1 Introduction

High frequency duration models in finance were first introduced by Engle and Russell (1995) with the Autoregressive Conditional Duration (ACD) model. In that paper, they did not take the traditional viewpoint of looking at the volatility of the price process, but worked instead on the duration (i.e. time elapsed) between two transactions of a stock. These durations exhibit clustering (short or large durations tend to be followed by durations of the same kind) and unconditional overdispersion (the ratio of standard deviation to mean is larger than one). The structure of ACD models is quite similar to the structure of GARCH processes, but rather than specifying an autoregressive structure on the conditional variance of the price process, they introduce an autoregressive structure on conditional durations. Engle and Russell applied their model to data for the IBM stock on the New York Stock Exchange (Engle and Russell, 1995) and also to data for the USD/DM exchange rate on the foreign exchange market (Engle and Russell, 1997).

Duration models are particularly suited for applications in a high frequency setting, as one deals with irregularly spaced data. In this framework, the time elapsed between two market events (for example a trade, a specific type of trade such as a large buy, a revision of bid/ask quotes ...) conveys meaningful information and is the object of modeling. The importance of taking into account the time dimension of the price process is particularly stressed by the recent market microstructure literature (O'Hara, 1995; Easley and O'Hara, 1992; Easley, Kiefer and O'Hara, 1997). Since the original paper by Engle and Russell (1995), several high frequency duration models have been put forward. For example, Bauwens and Giot (1997) introduced a logarithmic version of the ACD model (Log-ACD), which allows for more flexbility when conditioning variables are included in the model, e.g. in order to test a microstructure effect. Ghysels, Gouriéroux, and Jasiak (1997) proposed the Stochastic Volatility Duration model, which accounts for stochastic volatility in the durations.

All these models share a common feature: they only take into account the duration between market events, and do not incorporate information given by the price process (at most they include additional variables drawn from the microstructure literature). This is a main drawback, as the information on the price process is of primary importance. Combining information given by the price process and the duration between market events is an important extension. This issue was first addressed by Engle (1996), who proposed a first version of an ACD-GARCH model: a marginal ACD model is fitted for the durations, while the volatility of the price process is modelled by a GARCH process, conditionally on the duration. The ACD-GARCH model is also studied by Ghysels and Jasiak (1997), who introduce a joint structure on the ACD and GARCH parts. In this way, they can take into account cross effects between the durations and the volatility. More recent price/duration models are proposed by: Engle and Russell (1998), who introduce the Autoregressive Conditional Multinomial Model; Darolles and Lefol (1998), who work on a discrete diffusion process estimated with irregularly spaced data; Meddahi, Renault and Werker (1998) who propose a stochastic volatility model for unequally spaced observations, and also a generalization of ACD models.

In this paper we introduce an ACD model for durations between bid-ask price quotes by a market maker, which takes into account the direction (upward or downward) of the change of the average of the bid and ask prices. The model introduced is an asymmetric ACD model, which can be thought of as a two state transition model combined with an ACD model (the two states being the upward and downward price movements). We apply this new model to intraday data for the IBM stock. As in Bauwens and Giot (1997), we also look at the influence of the trade process on the bid/ask quotes revision process by introducing the buy/sell disequilibrium in the specification of the asymmetric model. The empirical results of this paper clearly indicate that working with an asymmetric ACD model is an improvement over the original ACD model.

The rest of the paper is organized in the following way. In Section 2, we briefly review ACD models, to set up the framework. In Section 3, we introduce our asymmetric ACD model, which combines features from transition models and duration models. In Section 4 we apply the asymmetric ACD model to the IBM stock and examine the way specialists revise their bid/ask quotes. Section 5 concludes.

2 A brief review of ACD models

Let x_i be the duration between two bid/ask quotes occuring at times t_{i-1} and t_i (a quote being a collection of data relative to a buy or a sell of a security on a stock exchange), i.e. $x_i = t_i - t_{i-1}$ (measured in seconds). The assumption introduced by Engle and Russell (1995) is that the time dependence in the durations can be subsumed in their conditional expectations $E(x_i|J_{i-1})$, in such a way that $x_i/E(x_i|J_{i-1})$ is independent and identically distributed. J_{i-1} denotes the information available at time t_i, which includes at the least the past durations.

The ACD model specifies the observed duration as the mixing process

$$x_i = \Psi_i \epsilon_i \tag{1}$$

where the ϵ_i are IID, with $E(\epsilon_i) = \kappa$, so that $E(x_i|J_{i-1}) = \Psi_i \kappa$.

A second equation specifies an autoregressive model for the conditional durations Ψ_i:[1]

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

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

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

A variant of the ACD model that does not require the positivity of the coefficients is the Log-ACD model of Bauwens and Giot (1997). It is obtained by defining

$$\Psi_i = e^{\psi_i} \tag{3}$$

in (1), and replacing (2) by

$$\psi_i = \omega + \alpha \epsilon_{i-1} + \beta \psi_{i-1} \tag{4}$$

i.e. the autoregressive equation bears on the *logarithm* of the conditional expectation. The autoregressive structure in equations (2) and (4) allows the model to account for the clustering of the durations.

In order to estimate the model by the maximum likelihood (ML) method, a parametric hypothesis on the distribution of ϵ_i is necessary. Possible choices, among others, are the gamma and the Weibull families, which nest the exponential one. Let us assume for example that ϵ_i is distributed as a Weibull(1,γ) random variable, such that $\kappa = E(\epsilon_i) = \Gamma(1 + 1/\gamma)$, where $\Gamma(.)$ is the usual Gamma function.

Equation (1) can then be viewed in the framework of a time-varying hazard model. The hazard function of ϵ (defined as the density divided by the survivor function) is given by

$$h(\epsilon_i) = \gamma \epsilon_i^{\gamma-1} \tag{5}$$

If $\gamma < 1$, the model exhibits a decreasing hazard function: long durations will be less likely. If $\gamma > 1$, the model shows an increasing hazard function: long durations will be more likely. Thus the hazard function of x_i is given by

$$h(x_i) = \frac{\gamma}{\Psi_i} \left(\frac{x_i}{\Psi_i} \right)^{\gamma-1} \tag{6}$$

When Ψ_i is small, the expected duration of x_i is small (by definition). In this case, as implied by equation (6), the corresponding hazard of x_i is high: the probability of witnessing the end of duration x_i is high. This is another way of looking at ACD models.

As such, these models are self contained: the durations between two consecutive bid/ask quotes, or more generally between two market events, are modelled conditionally on the past durations. The information given by features linked to the market events is not used. Such observable features may include the exchanged volume and the corresponding transaction price, the bid and ask prices quoted by the market makers, released news, or functions thereof. Other features, such as the volatility, may be relevant but unobservable.

A natural extension of ACD models consists of combining them with a model bearing on such features of the market. As briefly explained in the introduction, this has been proposed by several authors. At a high level of generality, such models can be described as follows. Let us consider the marked point process (x_i, y_i), where y_i is a set of variables (observed or latent) at the end of duration x_i, usually referred to as the "end state" of duration x_i. The ACD models described so far in this section can be conceived as marginal models of x_i. Durations can be modelled conditionally on a variable in y_i in order to introduce a microstructure effect, as in Engle and Russell (1997), and Bauwens and Giot (1997). The ACD-GARCH model of Engle (1996) is a model for a subset of y_i (the transaction price and its volatility) conditional on x_i (the duration between two market transactions), combined with a marginal ACD model for x_i. The model introduced in the next section is a joint model for a duration x_i and the direction of the price change (y_i, a binary variable) that occurs between the beginning and the end of x_i. Clearly, different choices of x_i and y_i, combined with different marginalization and conditioning operations can generate a vast range of statistical models. Future research is likely to focus on the specification of such models, and on their relevance for different objectives, such as (ex-post) market analyses, prediction, tests of microstructure hypotheses...

3 Introducing asymmetry in ACD models

In this section, we introduce a new model for durations between two bid/ask quotes, which incorporates the main features of ACD models but also uses information on the upside or downside movement of the (bid/ask) price process. Indeed, it turns out that the existing ACD or Log-ACD models can be extended into models that combine information on the duration and also information on the direction of the price movement. We begin by presenting a simple two state transition model that highlights the feature we want to incorporate in the model, in order to get the asymmetric ACD model (which is then presented in the second part of this section).

3.1 A two state transition model

A transition model has been applied to data from the Paris Bourse by Bisière and Kamionka (1998), who look at the possible transitions between the different types of orders that can be submitted by traders: each type of order represents a given state and the succession of these states can be modelled by a transition model.

We consider the marked point process consisting of the pairs (x_i, y_i), where x_i is the duration between two bid/ask quotes, and y_i is a variable indicating the direction of the average bid and ask price change, i.e.

- $y_i = 1$ if the mid price increased over duration x_i;

- $y_i = -1$ if the mid price decreased over duration x_i.

At the end of duration x_i, there are only two possible end states: $y_i = 1$ or $y_i = -1$. For simplicity, let us assume that the hazard function is constant, which is equivalent to an exponential distribution for x_i. The idea of the transition model is to let the hazard depend on the end state, in this case to have two hazards since y_i is binary. Moreover in each case the hazard can be a function of the past, in particular of the previous state y_{i-1}. If we define λ^+ (respectively λ^-) as the hazard of duration x_i when the end state is $y_i = 1$ (respectively -1), then

$$\begin{aligned} \lambda^+ &= \beta_1 I_{i-1}^+ + \beta_2 I_{i-1}^- \\ \text{and} & \\ \lambda^- &= \beta_3 I_{i-1}^+ + \beta_4 I_{i-1}^- \end{aligned} \tag{7}$$

where I_{i-1}^+ is an indicator variable which is equal to one if $y_{i-1} = 1$ and zero otherwise, and I_{i-1}^- is the complementary indicator variable (equal to one if $y_{i-1} = -1$ and zero otherwise).

At the end of duration x_i, either state $y_i = 1$ or state $y_i = -1$ is realized. The duration corresponding to the state that is not realized is truncated, since the observed duration is the minimum of two possible durations: the one which would realize if $y_i = 1$, and the one which would realize if $y_i = -1$. Thus the likelihood function[2] for duration x_i and state y_i, given the previous state, is equal to

$$L(x_i, y_i | y_{i-1}) = (\lambda^+)^{I_i^+} e^{-\lambda^+ x_i} (\lambda^-)^{I_i^-} e^{-\lambda^- x_i} \tag{8}$$

For example, if state $y_i = 1$ is observed ($I_i^+ = 1$ and $I_i^- = 0$), x_i contributes to the likelihood function via its density function $\lambda^+ e^{-\lambda^+ x_i}$ and via the survivor function $e^{-\lambda^- x_i}$.

Although this model is too simple to capture the dynamics of the duration and price processes inherent in high-frequency quotes data, it puts forward a rather simple way of combining a duration model (in this case a constant hazard model) with the information given by the direction of the price process (an increase or decrease in the price of the security).

[2] In the likelihood function given here, we jointly model the duration x_i and the end state y_i. The realized state contributes to the likelihood function via its density function, while the truncated state contributes to the likelihood via its survivor function.

3.2 An asymmetric ACD model

Instead of assuming a constant hazard for the two-state transition process, we can use a time-varying hazard of the ACD type. The asymmetric Log-ACD model is then defined by the following equations:

- if the end state of duration x_i is $y_i = 1$:

$$x_i = \Psi_i^+ \epsilon_i^+ = e^{\psi_i^+} \epsilon_i^+ \tag{9}$$

where the ϵ_i^+ are IID and follow a Weibull(1,γ^+) distribution. The expectation of x_i, conditional on the past and on $y_i = 1$, is given by $\Psi_i^+ \Gamma(1 + 1/\gamma^+)$. The autoregressive process on ψ_i^+ is defined as

$$\psi_i^+ = (\omega_1 + \alpha_1 \epsilon_{i-1}^+) I_{i-1}^+ + (\omega_2 + \alpha_2 \epsilon_{i-1}^+) I_{i-1}^- + \beta^+ \psi_{i-1}^+ \tag{10}$$

The hazard for x_i is given by

$$h(x_i | y_i = 1) = \frac{\gamma^+}{\Psi_i^+} \left(\frac{x_i}{\Psi_i^+} \right)^{\gamma^+ - 1} \tag{11}$$

- if the end state of duration x_i is $y_i = -1$:

$$x_i = \Psi_i^- \epsilon_i^- = e^{\psi_i^-} \epsilon_i^- \tag{12}$$

where the ϵ_i^- are IID and follow a Weibull(1,γ^-) distribution. In this case, ψ_i^- is defined as

$$\psi_i^- = (\omega_3 + \alpha_3 \epsilon_{i-1}^-) I_{i-1}^+ + (\omega_4 + \alpha_4 \epsilon_{i-1}^-) I_{i-1}^- + \beta^- \psi_{i-1}^- \tag{13}$$

The hazard for x_i is given by

$$h(x_i | y_i = -1) = \frac{\gamma^-}{\Psi_i^-} \left(\frac{x_i}{\Psi_i^-} \right)^{\gamma^- - 1} \tag{14}$$

This model exhibits the main features of the ACD model (i.e. autoregressive structure on the conditional expectation of the duration between two successive bid/ask quotes) while incorporating conditioning information on the underlying bid/ask mid price. More precisely, the duration process can be different according to the direction (upward or downward) of the mid price of the quotes.

The likelihood function for x_i and y_i, given the past information (that includes past states and durations) is given by

$$L(x_i, y_i|J_{i-1}) = \left[\frac{\gamma^+}{\Psi_i^+}\left(\frac{x_i}{\Psi_i^+}\right)^{\gamma^+-1}\right]^{I_i^+} e^{-\left(\frac{x_i}{\Psi_i^+}\right)^{\gamma^+}} \left[\frac{\gamma^-}{\Psi_i^-}\left(\frac{x_i}{\Psi_i^-}\right)^{\gamma^--1}\right]^{I_i^-} e^{-\left(\frac{x_i}{\Psi_i^-}\right)^{\gamma^-}} \tag{15}$$

Estimating the parameters of (15) is relatively straightforward. We can estimate them jointly by maximizing the likelihood (15), or equivalently, we can estimate the parameters relative to the price increase and decrease separately. For example, ω_1, α_1, ω_2, α_2, β^+ and γ^+ can be estimated by maximizing $L^+(x_i, y_i|J_{i-1}) = \left[\frac{\gamma^+}{\Psi_i^+}\left(\frac{x_i}{\Psi_i^+}\right)^{\gamma^+-1}\right]^{I_i^+} e^{-\left(\frac{x_i}{\Psi_i^+}\right)^{\gamma^+}}$.

In what precedes, we could assume an identical distribution for ϵ^+ and ϵ^-, i.e. that $\gamma^+ = \gamma^-$ (a testable hypothesis). In this case, we cannot split the log-likelihood function in two parts, each depending on a different subset of the parameters.

The asymmetry of the duration process with respect to the different end states of the price process could also be applied to the ACD model (rather than to the Log version).

3.3 Additional explicative variables

The specification for the conditional expectation of duration x_i as given by equations (10) and (13) does not use possible relevant information that may be included in the information set (as argued at the end of Section 2). Recently, several authors have proposed to include in such equations explicative variables drawn from the microstructure literature, in order to better specify the expected conditional duration. The variables that are taken into account are usually related to the information models originally developed by Glosten and Milgrom (1985). These models assume that the market maker faces non informed (or liquidity traders) and informed traders: as the informed traders have an informational advantage (they know the outcome of the price process), the market maker loses on the trades made with this type of traders. Some of the proposals that were made are:

- the spread. Easley and O'Hara (1992) extend the information model introduced by Glosten and Milgrom (1985) by focusing on the role of time in price adjustment. 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. This should lead to a negative coefficient for the spread when introduced in the specification of the expected conditional duration. This issue is investigated by Engle and Russell (1997).

- the disequilibrium between the volume traded on the buy side and the volume traded on the sell side. Bauwens and Giot (1997) introduced this disequilibrium as an explicative variable for the expected conditional duration. This variable captured the relative importance of the disequilibrium between the buy side and the sell side of the market: a large value for the disequilibrium should lead to a shorter expected duration as market makers revise their quotes more rapidly when faced with a large volume on one side of the market. As suggested by the literature on information models in market microstructure, this revision is consistent with the Bayesian (learning) behavior of the market makers. This issue is also related to the notion of one sided volume, as introduced by Engle and Lange (1997). It should however be kept in mind that inventory models can also explain a more rapid updating of the bid/ask quotes when the market makers are faced with a large disequilibrium in the trades process. Indeed, inventory models such as introduced by Ho and Stoll (1981) predict that market makers will update their quotes on both sides of the market when faced with an increase or decrease of their optimal inventory (for example, following a large rise in their inventory due to a succession of sell orders, market makers will decrease their bid and ask quotes to elicit buying from the traders). Excellent reviews of both types of models are given in Biais, Foucault et Hillion (1997) and O'Hara (1995).

With the new model proposed in the previous section, this last issue can be further investigated by including the signed disequilibrium[3] as an additional variable in equations (10) and (13). The signed disequilibrium is measured by

$$dis_i = \frac{\sum_k B_k vol_k - \sum_k S_k vol_k}{\sum_k vol_k} \tag{16}$$

where k spans all trades made between the two consecutive bid/ask quotes that occured at time t_{i-1} and t_i. B_k is an indicator variable that takes the value of 1 if the k- th trade is a buy and the value of 0 otherwise, S_k takes the value of 1 if the k-th trade is a sell and the value of 0 otherwise, vol_k is the volume attached to the k-th trade and the subscript i characterizes the successive bid/ask durations. Thus dis_i is a measure of the signed buy/sell disequilibrium during duration x_i, prior to the quote updating at t_i: dis_i is close to 0 if the disequilibrium is small, dis_i is close to -1 if the volume on the sell side is much larger than the volume on the buy side, while dis_i is close to 1 in the other case.

[3]Bauwens and Giot (1997) take the absolute value of the numerator of (16), so that their measure gives the relative importance of the disequilibrium, but does not indicate whether the disequilibrium comes from the buy side or from the sell side.

The new specification[4] for the expected conditional duration relative to the end state $y_i = 1$ is thus

$$\psi_i^+ = (\omega_1 + \alpha_1 \epsilon_{i-1}^+) I_{i-1}^+ + (\omega_2 + \alpha_2 \epsilon_{i-1}^+) I_{i-1}^- + \beta^+ \psi_{i-1}^+ + \eta^+ dis_i \quad (17)$$

while the specification for the end state $y_i = -1$ is

$$\psi_i^- = (\omega_3 + \alpha_3 \epsilon_{i-1}^-) I_{i-1}^+ + (\omega_4 + \alpha_4 \epsilon_{i-1}^-) I_{i-1}^- + \beta^- \psi_{i-1}^- + \eta^- dis_i \quad (18)$$

If market makers revise their quotes in a Bayesian fashion based on the evolution of the trade process, the additional explicative variable should be significant. Furthermore, as we now include the signed disequilibrium, η^+ should be significantly negative (a large disequilibrium on the buy side, i.e. dis_i close to 1, shortens the expected duration of an upward quote revision, or correspondingly increases the hazard of exiting to a +1 state). In a similar way, η^- should be significantely positive (a large disequilibrium on the sell side, i.e. dis_i close to -1, shortens the expected duration of a downward quote revision, or correspondingly increases the hazard of exiting to a -1 state).

These issues are investigated in the next section, where we apply our asymmetric model to actual bid/ask quotes data.

4 An application to the IBM stock

As in Bauwens and Giot (1997), we tested our model on data for the IBM stock. IBM is mainly traded at the New York Stock Exchange (we did not take into account the data from the other regional stock exchanges). At the NYSE, a specialist collects all trades and acts as the only market maker. With the existing limit orders, he is the only supplier of liquidity.

The data were extracted from the Trades and Quotes Database (TAQ Database), released by the NYSE. 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 selected the "regular" trade quotes. Trades and bid/ask quotes recorded before 9:30 AM and after 16:00 PM were also deleted. Some information about the data is given in Table 1.

Regarding the IBM data, the number of bid/ask quotes (34,321) is approximately half the number of trades (61,063). We filtered the bid/ask quotes durations and retained those leading to a significant cumulated change in the mid price

[4] As dis_i contains information on the trades process, ψ_i^+ and ψ_i^- are now conditioned on the information at $i-1$ and the information conveyed by the trades process (which includes the trades made during duration x_i, up to the last trade made before the end of the duration).

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 quotes 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).

When filtering the data and retaining the durations leading to at least a \$0.125 cumulated change in the mid price, the average number of trades over the filtered bid/ask quote duration is equal to 9.1, while the mode is at 4. Several other thresholds have been tried:

- no filtering: the average number of trades over the bid/ask quote duration is equal to 1.8, while the mode is at 0; thus, without filtering the data, the number of buys or sells occuring in the spell of duration x_i - see (16) - is too small to measure sensibly the state of the market, at least for a high proportion of the durations;

- retaining the durations leading to a cumulated \$0.25 change in the mid price: the average number of trades over the bid/ask quote duration is equal to 27.8, while the mode is at 8; although the results are not given in this paper, the computation of the empirical transition probabilities and the estimation of the asymmetric ACD model yield the same type of results as those given below for the filtering at \$0.125.

As Engle and Russell (1995), we transformed 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 is called 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:

$$X_i = x_i \, \phi(t_i) \tag{19}$$

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) standardized duration. The deterministic tod effect is defined

Table 1: Information on the data

	IBM
Number of trades	61,063
Number of b/a quotes	6,728
Mean of x_i and X_i	1 - 215.92
Standard deviation of x_i and X_i	1.51 - 365.46
Minimum value x_i and X_i	0.0030 - 1
Maximum value x_i and X_i	35.51 - 7170

Data extracted from the September, October, and November 1996 TAQ CD-ROM of NYSE. The given number of bid/ask quotes is the number obtained after filtering the data at \$0.125 (the original number of bid/ask quotes was equal to 34,321). x_i is a time-of-day standardized duration, see (19), and is measured in seconds. The mean of x_i is equal to 1 by construction, after the removal of the time-of-day effect. X_i is a (non standardized) filtered duration.

as the expected duration conditioned on time-of-day, where the expectation is computed by averaging the durations over thirty minutes intervals, no difference being made between the different trading days of the week.

Prior to estimating the model, we computed empirical conditional transition probabilities for the filtered bid/ask quotes process. The four conditional transition probabilities, given in Table 2, indicate that the succession of similar states (either price increase or decrease) is more likely than price reversals (price increase followed by a price decrease, or the opposite).

The results given in Table 2 are stable when other threshholds (for example filtering the data by retaining durations leading to a \$0.25 or \$0.5 change in the mid price) are chosen, and when subperiods of the original sample are selected (for example selecting the first half of the data). In all cases, the succession of similar states is more likely than price reversals. This is an important issue, as it implies that the data exhibits positive serial correlation when one deals with filtered durations[5].

Estimates of the model are reported in Table 3. Estimation has been performed by maximum likelihood (BHHH algorithm in GAUSS). The upper block of Table 3 is for the part of the model when the end state is a price increase

[5]This important issue has been extensively studied in the finance litterature, as it is related to the market efficiency hypothesis (see for example Campbell, Lo and MacKinlay (1997), chapter 2). However most studies have been conducted on low frequency data.

Table 2: Empirical conditional transition probabilities

	$y_i = +1$	$y_i = -1$
$y_{i-1} = +1$	57.48	42.52
$y_{i-1} = -1$	45.12	54.88

Empirical conditional transition probabilities computed for the filtered bid/ask quotes. $y_i = +1$ corresponds to an increase in the mid-price from $i - 1$ to i, $y_i = -1$ corresponds to an decrease in the mid-price from $i - 1$ to i.

- see equations (9) and (17) - and the lower block is for the part when the end state is a price decrease - see (12) and (18). The main features of these results are:

1) strong autoregressive effects (β coefficients) when the disequilibrium is not included; this result is similar to those of Bauwens and Giot (1997) for the symmetric ACD model. However in the symmetric version, the introduction of the disequilibrium did not change the value of β.

2) Without the disequilibrium variable, the estimates of the γ parameters are quite close to 1, implying exponential distributions for the innovations. With the disequilibrium variable, a Weibull distribution with a slightly decreasing hazard function is favoured. Moreover, the hypothesis that $\gamma^+ = \gamma^-$ is not rejected at conventional levels of significance in both cases.

3) When the end state is a price increase ($y_i = 1$), α_1 is much smaller than α_2: the expected duration of witnessing a transition from state +1 to state +1 is smaller than the expected duration of witnessing a transition from state -1 to state +1. Succession of +1 states are thus more likely than $-1, +1$ transitions, as was expected from the empirical transition probabilities given in Table 2.

4) When the end state is a price decrease ($y_i = -1$), α_3 is larger than α_4, at least in the case where the disequilibrium variable is included; the expected conditional duration of witnessing a transition from state -1 to state -1 is smaller than the expected duration of witnessing a transition from state +1 to state -1. Once again, succession of similar states (in this case -1 states) are more likely than $+1, -1$ transitions, in agreement with the proportions given in the last column of Table 2.

5) The estimates of η^+ and η^- are strongly significant, with η^+ taking a large negative value approximately opposite to η^- (the hypothesis that the two parameters are equal in absolute value is not rejected at the 1 par cent level). These coefficients have the expected sign, based on the discussion at the end of Section 3: a large positive disequilibrium (i.e. on the buy side) decreases the expected duration of an exit to state +1 (η^+ negative); a large negative disequilibrium (i.e. on the sell side) decreases the expected duration of an exit to state -1 (η^- positive).

Table 3: ML results for IBM
(asymmetric Log-ACD model)

	End state: price increase	
ω_1	-0.098 (0.013)	0.781 (0.044)
ω_2	-0.061 (0.012)	0.472 (0.062)
α_1	0.127 (0.021)	0.314 (0.034)
α_2	0.229 (0.027)	0.728 (0.071)
β^+	0.980 (0.005)	0.160 (0.027)
η^+	-	-1.242 (0.038)
γ^+	0.986 (0.012)	0.942 (0.012)
	End state: price decrease	
ω_3	-0.076 (0.012)	-0.112 (0.064)
ω_4	-0.090 (0.013)	0.654 (0.043)
α_3	0.206 (0.026)	0.818 (0.077)
α_4	0.167 (0.022)	0.375 (0.034)
β^-	0.987 (0.004)	0.322 (0.034)
η^-	-	1.159 (0.036)
γ^-	0.991 (0.013)	0.936 (0.012)

For a definition of the model and of the parameters, see equations (9) and (17) for the price increase and (12) and (18) for the price decrease. BHHH asymptotic standard errors are given inparantheses.

Using a likelihood ratio test, it is possible to test the asymmetric Log-ACD model (without the disequilibrium variable) introduced in this paper against the symmetric version of Bauwens and Giot (1997).[6] The asymmetric model introduced in Section 3.2 - given by equations (9), (10), (12) and (13)) - can be constrained into a symmetric one by setting $\omega_1 = \omega_2 = \omega_3 = \omega_4$, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$, $\beta^+ = \beta^-$, and $\gamma^+ = \gamma^-$, so that $\psi_i^+ = \psi_i^-$. The constrained model is then equivalent to the original Log-ACD model introduced by Bauwens and Giot (1997). Setting $\gamma^+ = \gamma^- = \gamma$, $\psi_i^+ = \psi_i^- = \psi_i = \omega + \alpha\epsilon_{i-1} + \beta\psi_{i-1}$ and $\Psi_i = e^{\psi_i}$, the likelihood of the constrained model is

$$L(x_i|J_{i-1}) = 2\frac{\gamma}{\Psi_i}\left(\frac{x_i}{\Psi_i}\right)^{\gamma-1} e^{-2\left(\frac{x_i}{\Psi_i}\right)^{\gamma}} \tag{20}$$

Once the parameters (ω, α, β and γ) of the model have been estimated, the corresponding log-likelihood can be computed using (20), and compared to the log-likelihood of the asymmetric model as given by equation (15). The estimated values of the parameters (and the standard errors) are $\omega = -0.091$ (0.007), $\alpha = 0.088$ (0.007), $\beta = 0.984$ (0.003), $\gamma = 0.987$ (0.009) and the log-likelihood is equal to -10343.48. The log-likelihood of the asymmetric model is equal to -10329.48. Thus the test statistic, which is a $\chi^2(8)$, is equal to 28: it is significant at the one percent level.

From the results given in Table 3, including the disequilibrium on the volume traded during duration x_i thus accounts for a large part of the dynamic structure of the filtered bid/ask quotes process. At the beginning of this section, a filtered duration was defined as the minimum amount of time needed for the mid-price to increase or decrease by at least \$0.125. The filtered durations are thus inversely related to the volatility of the mid- price, and the ACD model on filtered durations can be seen from a different viewpoint, mainly that of volatility clustering models. In this perspective, the results given in Table 3 indicate that one can partially explain the volatility of the mid-price by information related to the volume traded (through the disequilibrium variable). This kind of results has already been put forward for daily GARCH processes (on the transaction prices) by Lamoureux and Lastrapes (1990).

5 Conclusion

ACD models seem to be a useful and promising tool not only to model durations arising from irregular events occuring on financial markets, but also as a building block of larger models bearing on features of the market events, such as price, volume, volatility. This paper has proposed a two state transition model for

[6]The Log-ACD model of Bauwens and Giot (1997) with the *unsigned* disequilibrium variable (i.e. in (16), one takes the absolute value of the numerator) is not nested in the asymmetric Log-ACD model with the *signed* disequilibrium, proposed in this paper. A likelihood ratio test is thus not feasible in this case.

the mid bid/ask price of a stock, where the two states correspond either to a price increase or to a price decrease, and where the durations are modelled by an ACD process. Our empirical results favour the asymmetric specification as an enrichment of the symmetric ACD one, but more empirical research is needed to know whether this holds more generally.

Extensions of this model could consist of considering a more refined state space (more than two states), other distributions than the Weibull for the innovations (possibly a semi-parametric specification), and more refined parameterizations of the autoregressive equations for the conditional durations. More fundamentally, theoretical foundations for the determinants of the asymmetry revealed by our empirical results should be uncovered.

References

[1] Biais, B., Hillion, P. and Spatt, C. (1995). An empirical analysis of the limit order book and the order flow in the Paris Bourse. *Journal of Finance* 50, 1655-1689.

[2] Biais, B., Foucault, T. et Hillion, P. (1997). *Microstructure des marchés financiers: institutions, modèles et test empiriques.* Paris: Presses Universitaires de France.

[3] Bisière, C. and Kamionka, T. (1998). Timing of orders, orders aggressiveness and the order book at the Paris Bourse. Mimeo. GREMAQ, Toulouse.

[4] Bauwens, L. and Giot, P. (1997). The Logarithmic ACD model: an application to the Bayesian updating behavior of market makers. Revised version of CORE DP 9789. UCL, Louvain- La-Neuve.

[5] Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997). The econometrics of financial markets.. Princeton University Press.

[6] Darolles, S. and Le Fol, G. (1998). Discrete price and random time: transaction price dynamics. Mimeo. CREST, Paris.

[7] Easley, D. and O'Hara, M. (1992). Time and the process of security price adjustment. *The Journal of Finance* 19, 69-90.

[8] Easley, D., Kiefer, N.M. and O'Hara, M. (1997). The information content of the trading process. *The Journal of Empirical Finance* 4, 159-186.

[9] Engle, R. (1996). The econometrics of ultra-high frequency data. *Discussion Paper 96-15.* University of California, San Diego.

[10] Engle, R. and Lange, J. (1997). Measuring, forecasting and explaining time varying liquidity in the stock market. *Discussion Paper 97-12.* University of California, San Diego.

[11] Engle, R. and Russell, J. (1995). Autoregressive conditional duration: a new model for irregularly spaced time series data. *Mimeo.* University of California, San Diego.

[12] Engle, R. and Russell, J. (1997). Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model. *Journal of Empirical Finance* 4, 187-212.

[13] Engle, R. and Russell, J. (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.

[14] Ghysels, E. and Jasiak, J. (1997). GARCH for irregularly spaced financial data: the ACD-GARCH model. *Discussion Paper 97s-06.* CIRANO, Montréal.

[15] Glosten, L. and Milgrom, P. (1985). Bid, ask, and the transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 13, 71- 100.

[16] Ghysels, E., Gouriéroux, C. and Jasiak, J. (1997). Stochastic volatility duration models. Mimeo. CREST, Paris.

[17] Ho, T. and Stoll, H. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9, 47-73.

[18] Lamoureux, C.G. and Lastrapes, W.D. (1990). Heteroskedasticity in stock return data: volume versus GARCH effects. *Journal of Finance* 45, 220-229.

[19] Meddahi, N., Renault, E. and Werker, B. (1998). Modelling high-frequency data in continuous time. Mimeo.

[20] O'Hara, M. (1995). *Market microstructure theory.* Oxford: Basil Blackwell.