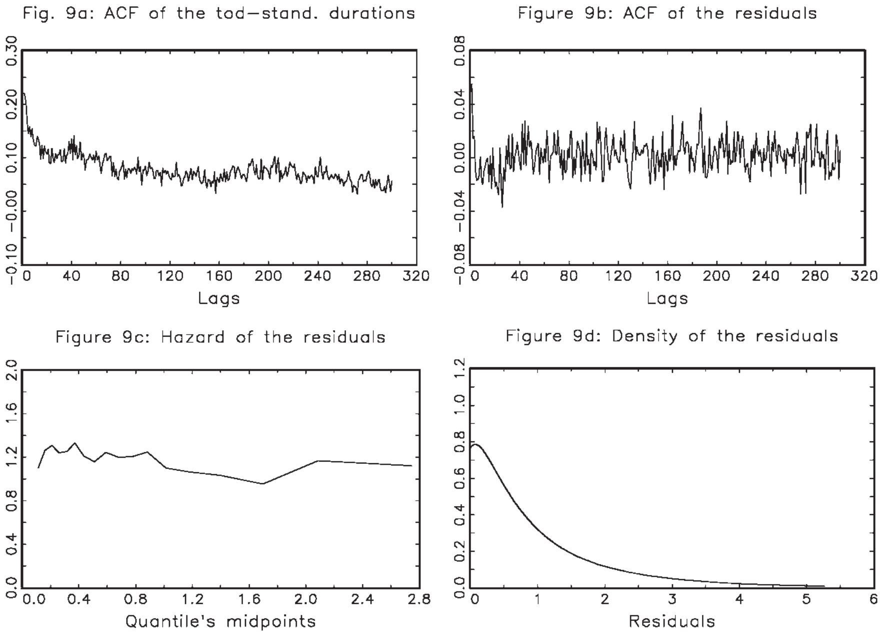
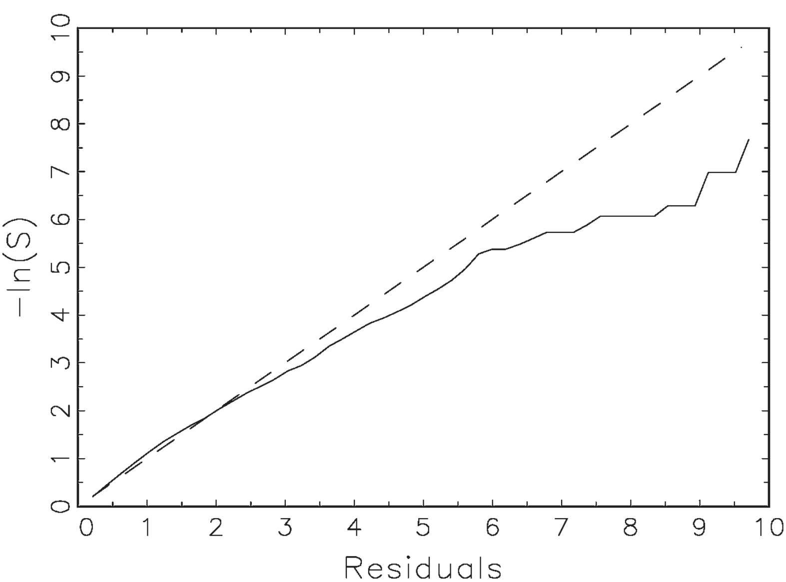


FIGURE 10

Negative Logarithm of the Survivor Function of the Residuals for the DISNEY Stock Plotted Against $(e_i\Gamma(1+1/\widehat{\gamma}))^{\gamma}$



suggested by ENGLE and RUSSELL [1998], this can be checked visually by plotting $-\ln(S)$ against $(e_i \Gamma(1 + 1/\widehat{\gamma}))^{\widehat{\gamma}}$. We illustrate this method in Figure 10 for the DISNEY stock (the other two stocks exhibit a similar relationship). We obtain similar results as those given in ENGLE and RUSSELL [1998]: while the negative logarithm of the survivor function is close to a linear function of the residuals for most of the range, there is a larger discrepancy for the larger residuals, indicating that these are overrepresented.

5) The autoregressive coefficient β is close to 1 for all stocks (and for all the specifications tested), while being significantly smaller than 1 at the 5 percent level. Regarding the additional market microstructure variables included in the model, the results given in Tables 7, 9 and 11 indicate that:

– With the lagged trading intensity ($tint_{i-1}$), lagged average volume per trade ($avol_{i-1}$), and lagged average spread (sp_{i-1}) included separately as additional explicative variables, the coefficients η_1, η_2 and η_3 are negative and strongly significant for all stocks. Thus, a higher trading intensity, a higher average volume per trade and a higher average spread all shorten the next expected duration. This is in agreement with the hypotheses put forward by EASLEY and O'HARA [1992].

– When all three lagged additional variables are included at the same time, the coefficients η_1, η_2 and η_3 are still negative, but η_2 is not significant for BOEING, while η_3 is not significant for DISNEY and IBM. For all stocks, the coefficient of the lagged trading intensity (η_1) is negative and strongly significant. From these results, it would seem that the trading intensity is the variable that has the most important impact on the bid-ask quote process. It is interesting to note that these results are quite similar to those of JONES, KAUL and LIPSON [1994]. Indeed, in their detailed study (using daily data) on a large sample of stocks traded on the NASDAQ, they found that it is the number of trades, and not the average volume per trade, that has the most important impact on the volatility of the bid-ask quote process.

Finally, we estimated the Log-ACD models on several subsample periods. More precisely, we split our data into three periods featuring the same number of observations and we estimated the models on each subsample. For all three periods, the results were very similar to those reported for the whole dataset. It should however be noted that the Q statistics of the tod adjusted durations for the IBM stock were much lower than for the three month period, and that the possible long memory effect was clearly not present in the smaller datasets.

Conclusion

The Engle and Russell [1998] paper introduced the ACD model, which presents a new way of modelling random durations arising from high-frequency data. In this paper, we introduced two specifications of a logarithmic version of the ACD model. While retaining the main characteristics of the ACD model, the Log-ACD is more flexible as no restrictions are required on the sign of its coefficients. The ACD and Log-ACD models were applied to the durations of the bid-ask quotes posted by specialists on the NYSE for three actively traded stocks. We highlighted the impact of characteristics of the trade process on the dynamics of the bid-ask quote process. More precisely, the influence of the trading intensity, the average volume per trade, and the average spread on the bid-ask quote process is negative and significant, which is compatible with the Easley and O'Hara [1992] model.

Extensions of the model are possible. Firstly, a more general distribution for the error term of the model could be used, for example the Burr distribution which has a non-monotonic hazard function (see Grammig and Maurer [1999]). Secondly, the specification of the models could include non-linear transformations of the explicative variables. Thirdly, while we focused on the variables suggested in the Easley and O'Hara [1992] model, the litterature on market microstructure suggests other possibilities, such as the depth at the bid and ask, and the changes in price or in spread (see Engle and Lunde [1998]). These variables could also be included in the model. Fourthly, it would be insightful to compare the ACD and Log-ACD models using some formal test procedures, information criteria, and from the viewpoint of predictive performance. This is the topic of a paper by Bauwens, Giot, Grammig and Veredas [1999], who use in particular density forecasts as suggested by Diebold, Gunther and Tay [1997], to compare ACD, Log-ACD, SCD, and SVD models. ■

• References

AMIHUD Y., MENDELSON H. (1980). – « Dealership Market: Market Making with Inventory », *Journal of Financial Economics* 8, pp. 31-53.

BAUWENS L., GIOT P. (1998). – « Asymmetric ACD Models: Introducing Price Information in ACD Models with a Two State Transition Model », CORE *Discussion Paper* 9844.

BAUWENS L., VEREDAS D. (1999). – « The Stochastic Conditional Duration Model: a Latent Factor Model for the Analysis of Financial Durations », CORE *Discussion Paper* 9958.

BAUWENS L., GIOT P., GRAMMIG J., VEREDAS D. (1999). – « A Comparison of Financial Duration Models *via* Density Forecasts », *Mimeo*.

BIAIS B., HILLION P., SPATT C. (1995). – « An Empirical Analysis of the Limit Order book and the order Flow in the Paris Bourse », *Journal of Finance* 50, pp. 1655-1689.

BIAIS B., FOUCAULT T., HILLION P. (1997). – *Microstructure des marchés financiers : institutions, modèles et tests empiriques*, Paris: Presses Universitaires de France.

BISIÈRE C., KAMIONKA T. (2000). – « Timing of Orders, Orders Aggressiveness and the Order Book at the Paris Bourse », *Annales d'Économie et de Statistique*, this issue, pp. 43-72.

BLUME L., EASLEY D., O'HARA M. (1994). – « Market Statistics and Technical Analysis: the Role of Volume », *The Journal of Finance* 49, pp. 153-181.

BOLLERSLEV T., DOMOWITZ I. (1993). – « Trading Patterns and Prices in the Interbank Foreign Exchange Market », *Journal of Finance* 48, pp. 1421-1443.

BROCK W.A., KLEIDON A.W. (1992). – « Periodic Market Closure and Trading Volume: A Model of Intraday Bids and Asks », *Journal of Economic Dynamics and Control* 16, pp. 451-489.

CAMPBELL J.Y., LO A.W., MACKINLAY A.C. (1997). – *The Econometrics of Financial Markets*, Princeton: Princeton University Press.

CHAN K.C., CHRISTIE W.G., SCHULTZ P.H. (1995). – « Market Structure and the Intraday Pattern of Bid-Ask Spreads for NASDAQ Securities », *Journal of Business* 68, pp. 35-60.

DIEBOLD F.X., GUNTHER T.A., TAY A.S. (1997). – « Evaluating Density Forecasts », *Mimeo*. Department of Economics, University of Pennsylvania.

EASLEY D., O'HARA M. (1992). – « Time and the Process of Security Price Adjustment », *The Journal of Finance* 19, pp. 69-90.

EASLEY D., KIEFER N.M., O'HARA M. (1997). – « The Information Content of the Trading Process », *The Journal of Empirical Finance*, pp. 159-186.

EDDELBUTEL D., MCCURDY T.H. (1996). – « The Impact of News on Foreign Exchange Rates: Evidence from very High Frequency Data », *Mimeo*.

ENGLE R. (2000). – « The Econometrics of Ultra-High Frequency Data », *Econometrica 68*, pp. 1-22.

ENGLE R., LANGE J. (1997). – « Measuring, Forecasting and Explaining Time Varying Liquidity in the Stock Market », *Discussion Paper* 97-12, University of California, San Diego.

ENGLE R., LUNDE A. (1998). – « Trades and Quotes: a Bivariate Point Process », *Discussion Paper* 98-07, University of California, San Diego.

ENGLE R., RUSSELL J. (1997). – « Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model », *Journal of Empirical Finance* 4, pp. 187-212.

ENGLE R., RUSSELL J. (1998). – « Autoregressive Conditional Duration; A New Model for Irregularly Spaced Transaction Data », *Econometrica* 66, pp. 1127-1162.

GARMAN M. (1976). – « Market Microstructure », *Journal of Financial Economics* 3, pp. 257-275.

GEWEKE J. (1986). – « Modelling the Persistence of Conditional Variances: A Comment ». *Econometric Reviews* 5, pp. 57-61.

GHYSELS E., GOURIÉROUX C., JASIAK J. (1997). – « Stochastic Volatility Duration Models », *Mimeo*, CREST, Paris.

GHYSELS E., JASIAK J. (1998). – « GARCH for Irregularly Spaced Financial Data: The ACD-GARCH model », *Studies in Nonlinear Dynamics and Econometrics* 2.

GIOT P. (1999). – « Time Transformations, Intraday Data and Volatility Models », CORE *Discussion Paper* 9944, Forthcoming in *Journal of Computational Finance*.

GLOSTEN L., MILGROM P. (1985). – « Bid, Ask, and the Transaction Prices in a Specialist Market with Heterogeneously Informed Traders », *Journal of Financial Economics* 13, pp. 71-100.

GOODHART C.A.E., O'HARA M. (1997). – « High Frequency Data in Financial Markets: Issues and Applications », *The Journal of Empirical Finance*, pp. 73-114.

GRAMMIG J., MAURER K.-O. (1999). – « Non-Monotonic Hazard Functions and the Autoregressive Conditional Duration Model », *Discussion Paper* 50, SFB 373, Humboldt University Berlin, Forthcoming in *The Econometrics Journal*.

GUILLAUME D.M., PICTET O.V., DACOROGNA M.M. (1995). – « On the Intra-Daily Performance of GARCH Processes », Olsen Preprint.

GUILLAUME D.M., DACOROGNA M.M., DAVÉ R.R., MULLER U.A., OLSEN R.B., PICTET O.V. (1997). – « From the Bird's Eye to the Microscope: A Survey of New Stylized Facts of the Intra-Daily Foreign Exchange Markets », *Finance and Stochastics* 1, pp. 95-129.

GWILYM O.A., BUCKLE M., THOMAS S. (1997). – « The Intraday Behavior of Bid-Ask Spreads, Returns, and Volatility for FTSE-100 Stock Index Options », *Journal of Derivatives* 4, pp. 20-32.

HAFNER C. (1996). – « Estimating High Frequency Foreign Exchange Rate Volatility with Nonparametric ARCH Models », *Journal of Statistical Planning and Inference* 68, pp. 247-269.

HASBROUCK J. (1988). – « Trades, Quotes, Inventories, and Information », *Journal of Financial Economics* 22, pp. 229-252.

HASBROUCK J. (1991). – « Measuring the Information Content of Stock Trades », *Journal of Finance* 46, pp. 179-208.

HAUSMAN J., LO A.W., MACKINLAY A.C. (1992). – « An Ordered Probit Analysis of transaction Stock Prices », *Journal of Financial Economics* 31, pp. 319-379.

HO T., STOLL H. (1981). – « Optimal Dealer Pricing Under Transactions and Return Uncertainty », *Journal of Financial Economics* 9, pp. 47-73.

JASIAK J. (1998). – « Persistence in Intertrade Durations », *Finance* 19, pp. 166-195.

JONES C.M., KAUL G., LIPSON M.L. (1994). – « Transactions, Volume, and Volatility. *Review of Financial Studies* 7, pp. 631-651.

KYLE A.S. (1985). – « Continuous Auctions and Insider Trading », *Econometrica* 53, pp. 1315-1365.

MADHAVAN A., SOFIANOS G. (1998). – « An Empirical Analysis of NYSE Specialist Trading », *Journal of Financial Economics* 48, pp. 189-210.

NELSON D.B. (1991). – « Conditional Heteroskedasticity in Asset Returns: A New Approach », *Econometrica* 59, pp. 349-370.

O'HARA M. (1995). – *Market Microstructure Theory*, Oxford: Basil Blackwell.

RUSSELL J., ENGLE R. (1998). – « Econometric Analysis of Discrete-Valued, Irregularly-spaced Financial Transactions Data using a New Autoregressive Conditional Multinomial Model », *Discussion Paper* 98-10, University of California, San Diego.

STOLL H. (1978). – « The Supply of Dealer Services in Securities Markets », *Journal of Finance* 33, pp. 1133-1151.

APPENDIX

We consider the ACD model defined by equations (1)-(3). In addition to the unconditional mean (μ) and variance (σ^2) of x_i - see (4) and (5) - ENGLE and RUSSELL [1995] give the variance of Ψ_i:

$$\text{Var}(\Psi_i) = \frac{\alpha^2 \kappa \mu^2}{1 - (\alpha + \beta)^2 - \alpha^2 \kappa}. \tag{22}$$

These results are useful in order to compute the autocovariance function of x_i. Let us call γ_j the autocovariance between x_i and x_{i-j}. To get γ_1 we need to compute $E(x_i x_{i-1})$, which can be done as follows:

$$\begin{aligned} E(x_i x_{i-1}) &= E[x_{i-1} E(x_i | I_{i-1})] = E(x_i \Psi_i) \\ &= \omega E(x_{i-1}) + \alpha E(x_{i-1}^2) + \beta E(x_{i-1} \Psi_{i-1}) \\ &= \omega E(x_{i-1}) + \alpha E(x_{i-1}^2) + \beta E(\Psi_{i-1}^2) \end{aligned} \tag{23}$$

since

$$\begin{aligned} E(x_{i-1} \Psi_{i-1}) &= E[E(x_{i-1} \Psi_{i-1} | I_{i-2})] \\ &= E[\Psi_{i-1} E(x_{i-1} | I_{i-2})] = E(\Psi_{i-1}^2). \end{aligned} \tag{24}$$

Therefore,

$$\gamma_1 = \omega\mu + \alpha a + \beta b - \mu^2, \tag{25}$$

where

$$\begin{aligned} a &= E(x_{i-1}^2) = \sigma^2 + \mu^2 \\ b &= E(\Psi_{i-1}^2) = \text{Var}(\Psi_i) + \mu^2. \end{aligned} \tag{26}$$

Starting with $p = 2$, a recursive law is available for γ_p. Indeed, for γ_2, we have:

$$\begin{aligned} \gamma_2 &= E[(x_{i-2} E(x_i | I_{i-1})] - \mu^2 \\ &= E[x_{i-2}(\omega + \alpha x_{i-1} + \beta \Psi_{i-1})] - \mu^2 \\ &= E\{x_{i-2}[\omega + \alpha x_{i-1} + \beta(\omega + \alpha x_{i-2} + \beta \Psi_{i-2})]\} - \mu^2 \\ &= \omega\mu + \alpha(\gamma_1 + \mu^2) + \beta\omega\mu + \alpha\beta a + \beta^2 b - \mu^2. \end{aligned} \tag{27}$$

In a similar way, γ_3 can be expressed as a function of γ_2 and γ_1:

$$\gamma_3 = \omega\mu + \alpha(\gamma_2 + \mu^2) + \omega\beta\mu + \alpha\beta(\gamma_1 + \mu^2) + \omega\beta^2\mu + \alpha\beta^2 a + \beta^3 b - \mu^2.$$

Generalizing this expression to γ_p, it is defined recursively by

$$\gamma_p = \omega\mu + \sum_{q=1}^{p-1} A_{p,q} + \beta^{p-1}\alpha a + \beta^p b - \mu^2 \tag{29}$$

where

$$A_{p,q} = \alpha\beta^{q-1}(\gamma_{p-q} + \mu^2) + \omega\mu\beta^q. \tag{30}$$