Persistence in Intertrade Durations

Joanna Jasiak*†

This version: March 14, 1999

(First draft: August 6, 1996)

Abstract

This paper examines long-term dependence in times between trades on financial markets. The autocorrelation functions of several intertrade duration series show a slow, hyperbolic rate of decay typical for long memory processes. For example, a shock to times between trades of the Alcatel stock on the Paris Stock Exchange (SBF Paris Bourse) may persist in the *transactions time* for a long period of 1000 or 2000 ticks. With an average duration of 52 seconds between transactions this may amount to sixteen or thirty two hours in calendar time. This paper introduces a fractionally integrated autoregressive conditional duration (FIACD) model for intertrade duration series. It also examines transformed duration processes representing times between consecutive returns to states of null, positive or negative returns. This approach captures the relationship between the duration persistence and return dynamics. The times elapsed between returns to various states feature very similar autocorrelation patterns and do not possess the long memory property. The persistence in durations is also determined by the times spent within specific states of returns. The average visiting time is state dependent, features intraday variation and may be considered as an instantaneous measure of state persistence. The long memory patterns are examined in data on the Alcatel and IBM stocks traded on the SBF Paris Bourse and NYSE.

*York University, *e-mail*: `jasiakj@yorku.ca`

†Research supported by the Natural Sciences and Engineering Research Council of Canada

The author thanks Christian Gourieroux for helpful comments

1 Introduction

The trade intensity on financial markets fluctuates within each day. It remains still unclear to what extent these movements are determined by past events rather than due to pure randomness. Potentially, any sudden acceleration or slowdown of trading activity may either have a temporary or permanent impact on future trades. This depends on the range of temporal dependence characterizing the trade intensity, or equivalently on the persistence in times between trades, called the intertrade durations. In high frequency duration data the notions of memory and shock persistence remain relatively little explored and require a particular interpretation. For example, a shock to times between trades of the Alcatel stock on the Paris Stock Exchange (SBF Paris Bourse) may persist in the *transactions time* for a long period of 1000 or 2000 ticks, i.e. have an impact on 1000 or 2000 future times to trade. This is due to a very slow, hyperbolic decay rate of the autocorrelation function of the duration series, typical for long memory processes. With an average duration of 52 seconds between trades, this amounts to sixteen or thirty two hours in the calendar time. Hence, in high frequency data the long range of temporal dependence may span several trading days and needs to be accounted for in models and forecasts of market activity.

In the financial literature, the long-term dependence and shock persistence have attracted a considerable interest in the context of return and volatility dynamics. This issue was first related to mean reversion in asset prices. Indeed, the *long memory* with range of temporal dependence amounting to several years has been reported in various studies based in general on daily and monthly data, and concerning stock returns [see, for example, Greene and Fielitz (1977), Aydogan and Booth (1988), Lo (1991), Jacobsen (1996)], exchange rates [see, for example, Booth, Kaen and Koros (1982) and Baillie and Bollerslev (1994)], and interest rates [see Shea (1991), Backus and Zin (1993), Crato and Rothman (1994)]. More recently, several papers drew attention to long memory in the volatility of financial assets sampled at daily or higher frequencies [see, for example, Taylor (1986), Ding, Granger and Engle (1993) and Dacorogna et. al. (1993)]. To accommodate high persistence in the conditional variances Baillie, Bollerslev and Mikkelsen proposed a fractionally integrated GARCH model [see, Baillie, Bollerslev and Mikkelsen, (1996)]. It was further investigated by Bollerslev and Mikkelsen (1996), McCurdy and Michaud (1996) and Comte and Renault (1996), this last work discussing the long memory of volatility in the context of continuous time processes [1].

The observed intertrade durations are determined by the speed of trading, and reveal information on the traders behaviour and their decision making. On the other hand, information on the empirical moments of durations, their dynamics and intraday variation may be considered as an

[1] Baillie (1996) provides an updated survey of literature on long memory processes.

important input for strategic trading. The research on predictable patterns in duration dynamics is quite recent. Early evidence on periodicities and duration clustering entailed by a short range temporal dependence is discussed in Engle and Russell (1996). The persistence range for various duration transforms is covered in Gourieroux, Jasiak (1998). The paper by Gourieroux, Jasiak and Lefol (1996) emphasizes the role of time to trade as a major market liquidity factor. This issue is essentially related to the tradeoff between fast trading, which implies a price change, and a slow sequential execution of the order to minimize the price impact of a transaction. Such problems often arise in block trading. The time also matters to a great extent in the limit order executions or allocations on markets characterized by different speeds of activity (trade intensity). In this context, despite a relatively long expected time to trade, some markets may be preferred for their lower variability of waiting times to trade, or otherwise less risk regarding the time of trade.

Empirically, the evidence for long range of time dependence in intertrade durations is revealed by a highly persistent pattern of the autocorrelations displaying a slow, hyperbolic rate of decay. As argued before, this feature needs to be accommodated in estimation and forecasting of market activity. The aforementioned recent work of Engle and Russell (1996) introduced a class of ARMA-type models called Autoregressive Conditional Duration (ACD) models for duration data. These models account for short serial dependence in expected durations and thus impose an exponential decline pattern on the autocorrelation function. In empirical applications of ACD models to high frequency intertrade durations the estimated coefficients on lagged variables sum up nearly to one. Such evidence indicates a potential misspecification that arises when an exponential decay pattern is fitted to a process showing an hyperbolic rate of decay. This would suggest that a more flexible structure allowing for longer term dependencies might improve the fit. This also is the motivation of the present paper for introducing a class of fractionally integrated ACD models (FIACD) to capture the long-term dependencies in the duration series.

The paper is organized as follows. In section 2 we introduce the FIACD model and discuss its properties. In section 3 the long memory fractional model is applied to the Alcatel and IBM data [2]. Further insights into the nature of long memory in the trade intensity are described in section 4 where durations are examined separately in different market regimes. Three basic market states of positive, null and negative returns are distinguished and series of durations are transformed into times between consecutive returns to these states. Besides the persistence in the return-to-state times, the average amount of time spent in each state arises as a complementary measure of persistence. Consequently, we study the visiting times spent by the trading process in the states of positive, null and negative returns. In particular we compare their intraday dynamics, and discuss information revealed by their varying means and variances. Section 5 concludes the paper.

[2] The series of IBM durations were also studied by Engle and Russell (1996).

2 Long Memory Autoregressive Conditional Duration Model

There exist several definitions of long-term dependence [3]. For example, we say that the process has long memory if the sum of absolute values of its autocorrelations is nonfinite. The short and long memory process can easily be distinguished in terms of the decay properties of the autocorrelation function. A stationary and invertible autoregressive moving average (ARMA) process has an exponentially decaying autocorrelation function (ACF) while fractionally integrated processes featuring long memory have ACF's decaying at a hyperbolic rate. Examples of such processes are ARFIMA models (Granger (1980) and Granger and Joyeux(1980)) for persistence in the conditional means, or FIGARCH (Baillie, Bollerslev, Mikkelsen (1996)) representing persistence in the conditional variances.

A slow decay rate characterizes the ACF of the two series of intertrade durations considered in this study which are: a) the IBM stock in November 1993, source: ISSM (Institute for the Study of Security Markets), and b) the Alcatel stock in July and August 1996, source: records of the Paris Stock Exchange (SBF Paris Bourse). The top panels of Figures 2.1 - 2.2 display the first 1000 autocorrelations of the data, corresponding to roughly 7 hours on the IBM market, and 15 hours on the Alcatel market. This means that the current intertrade duration is significantly related to the lagged durations up to and beyond the order 1000, although it spans a relatively short calendar time. However since the range of temporal dependence exceeds one day, this long memory property of high frequency duration data is relevant for both the intra- and interday trade dynamics. Hence, the approach requires large sets of high frequency data not necessarily covering long spans in calendar time.

Figure 2.1: Autocorrelations of Times Between Trades, IBM.

Figure 2.2: Autocorrelations of Times Between Trades, Alcatel.

2.1 The Fractionally Integrated ACD Model

This subsection extends the standard ACD model to include long range durations dependence. In the time series literature, the long memory adapted GARCH models are called fractionally integrated (FIGARCH) processes. Therefore by analogy this process is called FIACD, i.e. Fractionally Integrated ACD models.

Consider first the ACD model based on the exponential distribution. Let N be the number of events observed at random times. The N events are indexed by $i = 1, \dots N$ from the first observed event to the last. Let t_i be the time at which the i^{th} event occurs. Then $X_i = t_i - t_{i-1}$ is the time

[3] see for details section 2.2 in the survey by Baillie (1996).

between the $(i-1)^{\text{th}}$ event and the i^{th} event and X_i is called the i^{th} duration. The ACD(p, q) model proposed by Engle and Russell has a probability distribution function satisfying:

$$\varphi(X_i,\, \psi_i;\, \theta) = \frac{1}{\psi_i} f\left(\frac{X_i}{\psi_i}\right), \tag{2.1}$$

$$\psi_i = w + \alpha(L)X_i + \beta(L)\psi_i, \tag{2.2}$$

where: f is a given p.d.f. independent of i, with unitary mean, L denotes the lag operator, and θ is the parameter vector characterizing the polynomials $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \ldots + \alpha_q L^q$ and $\beta(L) = \beta_1 L + \beta_2 L^2 + \ldots + \beta_p L^p$. We denote by ψ_i the conditional expectation (and also the conditional standard deviation up to a scalar) of X_i given the past. The coefficients $\alpha_j,\ j = 1, ..., q,\ \beta_j,\ j = 1, ..., p$ are assumed to be nonnegative to ensure the positivity of ψ_i. This specification implies that the effect of past durations on the current conditional expected value decays exponentially with the lag length. Indeed, the ACD(p, q) process may be rewritten as an ARMA(m, p) process in X_i, where $m = \max(p,\, q)$, and:

$$[1 - \alpha(L) - \beta(L)]X_i = w + [1 - \beta(L)]v_i,$$

or equivalently,

$$[1 - \phi(L)]X_i = w + [1 - \beta(L)]v_i, \tag{2.3}$$

where $\phi(L) = \alpha(L) + \beta(L) = \phi_1 L + \phi_2 L + \ldots$, and $v_i = X_i - \psi_i$ is the linear innovation of the duration process. The stationarity and invertibility conditions require that the roots of $[1 - \alpha(L) - \beta(L)] = 1 - \phi(L)$ and $[1 - \beta(L)]$, respectively, lie outside the unit circle. Equivalently since α_j, β_j are nonnegative we need $\left(\sum_{j=1}^q \alpha_j + \sum_{j=1}^p \beta_j\right) = \sum_{j=1}^m \phi_j < 1$. The corresponding fractionally integrated process is obtained by introducing the fractional differencing operator:

$$[1 - \phi(L)](1 - L)^d X_i = w + [1 - \beta(L)]v_i. \tag{2.4}$$

The fractional differencing operator $(1 - L)^d$ is defined by its expansion which can be expressed in terms of the hypergeometric function H,

$$\begin{aligned}(1 - L)^d \ = \ H(-d,\, 1,\, 1;\, L) \ &= \ \sum_{k=0}^{\infty} \Gamma(k - d)\Gamma(k + 1)^{-1}\Gamma(-d)^{-1} L^k \\ &= \ \sum_{k=0}^{\infty} \pi_k L^k, \quad \text{say,}\end{aligned} \tag{2.5}$$

where Γ denotes the gamma function and $0 < d < 1$. For any $\delta > 0$ the gamma function is defined by:

$$\Gamma(\delta) = \int_0^\infty \exp(-x) x^{\delta - 1} dx,$$

In particular, for an integer δ, say $\delta = n$,

$$\Gamma(n) = (n-1)!.$$

Let us assume that all the roots of $[1 - \phi(L)]$ and $[1 - \beta(L)]$ lie outside the unit circle. By substituting $X_i - \psi_i$ for v_i in equation (2.4) we obtain the FIACD(p, d, q) model:

$$\begin{aligned} [1 - \beta(L)]\psi_i &= w + [1 - \beta(L) - [1 - \phi(L)](1 - L)^d]X_i && (2.6) \\ &= w + \lambda(L) X_i, && (2.7) \end{aligned}$$

where $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \ldots$. In order to guarantee the positive sign of expected durations all coefficients in the last equation have to be nonnegative, i.e. $\lambda_k \geq 0$ for $k = 1, 2 \ldots$ This representation is comparable to the FIGARCH(p, d, q) model of Baillie, Bollerslev and Mikkelsen (1996).

2.2 Stationarity, Ergodicity and Impulse Responses

This subsection covers the stationarity and ergodicity of the basic model as well as the impulse response functions. In this context, the results for FIACD can be compared to the properties and shock persistence of the Integrated ACD process (i.e. the FIACD with $d = 1$), called IACD, and the covariance stationary ACD processes. For $0 < d \leq 1$, the hypergeometric function evaluated at $L = 1$ equals 0, i.e. $H(-d, 1, 1; 1) = 0$. This means that the first unconditional moment of X_i is infinite and the FIACD process is not weakly stationary, a feature it shares with the class of integrated processes IACD. However, as shown by Bougerol and Picard (1992), the autoregressive models with nonnegative i.i.d. coefficients including the integrated processes are strictly stationary and ergodic. This result implies, for example, the stationarity and ergodicity of the integrated and fractionally integrated GARCH processes. By the same argument, the high order lag coefficients in the infinite autoregressive representation of any FIACD model may be shown to be dominated in an absolute value sense by the corresponding coefficients of the integrated IACD process. By an extension of the proof for the integrated ACD process presented in the Appendix 1, the FIACD(p, d, q) class of processes is strictly stationary and ergodic for $0 \leq d \leq 1$.

The persistence of shocks to the FIACD process can be studied in the context of the impulse response analysis. The impulse response function measures the time profile of the effect of a shock on the behaviour of the series [see, Koop, Pesaran and Potter (1996)]. The traditional impulse response function is designed to provide an answer to the question:"What is the effect of a shock of a given size hitting the system at time t on the state of the system at time $t+n$, given that no other shocks hit the system". In the linear framework, the impulse response functions satisfy several regularity conditions, like: a) the symmetry with respect to positive and negative shocks of the same size, b) the linearity, i.e. a shock of size 2 has exactly twice the effect of a shock of size 1, and c) the history independence property, which implies that the past does not affect the response. As an illustration consider a moving average representation of a second order stationary linear process : $X_i = c + A(L)v_i = c + a_0 v_i + a_1 v_{i-1} + a_2 v_{i-2} + ...$ An impact of a shock at time 0 evolves in time at a rate determined by the convergent sequence of moving average coefficients. Alternatively, if shocks are added to the system repeatedly at each point of time i, their impact becomes permanent and the impulse response weights become: $b_0 = a_0$, $b_1 = a_0 + a_1$, $b_2 = a_0 + a_1 + a_2$, ...The relation between the moving average coefficients and the cumulative impulse response weights can be written as: $(1-L)a_i = b_i$, implying $B(L) = (1-L)^{-1}A(L)$.

In the nonlinear framework, the impulse responses are much more complex, and in general violate the aformentioned conditions [4]. Various approaches to the impulse response analysis in nonlinear models have been proposed by Gallant, Rossi and Tauchen (1993), Koop, Pesaran and Potter (1996) and Gourieroux and Jasiak (1999). In the present paper we follow the approach developped in Gourieroux and Jasiak (1999). The impulse response is based on the Volterra decomposition [see, Gourieroux and Jasiak (1999), property 5], where:

$$Y_t = a_t(\epsilon_t, \epsilon_{t-1}, ..., \epsilon_1, \underline{\epsilon_0}), \tag{2.8}$$

and (ϵ_t) is a gaussian white noise with unitary variance, and $\underline{\epsilon_0}$ denotes the history of the process up to time 0. Since the distribution of ϵ_t is symmetric, the shocks δ and $-\delta$ have the same infinitesimal occurrence.

Following Gallant, Rossi and Tauchen (1993), the analysis is conditioned on the history of the process before the occurrence of shocks. Accordingly, for shocks hitting the process at date 1, past values of the process and the innovations are known, i.e. $\underline{\epsilon_0}$ is fixed. The analysis consists in finding, at date 0, the effect of a sequence of deterministic shocks $\delta_1, \delta_2, ..., \delta_t, ...$ occurring at future dates on the future profile of the process. These effects have to be measured with respect to a benchmark, i.e. the unperturbed path. The benchmark is random, since future innovations are unknown. We denote by: $\epsilon_1^s, \epsilon_2^s, ..., \epsilon_t^s$ a future path of innovations, where $\epsilon_1^s, \epsilon_2^s, ..., \epsilon_t^s$ are IIN (0,1) conditional on $\underline{\epsilon_0}$. The random benchmark is:

[4] The impulse response analysis for the GARCH model is discussed in Engle, Ito and Lin (1990)

$$Y_t^s(\underline{\epsilon_0}) = a_t(\epsilon_t^s, \epsilon_{t-1}^s, ..., \epsilon_1^s, \underline{\epsilon_0}), \tag{2.9}$$

whereas the profile subject to shocks is:

$$Y_t^s(\underline{\delta}, \underline{\epsilon_0}) = a_t(\epsilon_t^s + \delta_t, \epsilon_{t-1}^s + \delta_{t-1}, ..., \epsilon_1^s + \delta_1, \underline{\epsilon_0}), \tag{2.10}$$

where $\underline{\delta} = (\delta_1, ..., \delta_t, ...)$.

The entire effect of the sequence of shocks is summarized by the joint path distribution of:

$$[Y_t^s(\underline{\epsilon_0}), Y_t^s(\underline{\delta}, \underline{\epsilon_0}),\ t \geq 1].$$

Consider now nonlinear innovations $\epsilon_i = g^{-1}\left(\frac{X_i}{\psi_i}\right)$, where $g = \Phi^{-1} \cdot F$ and F is the c.d.f. of durations X_i. The ACD and FIACD models admit a representation $X_i = \psi_i g(\epsilon_i)$, where

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

for the ACD(1,1) models, and

$$\psi_i = \omega + \beta\psi_{i-1} + \lambda(L) g(\epsilon_i)\psi_i, \tag{2.11}$$

for the FIACD$(1, \mathrm{d}, 1)$[5]. Consequently, the impulse responses are defined by the joint path distribution of:

$$[X_i^s(\underline{\epsilon_0}), X_i^s(\underline{\delta}, \underline{\epsilon_0}),\ i \geq 1]. \tag{2.12}$$

For a stationary ACD(1,1) model with a negative Liapunov exponent, Gourieroux, Jasiak(1999) show that effects of temporary shocks vanish asymptotically in average and path- by- path. For nonstationary ACD processes, the shocks have explosive effects. The FIACD models are expected to show a behaviour similar to that of a stationary ACD model.

2.3 Inference

Before discussing statistical inference, it is necessary to complete the definition of the basic FIACD model by explaining how f, i.e. the baseline duration p.d.f., is selected. Two approaches can be followed. In the first one, the p.d.f. may be chosen from a given parametric family, leading to a fully parametric model for the duration process. In this approach, parameters can be estimated by maximum likelihood. In the second approach the p.d.f. remains unspecified except for the constraint of unitary expectation, yielding a semi-parametric model, which requires appropriate estimation methods.

[5] The first terms of the $\lambda(L)$ polynomial are: $\lambda_1 = (-\beta - \pi_1 + \phi), \lambda_2 = (-\pi_2 + \phi\pi_1), \lambda_k = (-\pi_k + \phi\pi_{k-1})$, where π_k denote the terms of the expansions of $(1-L)^d$.

i) Parametric Model

It is often assumed that the baseline f is an exponential distribution. In this case, the joint log-likelihood function is:

$$L(\theta) = -\sum_{i=1}^{N} \log(\psi_i) - \sum_{i=1}^{N} \frac{X_i}{\psi_i},$$

where

$$\theta = (w,\, \beta,\, \alpha,\, d)'.$$

As pointed out by Engle and Russell, alternatively a Weibull distribution can be considered:

$$\varphi(X_i,\, \psi_i;\, \theta) = \gamma \left(\frac{\Gamma(1+1/\gamma)}{\psi_i}\right)^{\gamma} X_i^{\gamma-1} \exp\left\{-\left(\frac{X_i\Gamma(1+1/\gamma)}{\psi_i}\right)^{\gamma}\right\}, \quad \gamma \geq 0,$$

which encompasses the exponential for $\gamma = 0$. In such a case the log-likelihood function becomes:

$$L(\theta) = \sum_{i=1}^{N} \left\{\log\left(\frac{\gamma}{X_i}\right) + \gamma \log\left(\frac{\Gamma(1+1/\gamma)X_i}{\psi_i}\right) - \left(\frac{\Gamma(1+1/\gamma)X_i}{\psi_i}\right)^{\gamma}\right\},$$

and has to be maximized with respect to all parameters including γ.

In the empirical work, however, duration data usually feature departures from either exponential or Weibull distributions. For this reason it may be preferable to leave the baseline p.d.f. f unspecified.

ii) Semi-parametric models.

In the semi-parametric framework, we only specify the form of the conditional mean ψ_i. Various estimation methods of the parameter $\theta = (\omega, \beta, \alpha, d)'$ can be proposed, especially the pseudo-maximum likelihood method. The idea is to select a priori a p.d.f. f_0 and to maximize the associated misspecified log-likelihood function:

$$L(\theta;\, f_0) = -\sum_{i=1}^{N} \log \psi_i + \sum_{i=1}^{N} \log f_0\left(\frac{X_i}{\psi_i}\right).$$

It is known that this procedure provides consistent estimators for an appropriate choice of the misspecified p.d.f. f_0. In particular, we can specify f_0 as either the standard normal distribution, which leads to the ordinary least squares criterion:

$$\min_{\theta} \sum_{i=1}^{N} (X_i - \psi_i)^2,$$

[This method is also called Quasi-Maximum Likelihood, henceforth QML], or select for f_0 the exponential distribution [see, Gourieroux, Monfort, Trognon (1984)]. Then the objective criterion is:

$$-\sum_{i=1}^{N} \log \psi_i - \sum_{i=1}^{N} \frac{X_i}{\psi_i}.$$

This second approach seems more relevant in an application to duration data, where the exponential distribution fits better a duration distribution than the gaussian one. The asymptotic properties of QML estimators of the FIACD(p, d, q) model with $d \in (0,1)$ can be obtained by extending the proof of consistency and normality of the QML estimators developped by Lee and Hansen (1994) for the Integrated GARCH(1,1) process and a gaussian, misspecified p.d.f.. Under regularity conditions, the QML estimator of θ is consistent and has a normal limiting distribution.

$$T^{1/2}(\hat{\theta}_T - \theta) \Rightarrow N\left\{0,\ J^{-1}\, I\, J^{-1}]\right\},$$

where:

$$J = E_0\left[-\frac{\partial^2 \log l_t(\theta)}{\partial\theta\partial\theta'}\right]$$

and

$$I = E_0\left[\frac{\partial \log l_t(\theta)}{\partial\theta}\frac{\partial \log l_t(\theta)}{\partial\theta'}\right]$$

where E_0 indicates that the expectation is taken with respect to the true distribution.

The optimizations are in practice performed numerically, and require at each iteration the computation of the conditional expectation ψ_i evaluated at the corresponding θ value. Therefore these optimizations will not only provide the estimator $\hat{\theta}_T$ of the parameter, but also the approximated conditional expectations $\hat{\psi}_i$, (say). These predictions can be compared with the observed durations. More precisely we can compute the corrected durations, also called the generalized residuals, $\hat{u}_i = \frac{x_i}{\hat{\psi}_i}$, and then consider the form of their empirical distribution to obtain information on the underlying baseline f.

3 Application to intertrade durations for the IBM and Alcatel stocks.

In the empirical application, the FIACD model was fitted to two data sets of times between trades on the Alcatel and IBM stocks introduced in section 2.

These two stocks are actively traded five days per week for seven hours daily although the Alcatel data also include observations on orders collected before the market opening (the so-called pre-opening period). The times between the market closures and the next day openings as well as the weekend gaps were omitted. In the Alcatel data, the observations corresponding to the market openings have been deleted since the matching procedures are different for the opening and for the intraday period. Moreover some trades may occur simultaneously, in particular when a large order is split. Such simultaneous trades have been aggregated. As a consequence of these two data adjustments, we retain in the Alcatel series strictly positive durations only. There is an advantage of considering such sample, which allows us to show that the long memory property is not due to the occurrence of runs of zero durations. In the IBM sample there are relatively few simultaneous trades recorded at the opening, while most of these trades are recorded during the trading day and correspond to zero returns. We removed from the sample only observations on simultaneously occurring zero durations and returns, what reduced the sample size by more then 50%. This suggests that simultaneously occurring trades represent a significant feature of this stock dynamics. Indeed, if all zero durations were removed, the sample size would decrease by 75% and the sample mean would double.

After these adjustments the two sets of data comprise 23704 observations on IBM and 20502 observations on Alcatel. The IBM stocks are traded on average every 19.32 sec. with a standard deviation 21.51, while Alcatel has a slower rate of trading activity with a mean duration 52.56 sec. and standard deviation 83.66.

All duration data under study display strong seasonal patterns. The periodic components can easily be detected in the intraday means. Table 3.1 shows the hourly averages of durations over the daily trading cycles defined in local times of New York and Paris.

Table 3.1: Hourly Means and Std. Dev. of Durations

	IBM		Alcatel	
Time	*mean*	*st.dev.*	*mean*	*st.dev.*
09 – 10	15.6699	23.7194	–	–
10 – 11	16.7597	18.4828	42.4058	54.5180
11 – 12	18.3727	19.8592	48.6615	61.1579
12 – 13	21.7982	21.9646	82.5951	111.4793
13 – 14	24.9141	27.7905	129.7182	170.3298
14 – 15	20.4415	20.4922	63.8132	94.4518
15 – 16	18.0713	18.8944	44.3188	68.7299
16 – 17	–	–	32.6070	53.1288

There is a high intraday variation in the Alcatel data due to a pronounced lunch time trough in trades. Prior to estimation the data were adjusted for periodic effects. The intraday periodic patterns were removed by computing the deterministic means conditioned on the time of the day and dividing each observation by this value [see Engle and Russell (1998)]. The seasonally adjusted sets of data comprise: a) 23704 observations on IBM durations adjusted for intraday seasonality with mean 0.99 and standard deviation 1.11 ; b) 20502 observations on Alcatel with mean 0.99 and standard deviation 1.43 . Clearly the seasonal adjustment procedure is more successful in the IBM data compared to Alcatel, where we still observe a significant overdispersion.

The persistence in the intertrade times is reflected by the slow decay rate of autocorrelations. The top panels of Figures 2.1 - 2.2 display the autocorrelations for unadjusted data up to the lag 1000. More detailed data are provided in the first two columns of Table 3.2 where twenty first autocorrelations are reported, supplemented by the value of the Liung-Box statistics for the entire samples.

Table 3.2: Autocorrelation Functions

	Durations		FIACD	
Lags	*IBM*	*Alcatel*	*IBM*	*Alcatel*
1	.0193	.2410	3.6e-03	-.0090
2	.0777	.2185	7.7e-03	.0031
3	.0988	.1988	3.4e-03	-.0124
4	.0962	.2126	-3.3e-03	.0083
5	.0930	.1999	-1.2e-03	-.0097
6	.0897	.1794	-6.9e-03	-.0231
7	.0836	.1705	1.0e-03	-.0064
8	.0841	.1767	-1.3e-03	-.0168
9	.0804	.1712	-9.9e-03	-.0051
10	.0859	.1737	6.5e-03	-.0038
11	.0805	.1743	-7.9e-03	-.0112
12	.0840	.1646	-7.2e-03	-.0079
13	.0730	.1542	-2.0e-03	-.0048
14	.0812	.1528	3.3e-03	-.0109
15	.0748	.1574	-1.0e-03	.0004
16	.0767	.1580	-7.2e-03	.0060
17	.0751	.1492	9.8e-03	-.0033
18	.0831	.1575	-2.6e-03	.0164
19	.0756	.1458	-6.9e-03	.0010
20	.0644	.1301	-3.7e-03	.0024
Ljung-Box	5364.	12767.	23.35	40.36

An argument often put forward against the long memory in return volatility is that it is spuriously created by breaks in intraday trend. To show that the seasonal adjustment does not remove the long range dependence, we plot in the bottom panels of Figures 2.1 and 2.2 the autocorrelation functions of the adjusted data.

The estimation of the Fractionally Integrated ACD involves approximation of a polynomial of infinite order and requires conditioning of the model on the pre - sample values of durations. The sample is augmented by the unconditional sample mean. The truncation point chosen to approximate the infinite autoregressive polynomial $\lambda(L)$ in equation (2.8) was set equal to 1000.

This approach imparts an approximation bias and has an impact on the estimated parameter values [for a discussion see Baillie, Bollerslev, Mikkelsen, (1996)]. Given the large sizes of samples under study this effect should be attenuated despite the long range dependencies.

For comparison we also report estimates for the ACD(1,1) model. We do not discuss the ACD(2,2) or WACD(2,2) model based on the Weibull likelihood function, as all of these models produce sums of α and β parameters in (2.2) close to one (Engle, Russell (1998))

The estimated parameters of the ACD(1,1) model and of the FIACD(1, d, 1) and FIACD(1, d, 0) model are reported in the Table 3.3.

Table 3.3: Estimation of the FIACD model

	IBM		*Alcatel*	
Parameters	*Est.*	*S.D.*	*Est.*	*S.D.*
		$ACD(1,1)$		
ω	0.0028	0.0055	0.0037	0.0003
β	0.0128	0.0064	0.0689	0.0016
α	0.9843	0.0089	0.9294	0.0015
		$FIACD(1,d,1)$		
ω	0.0850	0.0100	0.0229	0.0031
β	0.5401	0.0244	0.5753	0.0261
ϕ	0.2591	0.0191	0.2938	0.0219
d	0.2770	0.0144	0.4304	0.0170
		$FIACD(1,d,0)$		
ω	0.3795	0.1947	1.0068	0.4426
β	0.0508	0.1124	0.3242	0.0375
ϕ	-	-	-	-
d	0.1273	0.0666	0.4624	0.0353

All fractional parameters of the FIACD(1, d, 1) model are significant, what reveals the persistence phenomenon and suggests misspecification of the ACD(1,1). As for the FIACD(1, d, 0) model, we find that the parameter β in the IBM sample is not significant, while further diagnostics show that this specification does not significanly reduce the serial correlation in the Alcatel sample and hence has to be rejected.

The adequacy of the FIACD(1, d, 1) model may be verified in various ways. We plot in Figure 3.1 the duration series along with the residuals $\hat{u}_i = \frac{X_i}{\psi_i}$, i varying.

Figure 3.1. Sample paths of durations and residuals

We observe that the dynamics of the Alcatel series is not homogenous over the entire sampling period. Clearly, there are more extreme durations occurring in the second part of the data, corresponding to the month of August. This variation is not perfectly accommodated by the model.

In both data sets, the corrected duration (residuals) feature considerably less autocorrelation than the durations themselves. Their autocorrelation function is given in Figure 3.2:

Figure 3.2: Autocorrelation of residuals

The values of the first twenty autocorrelations, along with the value of Liung Box statistics for the residuals appear in the last two columns of Table 3.2 above. We accept the null hypothesis that the residuals of the FIACD(1, d, 1) model fitted to the IBM and Alcatel data are white noise at significance levels of 5% and 10 % respectively.

Finally we give in Figure 3.3, a kernel estimator of the baseline density function f.

Figure 3.3: Estimated density functions.

A simple test can be performed to verify whether the underlying dustribution is Weibull. In this case, the density of the noise term v could be written:

$$f(v) = \gamma a^{-\gamma} v^{\gamma-1} \exp\left[-\left(\frac{v}{a}\right)^{\gamma}\right].$$

The test is based on the pattern of the survivor function:

$$S(v) = \int_v^{\infty} f(u)du = \exp\left[-\left(\frac{v}{a}\right)^{\gamma}\right],$$

which satisfies the relation:

$$\log[-\log S(v)] = \gamma \log v + \gamma log\frac{1}{a}.$$

The test consists on verifying whether this relationship is approximately satisfied by the estimated survivor functions, at values corresponding to the residuals of the FIACD(1, d, 1) model. The Figure 3.4 displays the estimated survivor functions for the IBM and Alcatel data. The Table 3.4 below presents the estimated coefficients of the regression of $\log[-\log \hat{S}(\hat{v_j})]$ on $\log \hat{v}_j$ for $j = 1, ...$.

Table 3.4: Estimation of the survivor function

	IBM		*Alcatel*	
Parameters	*Est.*	*S.D.*	*Est.*	*S.D.*
const	0.3993	0.0269	0.3085	0.0292
γ	0.6312	0.0107	0.7078	0.0131

The estimated coefficients are statistically significant. Both γ's take values different from 1, and as such do not support the hypothesis of the exponential distribution of residuals. The parameters a can be computed from the constants. They take values 0.53 and 0.64 for IBM and Alcatel respectively, while we would expect them to be 1 if the intensity parameter was indeed unitary. Finally a and γ substituted into the formula of the unconditional mean: $\mu = a\Gamma(1 + 1/\gamma)$ yield 0.81 and 0.75, so that again we find evidence against the exponential distribution with intensity 1. However we need to remember that the accuracy of these results strongly depends on the precision of the approximated empirical density function used to compute survivor function and the number of x-coordinates retained for the density estimation. The density was estimated at 500 points, what represents only 2% of datapoints and was performed using a standard S+ routine. For this reason we need to interpret the regression outcomes with caution.

Finally, we investigate the dissipation of shocks in both samples. The impulse responses are given by the joint path distributions of $\hat{X}_i(\delta) = g(\epsilon_i(\delta))\hat{\psi}_i(\delta)$ for $i = T+1, T+2, T+3, T+4$, i.e. the distributions of out of sample, simulated durations subject to shocks, conditioned on the history of the process. The shocks to the system occurr only at $i = T+1$, i.e. $\underline{\delta} = (\delta_{T+1}, 0, 0, 0)$. We consider a positive shock of size 1.0, a negative shock of size -1.0, as well as the unperturbed baseline profile. The values of $\hat{\psi}_i(\delta)$ are evaluated by solving recursively equation (2.11), and used in the next step to compute $\hat{X}_i(\delta)$. The results are based on 10 000 replications. The distributions of impulse responses are plotted in Figures 3.5a and 3.5b for IBM and Alcatel, respectively.

Figure 3.5a: Impulse responses, IBM

Figure 3.5b: Impulse responses, Alcatel

The shock effects can be assessed by comparing the perturbed paths to the unperturbed one. The mean deviations from baseline are $-0.788, +1.5167$; $+0.0032, -0.0062$; $-0.0205, +0.0396$; $-0.0360, +0.0692$ over four horizons in the IBM sample and $-0.2399, +0.4596$; $-0.0353, +0.0679$; $-0.0252, 0.0481$; $-0.0249, +0.0475$ over four horizons in the Alcatel sample. We find that mean deviations from baseline are not symmetric and decline very slowly with horizon.

The distributions of future duration paths can be compared using their empirical moments. The Table 3.5 below shows the means and variances of the impulse response distributions.

Table 3.5: Means and variances of impulse responses

	IBM		*Alcatel*	
Shocks	*mean*	*var*	*mean*	*var*
horizon 1				
negative	0.4438	0.4063	0.1366	0.0380
baseline	1.2320	1.5798	0.3765	0.1477
positive	2.7488	4.4049	0.8362	0.4109
horizon 2				
negative	1.2740	1.6390	0.7061	0.4982
baseline	1.2707	1.6303	0.7415	0.5527
positive	1.2645	1.6140	0.8094	0.6673
horizon 3				
negative	1.2394	1.4711	0.6763	0.4874
baseline	1.2599	1.5214	0.7015	0.5289
positive	1.2996	1.6217	0.7497	0.6127
horizon 4				
negative	1.2296	1.5096	0.6591	0.4644
baseline	1.2656	1.6051	0.6841	0.5039
positive	1.3349	1.7978	0.7316	0.5837

4 Duration Persistence in Varying Market Regimes

It may be interesting to study if the persistence found in intertrade durations also characterizes some other duration processes deduced from the trading data, and in particular obtained by endogenously aggregating a number of intertrade durations. As an illustration and a potential contribution to technical analysis, where the analysts are often concerned with sequences of upward and downward price jumps, we distinguish three basic regimes of the market, corresponding to positive, null and negative intertrade returns, respectively.

Under the simplifying assumption of independent price movements of equal size, this approach can be illustrated by a trinomial tree, where we distinguish transitions to the state of positive returns, identified as price upswings, transitions to the state of negative returns, i.e. price downswings, and transitions to the stable state of 0 returns. Moreover, the process is allowed to remain in each of these states, producing a series of durations of visiting times. The probability that the movement is an upswing conditional on the occurrence of a price movement is denoted by p, the probability that a movement is a downswing is $1-p$, while the probability of remaining in the state 0 is α. Hence if we consider the state 0 as the initial state, the process makes a transition to the up state u with probability $p(1-\alpha)$, it makes a transition to the down state d with probability $(1-p)(1-\alpha)$, and it remains in 0 with probability α.

In this simplified framework we can analytically derive the densities of durations. Consider first the durations of returns to the up states denoted D_{uu}. We have:

$$\begin{aligned} P[D_{uu} = n] &= P[\text{last move was up}]P[\text{before (n-1) moves were not}] \\ &= p(1-\alpha)[1-p(1-\alpha)]^{n-1} \sim \mathrm{G}(p(1-\alpha)), \quad n > 1, \end{aligned}$$

where G denotes the geometric distribution, the discrete time counterpart of the exponential distribution. For the return durations to falling prices, denoted D_{dd}, we find that:

$$P[D_{dd} = n] = (1-p)(1-\alpha)[1-(1-p)(1-\alpha)]^{n-1} \sim \mathrm{G}((1-p)(1-\alpha)).$$

Accordingly the durations of returns to 0, denoted D_{00} satisfy:

$$D_{00} \sim \mathrm{G}(\alpha),$$

while the durations of visiting times in 0 are:

$$\bar{D}_0 \sim \mathrm{G}(1-\alpha).$$

Similarly,the visiting times in the u state are:

$$\bar{D}_u \sim \mathrm{G}(1-p(1-\alpha)),$$

and the durations spent in the d state are:

$$\bar{D}_d \sim \mathrm{G}[1-(1-p)(1-\alpha)].$$

Hence, the expected value of a price movement at trade n after $n-1$ trades not involving any price changes is:

$$EY = \frac{1}{p}$$

We also can summarize the expected durations of returns to various states and the durations of the visiting times:

$$\begin{gathered} E(D_{uu}) = \frac{1}{p(1-\alpha)}, \quad E(D_{dd}) = \frac{1}{(1-p)(1-\alpha)}, \quad E(D_{00}) = \frac{1}{\alpha}, \\ E\bar{D}_u = \frac{1}{1-p(1-\alpha)}, \quad E\bar{D}_d = \frac{1}{1-(1-p)(1-\alpha)}, \quad E\bar{D}_0 = \frac{1}{1-\alpha}, \end{gathered}$$

The empirical results presented later in this section allow us to test the validity of this model. Implicitely we also verify the hypothesis of independence of price movements, based on the evidence from the duration behavior. In this section only the IBM data were examined.

4.1 Return-to-state Times

For each regime, we consider the return-to-state times, which measure durations between consecutive positive (null or negative) returns. Such return-to-state times are obtained from an aggregation of intertrade durations over a random number of trades depending on the price evolution.

At this point it might also be interesting to distinguish states determined by seller or buyer initiated trades [see Darolles, Gourieroux, LeFol (1998)], and therefore examine a six state regimes. This analysis would require however more detailed data.

Empirical results show that the average time to return to a positive or a negative state differ very little. The fact that traders expect to wait almost the same time to return to these states can be interpreted as a market efficiency condition, whenever the positive and negative price movements have about the same size and their occurrences are independent.

To examine this issue, let us consider a simple model representing only the transition durations between the up and down states, and disregarding the visiting times and the presence of the 0 return state. This framework requires some additional assumptions. Let us assume that at each trade, two possible price movements are possible and denote by u the upswing and by d the downswing. Under the assumption A.1:

A.1. The upward and downward movements are independent,

the probability of u is p, while the probability of d is $(1-p)$. We find that the price forecast h trades ahead is:

$$p_{t+h} = p_t + hd + (u-d)\mathcal{B}(h,p),$$

where $\mathcal{B}(.,.)$ denotes the binomial distribution.
By imposing an additional assumption:

A.2 The movements are of equal size : $d = -u$,

The prices follow a random walk if $p = \frac{1}{2}$, since $E_t p_{t+h} = p_t - hu + 2u\frac{h}{2} = p_t$. In fact we have 8378 returns to positive returns and 8304 returns to negative returns indicating almost equal probabilitites of a return to state and a move out of state. In this framework, the durations between two successive states are geometrically distributed with parameter p for upswings and $1-p$ for downswings. The means of return durations are equal only if $p = 1-p$, so that $p = \frac{1}{2}$. Hence the empirical findings in Table 4.1 below confirm the hypothesis of a random walk price process.

Table 4.1: Means and St.Dev. of IBM Return Times

	ret > 0	ret < 0	ret $= 0$
mean	54.1346	54.6929	65.2266
st.dev.	56.3897	56.8361	56.85664

We can also verify whether the data satisfy the more complex model presented at the beginning of this section. We have $\frac{1}{p(1-\alpha)} \approx 54$, $\frac{1}{(1-p)(1-\alpha)} \approx 54$, while $\frac{1}{\alpha} \approx 65$. Hence by substituting $p = \frac{1}{2}$ in the first formula we find that $(1-\alpha) = \frac{1}{108}$, what suggest that the simple model is too restrictive.

The series of return-to-state times feature as well a very similar pattern of autocorrelations displayed in Figure 4.1.

Figure 4.1: Autocorrelations of return times

We observe a short range of persistence of the return-to- state durations while the entire series features long memory. The autocorrelation patterns are amazingly similar indicating that shocks to times between consecutive price increases dissipate at the same rate as shocks to times between price decreases. The Figure 4.2 shows the density functions of return times. We note that the return times to negative returns quite often admit extreme values, as indicated by the long tail of the distribution.

Figure 4.2: Densities of return times

4.2 State Visiting Times

A natural nonparametric measure of persistence is the expected time to remain in a given state. Let us again isolate the states of positive, negative or null returns and examine for how long they are expected to last. The Table 4.2 shows that on average the time to remain in the negative return state is very close to the time spent in the positive return state, since the difference is negligible.

Table 4.2: Means and St.Dev. of Visiting Times, IBM

	ret > 0	ret < 0	ret $= 0$
mean	14.0679	13.9782	55.9958
st.dev.	21.3203	22.2584	46.5364

There is also a close accordance of the number of visits to the positive returns (7700) and negative (7669) states.

The Figure 4.3 below shows the amazing resemblance of the empirical densities of visiting times in the states of positive and negative returns.

Figure 4.3: Densities of Visiting Times.

Note however that there is some intraday variation in the expected state visiting times and their variances. Indeed, a trader may optimize in terms of maximizing the expected amount of time he makes money and keep trading IBM between 1:00 and 2:00 p.m. or eventually minimize the variance and hence chose to make transactions in the morning between 10 and 11 a.m..

Table 4.2: Hourly Means and Variances of Visiting Times, IBM

	ret < 0		ret > 0	
hour	mean	st.dev.	mean	st.dev
09–10	10.0077	21.8004	11.8036	26.1404
10–11	11.7762	19.5450	11.4661	17.2447
11–12	12.9509	19.7635	12.6744	18.9946
12–13	17.0056	23.0073	17.0085	22.8214
13–14	18.6678	30.4269	20.1929	26.6540
14–15	14.7436	21.7993	15.3199	21.5434
15–16	13.5446	19.9151	12.1284	17.3751

The temporal dependence of the return durations represents the memory of the inter-state transitions. On the other hand the memory of the visiting states corresponds to the persistence of states, themselves. There is potentially a relationship between these two types of persistence and their outcome which is the persistence of the entire series of unconditional durations.

5 Conclusions

This paper presented empirical evidence for the presence of long memory, and proposed a fractionally integrated model for high frequency duration data. Although the empirical part of this work involved only a specific series of intertrade durations, there also exist other potential applications, especially in finance, insurance and computer science.

The paper also investigated transformed duration series obtained by aggregation of times between returns to a given stock price-based state of the market and times spent within these states.

Almost identical autocorrelation patterns were found in the return-to-state durations which also feature quite long ranges of temporal dependencies but do not possess the long memory property.

This approach also revealed the expected time to revisit a given state and its variation. We found that the expected times to return to positive and negative states are almost equal.

A very close accordance was also found in values of the expected state visiting times. Given that the probabilities of visits in these states are also similar, there are potentially conclusions on regularities in times to trade on financial markets. In general, the analysis of durations provides insights on some, yet unexplored issues, like the efficiency of the speed of trading and its relationship to the arbitrage opportunities.

Appendix

Stationarity and Ergodicity Properties of ACD and FIACD Models

This Appendix presents the conditions for stationarity and ergodicity of ACD and FIACD models. As noted in section 2.2 the FIACD(p, d, q) process is strictly stationary and ergodic for $0 \leq d \leq 1$ if strict stationarity and ergodicity for an integrated ACD process, (henceforth IACD). can be demonstrated. The demonstration relies on the proof by Bougerol and Picard (1992) of strict stationarity for autoregressive models with nonnegative i.i.d. coefficients, including the integrated processes. This result is adapted to ACD processes with parameters lying outside the second-order stationarity region. By the theory of products of random matrices it is shown that the necessary and sufficient condition for stationarity is the negativity of an associated Lyapunov exponent (see Bougerol and Picard (1992)).

Let $\| \cdot \|$ denote any norm on $\mathcal{R}^m$. We define an operator norm on the set $\mathcal{M}(m)$ of $m \times m$ matrices by:

$$\| M \| = \sup \{ \| MX \| / \| X \|;\ X \,\epsilon\, \mathcal{R}^m,\ X \neq 0 \}$$

for any M in $\mathcal{M}(m)$. The top Lyapunov exponent associated to a sequence $\{A_n,\, n \,\epsilon\, \mathcal{Z}\}$ of i.i.d. random matrices

$$\gamma = \inf_{n \geq 0} \left\{ E \left(\frac{1}{n+1} \log \ \| A_0 \, A_{-1} \, \ldots \, A_{-n} \ \| \right) . \right\}$$

If the exponent γ exists and is finite, we know that $\gamma \leq E(\log \| A_0 \|)$ with equality when $m = 1$, and by the subadditive ergodic theorem [see Kingman (1973), Theorem 6] we have almost surely

$$\gamma = \lim_{n \to \infty} \frac{1}{n} \log \| A_0 \, A_{-1} \, \ldots \, A_{-n} \| \, . \tag{A.1}$$

Suppose that p, $q \geq 2$ for notational convenience, by adding some α_i or β_i equal to 0 if needed. For any $n \,\epsilon\, \mathcal{Z}$, let $\eta_n = \frac{X_n}{\psi_n}$ and

$$\begin{aligned} \tau_n &= (\beta_1 + \alpha_1 \, \eta_n,\, \beta_2,\, \ldots,\, \beta_{p-1}) \,\epsilon\, \mathcal{R}^{p-1} \\ \xi_n &= (\eta_n,\, 0,\, \ldots,\, 0) \,\epsilon\, \mathcal{R}^{p-1} \\ \alpha &= (\alpha_2,\, \ldots,\, \alpha_{q-1}) \,\epsilon\, \mathcal{R}^{p-2} \end{aligned}$$

We next define the $(p+q-1) \times (p+q-1)$ matrix A_n:

$$A_n = \begin{bmatrix} \tau_n & \beta_p & \alpha & \alpha_q \\ I_{p-1} & 0 & 0 & 0 \\ \xi_n & 0 & 0 & 0 \\ 0 & 0 & I_{q-2} & 0 \end{bmatrix} \qquad (A.2)$$

where I_{p-1} and I_{q-2} are the identity matrices of sizes $p-1$ and $q-2$, respectively. Since the random variables $\{\eta_n, n \,\epsilon\, \mathcal{Z}\}$ are independent with baseline distribution f, the random matrices $\{A_n, n \,\epsilon\, \mathcal{Z}\}$ are i.i.d. Since f, which can either be the exponential or Weibull distribution, has a finite variance, all the coefficients of these matrices are square integrable. This implies that $E(\max(\log \parallel A_0 \parallel, 0))$ is finite. Thus, the top Lyapunov exponent γ of the sequence $\{A_n, n \,\epsilon\, \mathcal{Z}\}$ is well defined.

Theorem A.1. When $w > 0$, the ACD eq (2.2) has a strictly stationary solution if and only if the top Lyapunov exponent γ associated with the matrices $\{A_n, n \,\epsilon\, \mathcal{Z}\}$ is strictly negative. Moreover, this stationary solution is ergodic. It is the only strictly stationary solution when the η_n's are given.

When $p = q = 1$ we have $\gamma = E(\log(\beta_1 + \alpha_1 \eta_n))$. By analogy to the Bougerol and Picard proof of strict stationarity of GARCH processes, we relate the ACD process to the following multivariate model: Let $\mathcal{M}^+(m)[\text{resp}(\mathcal{R}^+)^m]$ be the set of $m \times m$ matrices (resp m-dimensional vectors) with nonnegative coefficients.

Definition A generalized autoregressive equation with nonnegative i.i.d. coefficients is:

$$Y_{n+1} = A_{n+1} Y_n + B_{n+1}\,, \qquad n \,\epsilon\, \mathcal{Z}$$

where $\{(A_n, B_n), n \,\epsilon\, \mathcal{Z}\}$ is a given sequence of independent, identically distributed random variables with values in $(\mathcal{M}^+(m) \times (\mathcal{R}^+)^m$ and Y_n in $\mathcal{R}^m$.

We will now prove Theorem A.1 by showing that a necessary and sufficient condition for existence of a strictly stationary, nonnegative solution of this equation is the strict negativity of the top Lyapunov exponent associated with the sequence $\{A_n, n \,\epsilon\, \mathcal{Z}\}$.

Lemma A.1. Let $\{A_n, n \,\epsilon\, \mathcal{Z}\}$ be a sequence of i.i.d. random matrices such that $E(\max(\log \parallel A_0 \parallel, 0))$ is finite. If almost surely

$$\lim_{n \to \infty} \parallel A_0 A_{-1} \ldots A_{-n} \parallel = 0,$$

then the top Lyapunov exponent associated with this sequence is strictly negative.

Proof: See Bougerol and Picard (1990), Bougerol (1987).

We consider an ACD(p, q) process as defined in eq. (2.1 – 2.2). Let $\eta_n = \frac{X_n}{\psi_n}$. The distribution of η_n conditional on the history $\mathcal{F}_{n-1}$ is f. Therefore, η_n is independent of $\mathcal{F}_{n-1}$. Since η_n is $\mathcal{F}_n$–measurable, the random variables $\{\eta_n,\, n \,\epsilon\, \mathcal{Z}\}$ are independent with distribution f.

Let
$$B = (\delta,\, 0,\, \ldots\, 0)' \,\epsilon\, \mathcal{R}^{p+q-1}$$
and
$$Y_n = (\psi_{n+1},\, \ldots,\, \psi_{n-p+2},\, X_n,\, \ldots,\, X_{n-q+2}) \tag{A.4}$$

then X_n is a solution of (2.1 – 2.2) if and only if Y_n is a solution of

$$Y_{n+1} = A_{n+1} Y_n + B \qquad n \,\epsilon\, \mathcal{Z} \tag{A.5}$$

where the A_n matrices are defined in (A.2).

Proof of Theorem A.1. We suppose that there exists a strictly stationary solution $\{X_n,\, n \,\epsilon\, \mathcal{Z}\}$ of equation (2.2). Consider the process $\{Y_n,\, n \,\epsilon\, \mathcal{Z}\}$ defined by (A.4). For $n > 0$, using (A.5), we have

$$\begin{aligned} Y_0 &= A_0\, Y_{-1} + B \\ &= A_0\, A_{-1}\, Y_{-2} + B + A_0\, B \\ &= A_0\, A_{-1}\, \ldots\, A_{-n}\, Y_{-n-1} + B + \sum_{k=0}^{n-1} A_0\, \ldots\, A_{-k}\, B \end{aligned}$$

All the coefficients of A_n, Y_n and B are nonnegative. Thus for $n > 0$

$$\sum_{k=0}^{n-1} A_0\, \ldots\, A_{-k}\, B \le Y_0$$

This shows that the series $\sum_{k=0}^{n-1} A_0\, \ldots\, A_{-k}\, B$ converges a.s. Hence, $A_0\, \ldots\, A_{-n}\, B$ converges a.s. to zero when $n \to +\infty$. This result is proven by Bougerol and Picard [see Bougerol, Picard (1992)]. The next step is to use Lemma A.1 to conclude that the top Lyapunov exponent associated with matrices A_n is strictly negative.

We first suppose that the exponent γ is strictly negative. Then (A.1) implies that the series

$$\sum_{k=0}^{\infty} A_n\, A_{n-1}\, \ldots\, A_{n-k}\, B$$

converges a.s. for any n. We define a sequence $\{Y_n,\, n \,\epsilon\, \mathcal{Z}\}$ by

$$Y_n = B + \sum_{k=0}^{\infty} A_n\, A_{n-1}\, \ldots\, A_{n-k}\, B.$$

This sequence is a nonnegative solution of (A.5). Let $\psi_n = Y^1_{n-1}$ where Y^1_{n-1} is the first component of the vector Y_{n-1}. Then $X_n = \psi_n \eta_n$ is a solution of the ACD model in (2.2). The process of $\{(A_n, \eta_n), n \epsilon \mathcal{Z}\}$ is strictly stationary and ergodic, and we can write $X_n = G(\eta_n, A_n, A_{n-1}, A_{n-2} \ldots)$ for some measurable function G independent of n_0. Therefore, $\{X_n, n \epsilon \mathcal{Z}\}$ is a strictly stationary and ergodic process, solution of (2.1 – 2.2).

Assume that (Z_n) is another strictly stationary solution of (6). Hence, for $n > 0$

$$\begin{aligned} \| Y_0 - Z_0 \| &= \| A_0 \ldots A_{-n}(Y_{-n-1} - Z_{-n-1}) \| \\ &\leq \| A_0 \ldots A_{-n} \| \| Y_{-n-1} - Z_{-n-1} \| . \end{aligned}$$

Since $\| A_0 \ldots A_{-n} \|$ converges to 0 a.s. and the law of $(Y_{-n-1} - Z_{-n-1})$ is independent of n imply that $Y_0 - Z_0$ converges to 0 in probability. This means that $Y_0 = Z_0$. Hence, (2.2) has a unique solution, one the η_n's are given.

Let us prove now the existence and uniqueness of a strict stationary solution of integrated ACD.

Corollary A.1. Suppose that the support of f is not bounded, $f(\{0\}) = 0$ and that all the coefficients α_i and β_i are positive. Then if $\sum_{i=1}^{p} \beta_i + \sum_{j=1}^{q} \alpha_j = 1$, the ACD defined in (2.2) has a unique stationary solution.

Proof. Following Bougerol and Picard (1992) we run recursions on q and expand the determinant with respect to the last column. We can see that

$$\mathrm{Det}(z\, I_m - E(A_1)) = z^{p+q-1} \left(1 - \sum_{j=1}^{q} \alpha_j\, z^{-j} - \sum_{i=1}^{p} \beta_i\, z^{-i} \right)$$

From the inequality $| a - b | \geq |(| a | - | b |)|$ we have in the case where $| z | > 1$

$$| \mathrm{Det}(z\, I_m - E(A_1)| > 1 - \sum_{j=1}^{q} \alpha_i - \sum_{i=1}^{p} \beta_i \tag{A.6}$$

Since the right-hand side is zero and since $\mathrm{Det}(I_m - E\, A_1)) = 0$, this equation implies that the spectral radius ρ of the matrix $E(A_1)$ is 1. Also, almost surely, all the coefficients of the matrix $A_2\, A_1$ are positive and A_1 has no zero column or zero row. Since A_1 is not a.s. bounded, these properties imply by Kester and Spritzer (1984, Theorem 2) that the top Lyapunov exponent γ satisfies $\gamma < \log \rho$. Consequently, $\gamma < 0$ and the corollary follows from Theorem A.1.

Figure 2.1: Autocorrelations of Times between Trades, IBM

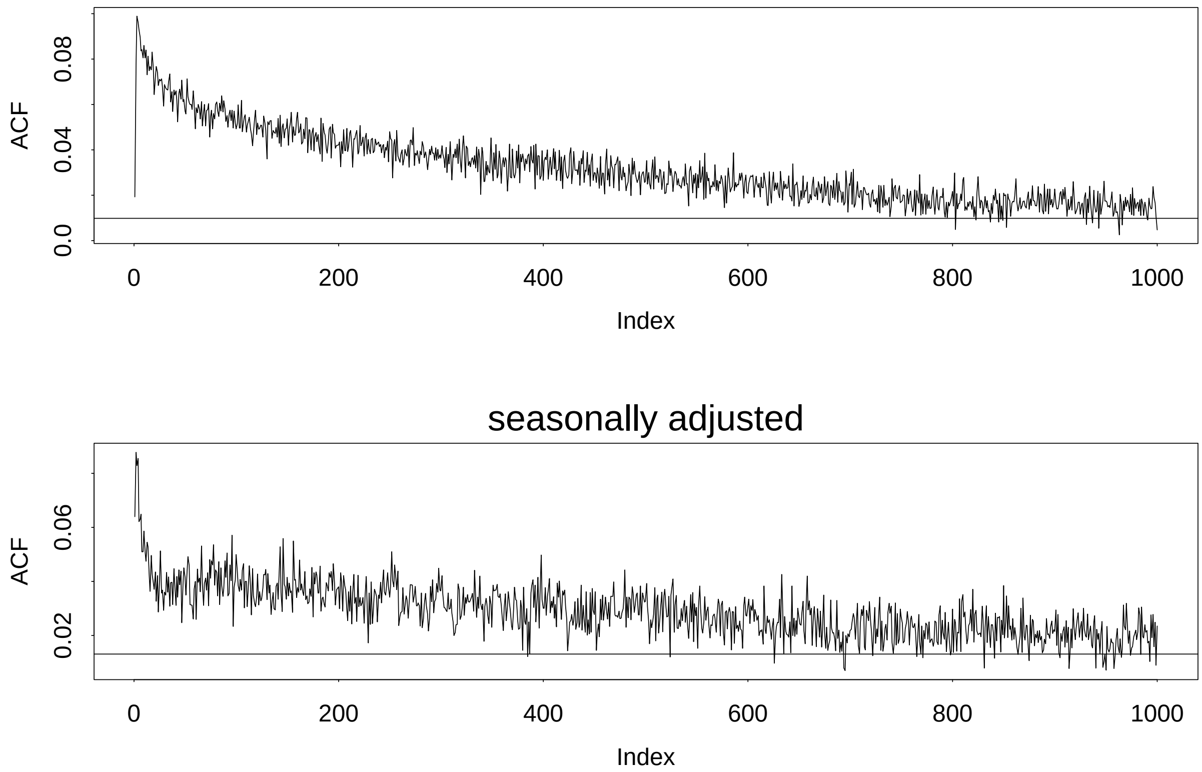


Figure 2.2: Autocorrelations of Times between Trades, ALCATEL

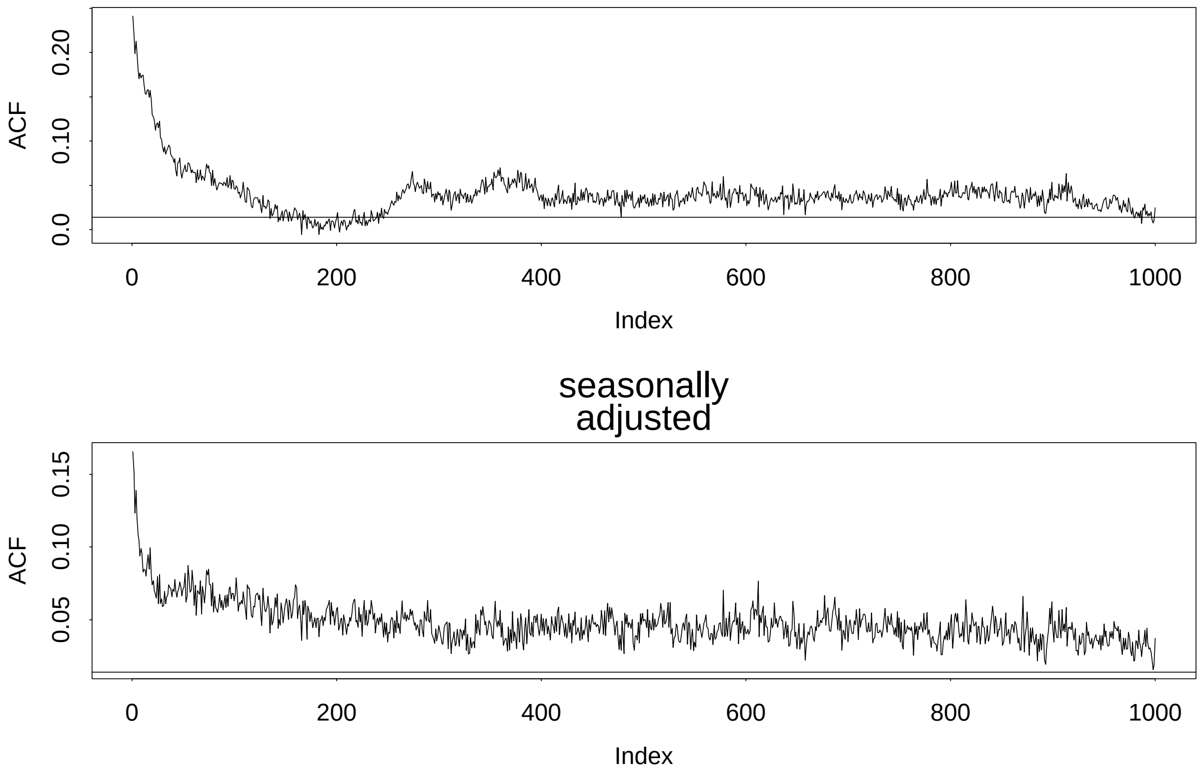


Figure 3.1: Sample paths of durations and residuals

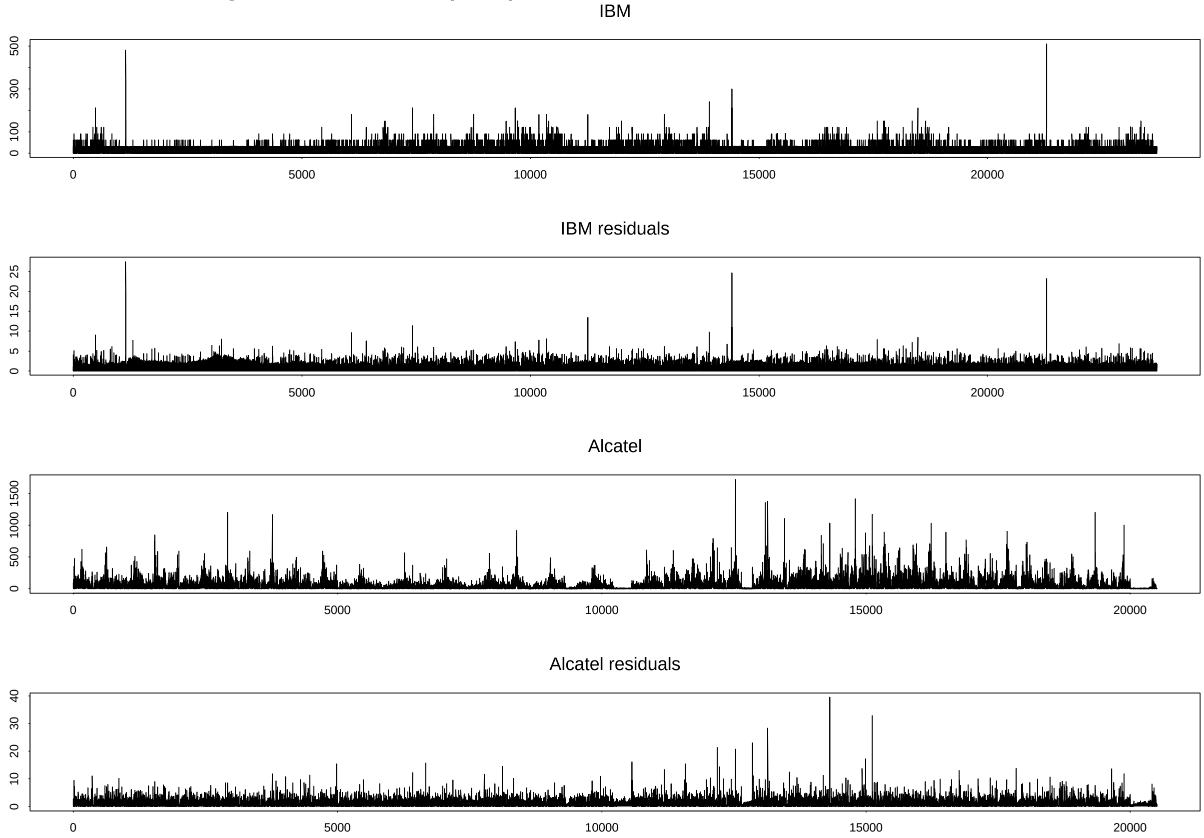


Figure 3.2: Autocorrelation of residuals

IBM

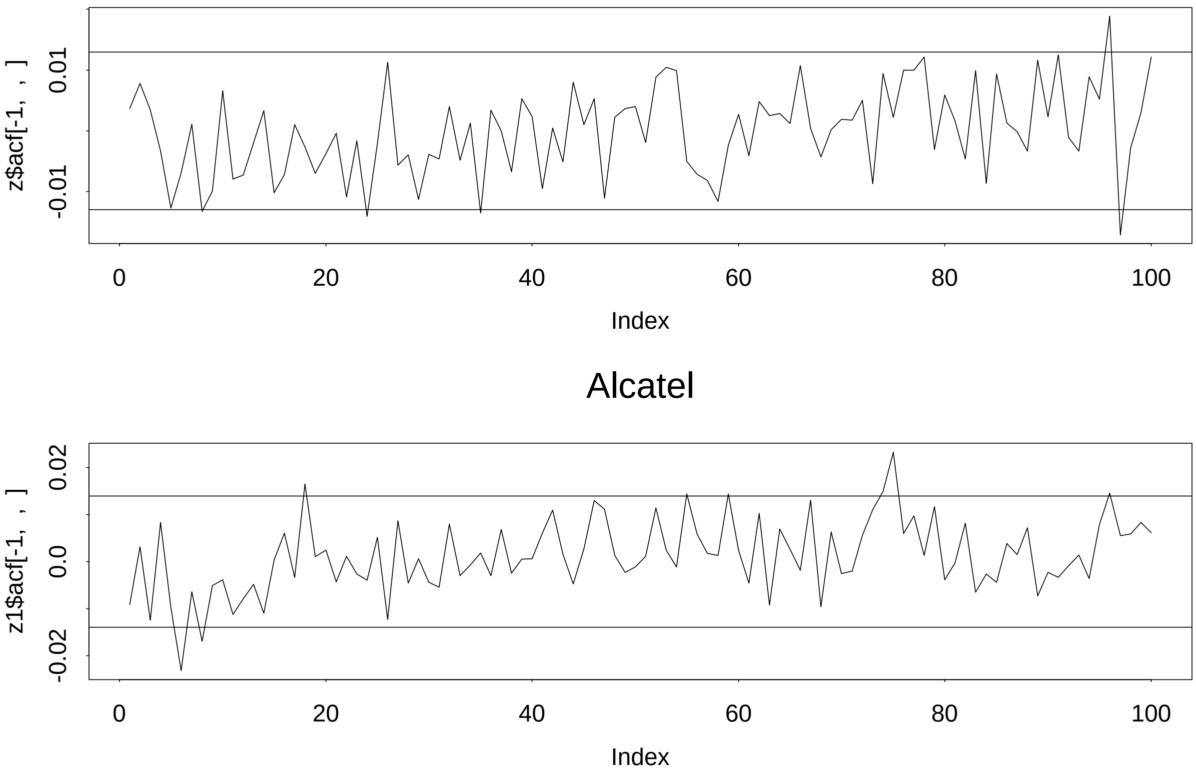


Figure 3.3: Estimated density functions

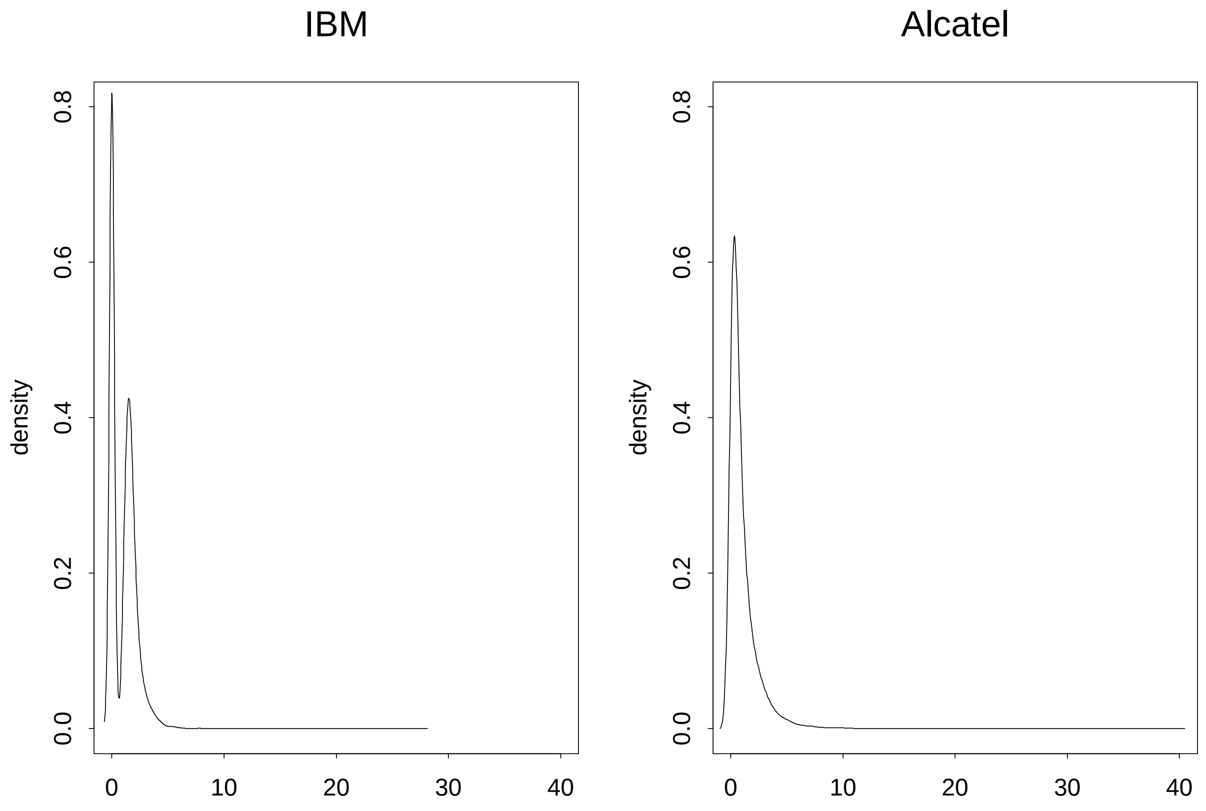


Figure 3.4: Estimated survivor functions

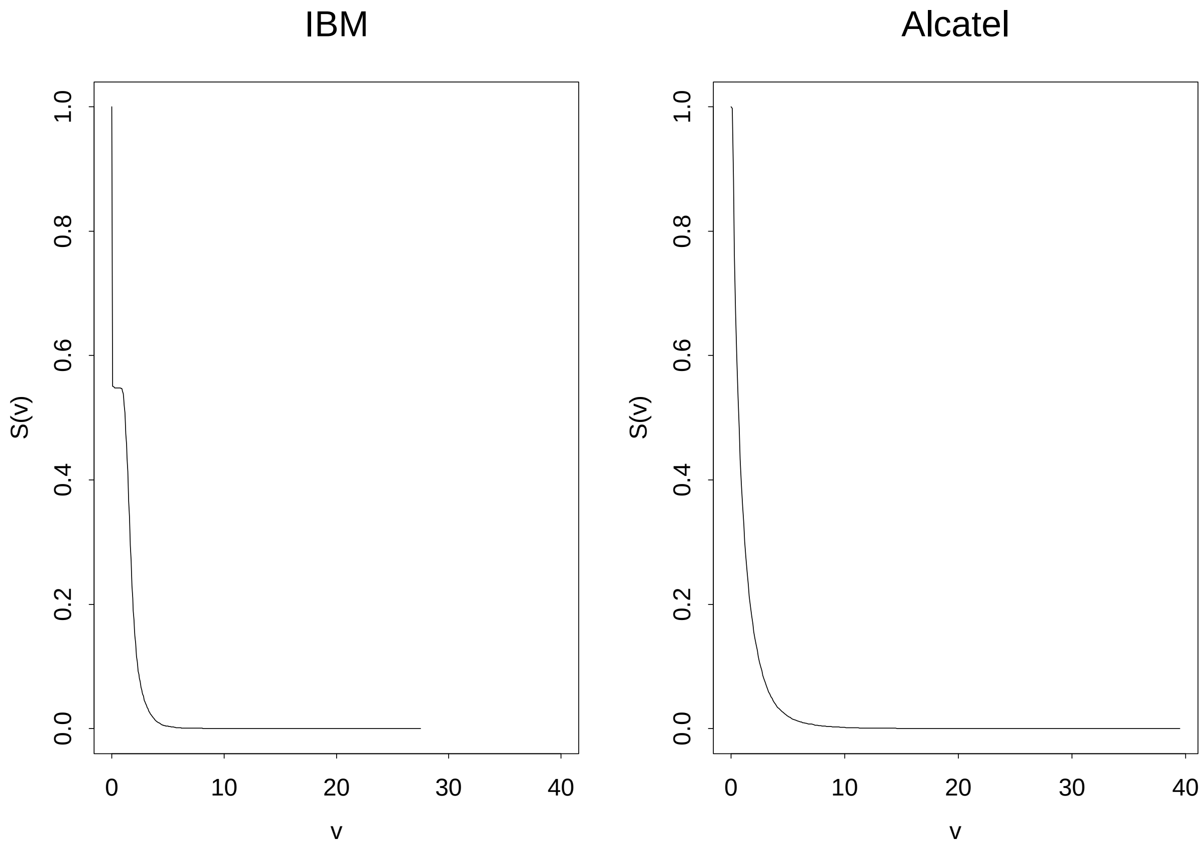


Figure 3.5a: Impulse responses, IBM
___baseline, - - - negative, ...positive

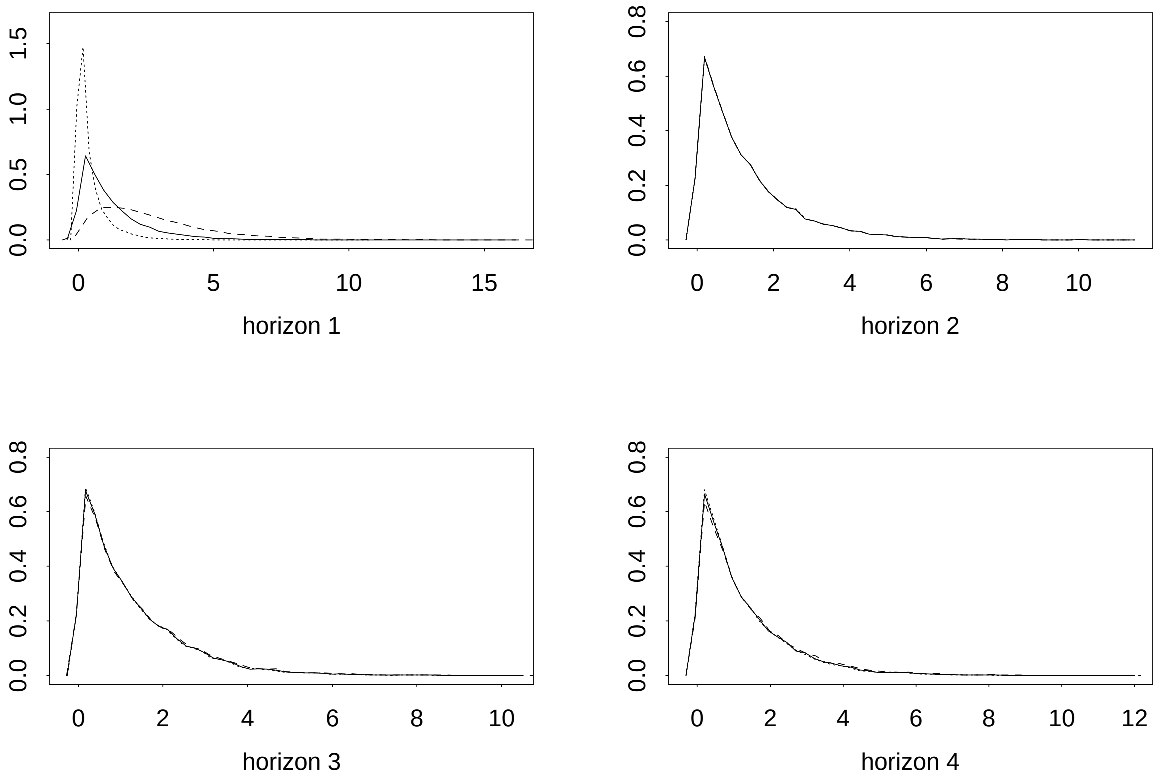


Figure 3.5b: Impulse responses, Alcatel

___baseline, - - - negative, ...positive

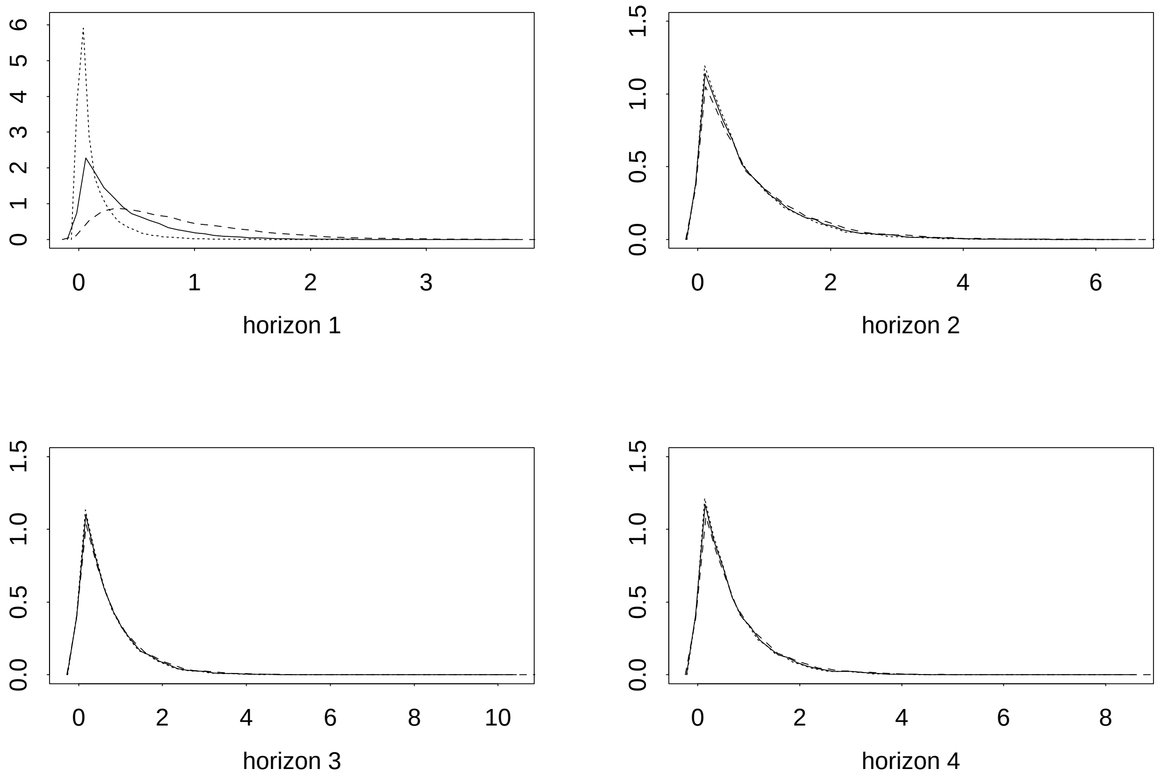


Figure 4.1: Autocorrelations of Return Times

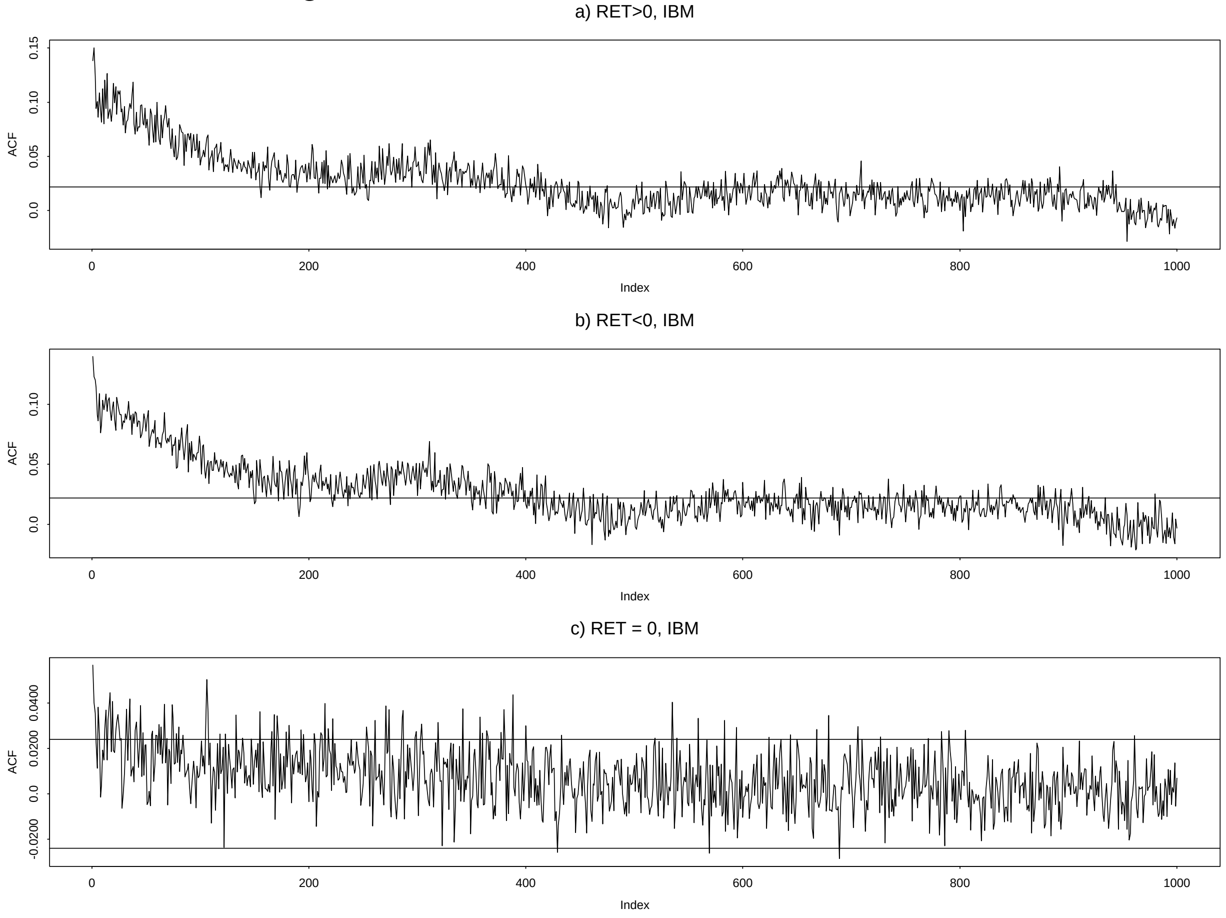


Figure 4.2: Densities of Return Times

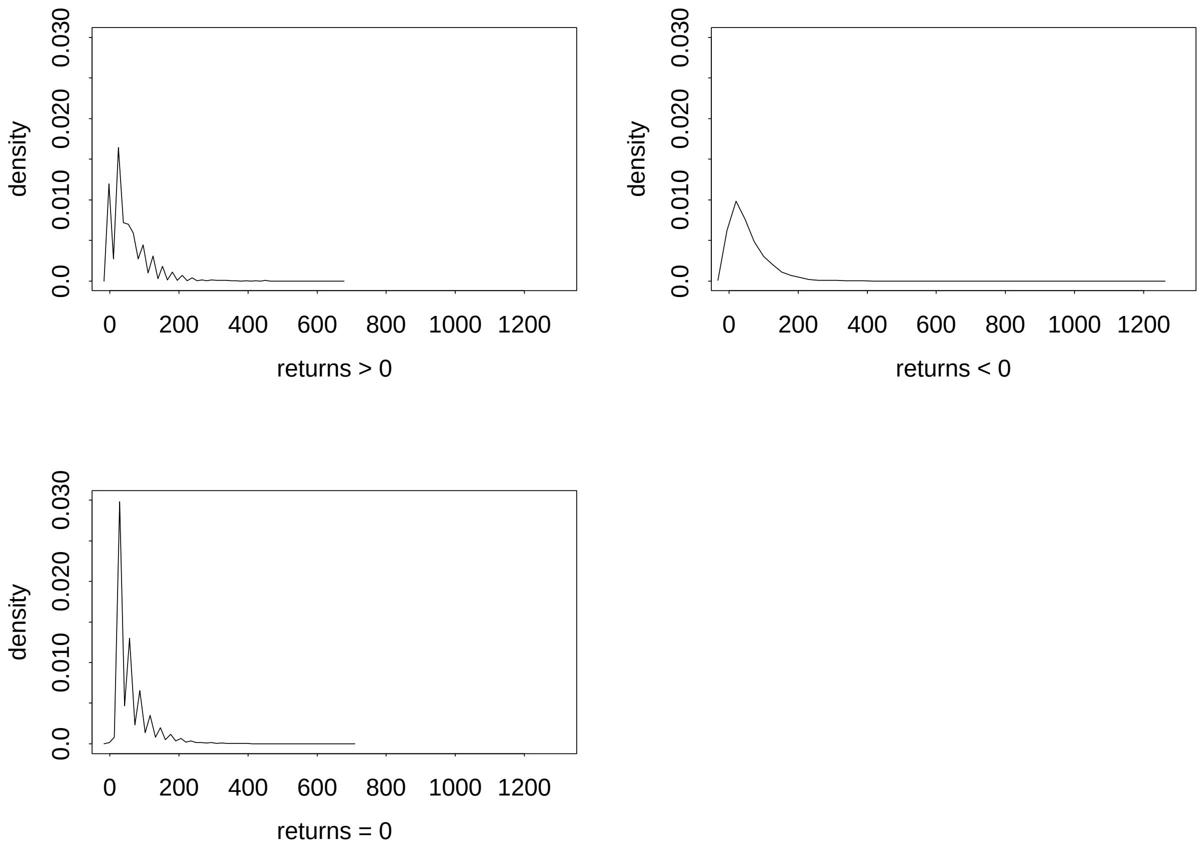
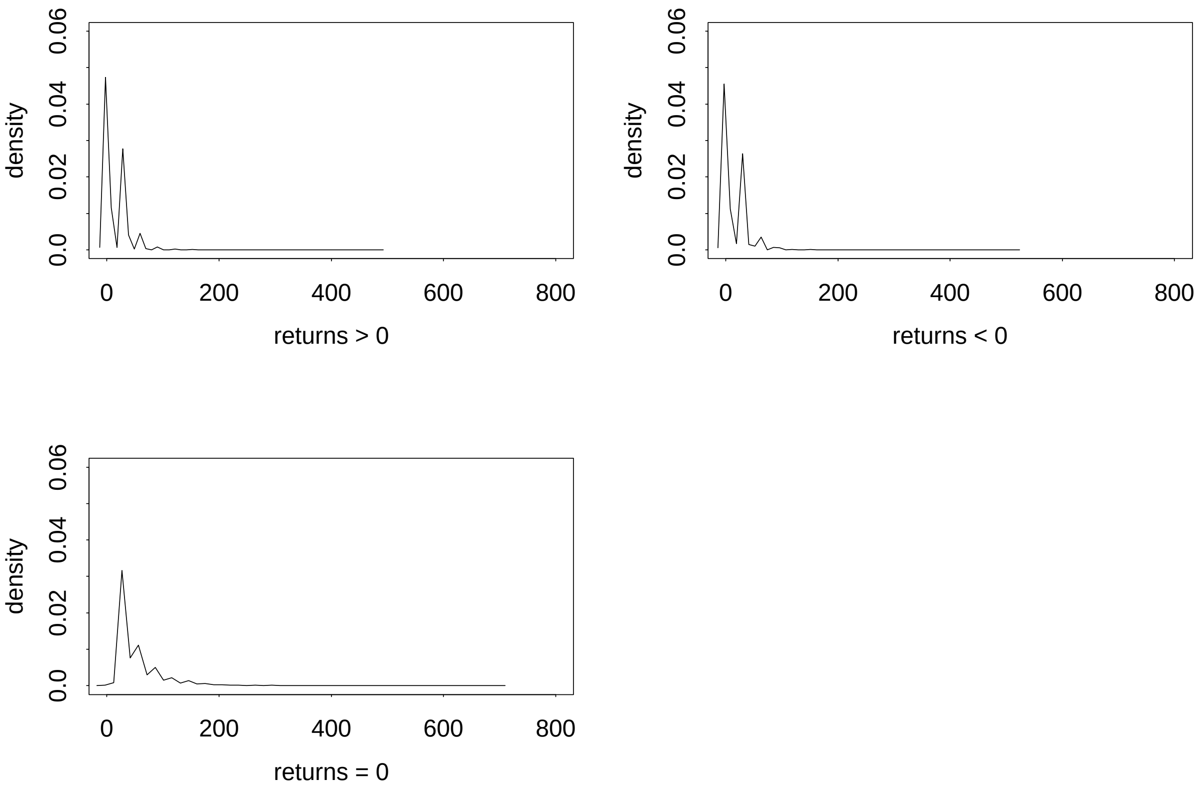


Figure 4.3: Densities of Visiting Times



References

[1] Akonom, J. and C. Gouriéroux (1988) "A Functional Limit Theorem for Fractional Processes" *CREST Discussion Paper no. 8801.*

[2] Aydogan, K. and G.G. Booth (1988) "Are There Long Cycles in Common Stock Returns?" *Southern Economic Journal*, **55**, 141 – 149.

[3] Backus, D.K. and S.E. Zin (1993) "Long Memory Inflation Uncertainty: Evidence from the Term Structure of Interest Rates," *Journal of Money, Credit and Banking*, **44**, 681 – 700.

[4] Baillie, R.T. (1996) "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, **73**, 5 – 59.

[5] Baillie, R.T. and T. Bollerslev (1991) "Intra Daily and Inter Market Volatility in Foreign Exchange Rates," *Review of Economic Studies*, **58**, 565 – 585.

[6] Baillie, R.T. Bollerslev, T. and H.O. Mikkelsen (1996) "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, **74**, 3 – 30.

[7] Beran, J. (1992) "Statistical Methods for Data with Long-Range Dependence," *Statistical Science*, **7**, 404 – 416.

[8] Bollerslev, T. and H. O. Mikkelsen (1996) "Modeling and Pricing Long Memory in Stock Market Volatility," *Journal of Econometrics*, **73**, 151 – 184.

[9] Booth, G.G., F.R. Kaen and P.E. Koveos (1987) "R/S Analysis of Foreign Exchange Rates under Two International Money Regimes," *Journal of Monetary Economics*, **10**, 407 – 415.

[10] Bougerol, P. (1987) "Tightness of Products of Random Matrices and Stability of Linear Stochastic Systems," *Annals of Probability*, **15**, 40 – 74.

[11] Bougerol, P. and N. Picard (1990) "Strict Stationarity of Generalized Autoregressive Processes," *Annals of Probability.*

[12] Bougerol, P. and N. Picard (1992) "Stationarity of GARCH Processes," *Journal of Econometrics*, **52**, 115 – 127.

[13] Cheung, Y.W. and K.S. Lai (1993) "A Fractional Cointegration Analysis of Purchasing Power Parity", *Journal of Business and Economic Statistics*, **11**, 102-112.

[14] Comte, F. and E. Renault (1996) "Long Memory Continuous Time Models," *Journal of Econometrics*, **73**, 101 – 150.

[15] Crato, N. and P. Rothman (1994) "Fractional Integration Analysis of Long-Run Behavior for US Macroeconomic Time Series," *Economic Letters*, **45**, 287 – 291.

[16] Dacorogna, M.M., U.A. Muller, R.J. Nagler, R.B. Olsen and O.V. Pictet (1993) "A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market," *Journal of International Money and Finance*, **12**, 413 – 438.

[17] Darolles, S., C. Gourieroux and G. LeFol (1998) "Intraday Transaction Price Dynamics", D.P. CREST.

[18] Ding, Z., C.W.J. Granger and R.F. Engle (1993) "A Long Memory Property of Stock Returns and a New Model," *Journal of Empirical Finance*, **1**, 83 – 106.

[19] Engle, R. and T. Bollerslev (1986) "Modelling the Persistence of Conditional Variances", *Econometric Reviews*, **5**, 1-50.

[20] Engle, R., Ito, T., and W.L. Lin (1990) "Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market," *Econometrica*, **58** , 525-542.

[21] Engle, R. and G.G.J. Lee (1993) "A Permanent and Transitory Component Model of Stock Return Volatility", manuscript, University of California, San Diego.

[22] Engle, R. and J.R. Russell (1998) "The Autoregressive Conditional Duration Model," *Econometrica*, **66**, 1127-1163.

[23] Galbraith, J.W. and V. Zinde-Walsh (1997) "Time-Domain Methods for the Estimation of Fractionally-Integrated Time Series Models", Working Paper, McGill University.

[24] Gallant,A.R., Rossi, P.E. , and G. Tauchen (1993) "Nonlinear Dynamic Structures", *Econometrica*, **61**, 871-907.

[25] Geweke, J. and S. Porter-Hudak (1983) "The Estimation and Application of Long Memory Time Series Models", *Journal of Time Series Analysis*, **4**, 221-238.

[26] Goncalves, E. (1987) "Une Généralisation des Processus ARMA", *Annales d'Economie et de Statistique*, **5**, 110-145.

[27] Goncalves, E. and C. Gouriéroux (1988) "Agrégation de Processus Autorégressifs d'ordre 1", *Annales d'Economie et de Statistique*, **12**, 128-149.

[28] Gouriéroux, C. and J. Jasiak (1998) "Nonlinear Autocorrelograms; an Application to Intertrade Durations", D.P. CREST 9841.

[29] Gouriéroux, C. and J. Jasiak (1999) "Nonlinear Innovations and Impulse Responses",manuscript, York University.

[30] Gouriéroux, C., J. Jasiak and G. LeFol (1996) "Intraday Market Activity", *Journal of Financial Markets*, forthcoming.

[31] Gouriéroux, C. and A. Monfort (1997), " Time Series and Dynamic Models", Cambridge University Press.

[32] Gouriéroux, C., A. Monfort and A. Trognon (1984), " Pseudo-Maximum Likelihood Methods: Theory", *Econometrica*, **52**, 681-700.

[33] Granger, C.W.J. (1980) "Long Memory Relationships and the Aggregation of Dynamic Models", *Journal of Econometrics* , **14**, 227-238.

[34] Granger, C.W.J. and R. Joyeux (1980). "An Introduction to Long Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis*, **1**, 15-39.

[35] Greene, M. and B. Fielitz (1977) "Long-Term Dependence in Common Stock Returns," *Journal of Financial Economics*, **4**, 339 – 349.

[36] Harvey, C.R. and R.D. Huang (1991) "Volatility in the Foreign Currency Futures Market," *Review of Financial Studies*, **4**, 543 – 569.

[37] Hurst, H.E. (1951) "Long-Term Storage Capacity of Reservoirs," *Proceedings of the Institute of Civil Engineers*, **116**, 770 – 799.

[38] Jacobsen, B. (1996) "Long Term Dependence in Stock Returns," *Journal of Empirical Finance*, **3**, 393 – 420.

[39] Kester, M. and F. Spitzer (1984) "Convergence in Distribution of Products of Random Matrices," *Zeitschrift für Wahrscheinlichkeits Theorie und Verwandte Gebiete*, **67**, 363 – 386.

[40] Kingman, J.F.C. (1973) "Subadditive Ergodic Theory," *Annals of Probability*, **1**, 883 – 909.

[41] Koop, G., Pesaran, M.H., and S. Potter (1996) "Impulse Response Analysis in Nonlinear Multivariate Models", *Journal of Econometrics*, **74**, 119-149.

[42] Lee, S.W. and B.E. Hansen (1994) "Asymptotic Theory for the GARCH(1,1) Quasi Maximum Likelihood Estimator", *Econometric Theory*, **10**, 29-52.

[43] Lin, Jin-Lung (1991) "Generalized Integrated Process and the Aggregation of Dynamic Time Series", *Academia Economic Papers*, **19**, 207-226.

[44] Lo, A.W. (1991) "Long Term Memory in Stock Market Prices," *Econometrica*, **59**, 1279 – 1313.

[45] McCurdy, T.H. and P.K. Michaud (1996) "Capturing Long Memory in the Volatility of Equity Returns: a Fractionally Integrated Asymmetric Power ARCH Model", *Working Paper University of Toronto.*

[46] Shea, G.S. (1991) "Uncertainty and Implied Variance Bounds in Long Memory Models of Interest Rate Term Structure," *Empirical Economics*, **16**, 287 – 312.

[47] Taylor, S. (1986) *Modeling financial time series*, (Wiley, Chichester).

[48] Wood, R., T. McInish and J.K. Ord (1985) "An Investigation of Transaction Data for NYSE Stocks," *Journal of Finance*, **40**, 723 – 739.