
ACD models: Models for data irregularly spaced in time

A gentle introduction, April 2008

Guillaume Weisang, *Doctoral Student,*
Department of Mathematical Sciences,
Bentley College

Abstract: This paper has two goals. First, this paper reviews the family of Augmented Autoregressive Conditional Duration (AACD) models. Second, I provide source code for the estimation of some linear ACD models as well as examples of empirical applications. A warning is needed however: this is part of an ongoing work. Please report to the author any mistake. Your comments will be most appreciated as well.

Introduction

Since the beginning of the 1990s, with the availability of new data storage, the finance industry has had a growing interest in what is called high frequency data analysis (also sometimes called ultra-high-frequency data). This has led to several important contributions in the domain of econometrics, particularly dealing with irregularly spaced data. This paper focuses on the Autoregressive Conditional Duration (ACD) model, introduced by Engle and Russell in 1998. In substance, the paper aims first at providing an exposé sufficient for any non expert reader interested in the subject to grasp the origins of the ACD model, i.e. the problem it purports to solve, as well as the fundamental properties and results tied to it. Secondly, it contributes a number of source programs in *R* which can then be used by the reader to estimate her own model. I must warn the reader however that it is still ongoing work, and therefore there is a lot of room for improvement, especially on the implementation side.

I will start by providing the reader background information that should give enough light to carry through the mathematical equations of the second part, where the linear ACD model is presented. In a third part, I try to give the reader enough information and knowledge to use a broader ACD model: the Augmented ACD model. Finally in a fourth part, I provide the estimation functions and some examples of the simplest models, and I discuss the results of my estimations in an effort to provide the reader with an example of what can easily be achieved.

Background

Understanding some of the most interesting features of ACD models calls for a historical perspective on the motivations that lead to their formulation and subsequent extensive study. The ACD model was first published in the seminal paper of Engle and Russell in 1998 (Engle and Russell 1998) as an answer to

a peculiar problem that the recently new availability of data in finance had created. By making accessible to researchers the records of every single operation, including quotes and transactions, at the beginning of the decade, the stock exchanges – NYSE in the lead – had given to the finance research community a new field of inquiry as well as a true challenge to solve. The reason was twofold. On the one hand, access to this tremendous amount of real data for analysis meant potentially enhancing our understanding of the market’s microstructure and its influence on the trading process. From a business perspective, for the professionals intervening on those markets, better understanding meant potential better risk management or better reading of market’s movements. On the other hand, the very nature of these data was a challenge for classic time series analysis. Throughout the paper, when opportune, I will try to point to the contribution of ACD models to answering those various endeavors, but first I will explain the nature of the challenge to which ACD models meant at first to answer.

The problem of irregularly spaced data

In classical econometric techniques, time series are often taken to be sequences of data points separated by uniform time intervals. As Engle (2000) reports, one measure of progress in econometrics is the frequency of data from Liu’s quarterly model of the US economy in 1963 to the level that is the very subject of this paper. The frequency of a time series is a measure of the fineness of the available information about the postulated underlying reality that creates it. In a way, it represents the rate at which an economic phenomenon, like a price, is sampled. The more often it is observed, the better knowledge we have of it, and the better control we can have of the eventual consequences of the decisions taken with respect to it, e.g. economic policies or investment decisions. Hence, the quest for ever higher frequencies was a quest for better or fuller information. In a way, both data frequency and modeling techniques were developed according to the advancement of information technologies, essential for keeping records and processing the scores of pieces of information involved. Yet, with the advent of transactions and quotes databases, researchers stumbled on a problem that required a new kind of modeling.

Somewhat counter to intuition, one cannot sample at any arbitrary frequency. Reaching a certain level, as shown on figure 1, one encounters what could be called “atomic” events that are irregularly spaced in time. Those events represent the points in time where new information on the phenomenon is released. Let’s take an example to help ground those ideas. Consider a trading process. One important measure associated with trade is the price at which transactions are made. Even if we suppose that a good always has a price, it is never observed unless a quote, i.e. an offer to buy or sell at a certain price, is made. To track the price of a good, one can observe it at regular time intervals, like every day or every hour, but one only observes the last quote that was made. This type of regular observations is sometimes referred to as “aggregated” because the change from one observation to another is the result of an aggregation of all the quotes made in the time spell between. In the case of trade, the quotes and the transactions represent the “atomic” events that reveal the price of a good. As we augment the frequency of observations we soon arrive at a time scale commensurate with the time scale of those “atomic” events. If we were to go further, nothing new would be learnt, for in between the “atomic” events, no new information is released. Furthermore, in some cases, the time elapsed between those events, the duration, may be revealing in itself of the dynamic of the underlying process. In the case of trade, sequences of successive short durations are synonymous with higher activity than usual and may signal certain behavior from knowledgeable investors. Researchers, and business people alike, are interested in getting at this information, and therefore push for models to extract it.

INSERT FIGURE 1 AROUND HERE

Let's note an atomic event (t, P_t) , i.e. a couple formed by the time of the event and its "mark," a price for example. Let's note the data: $\{(t_i, P_{t_i}), i = 1, \dots, N\}$ where the i^{th} observation has a joint density conditional on the past available information given by

$$(1) \quad (t_i, P_{t_i})|F_{i-1} \sim f(t_i, P_{t_i}|\check{t}_{i-1}, \check{P}_{t_{i-1}}; \theta_i)$$

where $\check{z}_i = \{z_i, z_{i-1}, \dots, z_1\}$ denotes the past of z up to time t_i , and θ_i are parameters potentially different for each observation.

Traditional econometric techniques have been designed to model marks observed at homogenous time intervals, and therefore are not immediately adequate to model these data because they fail to extract the time information. One remark made *inter alia* by Engle (2000) eventually lead to the development of ACD models. The data are irregularly spaced in times, but with respect to the "atomic" events, the marks are in fact regularly spaced: exactly every event. In mathematical terms, this is can be seen by rewriting (1), without loss of generality, under the form

$$(2) \quad f(t_i, P_{t_i}|\check{t}_{i-1}, \check{P}_{t_{i-1}}; \theta_i) = g(t_i|\check{t}_{i-1}, \check{P}_{t_{i-1}}; \theta_{1i})q(P_{t_i}|t_i, \check{t}_{i-1}, \check{P}_{t_{i-1}}; \theta_{2i})$$

where g denotes the conditional density function of the time t of an "atomic" event given the past and q the conditional density function of the mark P_t given the past and the time t . In other words, in the "subjective" time related to the "atomic" events, one can use traditional econometric techniques to model the marks, i.e. model the function q . ACD models attempt to model the transition from our "objective" time to the "subjective" time of the information process. In fact, ACD models are concerned with durations, i.e. the time elapsed between events. Let's define $x_i = t_i - t_{i-1}$ the duration between events i and $i - 1$. One can rewrite (1) and (2) under the following form

$$(x_i, P_i)|F_{i-1} \sim f(x_i, P_i|\check{x}_{i-1}, \check{P}_{i-1}; \theta_i)$$

$$f(x_i, P_i|\check{x}_{i-1}, \check{P}_{i-1}; \theta_i) = g(x_i|\check{x}_{i-1}, \check{P}_{i-1}; \theta_{1i})q(P_i|x_i, \check{x}_{i-1}, \check{P}_{i-1}; \theta_i)$$

To sum up, the idea is to produce a type of recursive models, where a traditional model is used on top of a layer modeling time. Recent examples include the Autoregressive Conditional Multinomial-Autoregressive Conditional Duration (ACM-ACD) model of Russell and Engle (2005), the Intradaily VaR (IVaR) model of Dionne, Duchesne and Pacurar (2005) or the ACD-GARCH model of Ghysels and Jasiak (1998). However, ACD models are models in their own right, and belong to the category of duration models which are used primarily in Survival Analysis. In the next part I introduce the ACD models and discuss their relation to more classic duration models.

Modeling Durations and ACD models

Modeling durations is not a new endeavor. Duration models have been introduced in survival analysis for a few decades now. However, ACD models are made for situations not often encountered in survival analysis. As Engle and Russell (1998) noticed, even when time series are considered, the typical survival analysis dataset include short time series on many individuals, whereas in the case of trade for example, one considers long time series with at least one idiosyncrasy: a clustering feature. In the section below, I will first explain this particular characteristic, before introducing formally the ACD model and discuss its relation with mainstream duration models. Finally, I examine some simple implementations and give the corresponding Log-likelihood functions.

The problem of Clustering

Typically, the type of data that call for the use of ACD models exhibits a high dependency from one data point to another. In the finance literature (e.g. Tsay 2005), this feature is often referred to as clustering. Figure 2 presents a display of this clustering feature for the durations related to a change in price of the IBM stock at the New York Stock Exchange on Sep 1st, 2004. Clusters of short durations are

clearly visible. It means that short durations tend to be followed by short durations. Similarly, long durations tend to be followed by long durations, even though it is harder to see on figure 2. This is further confirmed when looking at the empirical autocorrelation function on figure 3. The autocorrelation coefficients decrease very slowly, being still significant after 500 lags. This clustering feature is not without resemblance to the clustering feature of volatility. Indeed, the notion of volatility bears a tight relationship with ACD models (Engle and Russell 1998; Ghysels and Jasiak 1998; Engle 2000; Grammig and Wellner 2002; Pacurar 2006). Furthermore, the equations of the ACD model are almost identical to the GARCH model (Tsay 2005).

INSERT FIGURES 2 AND 3 AROUND HERE

The ACD model

ACD model stands for Autoregressive Conditional Duration model. In the following section, I present the linear ACD model in its simplest form as it was introduced in the seminal paper of Engle and Russell (1998). It is easily generalizable to higher orders. As its name explicitly states, the ACD model imposes a generalized autoregressive structure to the conditional duration. Let's note ψ_i the conditional duration:

$$\psi_i = E_{i-1}(x_i | \check{x}_{i-1}, \check{P}_{i-1})$$

The linear ACD(1,1) is given by

$$(3) \quad \begin{cases} x_i = \psi_i \varepsilon_i & \text{(Mean equation)} \\ \psi_i = \omega + \alpha x_{i-1} + \beta \psi_{i-1} & \text{for } \alpha, \beta \geq 0, \omega > 0 \quad \text{(ACD equation).} \end{cases}$$

In a slightly more general fashion, Grammig and Maurer (2000) rewrite the mean equation of the model so that the relationship between the density function of ε_i and the conditional duration is more explicit:

$$(4) \quad \varepsilon_i = \frac{x_i}{f(\psi_i)} \quad \text{and} \quad f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$$

with the ε_i independent and identically distributed and their density satisfies

$$g\left(\frac{x_i}{f(\psi_i)} | \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right) = g\left(\frac{x_i}{f(\psi_i)}; \theta_g\right).$$

The last equation simply implies that the time dependence of the duration process is summarized by the conditional expected duration sequence. θ_g is a parameter vector. f is a strictly positive function with positive support and follows from the choice of g to the extent that the expected value of standardized duration ε must be equal to 1. Several possible choices have been considered in the literature, and I will present below three of the simplest ones along with an empirical application for each.

One last remark is that (3) has a convenient form because it allows various moments of x_i to be calculated by expectation regardless of the form of the baseline hazard function – which concept I introduce below. The unconditional mean and variance then have the form (Engle and Russell 1998):

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

$$Var(x_i) = \sigma^2 = \mu^2 \left(\frac{1 - \beta^2 - 2\alpha\beta}{1 - \beta^2 - 2\alpha\beta - 2\alpha^2} \right).$$

Note first that these two moments give stationarity conditions for the coefficients α and β : $\alpha + \beta < 1$ and $\beta^2 + 2\alpha\beta + 2\alpha^2 < 1$. Second, if $\alpha > 0$, the unconditional standard deviation σ will exceed the unconditional mean resulting in “excess dispersion” as so often noticed in duration data sets.

Relation to the classical survival analysis framework

In order to relate the ACD model to the classical survival analysis framework, let's take the example of the trade process on a financial market. Let's consider the process $N(t)$ that counts the number of transactions that have occurred at time t since the beginning of the day; and let's define the conditional intensity λ as

$$(5) \quad \lambda_i(t, \tilde{x}_{i-1}, \tilde{P}_{i-1}) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(N(t + \Delta t) > N(t) | \tilde{x}_{i-1}, \tilde{P}_{i-1})}{\Delta t} \\ = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t_{i-1} < T_i < t_{i-1} + \Delta t | t_0, t_1, \dots, t_{i-1}, \tilde{P}_{i-1})}{\Delta t}$$

where $T_i > t_{i-1}$ is the arrival time of the next event. That is, T_i is a random variable whose realization is t_i the time of the next event. The conditional intensity summarizes future expected arrival rates of the transactions and how these depend upon observables¹. It is therefore a natural quantity to look at. Let's now define the survival function S associated with the errors ε_i by: $S(t; \theta_g) = \int_t^\infty g(s; \theta_g) ds$. The *baseline hazard* is given by

$$\lambda_0(t; \theta_g) = \frac{g(t; \theta_g)}{S(t; \theta_g)}.$$

Understanding the role of the baseline hazard function is easier in the classic context of the Cox proportional hazard model. In the latter, the intensity of the process is given by

$$\lambda(t, \mathbf{z}) = \lambda_0(t) \exp(\mathbf{z}' \boldsymbol{\beta})$$

with \mathbf{z} a vector of covariates and $\boldsymbol{\beta}$ a vector of parameters. The baseline hazard λ_0 corresponds to the case where all predictors \mathbf{z} are equal to zero (Cox 1972; Haughton and Haughton *Forthcoming*). In the case of ACD models, the baseline hazard function is defined with respect to the errors ε_i only, and takes on a slightly different meaning. It represents the natural intensity of the process in its own time scale. In fact, one can easily show the following relationship between the conditional intensity, the baseline hazard and the conditional expected duration:

$$(6) \quad \lambda_i(t, \tilde{x}_{i-1}, \tilde{P}_{i-1}) = \frac{1}{f(\psi_i)} \lambda_0\left(\frac{t - t_{i-1}}{f(\psi_i)}; \theta_g\right).$$

In equation (6), $(f(\psi_i))^{-1}$ acts as a time accelerating factor whereas the baseline hazard intervenes in terms of the "standardized" duration $(t - t_{i-1})(f(\psi_i))^{-1}$. An ACD model "can use but does not require auxiliary data or assumptions on the causes of time flow" (Engle and Russell 1998). It is where the difference with the Cox proportional hazard model – in which the time accelerating factor depends on the value of covariates \mathbf{z} —lies. The ACD formulation is a pure time series model of time.

Note that some authors (cf. Grammig and Maurer 2000; Grammig and Wellner 2002) sometimes introduce pre-determined variables z_i in the ACD equation of (3), thus becoming

$$\psi_i = \omega + \alpha x_{i-1} + \beta \psi_{i-1} + \zeta z_i.$$

This change in specification is primarily made in order to test for some hypotheses on the market's microstructure. I will not develop this topic any further in this mostly expository paper but numerous papers are readily available to the interested reader (e.g. Engle and Russell 1998; Grammig and Maurer 2000; Zhang et al. 2001; Grammig and Wellner 2002; Pacurar 2006 and the references therein).

Three possible distributions for the standardized durations

In this section, I introduce three simple choices for the distribution of the standardized durations $\varepsilon_i = \frac{x_i}{f(\psi_i)}$ and give the expressions of their related hazard functions and conditional log-likelihood functions. The first two distributions, the Exponential-ACD (EACD) and the Weibull ACD (WACD), were

¹The conditional intensity is sometimes called the conditional hazard function, and the terms will be used interchangeably throughout the remaining of the paper.

introduced by Engle and Russell in their seminal paper in 1998. The last one, the generalized Gamma ACD (GACD) was introduced by Zhang, Russell and Tsay (2001) as a generalization of the WACD model (itself a generalization of the EACD model).

The Exponential ACD model

If one assumes that $f(\psi_i) = \psi_i$ and that the density of the standardized durations is an exponential distribution with parameter equals to 1,

$$(7) \quad g\left(\frac{x_i}{\psi_i} \mid \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right) = \exp\left(-\frac{x_i}{\psi_i}\right)$$

we obtain the Exponential-ACD model. One can easily derive that

$$S\left(\frac{x_i}{\psi_i}; \theta_g\right) = \int_{\frac{x_i}{\psi_i}}^{\infty} \exp(-s) ds = \exp\left(-\frac{x_i}{\psi_i}\right)$$

and therefore, $\lambda_i(t, \check{x}_{i-1}, \check{P}_{i-1}) = \psi_i^{-1}$. In other words, the conditional intensity of the i th duration is flat and does not depend on the time t . From (7), one also easily derive the conditional density of x_i as

$$g(x_i \mid \check{x}_{i-1}, \check{P}_{i-1}; \theta_g) = \frac{1}{\psi_i} \exp\left(-\frac{x_i}{\psi_i}\right).$$

Note that this is an exponential distribution with parameter $(\psi_i)^{-1}$. The likelihood function associated with any linear ACD(1,1) model for a set of durations $\mathbf{x}_T = (x_1, \dots, x_T)'$ is given by

$$(8) \quad G(\mathbf{x}_T \mid \theta_g) = \left[\prod_{i=2}^T g(x_i \mid \check{x}_{i-1}, \check{P}_{i-1}; \theta_g) \right] \times g(x_1 \mid \theta_g).$$

The conditional log-likelihood is then

$$L(\mathbf{x}_T \mid \mathbf{x}_1, \theta_g) = - \sum_{i=2}^T \left(\ln \psi_i + \frac{x_i}{\psi_i} \right).$$

The Weibull ACD model

The problem of a flat conditional intensity was already raised by Engle and Russell as not having a good fit with some semiparametric estimate of the baseline hazard of the data (Engle and Russell 1998), and they therefore propose to extend the EACD model by generalizing the exponential density of the standardized durations to a Weibull(1, γ) density. In this case, $f(\psi_i) = \phi_i = \psi_i \left\{ \Gamma\left(1 + \frac{1}{\gamma}\right) \right\}^{-1}$ and the density of the standardized durations is given by

$$(9) \quad g\left(\frac{x_i}{\phi_i} \mid \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right) = \frac{\gamma \phi_i}{x_i} \left(\frac{x_i}{\phi_i}\right)^{\gamma} \exp\left(-\left(\frac{x_i}{\phi_i}\right)^{\gamma}\right).$$

As for the exponential case, it is straightforward to derive from (9) the conditional density of the duration

$$g(x_i \mid \check{x}_{i-1}, \check{P}_{i-1}; \theta_g) = \frac{\gamma}{x_i} \left(\frac{x_i}{\phi_i}\right)^{\gamma} \exp\left(-\left(\frac{x_i}{\phi_i}\right)^{\gamma}\right).$$

The conditional hazard function implied by the WACD is given by $\lambda_i(x_i \mid \check{x}_{i-1}, \check{P}_{i-1}; \theta_g) = \phi_i^{-\gamma} x_i^{\gamma-1} \gamma$. If $\gamma = 1$ in the expressions above, the WACD model reduces to the exponential model. The WACD model allows therefore for more flexibility in the conditional hazard function that is either increasing, if $\gamma > 1$, or decreasing if $0 < \gamma < 1$. The conditional log-likelihood function of a WACD(1,1) model on a set of observed durations $\mathbf{x}_T = (x_1, \dots, x_T)'$ has expression

$$L(\mathbf{x}_T \mid \mathbf{x}_1, \theta_g) = \sum_{i=2}^T \left[\ln\left(\frac{\gamma}{x_i}\right) + \gamma \ln\left(\frac{x_i}{\phi_i}\right) - \left(\frac{x_i}{\phi_i}\right)^{\gamma} \right].$$

The Generalized Gamma ACD model

Similarly to the case of the EACD model, one can regret a lack of flexibility from the WACD model since it allows only for monotonic hazard functions. Therefore, Zhang, Russell and Tsay (2001), as well as Tsay (2005), make use of a generalized gamma distribution to characterize the standardized durations because one can then obtain a **U**-shaped and inverse **U**-shaped hazard function.

In this case, let's choose $f(\psi_i) = \xi_i = \psi_i \frac{\Gamma(\kappa)}{\Gamma(\kappa + \frac{1}{\gamma})}$, and let's note the conditional density of the standardized duration by

$$(10) \quad g\left(\frac{x_i}{\xi_i} | \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right) = \frac{\gamma \xi_i}{\varpi \Gamma(\kappa) x_i} \left(\frac{x_i}{\xi_i}\right)^{\kappa \gamma} \exp\left(-\left(\frac{x_i}{\xi_i}\right)^\gamma\right) \sim f_\Gamma\left(\frac{x_i}{\xi_i}; \gamma, \kappa\right)$$

where $\varpi = \frac{\Gamma(\kappa)}{\Gamma(\kappa + \frac{1}{\gamma})}$ and γ, κ are shape parameters (equation (5.56) in Tsay 2005). Note that if $\kappa = 1$, the distribution in (10) then reduces to the one of the WACD and if further $\gamma = 1$, it reduces further to an EACD model. The conditional density of the duration is again deduced from (10) and follows

$$g(x_i | \check{x}_{i-1}, \check{P}_{i-1}; \theta_g) = \frac{\gamma}{\varpi \Gamma(\kappa) x_i} \left(\frac{x_i}{\xi_i}\right)^{\kappa \gamma} \exp\left(-\left(\frac{x_i}{\xi_i}\right)^\gamma\right).$$

To my knowledge, there is no simple expression of the conditional hazard in the case of the GACD model. However, one can recognize a gamma distribution for $\left(\frac{x_i}{\psi_i}\right)^\gamma$ with parameters κ and $\beta = \varpi^\gamma$ as given below:

$$f_\Gamma(x | \kappa, \beta) = \begin{cases} \frac{1}{\beta^\kappa \Gamma(\kappa)} x^{\kappa-1} \exp(-x/\beta) & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

In fact, $g\left(\frac{x_i}{\xi_i} | \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right) = \gamma f_\Gamma\left(\frac{x_i}{\psi_i} | \kappa, \varpi^\gamma\right)$. One can then deduce $G\left(\frac{x_i}{\xi_i} | \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right)$ the cumulative distribution function associated with $g\left(\frac{x_i}{\xi_i} | \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right)$:

$$G\left(\frac{x_i}{\xi_i} | \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right) = \gamma \frac{\Gamma_l\left(\left(\frac{x_i}{\psi_i}\right)^\gamma; \kappa, \varpi^\gamma\right)}{\Gamma(\kappa)} \text{ and } S\left(\frac{x_i}{\xi_i} | \check{x}_{i-1}, \check{P}_{i-1}; \theta_g\right) = \gamma \frac{\Gamma_u\left(\left(\frac{x_i}{\psi_i}\right)^\gamma; \kappa, \varpi^\gamma\right)}{\Gamma(\kappa)}$$

where $\Gamma_l(x; \kappa, \beta) = \int_0^x x^{\kappa-1} \exp(-x) dx$ is the lower incomplete Gamma function and $\Gamma_u(x; \kappa, \beta) = \int_x^\infty x^{\kappa-1} \exp(-x) dx$ is the upper incomplete Gamma function, which are well-known functions with various known numerical algorithms for approximation. The baseline hazard is then given by

$$\lambda_0(\varepsilon_i) = \Gamma(\kappa) \frac{f_\Gamma\left(\left(\frac{x_i}{\psi_i}\right)^\gamma | \kappa, \varpi^\gamma\right)}{\Gamma_u\left(\kappa, \left(\frac{x_i}{\psi_i}\right)^\gamma / \varpi^\gamma\right)}.$$

Making use of (6), one can then derive an expression for the conditional hazard function for the GACD model for which one can find a numerical approximation. Furthermore, the shape of the conditional hazard function can be derived from its parameters values (cf. Fernandes and Grammig 2005). If $\kappa \gamma < 1$, the hazard rate is decreasing for $\gamma \leq 1$, and U-shaped for $\gamma > 1$. Conversely, if $\kappa \gamma > 1$, the hazard rate is increasing for $\gamma \geq 1$, and inverted U-shaped for $\gamma < 1$. Finally, if $\kappa \gamma = 1$, the hazard is decreasing for $\gamma < 1$, constant for $\gamma = 1$, (exponential case), and increasing for $\gamma > 1$.

Moreover, the conditional log-likelihood of the GACD(1,1) model on a set of observed durations $\mathbf{x}_T = (x_1, \dots, x_T)'$ is then given by

$$L(\mathbf{x}_T | \mathbf{x}_1, \theta_g) = \sum_{i=2}^T \left[\ln(\gamma) + (\kappa \gamma - 1) \ln\left(\frac{x_i}{\xi_i}\right) - \ln(\Gamma(\kappa) \xi_i) - \left(\frac{x_i}{\xi_i}\right)^\gamma \right].$$

In the empirical section of the paper, I provide some applications of these models on IBM stock data. Before turning to those empirical matters, I want to discuss most of the multiple versions that can be found in the literature by discussing succinctly the family of models that were introduced by Fernandes and Grammig (2006). The purpose of the next section is not to provide an extensive study of the properties of the Fernandes and Grammig family of ACD models – that can be found in their paper— but rather to provide to a researcher new to the subject the rationale for using them.

Extending the ACD models: Augmented ACD models

The attentive and expert reader will have easily recognized in the structure of the ACD model similarities with the GARCH models. I have discussed in the background section the fact that the typical data that call for the use of an ACD model share some characteristics with volatility data, which is the typical data modeled with GARCH models in finance. In the background section, I talked only of clustering to explain why it requires a type of duration model different from the classic survival analysis framework. Yet, there are other features that are shared by both ACD and GARCH datasets.

Persistence

First, as indicated by figure 3, there is a phenomenon of long memory or persistence of the effect of an observation long after it is past. This is characteristic of processes with unit roots, i.e. processes such that $\alpha + \beta = 1$. However, this causes estimation and specification problems as we have seen from the stationarity conditions resulting from the expression of the unconditional expected duration. Furthermore, the coefficients of the linear ACD model are constrained to be positive. To extend the use of the ACD models, and following the lead of the GARCH literature researchers have provided multiple new specifications for ACD models. Most of the new specifications can be reduced to a shrewd choice of $f(\psi_i)$ (e.g. Bauwens and Giot 2000; Dufour and Engle 2000). In way similar to Hentschel (1995) in the GARCH literature, Fernandes and Grammig (2006) nest most of these models using the Box-Cox transformation,

$$f_\lambda(\psi_i) = \frac{\psi_i^\lambda - 1}{\lambda} \text{ for a parameter } \lambda \geq 0.$$

Asymmetry

A second transformation eventually leads from the linear ACD model to the family introduced by Fernandes and Grammig. It is linked to the remark that the impact of a “shock,” i.e. an unusually small or large value of ε_i , is generally asymmetrically distributed around their expected value. Hence, a shorter than expected duration may not have the same signification than a longer one. Furthermore, the shock impact may not be uniform with respect to its size. Intuitively, this is more readily understandable in terms of conditional intensity. For example on a market, it is saying that as the time since the last trade becomes anomalously long with respect to the usual range of its expected value, one may become less and less sensitive to the passage of time. Thus, the associated conditional intensity flattens with respect to time and similarly for the impact on the conditional duration. To model this feature of the data, Fernandes and Grammig introduce a shocks impact function whose shape can be controlled by a shift parameter b and a rotation parameter c :

$$\text{Shocks}(\varepsilon_i) = [|\varepsilon_i - b| - c(\varepsilon_i - b)]^\nu.$$

The parameter ν is a shape parameter which determines whether the shocks impact curve is concave, for $\nu \leq 1$, or convex, for $\nu \geq 1$. In figures 4 and 5, I present the effect of the parameters b , c and ν .

INSERT FIGURES 4 and 5 AROUND HERE

Finally, the Augmented ACD model (AACD) ensues from the two transformations above by

$$f_\lambda(\psi_i) = \omega_* + \alpha_* \psi_{i-1}^\lambda \text{Shocks}(\varepsilon_{i-1}) + \beta \frac{\psi_{i-1}^\lambda - 1}{\lambda}$$

that is to say, $\psi_i^\lambda = \omega + \alpha \psi_{i-1}^\lambda [|\varepsilon_{i-1} - b| - c(\varepsilon_{i-1} - b)]^v + \beta \psi_{i-1}^\lambda$

where $\omega = \lambda\omega_* - \beta + 1$ and $\alpha = \lambda\alpha_*$. In table 1, I report the typology of some ACD models nested in this family.

INSERT TABLE 1 AROUND HERE

Fernandes and Grammig (2006) provides far more general conditions for ergodicity, stationarity and existence of moments of the AACD than the ones I provided for the linear ACD model. I refer the reader to their paper, with a word of advice. Fernandes and Grammig base their results on the excellent work of Carrasco and Chen (2002), but a recent working paper by Meitz and Saikkonen (2007) which give similar results working for AACD models implies that there are subtle mistakes in this work. I therefore advise the interested reader to read all three papers in conjunction to avoid painful surprises.

Empirical applications

This section is divided in two parts. In a first part, I contribute and explain the code of the functions I used for the estimation of linear ACD models with respectively exponential, Weibull or generalized gamma densities on the standardized durations. In the second part, I present the result of the estimation procedures on a dataset of trade transactions on IBM stock during the month of September 2004.

Estimation procedure and source code

The estimation procedure maximizes the conditional likelihood of the model with respect to its parameters. It therefore uses constrained optimization to ensure stationarity. The estimation procedures have been written in *R* version 2.5.1 (R_Development_Core_Team 2007).

The core of the estimation procedure relies on the `constrOptim` function from the `stats` package of the *R* environment for optimization under linear constraints. The linear constraints are specified with \mathbf{U}_i and \mathbf{c}_i (respectively a $C \times k$ matrix and a C vector with C the number of linear constraints and k the total number of parameters in the model) such that $\mathbf{U}_i \cdot \boldsymbol{\theta} - \mathbf{c}_i \geq \mathbf{0}_C$. The model's parameters are stacked in the $\boldsymbol{\theta}$ vector. So far, the algorithm used is the Nelder-Mead algorithm used by default by the `constrOptim` function. It is a robust though slow algorithm.

In Appendix A, I give the source code of the functions given as an argument to the `constrOptim` function for the estimation. There are three types of functions. The first, `psi_model`, is simply the iterative procedure that gives the sequence $\{\psi_i\}_{i \in (1,T)}$ given the parameters $\boldsymbol{\theta}$ and the observations $\mathbf{x}_T = (x_1, \dots, x_T)'$. The second type of functions is the log-likelihood functions for each distribution. These functions take as arguments the data, the orders of the linear ACD model and the parameters. The output is the value of the log-likelihood function of the associated model on the data given the parameters values. Finally, one needs an envelope function around the log-likelihood function that takes only the parameters as input. This envelope function is then passed on to the `constrOptim` function for minimization.

The primary feature of this source code is that it works and gives the right results. However, multiple enhancements are possible and left for future developments. For example, the optimization procedure could be sped up by providing the analytical gradients for each of the log-likelihood functions. The idea would then be to use the faster BFGS algorithm of the `constrOptim` function. For now, we are limited to the Nelder-Mead algorithm because we are using only numerical gradients procedures. Secondly, the whole procedure estimation could be bundled into a single *R* function. As I will explain the ultimate goal is to implement, if possible, an *R* package for ACD models estimations.

Empirical Applications

I present here the results of the estimation procedures just presented on a dataset of price durations. I first present the data and discuss its idiosyncrasies. I then present and discuss 5 different models.

The Data

The dataset has been compiled from the Transactions and Quotes (TAQ) database from the New York Stock Exchange. It represents the durations between price changes in the prices of transactions on the IBM stock in September 2004. I considered only strictly positive durations making the classic assumption (Pacurar 2006), often verified in this dataset, that transactions happening in the same second are the results of large orders split into smaller ones so as to be passed more easily. The dataset is therefore composed of 47611 observations. Typically, market transactions follow a diurnal pattern on any trading day. In the morning, upon opening of the market, the trading activity is high as the market assimilates the events that occurred since the market last closed. The activity then gradually subsides to a typical low in intensity around noon, before picking up again for another peak of activity just before closure at 4:30pm. In the upper half of figure 6, I report a nonparametric estimate² of this diurnal pattern. In order to be modeled using ACD, durations need to be adjusted for this deterministic component in their evolution. I therefore adjusted the data for the diurnal pattern estimated as a spline function with knots every half-hour. Since Engle and Russell (1998) have reported little difference in their estimation results between a joint log-likelihood estimation of the ACD model and the diurnal spline, and a two-steps estimation procedure, I opted for the latter and simpler one. In the lower half of figure 6, I display the resulting adjusted durations. An important fact to notice is that adjusting for this diurnal pattern does not change the features of the dataset which call for the use of ACD modeling, as I have presented above and in figures 2 and 3. In figure 6, one can note some temporal clustering. Moreover, when looking at the descriptive statistics of both the plain and the adjusted data as reported in table 2, one can still see some overdispersion in the adjusted data, as well as a slow diminishing sample autocorrelation function, just like the plain dataset.

INSERT FIGURE 6 AROUND HERE

The models: results and discussion

I estimated 5 different models: an EACD(1,1); a WACD(1,1); a GACD(1,1); an EACD(2,2) and a WACD(2,2). Considering the time required for each estimation, and the fact that it increases dramatically with the number of parameters and observations, the estimation of a GACD(2,2) was not considered. In table 3, I present the parameters for the different models as well as a few statistics to check the models' adequacy. The unconditional expected value of the adjusted duration $E(x_i)$ is in all cases close to the

² The nonparametric estimate was obtained using the `supsmu` procedure in *R*, which is the implementation of the super smoother of Jerome Friedman. A similar estimate was presented by Engle and Russell (1998).

sample mean reported in table 2. Following Engle and Russell, the ACD process assumes that a particular transformation of the data is i.i.d, and this assumption can be tested using a Ljung -Box test statistics

$$\hat{\varepsilon}_i = \frac{x_i}{f(\psi_i)} \sim \text{i.i.d.}$$

Note that the Ljung -Box statistics of no autocorrelation in a time series is not valid for any distribution function. However, it is reported in Pacurar (2006) among others to be asymptotically valid for the exponential family. Considering the size of the dataset, it is therefore reasonable to calculate it for the EACD models. It is reported for the other models because it is widely used in the ACD literature, and still can be used for comparison purpose. Another point of notice is that the assumption of ε_i being i.i.d implies that higher powers of ε_i are i.i.d as well. I therefore report the Ljung -Box statistic of the squares as well. All 5 models fail to pass the Ljung -Box test statistics at a 5% level. Note however the huge improvement of the score of the statistics as compared to the one of the adjusted durations time series, $Q(10) = 2066.998$. Failing to pass the test is an indication that there remains some autocorrelation, in other words some ACD effect, in the estimated residuals. Another striking feature that relates to that latter point is the fact that the sum of the coefficients is always extremely close to 1, which suggest as well a long memory effect. It is in part because of these two facts that the linear ACD model was supplemented by new specifications which are nested within the Fernandes and Grammig ACD family.

INSERT TABLE 2 AROUND HERE

I must admit that the results for the GACD models are quite unsatisfying. It is reported elsewhere in the literature that this type of density functions leads to significantly harder models to estimate (cf. Tsay 2005). One of the main disappointing features of these results is that with these values of κ and γ the numerical algorithms to approximate the conditional hazard yields nothing. The values are too small for the precision of normal machines. It is interesting to note that EACD and WACD models yields close results in terms of parameters, but as I report in figure 7 the conditional hazard functions are quite different. Note that the conditional hazard functions of the ACD models are flat (here estimated at $\psi_i = 2$), whereas for both WACD models they are increasing. One could note as well, for the GACD(1,1) model, that $\kappa\gamma > 1$ and $\gamma < 1$ which suggests that the conditional hazard function has an inverted U-shape, i.e. increasing for small values of t and then decreasing.

INSERT FIGURE 7 AROUND HERE

Conclusion

To conclude, I have provided an exposé on the rationale behind the use of an ACD model as well as its main characteristics and results. It should empower the reader with sufficient knowledge to choose an ACD model, between linear or more complex specification, to apply an ACD estimation procedure and to some extent discuss the adequacy of the results. What is lacking in the expository part of this paper is a full understandable exposé of the very recent literature on the evaluation of ACD models (Fernandes and Grammig 2005; Meitz and Teräsvirta 2006; Pacurar 2006; Meitz and Saikkonen 2007). Yet, it would require both more time to explain, and more elaborate mathematical skills than the ones needed in this paper.

In terms of contribution, the source code provided in this paper is a good start toward spreading the use of ACD models. It constitutes the basis, however incomplete, of a full R package that will hopefully be contributed to the community of researchers in the near future. Future developments include implementing the estimation procedures for general and special cases of the AACD specification. It includes also developing the test routines associated with the literature not exposed here, for as most statistical tests, the inner mechanisms are less important than an available test procedure. Since the estimation procedures are extremely slow in the R environment, I am considering developing the

product in C++, and providing an interface with *R*. As I said, this is an ongoing work, but the path that lays ahead is clear, only time is the constraining factor.

- BAUWENS, L. and P. GIOT (2000): The Logarithmic ACD-Model: An application to the Bid-Ask Quote Process of three NYSE Stocks. *Annales d'économie et de Statistiques* 60, 117-149.
- CARRASCO, M. and X. CHEN (2002): Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models. *Econometric Theory* 18, 17-39.
- COX, D. R. (1972): Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 34(2), 187-220.
- DIONNE, G., P. DUCHESNE and M. PACURAR (2005): Intraday Value at Risk using Tick-by-Tick Data with Application to the Toronto Stock Exchange. *Les Cahiers du CREF* 05(10).
- DUFOUR, A. and R. F. ENGLE (2000): "The ACD Model: Predictability of the Time Between Consecutive Trades," in *Discussion Papers in Finance*. ISMA Centre, 59.
- ENGLE, R. F. (2000): The Econometrics of Ultra-High-Frequency Data. *Econometrica* 68(1), 1-22.
- ENGLE, R. F. and J. R. RUSSELL (1998): Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica* 66(5), 1127-1162.
- FERNANDES, M. and J. GRAMMIG (2005): Nonparametric specification tests for conditional duration models. *Journal of Econometrics* 127(1), 35-68.
- FERNANDES, M. and J. GRAMMIG (2006): A family of autoregressive conditional duration models. *Journal of Econometrics* 130(1), 1-23.
- GHYSELS, E. and J. JASIAK (1998): GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model. *Studies in Nonlinear Dynamics and Econometrics* 2(4), 133-149.
- GRAMMIG, J. and K.-O. MAURER (2000): Non-monotonic hazard functions and the autoregressive conditional duration model. *Econometrics Journal* 3(1), 16.
- GRAMMIG, J. and M. WELLNER (2002): Modeling the interdependence of volatility and inter-transaction duration processes. *Journal of Econometrics* 106(2), 369-400.
- HAUGHTON, D. and J. HAUGHTON (Forthcoming): "Duration Models," in *Statistical Techniques for the Analysis of Living Standards Survey Data*.
- HENTSCHEL, L. (1995): All in the family: Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics* 39(1), 71-104.
- MEITZ, M. and P. SAIKKONEN (2007): "Ergodicity, mixing, and existence of moments of a class of Markov models with applications to GARCH and ACD models," in *SSE/EFI Working Paper Series in Economics and Finance*. 37.
- MEITZ, M. and T. TERÄSVIRTA (2006): Evaluating Models of Autoregressive Conditional Duration. *Journal of Business & Economic Statistics* 24(1), 20.
- PACURAR, M. (2006): "Autoregressive Conditional Duration (ACD) Models in Finance: A Survey of the Theoretical and Empirical Literature " in.: SSRN.
- R_DEVELOPMENT_CORE_TEAM (2007): "R: A Language and Environment for Statistical Computing," in. Vienna, Austria: R Foundation for Statistical Computing.
- RUSSELL, J. R. and R. F. ENGLE (2005): A Discrete-State Continuous-Time Model of Financial Transactions Prices and Times: The Autoregressive Conditional Multinomial-Autoregressive Conditional Duration Model. *Journal of Business & Economic Statistics* 23(2), 166.
- TSAY, R. (2005): *Analysis of Financial Time Series*: Wiley-Interscience.
- ZHANG, M. Y., J. R. RUSSELL and R. S. TSAY (2001): A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics* 104, 179-207.

Appendix A

```
#  
# Definition of log-likelihood functions for the EACD(r,s), WACD(r,s) and GACD(r,s) models  
# Note that the parameter beta is set to 1.  
#
```

- *psi_model(x, r, s, omega0, Lambda, Omega)*

```
psi_model <- function(x, r, s, omega0, Al, Be, ...){  
  N <- length(x)  
  
  if (max(r,s)>r) {  
    Al = c(Al, rep.int(0, max(r,s)-r))  
    r = max(r,s)  
  }  
  psim <- filter(x, Al, method="convolution", sides=1)  
  
  if (r > 1) psim <- psim[-(1:(r-1))] # getting the NA values out  
  psim <- ts(c(0, psim)) # ajout du zero ~ vrai lag 1  
  psim[-1] <- psim[-1] + omega0 # careful with adding omega0 to the starting values  
  
  # Correcting the start values  
  if (s > 1) {  
    start.values = rep.int(0,s)  
    sv = 1  
    for (k in 2:s){  
      sv = c(sv, sum(Be[k:s]))  
    }  
    print(sv)  
    psim[1:s] = psim[1:s] + sv  
  } else {  
    if (Be > 0) start.values = 1/Be else start.values = 0  
  }  
  
  # Recursive filter  
  psim <- filter(psim[-length(psim)], Be, method="recursive", init=start.values)  
  psim <- psim[time(psim)]  
  psim[1] <- psim[1] + omega0 # now you can update the starting values  
  
  psim <- c(rep.int(1,r), psim[-1])  
  #stopifnot(length(psim)==N)  
  psim  
}
```

- *Inlikelihood_GACD(x, garch.order = r, arch.order = s, theta = parameters)*

```
#  
## Inlikelihood_GACD(x, garch.order = r, arch.order = s, theta = parameters)  
#  
Inlikelihood_GACD <- function(x, garch.order = r, arch.order = s, theta = parameters){  
  if((garch.order+arch.order+3)!=length(theta)){  
    stop("Wrong number of parameters!") # theta = (omega0, lambda1, ..., lambda_r, omega1, ..., omega_s, alpha, kappa)  
  }  
  
  # Constraints GACD
```

```

ui = array(0,dim = c(r+s+4,r+s+3))
ui[1:(r+s+3),1:(r+s+3)] = diag(1, ncol = r+s+3, nrow = r+s+3)
ui[(r+s+4),] = t(-c(0, rep.int(1,r+s),0,0))
ci = c(rep.int(0,r+s+3),-1)

if (any(ui %*% theta - ci <= 0)){
  LnL = NA
}
else{
  # Local variables
  r = garch.order
  s = arch.order
  N = length(x)
  omega0 = theta[1]
  Lambda = theta[2:(r+1)]
  Omega = theta[(r+2):(r+s+1)]
  alpha = theta[length(theta)-1]
  ka = theta[length(theta)]

  # Initialisation of psi = omega0 + sum(Omega*x[t-r,t]) + sum(Gamma*psi[length(psi)-s],length(psi))
  psi = psi_model(adj_dur,r,s,omega0,Lambda,Omega)
  gka = gamma(ka)
  glam = gka / gamma(ka + 1/alpha)

  LnL = log(alpha / gka) + (ka*alpha - 1) * log(x) - ka*alpha*log(glam * psi) - (x / (glam * psi))**alpha
}
LnL
}

• Inlikelihood_EACD(x, garch.order = r, arch.order = s, theta = parameters)
Inlikelihood_EACD <- function( x, garch.order = r, arch.order = s, theta = parameters){
  if((garch.order+arch.order+1)!=length(theta)){
    print(paste("length(theta): ",length(theta),sep=""))
    cat(theta)
    stop("Wrong number of parameters!") # theta = (omega0, lambda1,..., lambdar,omega1,...omegas)'
  }
  theta = c(theta,1,1)
  Inlikelihood_GACD(x, garch.order, arch.order, theta)
}

• Inlikelihood_WACD(x, garch.order = r, arch.order = s, theta = parameters)
Inlikelihood_WACD <- function(x, garch.order = r, arch.order = s, theta = parameters){
  if((garch.order+arch.order+2)!=length(theta)){
    stop("Wrong number of parameters!") # theta = (omega0, lambda1,..., lambdar,omega1,...omegas,alpha)'
  }
  # Constraints WACD
  ui = array(0,dim = c(r+s+2,r+s+2))
  ui[1:(r+s+1),1:(r+s+1)] = diag(1, ncol = r+s+1, nrow = r+s+1)
  ui[(r+s+2),] = t(-c(0, rep.int(1,r+s),0))
  ci = c(rep.int(0,r+s+1),-1)

  if (any(ui %*% theta - ci <= 0)){
    LnL = NA
  }
  else{
    # Local variables
    r = garch.order
    s = arch.order
    N = length(x)

```

```

omega0 = theta[1]
Lambda = theta[2:(r+1)]
Omega = theta[(r+2):(r+s+1)]
alpha = theta[length(theta)]

# Initialisation of psi = omega0 + sum(Omega*x[t-r,t]) + sum(Gamma*psi[length(psi)-s],length(psi)-1]
psi = psi_model(x,r,s,omega0,Lambda,Omega,psi)
gma = lgamma(1 + 1/alpha)

LnL = alpha*gma + log(alpha) - log(x) + alpha*log(x/psi) - (exp(gma)*x/psi)**alpha
LnL = LnL[-1:(max(r,s)+1)]
}
LnL
}

```

- ***Optimized function for EACD model***

```

fn_EACD <- function(par_){
as.vector(lnlikelihood_EACD(adj_dur, garch.order = r, arch.order = s, theta = par_))
}

```

- ***Optimized function for WACD model***

```

fn_WACD <- function(par_ = parameters){
as.vector(lnlikelihood_WACD(adj_dur, garch.order = r, arch.order = s, theta = par_))
}

```

- ***Optimized function for GACD model***

```

fn_GACD <- function(par_ = parameters){
as.vector(lnlikelihood_GACD(adj_dur, garch.order = r, arch.order = s, theta = par_))
}

```

Table 1 - Source Fernandes and Grammig (2006)

Typology of ACD models	
Augmented ACD	
$\psi_i^\lambda = \omega + \alpha\psi_{i-1}^\lambda[\varepsilon_{i-1} - b - c(\varepsilon_{i-1} - b)]^v + \beta\psi_{i-1}^\lambda$	
Asymmetric power ACD ($\lambda = v$)	
$\psi_i^\lambda = \omega + \alpha\psi_{i-1}^\lambda[\varepsilon_{i-1} - b - c(\varepsilon_{i-1} - b)]^\lambda + \beta\psi_{i-1}^\lambda$	
Asymmetric logarithmic ACD ($\lambda \rightarrow 0$ and $v = 1$)	
$\log(\psi_i) = \omega + \alpha[\varepsilon_{i-1} - b - c(\varepsilon_{i-1} - b)] + \beta \log(\psi_{i-1})$	
Asymmetric ACD ($\lambda = v = 1$)	
$\psi_i = \omega + \alpha\psi_{i-1}[\varepsilon_{i-1} - b - c(\varepsilon_{i-1} - b)] + \beta\psi_{i-1}$	
Power ACD ($\lambda = v$ and $b = c = 0$)	
$\psi_i^\lambda = \omega + \alpha x_{i-1}^\lambda + \beta\psi_{i-1}^\lambda$	
Box-Cox ACD ($\lambda \rightarrow 0$ and $b = c = 0$)	
$\log(\psi_i) = \omega + \alpha\varepsilon_{i-1}^v + \beta\log(\psi_{i-1})$	
Logarithmic ACD type I ($\lambda, v \rightarrow 0$ and $b = c = 0$)	
$\log(\psi_i) = \omega + \alpha\log(x_{i-1}) + \beta\log(\psi_{i-1})$	
Logarithmic ACD type II ($\lambda \rightarrow 0, v = 1$ and $b = c = 0$)	
$\log(\psi_i) = \omega + \alpha\varepsilon_{i-1} + \beta\log(\psi_{i-1})$	
Linear ACD ($\lambda = v = 1$ and $b = c = 0$)	
$\psi_i = \omega + \alpha x_{i-1} + \beta\psi_{i-1}$	

Table 2 - Descriptive Statistics

IBM Price Durations	Plain	Adjusted
Sample size	47611	47611
Mean	10.05	1.051
Maximum	268	20.830
Minimum	1	0.078
Overdispersion	1.223	1.162
nth order sample autocorrelation		
$n = 1$	0.121	0.095
$n = 2$	0.111	0.078
$n = 3$	0.097	0.069
$n = 4$	0.102	0.072
$n = 8$	0.083	0.051
$n = 12$	0.080	0.048
$n = 16$	0.063	0.031
$n = 20$	0.061	0.029
$n = 24$	0.066	0.034
$n = 28$	0.067	0.037
$n = 32$	0.064	0.036
$n = 36$	0.063	0.028

Table 3 - Estimated models

Parameters		EACD(1,1)	WACD(1,1)	GACD(1,1)	EACD(2,2)	WACD(2,2)
ω_0		0.0245	0.0244	0.0317	0.0293	0.0316
α_1		0.0438	0.0438	0.0387	0.0000	0.0000
α_2		-	-	-	0.0532	0.0596
β_1		0.933	0.933	0.9320	0.2151	0.416
β_2		-	-	-	0.7039	0.495
γ		1	1.02	0.0768	-	1.023
κ		-	-	158.5837	-	-
Ljung -Box	$Q(10)$	43.1458	43.2054	61.5624	28.0078	30.8341
on $\hat{\varepsilon}_i$	p-value	0.0000	0.0000	0.0000	0.0018	0.0006
Ljung -Box	$Q(10)$	6.6084	6.6125	14.0382	7.0673	13.3598
on $\hat{\varepsilon}_i^2$	p-value	0.7618	0.7614	0.1713	0.7191	0.2042
$\sum_i(\alpha_i + \beta_i)$		0.9768	0.9768	0.9708	0.9722	0.9702
$E(x_i)$		1.0552	1.0517	1.0852	1.0546	1.0593

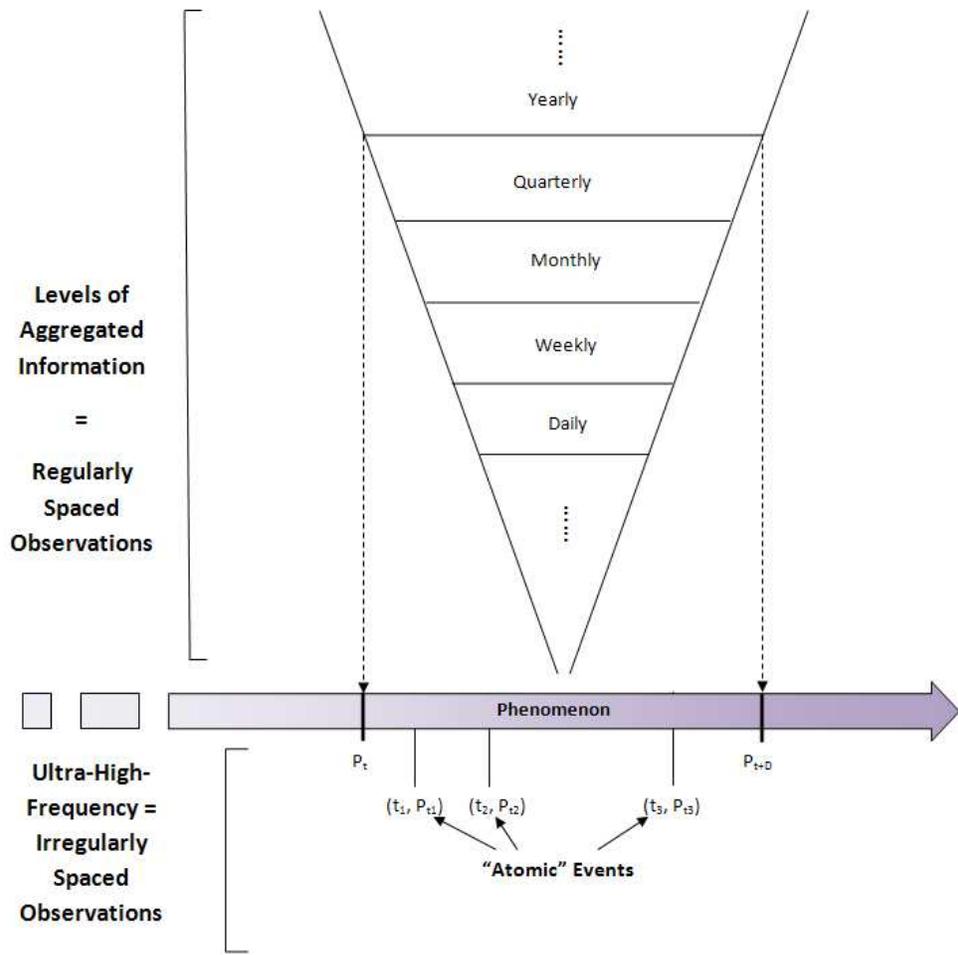


Figure 1 - The econometric paradigm and Ultra-high-frequency data

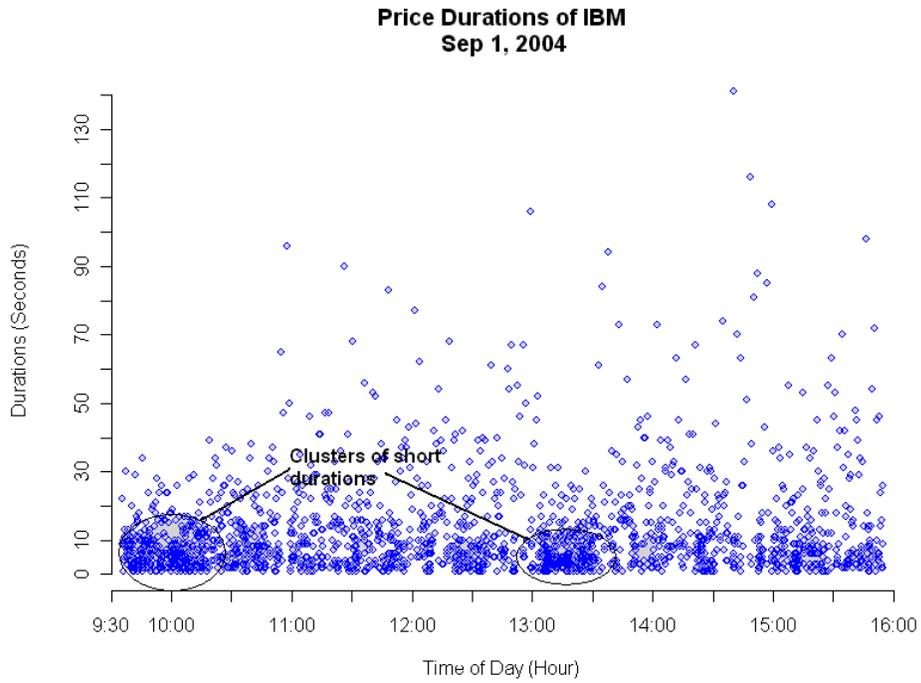


Figure 2 - Clustering of durations: an example
Prices Durations - IBM
September 2004

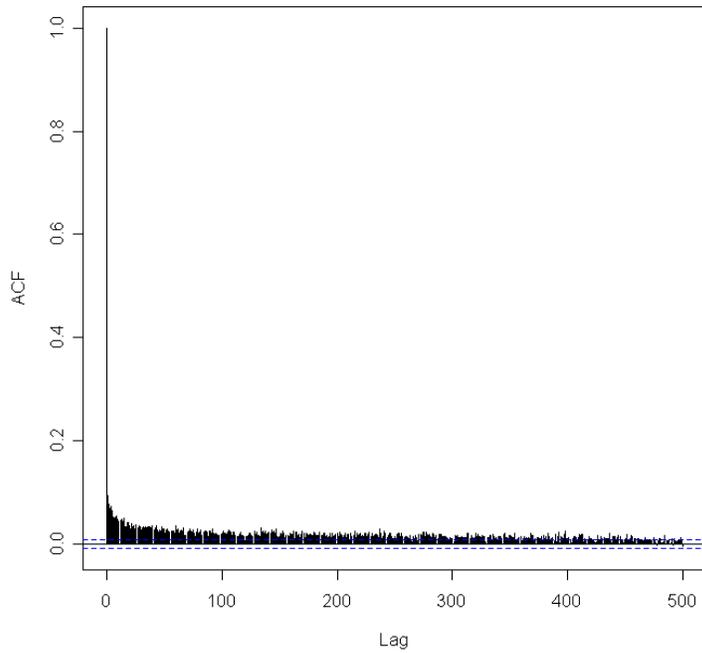


Figure 3 - Empirical Autocorrelation Function of the Prices duration of the IBM stock - September 2004

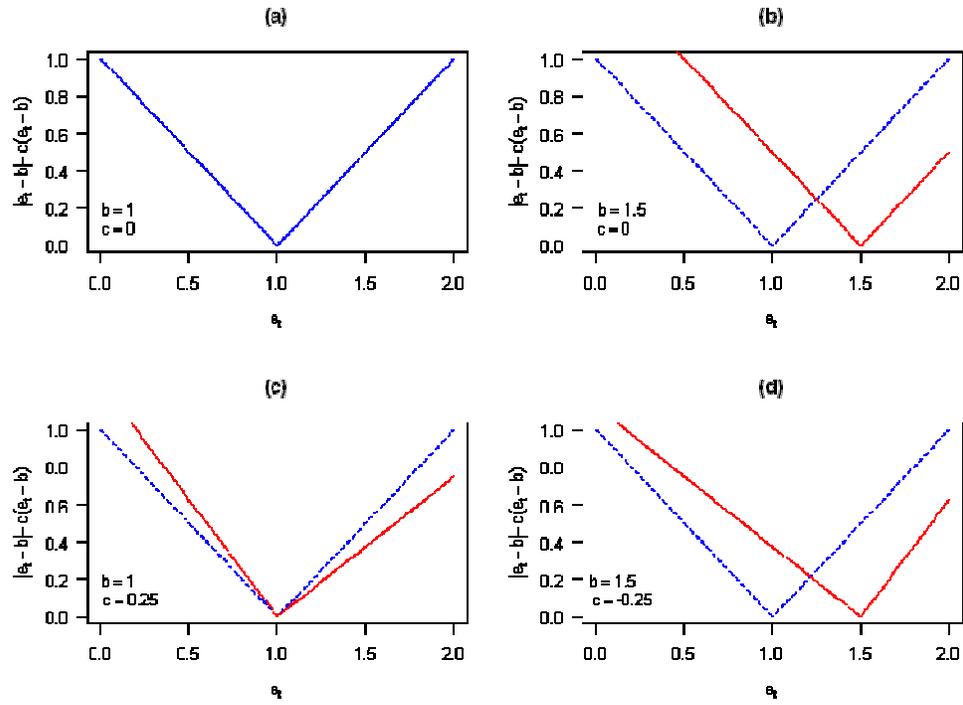


Figure 4 - Shocks Impact Curve - Effect of the shift and rotation parameters

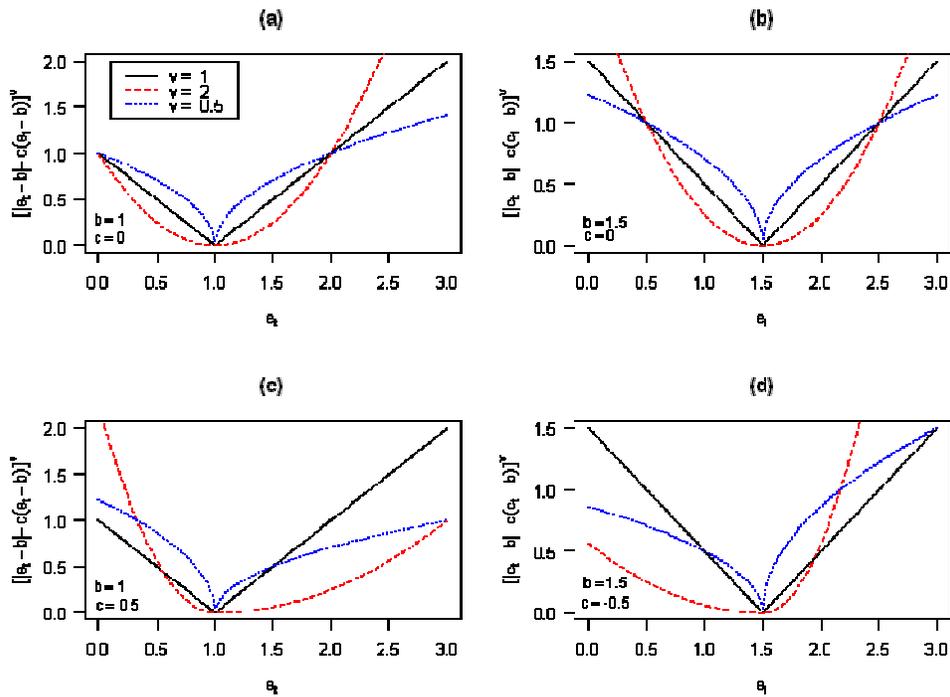


Figure 5 - Shocks Impact Curve - Combined Effects of the shift and rotation and shape parameters

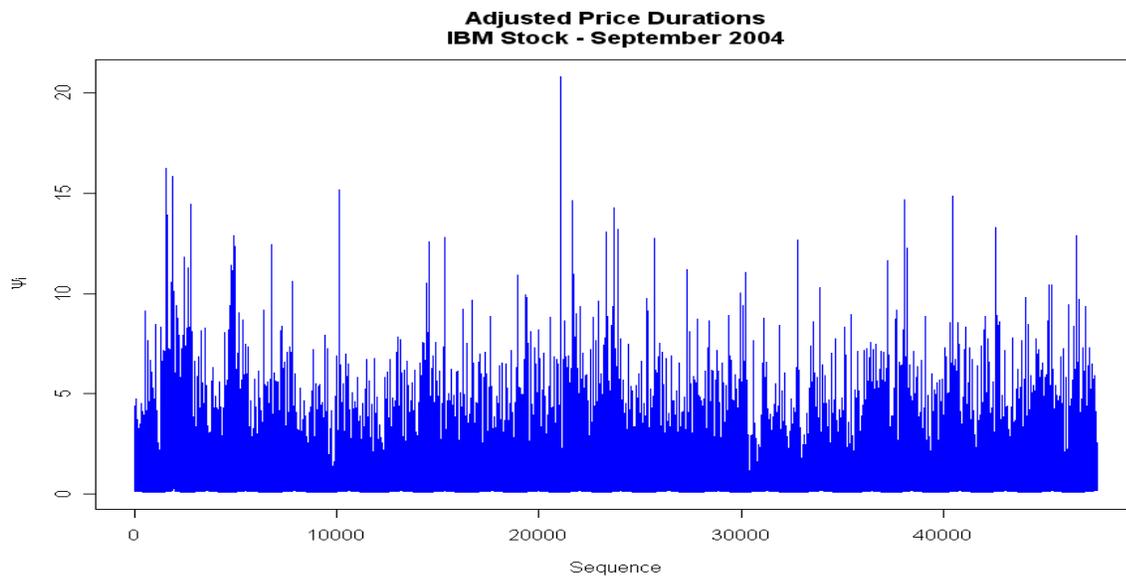
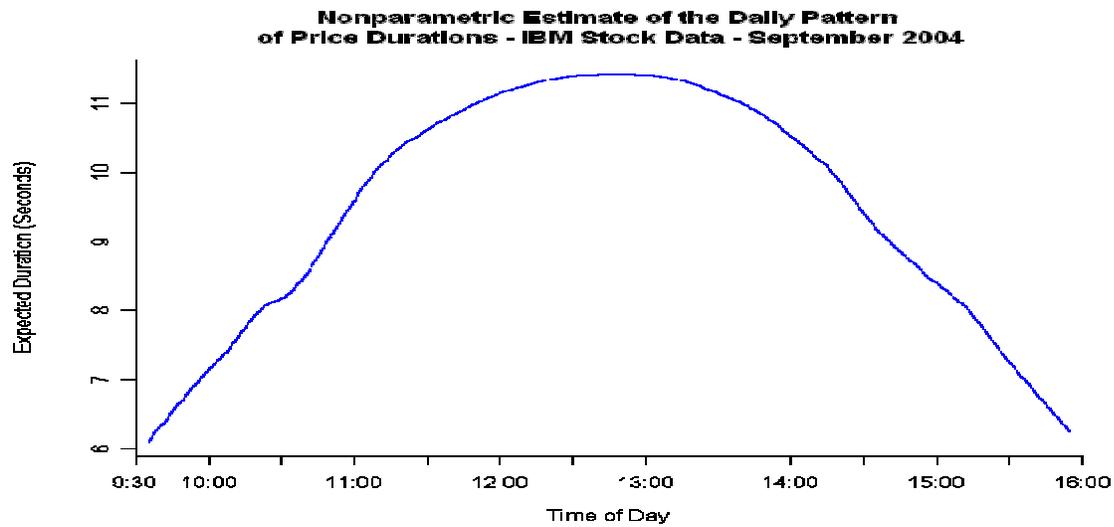


Figure 6 - Nonparametric Estimate of the daily pattern (upper figure) and IBM price durations adjusted for the daily pattern - September 2004

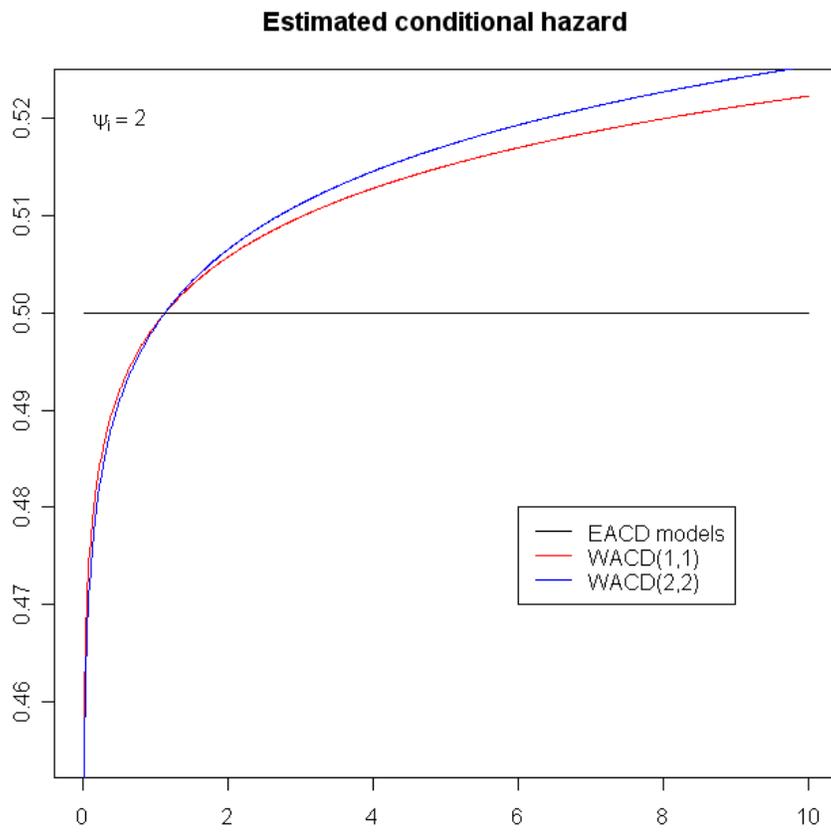


Figure 7 - Estimated Conditional Hazard functions for the EACD models and the WACD(1,1) and WACD(2,2) models