

MEDDELANDEN FRÅN
SVENSKA HANDELSHÖGSKOLAN
SWEDISH SCHOOL OF ECONOMICS
AND BUSINESS ADMINISTRATION
WORKING PAPERS

485

Bernard Ben Sita

AN EMPIRICAL STUDY OF THE MIXTURE
OF TIME AND MOVEMENTS IN PRICES

DECEMBER 2002

An Empirical Study of the Mixture of Time and Movements in Prices

Key words: Mixture, Persistence, Duration; GARCH

JEL Classification: G15

© Swedish School of Economics and Business Administration & Bernard Ben Sita

Bernard Ben Sita
Department of Finance and Statistics
Swedish School of Economics and Business Administration
P.O.Box 479
00101 Helsinki, Finland

Distributor:

Library
Swedish School of Economics and Business Administration
P.O.Box 479
00101 Helsinki
Finland

Phone: +358-9-431 33 376, +358-9-431 33 265
Fax: +358-9-431 33 425
E-mail: publ@shh.fi
<http://www.shh.fi/services/biblio/papers/index.htm>

SHS intressebyrå IB (Oy Casa Security Ab), Helsingfors 2002

ISBN 951-555-767-4
ISSN 0357-4598

An Empirical Study of the Mixture of Time and Movements in Prices

Ben Sita, Bernard*

31.07.2002

Abstract

This paper investigates the persistent pattern in the Helsinki Exchanges. The persistent pattern is analyzed using a time and a price approach. It is hypothesized that arrival times are related to movements in prices. Thus, the arrival times are defined as durations and formulated as an Autoregressive Conditional Duration (ACD) model as in Engle and Russell (1998). The prices are defined as price changes and formulated as a GARCH process including duration measures. The research question follows from market microstructure predictions about price intensities defined as time between price changes. The microstructure theory states that long transaction durations might be associated with both no news and bad news. Accordingly, short durations would be related to high volatility and long durations to low volatility. As a result, the spread will tend to be larger under intensive moments. The main findings of this study are 1) arrival times are positively autocorrelated and 2) long durations are associated with low volatility in the market.

JEL classification: G15

Keywords: Mixture, Persistence, Duration; GARCH

* Swedish School of Economics and Business Administration, Hietaniemenkatu 7 A, 5th floor, P.O. Box 479, FIN-00101 Helsinki, Finland, Tel. +358 (0) 9 4313 3380, E-mail: bernard.bensita@shh.fi. I express my deep gratitude to Professor Robert Engle who made available his algorithm for the estimation of the ACD model. I am grateful for the detailed comments of my supervisors, Professors Eva Liljeblom and Anders Löflund. The financial support from OKOBANK Group Research Foundation is gratefully acknowledged.

1. Introduction

An increasing number of studies¹ consider the dynamic of intra-trade duration within the Autoregressive Conditional Duration (henceforth, ACD) framework proposed by Engle and Russell (1998). The ACD model treats the fundamental questions of irregularly spaced arrivals in the context of high frequency financial data. The ACD model shares with GARCH and proportional hazard models² important features. While GARCH models tend to capture the lag variance effect of clustered quantities and proportional hazard models the intensity determined by a set of covariates, the ACD model measure and forecast the intensity of transactions arrivals conditionally.

Ordinarily, one will first transform inhomogeneous arrivals to homogeneous ones before applying standard econometric models. The transformation can be achieved by aggregating short intervals of time or long intervals of time. While the aggregation of short intervals of time will cause a certain type of heteroskedasticity, the aggregation of long intervals of time will cause a loss of important microstructure features of the data. A middle way that would maximize the benefits with irregularly spaced data is to treat the arrival times as random. This is the research avenue proposed in Engle and Russell (1998). Indeed, within the ACD framework a certain number of financial market microstructure predictions can be tested.

Easley and O'Hara (1992) argue that the length of time elapsed between consecutive transactions is not hazardous. Traders choose when to enter and to exit the market. Thus, the intra- trade durations constitute a source of information for quote submitters who by assumption have an information disadvantage in the presence of privately informed traders. In this paper, I treat the question of price formation in connection with intra-trade durations and formulate questions such as do intra-trade durations provide information to quote submitters or better, do intra-trade durations content information in the sense that quote submitters are illuminated about the type of traders in presence?

¹ See e.g. Ghysels and Jasiak (1997), Dufour and Engle (2000), Gerhard and Hautsch (2002)

² Readers interested in the structure and econometric of hazard models are referred to among others Cox (1972), Heckman and Walker (1990) and Lancaster (1990).

It takes time for market participants to learn about profit opportunities and to implement superior strategies. Easley and O'Hara (1992) associate long durations with no news while Diamond and Verrecchia (1987) associate them with bad news. In this paper, I examine those market microstructure predictions. The relation between the duration and the spread is expected to be negative to capture the learning element of trading. In terms of volatility, a positive relation is expected to signify that long durations are periods of low volatility in the market. Such a result will reinforce the belief that quote submitters formulate their questions not just in terms of "how much" or "how big" but also in terms of "how long".

To investigate the relationship between time arrivals and price movements, the concept of time deformation or subordinated process as developed in Clark (1973) is used as in Engle (1996). By definition, a subordinated process is a joint process in which one variable directs another one. This is a kind of joint interaction between two distinct variables. Clark (1973) uses price and volume, where volume is subordinated to movements in prices. In the case of Clark (1973), volume captures the rate of information arrivals. Thus, since the number of transactions is random, it might be receivable to view asset price movements as the realization of a process such as $y_t = y_{z_t}^*$, where y_t equals price and z_t is a non-decreasing, directing process. It can be then easily thought that z_t is related to transaction times, thus to arrival of information.

Building on the original idea as given in Clark (1973) and expressed in terms of time arrivals in Engle (1996), I investigate 1) the persistence pattern in the market, 2) the microstructure effects in the process of price formation and 3) the relation between time and price movements. To be specific, 1) I estimate the ACD model of Engle and Russell (1998) using transaction data from a limit order book, and 2) I extend the ACD model by including market microstructure effects. Finally, 3) I formulate an ARMA-GARCH model with a set of duration measures in order to test if the volatility increases with short durations. The third step achieves somehow the total effort put in to test all the formulated questions. To minimize the undesirable effects of market imperfections, prices are defined as midquotes, which are the average of the bid and ask prices at the time of transactions.

The benefits of this study are countable. If factors that determine the timing of transactions are related to the distribution of information across market participants, then the forecasts of the time between transactions may give added insight into the behavior of liquidity providers and seekers. Indeed, intensive moments as a result of short durations bring the good and the bad. From the perspective of quote submitters, the good comes with liquidity traders while the bad with informed traders. If short durations reveal informed traders then larger spreads would be taxed with low liquidity³. If informed traders split their trades for camouflage then the total volume would reveal them. Furthermore, this study constitutes a link between market system as it produced data and investors. Both the system manager and the investors realize the importance of monitoring each other in such thing as time and price priority.

The findings of this study are rather encouraging. The persistence pattern using time arrivals is clearly identified in the LOB context. The information arrival seems to be positively autocorrelated. By including microstructure effects such as spread and trading volume, it comes out that volume is negatively related to transaction times, indicating that volume conveys information to traders. The extended GARCH model reveals that 1) long durations are associated with low volatility as predicted by Easley and O'Hara (1992), 2) the spread is positively related to the volatility as predicted in microstructure empirical studies and 3) the trading volume is unexpectedly negatively related to the volatility. The last result is puzzling. The negative sign signifies that the relation between volume and volatility is quite complex as noted in Goodhart and O'Hara (1997).

This paper is organized as follows. Section 2 contains a brief introduction to the Helsinki Exchanges, which is an electronically driven limit order book market. The section pursues with a brief presentation of prior studies, the intra-data and the model to be estimated. It is worth to point out that only the useful part of the ACD model is presented. The reader who wishes to learn more about the model is strongly recommended to Engle and Russell (1998). Section three contains the empirical result at descriptive and parametric level. Section 4 summarizes the paper.

³ There are a variety of definitions on liquidity. In this context, liquidity is defined as the ability to perform a transaction without costs.

2. Empirical agenda

This section contains a brief description of the Helsinki Exchanges as it operates in 1995, an introduction to the concept of duration that is more and more used with the advent of high frequency data and a presentation of the data set. The section pursues with a formal presentation of the basic ACD model, its extension to include market microstructure effects and as related to the GARCH model. The extended GARCH model is formulated as in Engle (1996) and includes duration measures, trading volume and the bid-ask spread.

2.1 Trading practices in the Helsinki Exchanges

In 1989, the Helsinki Exchanges introduced a computerized, order-driven system. The system treats limit orders submitted by the members of the Exchanges who are dual capacity broker-dealers but no have the obligation to provide liquidity. Practically, the system registers and displays submitted orders in price and time priority. In 1995 there were three major sessions of trade: 1) the opening of trading between 8:30 and 9:50 a.m., 2), the continuous trading between 10:00 a.m. and 4:00 p.m. and 3) the after market trading between 4:05 and 5:05 p.m. In this paper, the focus lies on the data from the continuous trading session. Hedvall (1994) offers a thorough description of the market structure of the Helsinki Exchanges at its earlier and development stage.

2.2 Data

The data set consists of limit order book intraday data of the top four listed companies ranked by their market values at the end of 1995 in the Helsinki Exchange. The companies are Nokia with a market value of 51.3 billion FIM⁴, Repola with a market value of 12.3 billion FIM, Kymmene with a market value of 9.4 billion FIM and Outokumpu with a market value of 8.6 billion FIM. The data ranges from April 4, to June 30, 1995 for a total of 58 trading days. The data includes quote prices, quote volumes, transaction prices, transaction volumes and time stamps. Virtually, the data set is treated as in Engle (1996). The data preparation has reduced considerably the

⁴ FIM is an abbreviation for Finnish Markka, the currency at the sample time. Nowadays, the Euro is the current currency. At the sample time, 1 FIM equals approximately \$ US 0.84.

number of observations for the estimation. Following steps were taken in the preparation of the data. Firstly, I search for transaction and quote price changes by comparing price at time $t=0$ and price at time $t=1$. If the price at time $t=0$ equals the price at time $t=1$, I select price, time stamp and trading volume at time $t=1$ and ignore price and time stamp at time $t=0$ and add in such a case the trading volume at time $t=0$ to the trading volume at time $t=1$. Secondly, I match transaction and quote prices in time and date order. Thirdly, I selected bid and ask prices in terms of transaction prices. If the price at $t=0$ is a transaction price, I selected the prevailing bid and ask prices. Thus, the number of observations is determined in terms of the number of transaction price observations.

2.3 The series

Three series are fundamental for this study. The intra-return series (r_i) to measure the revision process of price. The intra-return is computed as the log difference between subsequent midquote prices obtained as $(q_i^a + q_i^b)/2$, where q_i^a is the ask quote and q_i^b is the bid quote. The intra-spread series to capture the asymmetry element of revision is computed as the log difference between q_i^a and q_i^b . The intra-trade duration series (x_i) is computed as the difference in seconds between subsequent transaction times.

Engle and Russell (1998) point out that the expected duration varies over the time of day, thus duration should be viewed as consisting of a stochastic component that should be modeled in an ACD model and a deterministic component that should be removed. Different techniques can be used to remove the seasonal effects in the series. Engle (1996) proposes a technique according to which actual durations are regressed on a spline function. The diurnally adjusted series is then the ratio of actual to fitted values. The same technique can be applied to diurnally adjust the return, the spread and the trading volume series. Thus the series are diurnally adjusted following Engle (1996).

2.4 Prior studies

Long before the ACD model proposed in Engle and Russell (1998), the concept of duration prospered in social science. Many of those models build on Cox (1972). Different social perspectives⁵ have been examined using social science duration models that are in practice different to “speculative” duration models but not in spirit. The difference lies basically in the data span. While “social” duration models look at durations given in month or even in year, speculative duration models use durations given in seconds. The points of convergence between the two types of models are quite many. Both models use about the same type of distribution since the data is typically a Poisson point process in both cases. To illustrate the point, Jaggia (1991) provides a model with the common distributions used in ACD models,

$$f(t | X) = \mu^k \alpha^{k-1} \frac{\exp(-\mu t^\alpha)}{\Gamma(k)}, \quad (1)$$

where

$$\mu = \exp(X\beta) = \exp(\beta_0 + \beta_1 X). \quad (2)$$

The mean, μ , is exponential to insure its non-negativity, X is a vector of explanatory variables, Γ is a symbol for the gamma function and β_i are the parameters to be estimated. From (1), we can get an exponential distribution by letting $\alpha = k = 1$, a Weibull distribution by letting $k = 1$ and a gamma distribution by letting $\alpha = 1$. In fact, Engle and Russell (1998) use both the exponential and the Weibull distribution to estimate their ACD model. The Bauwens and Veredas (1999) model is estimated with a Weibull and a Gamma distribution. Bauwens et al (2000) evaluating all the ACD models that have been proposed so far used an Exponential, a Weibull, and a Gamma distribution.

⁵ Among others, Lancaster (1979) examines the variation between unemployed job seekers using a parametric duration model. Heckman and Walker (1990) use as well a duration model to examine the relationships between wages and spacing of births in Sweden. Favoro et al (1994) use a duration model to examine the problem of an operator who owns a license to develop and extract oil from a field of known capacity.

The salient point with ACD models is that they provide an intensity measure for transaction prices. The intensity process of prices is obtained as a ratio of duration at time $t + 1$ and duration at $t+0$ as it tends to its zero limit. This feature has been proved to be one of the characteristics of transaction prices in stock markets. Indeed different results obtained from different ACD models show a clustering pattern in information arrivals, indicating that short durations are more likely to be followed by short durations as do long durations. This result is found among others in Ghysels and Jasiak (1997), Engle and Russell (1998) and Bauwens and Veredas (1999).

The clustering pattern of transaction durations has been detected in different type of ACD models. Since Engle and Russell (1998), the empirical analysis of durations has integrated some aspects of the market microstructure theory of financial markets as given for instance in Easley and O'Hara (1992). An interesting extension with respect to this study is the one relating transaction times to movement in prices. Engle (1996) provides a framework, which combines a GARCH model for returns, and duration measures obtained by estimating an ACD model. The mixture of transaction times and movements in prices follow from the original idea developed among other in Clark (1973). In this study, the mixture model proposed in Engle (1996) is examined using transaction data from a limit order book (LOB) market.

2.5 Basic model for intra-durations

A general assumption that is made when analyzing a financial series is that the observed series, $(x_1, x_2, x_3, \dots, x_T)$ is a stochastic realization. In that order, the intra-return, the intra-spread and the intra-trade duration series are treated as stochastic series. If there is not a shadow of doubt about the intra-return and the intra-spread series, the intra-time series have only lately been treated as such. In fact, Engle and Russell (1998) treat the arrival times as random variables, which follows a point process with a probability distribution defined in accordance with the degree of dispersion of the realizations. To account for the two moments that fundamentally characterize any series, Engle and Russell (1998) propose an autoregressive conditional model for intra-durations (ACD). The model stems from this basic specification:

$$x_i = \psi_i \varepsilon_i, \quad (3)$$

where ε_i are independent and identical increments with a distribution associated to the hazard type of functions. Depending on the adopted distribution of ε_i and the specifications of ψ_i , equation (3) can be expressed differently. Assuming for instance an exponential distribution, the essential features of equation (3) can be uniquely expressed in term of the mean equation as follows

$$\psi_i = E[x_i | x_{i-1}, \dots, x_1], \quad (4)$$

where $E[x_i | x_{i-1}, \dots, x_1]$ = the expected conditional mean. The mean function captures all the temporal dependence in duration. Therefore, the standardized durations obtained as x_i/ψ_i should be independently and identically distributed to allow proper estimations by maximum likelihood. By analogy to GARCH equation (4) can be expressed as follows

$$\psi_i = \varphi_0 + \sum_{j=1}^m \varphi_j x_{i-j} + \sum_{j=1}^q \vartheta_j \psi_{i-j}, \quad (5)$$

for $j = 1, \dots, m, q; i = 1, \dots, N$ and $\varphi_0 > 0$, $\varphi_j \geq 0$ and $\vartheta_j \geq 0$. Equation (5) is called ACD (m, q) with an infinite memory. Equation (5) implies that future durations are function of past durations. Due to the clustered nature of intra-trade durations and the assumption about the distribution of ε_i , equation (5) can take an exponential, a Weibull, or a gamma form.

2.6 Estimation steps

Since the expected durations vary over the time of day and the intra-returns exhibit a daily shape, a seasonal adjustment with respect to the time-of-day effects is required. Engle (1996) provides a technique by which an adjustment can be done. The technique consists of making a regression of the intra-durations on daytime defined as spline functions. The diurnally adjusted series is then the ratio of actual durations on the fitted durations.

The next step is to estimate an ACD model as given in (5). Practically, to estimate an ACD as a GARCH model, the diurnally adjusted intra-trade durations, \tilde{x} , should be expressed as $\sqrt{\tilde{x}}$ and the mean equation put to zero. The modified GARCH-model is then converted to an ACD-model as follows:

$$\psi_i = \varphi_0 + \varphi_1 \sqrt{\tilde{x}_{i-1}} + \vartheta \psi_{i-1}. \quad (6)$$

From (6), we get the conditional variance ψ_i that will be used late on. The third step is about examining the relationship between intra-trade durations and intra-returns. Engle (1996) proposes a model that is tractable and in accord with hypotheses in this study. The model combines the market volatility as captured by the intra-return with the market intensity as captured by the intra-durations. The conditional intra-return volatility may be expressed as

$$V_{i-1}(r | x_i) = h_i, \quad (7)$$

where the variance is conditioned on the current intra-trade durations as well as the past intra-returns and intra-trade durations. The diurnally adjusted intra-returns are appropriately expressed per unit of time. Dividing the intra-returns by the square root of intra-trade durations modifies slightly equation (7) as follows:

$$V_{i-1}\left(\frac{r_i}{\sqrt{x_i}} | x_i\right) = \sigma_i^2. \quad (8)$$

The ratio of intra-returns on the square root of intra-trade durations is then diurnally adjusted. The two conditional variances are related to the intra-trade durations. This relation can be expressed in term of the conditional variance of past intra-returns and intra-durations as follows:

$$E_{i-1}(h_i) = E_{i-1}(\tilde{x}_i \sigma_i^2), \quad (9)$$

where E_{i-1} is an expectation operator to signify that the conditional intra-return variance is a function of past intra-returns and intra-durations. Following Engle (1996), equation (7) can be given this testable form:

$$\tilde{r}_i = \phi \tilde{x}_i + \phi \tilde{r}_{i-1} + \theta \varepsilon_{i-1} \quad (10)$$

Equation (10) is an ARMA (1,1) with an exogenous variable, \tilde{x} intended to capture as predicted by Diamond and Verrecchia (1987) a negative relation between long intra-trade durations and intra-returns. The negative sign is an expression for bad news associated with long durations. Furthermore, equation (10) is the mean equation for the following GARCH-model:

$$\sigma_i^2 = \omega + \phi_1 \varepsilon_{i-1}^2 + \phi_2 \sigma_{i-1}^2 + \gamma_1 \tilde{x}_i^{-1} + \gamma_2 \frac{\tilde{x}_i}{\psi_i} + \gamma_3 \xi_{i-1} + \gamma_4 \tilde{s}_{i-1} + \gamma_5 \tilde{z}_{i-1}. \quad (11)$$

Equation (11) can be portioned in different models. For clarity, I stick to the estimation of (11) as such. Three factors are examined in (11). The price factor by the GARCH process, the time factor to capture the impact of private news and the size factor as a detection device. Following Easley and O'hara (1992) a positive sign is expected on γ_1 to signify that long durations are associated to no news. The diurnally adjusted volume \tilde{z} captures periods related to information-based transactions. A positive relation between the spread \tilde{s} and volatility will signify that long durations correspond to low volatility. The long run effect is measured by ξ_i computed by exponentially smoothing as $\xi_i = 0.005(\tilde{r}_{i-1}^2/\tilde{x}_{i-1}) + 0.995\xi_{i-1}$.

3. Empirical results

The central question of this study is about the impact of time in the process of price formation. Using transaction data, I present in this section, the main result of the relation between time and prices. To minimize the undesirable effects of bid ask spread bounce and price discreteness, the relation between time and prices is measured in term of midquote prices.

3.1 Intraday patterns

One of the findings about intra-periodicity is that return, volume, volatility and bid-ask spread exhibit an U-shaped pattern over the trading day. This particular pattern could not be fully detected with diurnally adjusted data as shown below:

Figure (1): Duration (Nokia)

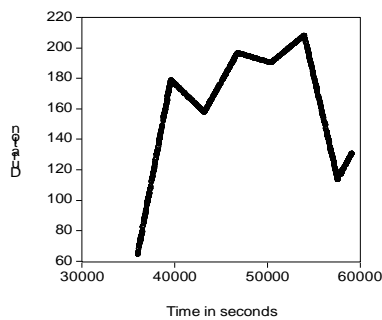


Figure (2): Spread (Nokia)

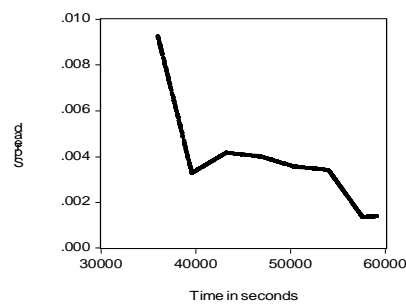


Figure (3): Volatility (Nokia)

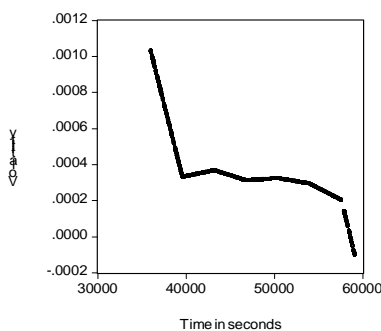
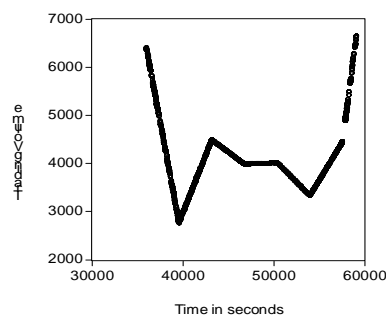


Figure (4): Volume (Nokia)



Note: figure (1), (2), (3) and (4) are obtained as a scatter with nearest neighbor fit. The y-axis gives values for durations, spreads, volatilities and volumes. Durations are obtained as the difference in seconds between successive arrival times. Spreads are obtained as log ask price minus log bid price. Volatilities are proxy for the absolute diurnally adjusted returns. Volumes are transaction volumes diurnally adjusted. The x-axis gives the number of seconds since midnight. The market open is at 10:00 a.m. thus at the point of 36,000 seconds since midnight. Figure (1) shows the duration pattern across the trading day whereas figure (2) shows the spread pattern across the trading day. Figure (3) shows the volatility pattern across the trading day whereas figure (3) shows the volume pattern across the trading day.

Figure (1), (2), (3) and (4) illustrate a daily pattern of prices as measured by the durations, spreads, volatility and volume. The x-coordinate' values are seconds about the opening and closing moments. The official opening time in the Helsinki Exchange was at that time fixed at 10:00 a.m. or 36,000 seconds from midnight. Figure (1) shows that durations are lower at the market open, goes up sometime after the market open and falls down at about 3:00 p.m. The shortest durations are observed at the market open.

Figure (2) shows that the spread is higher at the market open than at any other time, indicating that traders are much more uncertain about the true value of the asset at the market open. As trading goes on, traders are more and more aware of the true value of the asset. Figure (3) shows that volatility is higher at the market open, indicating that the process of price discovery is launched at the market open. Volatility decreases until it reaches its lowest point toward the end of the trading day. Figure (2) about spread and Figure (3) about volatility show about the same pattern. Figure (4) is at the odds of figure (2) and (3). The trading volume displays a U-shape pattern. Higher trading volumes at the market open and close are observed as predicted in Brock and Kleidon (1992). This pattern converges with the inverse of duration pattern displayed in figure (1). In sum, the figures show that the uncertainty about the asset true value decrease as time passes.

3.2 Price frequencies

The data preparation has consumed an important number of observations. Since I define the bid and the ask quotes in term of trades, the usable number of observations for estimations is about the number of transactions. Table (1) reports the total number of possible observations from the morning session to the evening session for a total of 58 trading days.

**Table 1:
Number of quote and trade prices**

Stocks	Bids	Asks	Trades	Total
Nokia	13 269 (33.73)	14 808 (37.64)	11 265 (28.63)	39 342 (100)
Repola	9 878 (37.15)	10 638 (40.01)	6 074 (22.84)	26 590 (100)
Kymmene	5 078 (33.64)	6 458 (42.78)	3 359 (23.58)	15 095 (100)
Outokumpu	3 400 (36.90)	3 995 (43.35)	1 820 (19.75)	9 215 (100)

Quotes and trades are given in the table as the total number of quotes and trades from 10:00 a.m. to 4:00 p.m. for a total of 58 trading days from April 3, to June 30, 1995. Within the parentheses are the percentages computed as Bids, Sells and Trades divided by Total.

The higher the market value of the stock, the lower the percentage of asks. Nokia, with the highest market value, shows consequently the lowest percentage of asks and Outokumpu, with the lowest market value, shows logically the highest percentage of asks, which means that traders are reluctant to sell highly valued asset. In a hypothetical market of only four stocks, trades from Nokia will account for about 50%⁶ of the whole trades in the market. In such a market, Nokia will affect beliefs positively by 41.96%⁷ and negatively by 41.25%⁸.

3.3 Descriptive statistics

The descriptive statistics are reported in this bisection. The diurnally adjusted series are expressed as fractions above or below normal. The mean and the standard deviation for diurnally adjusted intra-trade durations are expected to be 1. Nokia, Repola, Kymmene and Outokumpu show a mean about 1. In the case of Nokia, the range runs from 0.0034 to 38.3 or 1/294 below the mean and 38 times above the mean. The standard deviation for the four stocks is over 1, indicating an overdispersion phenomenon⁹. The unadjusted statistics are given in table (2).

⁶ obtained as 11265/22518

⁷ obtained as 13269/31625

⁸ obtained as 14808/35899

⁹ Statistics for the diurnally adjusted durations

Stocks	Mean	Standard deviation	Minimum	Maximum
Nokia	0.995	1.99	0.00340	38.3000
Repola	0.998	1.77	0.00256	17.7139
Kymmene	1.000	1.07	0.00015	8.46921
Outokumpu	1.003	1.84	0.00091	21.5899

Kymmene shows less dispersion; the mean is about equal to the standard deviation as expected. Nokia, Repola and Outokumpu show overdispersion.

Table 2:
Mean and standard deviation statistics

	Duration	Return	Spread	Volume
Nokia	161*** (332)	0.00446% (0.4696%)	0.3823%*** (1.0704%)	3,400*** (7,340)
Repola	332*** (609)	0.0045% (0.3213%)	0.2572%*** (0.6570%)	5,057*** (10,280)
Kymmene	472*** (895)	0.00722% (0.5343%)	0.4931%*** (0.990%)	4,173*** (14,489)
Outokumpu	813*** (1,535)	-0.0116% (0.6288%)	0.6569%*** (1.311%)	5,790*** (11,523)

Note: within the parentheses are given the standard deviation statistics. Duration is obtained as the difference in seconds between successive transactions times, Return as the log difference between successive midquote prices and Spread as the log difference between the ask and bid quote. Volume is the raw (no adjusted) transaction volume. The three asterisks indicate a significance level of one percent. I test for the null hypothesis that duration, return, spread and volume equal zero, respectively.

Table (2) reports the mean and the standard deviation statistics for duration, return, spread and volume. Outokumpu shows the highest average duration with its 813 seconds and average trading volume with its 5,790 number of shares. Nokia with the highest market value shows as expected the lowest average duration with its 161 seconds and the lowest average trading volume with its 3,400 number of shares. It is interesting to note on the one hand that the lower the market value of the stock, the higher the number of share traded and on the other hand that the higher the market value, the lower the waiting times between transactions.

The mean for the intra-return is largely under 1% and non-significant. Kymmene shows the highest mean with its 0.007% and Outokumpu shows the lowest mean with its -0.01%. The average spread is under 1% but significant different from zero. Ordinary, the highest the market value of the stock, the lowest the spread is, and vice versa. By Jacque-Bera test statistics (not reported) the null for a normal distributed intra-return and intra-spread series is rejected at 1% significance level. High kurtosis hinders the series to exhibit a normal shape¹⁰.

¹⁰ The Jarque-Bera is a test for testing whether the series is normally distributed. The test statistic measures the difference of the skewness and kurtosis of the series with those from the normal distribution. The statistic is computed as:

$$Jarque - Bera = \frac{N - k}{6} \left(S^2 + \frac{(K - 3)^2}{4} \right),$$

where S is the skewness, K is the kurtosis, and k represents the number of estimated coefficients used to create the series. Under the null hypothesis of a normal distribution, the Jarque-Bera statistic is distributed as χ^2 with 2 degrees of freedom.

Using diurnally adjusted data, the four stocks show a positive correlation between spread and return. Although the correlation is not strong; the highest correlation coefficient from Kymmene is 13.34% and the lowest from Outokumpu is 0.45%. Furthermore, duration is negatively related to spread but positively related to return. Between duration and size, there is a negative correlation. In sum, the adjusted series are not strongly interrelated. The highest correlation coefficient about 14.28% is between size and spread. However, the adjusted series show autocorrelation between the intra-trade durations at longer lag.

Table 3:
Autocorrelation of the raw and adjusted durations

	Nokia		Repola		Kymmene		Outokumpu	
	ARD	AAD	ARD	AAD	ARD	AAD	ARD	AAD
lag 1	0.07	0.05	0.01	-0.00	0.07	0.05	0.02	0.01
lag 2	0.13	0.11	0.09	0.08	0.07	0.05	0.05	0.05
lag 3	0.06	0.04	0.01	0.02	0.06	0.06	0.05	0.07
lag 4	0.06	0.06	0.07	0.06	0.05	0.04	0.03	0.06
lag 5	0.08	0.03	0.02	0.03	0.06	0.07	0.03	0.05
lag 10	0.06	0.05	0.02	0.03	0.04	0.05	0.04	0.07
lag 15	0.04	0.03	0.03	0.03	0.06	0.05	0.03	0.03
Q(15)	508***	368***	102***	114***	82***	93***	29***	45***

Note: ARD = Autocorrelation coefficient for raw durations, AAD = autocorrelation for diurnally adjusted durations, Q (15) = the Ljung-Box test statistics up to the 15th lag. The three asterisks indicate a one percent level of significance.

With 15 lags, the Ljung-Box statistics for the null hypothesis of temporal independence in duration is easily rejected. The autocorrelation coefficients are in general significant and positive. This result corroborates with earlier findings about significant clustering of price movements. In sum, the autocorrelation structure for duration is damping exponential (declining rapidly).

3.4 ACD-estimation

Engle and Russell (1998) show that the dynamic of trade can be captured by the durations between successive transaction times. The proposed ACD-model can be evaluated by estimating a variance equation as given in (5). Clearly, this is a mean process that is evaluated since the error that appears in (3) is assumed to have an exponential density equal to one.

Table 4:
Autoregressive Conditional Duration Estimates

	Nokia	Repola	Kymmene	Outokumpu
φ_0	0.0018**	0.00945***	0.0177***	0.0194**
φ_1	0.0159***	0.03241***	0.0394***	0.0631***
ϑ	0.9822***	0.95771***	0.9432***	0.9188***
$\varphi_1 + \vartheta$	0.9981	0.9910	0.9825	0.9819
NOB	8,099	3,554	2,522	1,233

Note: Equation estimated is $\psi_i = \varphi_0 + \varphi_1 \sqrt{\tilde{x}_{i-1}} + \vartheta \psi_{i-1}$, where ψ_i = the expected duration at time t=1 and \tilde{x} = diurnally adjusted duration. Equation is estimated by maximum likelihood. The adjusted duration is obtained as a ratio of raw duration and fitted duration from the spline function estimation. NOB = the number of observations after adjustments. The estimated parameters φ_0 , φ_1 and ϑ are robust to heteroskedasticity and autocorrelation. The two and three asterisks indicate a significance level of five and one percent, respectively.

The sum of φ_1 and ϑ is under unity although close enough to signify a persistent pattern in durations. The well-known IGARCH phenomenon does appear in the underlying process as the sum of the parameters (φ_1 and ϑ_1) is about 1 for the four stocks. Indeed, the persistence is much more pronounced for Nokia and Repola, the most traded among the four stocks. The influences from past intra-durations are measured by φ_1 while those from past duration expectations by ϑ . As assumed the three coefficients in table (4) are positive. The expected duration ψ_i absorbs all the persistence.

Furthermore, following Lee and Hansen (1994), the standardized intra-duration obtained as x_i/ψ_i is stationary and ergodic by unit root test. The diagnostic checks of the standardized intra-durations reveals that they are identically and independently distributed. The Ljung-Box statistics at 30 lag for simple x_i/ψ_i is only 20.5 for Nokia, 25.7 for Repola, 30.7 for Kymmene and 25.5 for Outokumpu. The Ljung-Box statistics at 30 lag for squared x_i/ψ_i is only 7.53 for Nokia, 22.5 for Repola, 12.4 for Kymmene and 14.3 for Outokumpu. It is clear that the standardized durations are cleaned enough to justify the estimation by maximum likelihood. The range for the standardized series runs from 0.0024 to 47.43 in the case of Nokia.

3.5 Microstructure effects

In stock markets there are variables that influence the frequency of quote revisions. The spread and the volume are such variables. The mechanism of price revision evolves by assumption as follows. Quote submitters hold the spread closer to the real value of the asset when duration is long but widen it as a response to a high total volume during a short period of time.

Table 5:
ACD estimation with microstructure effects

	Nokia	Repola	Kymmene	Outokumpu
φ_0	0.0033***	0.0072	0.0198**	0.0303**
φ_1	0.0161***	0.0268***	0.0429***	0.0706***
ϑ	0.9814***	0.9666***	0.9363***	0.9047***
γ_1	0.0005	0.0001	0.0004	0.0059
γ_2	-0.0014**	-0.0003	-0.0027	-0.0088**
NOB	8,099	3,554	2,522	1,233

Note: Estimated equation: $\psi_i = \varphi_0 + \varphi_1 \sqrt{\tilde{x}_{i-1}} + \vartheta \psi_{i-1} + \gamma_1 \tilde{s}_{i-1} + \gamma_2 \tilde{z}_{i-1}$, where \tilde{s} = diurnally adjusted spread and \tilde{z} = diurnally adjusted trading volume. The adjusted spread (volume) is obtained as a ratio of raw spread (volume) and fitted spread (volume) value from the spline function estimation. The variance equation is estimated by maximum likelihood. NOB = the number of observations after adjustments. The estimated coefficients are according to Bollerslev-Wooldrige robust standard errors & covariance. The two and three asterisks indicate five and one percent significance level, respectively.

Trading volume is negatively related to the expected duration, indicating that volume conveys information. Thus, short durations are related to large trading volume. The spread is positively related to the expected duration, even though the parameters are not significant. The microstructure effects have slightly scaled the ACD parameters. This result shows that even in LOB markets in which trades are executed against outstanding limit orders, prices are revised under asymmetric impulsion. Thus, market participants do learn from watching transaction time and do make difference between information-based and liquidity-based moments.

3.6 Mixture of time and prices

Time is not informative in itself. It must be marked. By simple analogy, each mark highlights the time interval since time runs unnoticeably to infinity. Therefore, prices and volumes pose limits to time. Thus, time becomes informative by association. Engle and Russell (1998) show that durations are random and empowered through trading. Indeed, Easley and O'Hara (1992) predict that traders despite periodicity intentionally choose the waiting times. This statement follows from the asymmetry theory about quote submitters protecting themselves against better-informed traders. Therefore long durations have been given two interpretations. Easley and O'Hara (1992) associate long durations with no news while Diamond and Verrecchia (1987) associate them with bad news. Let us examine those two predictions with a volatility model designed for prices.

Table 6:
GARCH equation with duration measures

	Nokia	Repola	Kymmene	Outokumpu
Mean equation				
φ	0.0001	-0.0049	0.0042	0.0696***
ϕ	0.2304	0.6140***	0.3079***	0.7906***
θ	-0.3121	-0.5487***	-0.5331***	-0.5387***
Variance equation				
ω	0.1079	0.1662	0.0609	0.1182
ϕ_1	0.0005	0.2556***	0.2015***	0.1996***
ϕ_2	0.5009***	0.4759***	0.4732***	0.4723***
γ_1	0.4948***	0.4859***	0.4708***	0.4097***
γ_2	-0.1185***	-0.1419***	-0.1467***	-0.1299***
γ_3	0.1003***	0.0726***	0.0579***	0.0676**
γ_4	0.7454**	-0.8385	-1.1583	-0.4881
γ_5	-0.1287**	-0.0822***	-0.0864***	-0.1696***
NOB	8,099	3,554	2,522	1,233

Note: 1) mean equation = $\tilde{r}_i = \varphi\tilde{x} + \phi\tilde{r}_{i-1} + \theta\epsilon_{i-1}$, where \tilde{r}_i = diurnally adjusted return and 2) the variance equation = $\sigma_i^2 = \omega + \phi_1\epsilon_{i-1}^2 + \phi_2\sigma_{i-1}^2 + \gamma_1\tilde{x}_i^{-1} + \gamma_2\tilde{x}_i/\psi_i + \gamma_3\xi_{i-1} + \gamma_4\tilde{s}_{i-1} + \gamma_5\tilde{z}_{i-1}$, where ξ = long run volatility, \tilde{s} = diurnally adjusted spread and \tilde{z} = diurnally adjusted volume. The adjusted return is obtained as a ratio of raw return and squared root of duration. The adjusted spread (volume) is obtained as a ratio of raw spread (volume) and fitted spread (volume) value from the spline function estimation. The mean and the ACD equations are estimated by maximum likelihood. The coefficients are according to Bollerslev-Wooldrige standard errors and covariance matrix. The two and three asterisks indicate a five and one percent significance level, respectively.

The drift coefficient φ that should capture the long duration effect on prices is globally positive and non-significant in all cases. Repola shows a negative and a non-significant drift coefficient. It seems like the data does not support the hypothesis according to which long durations are associated with bad news as predicted in Diamond and Verrecchia (1987). Instead, the data seems to support the hypothesis according to which long durations are associated with no news as predicted by Easley and O'Hara (1992). Furthermore, the AR coefficients are positive and significant while the MA coefficients are negative and significant for all stocks except for Nokia.

The variance equation gives a clear pattern about the real influence of duration, spread and volume on volatility. The result with respect to the magnitude and to the sign of coefficients is robust across the four stocks. Parameters γ_1 and γ_2 are designed to diversely capture the influence of duration on volatility. The first duration parameter shows a positive and significant relation, indicating that long durations correspond to calm period in the market. Thus, as predicted by Easley and O'Hara (1992) long durations are associated with the absence of news. The other duration parameter shows a negative and significant relation, indicating that the surprise component of duration increases volatility in the market. The parameter for the long run volatility shows a positive relation to the market volatility. The spread parameter γ_4 shows globally a positive relation to the market volatility, indicating that an increase in volatility leads to an increase in the spread. The volume parameter γ_5 shows a negative relation to the market volatility, indicating that the total volume of short-lived transactions increases the market volatility.

4 SUMMARY

This paper investigates the persistence pattern in the Helsinki Exchanges. The pattern is examined using a time and a price approach. Following Engle and Russell (1998), arrival times are modeled as following an Autoregressive Conditional Duration (ACD) process. The ACD model is designed for irregularly spaced transaction data. Since the time interval between successive transactions can be measured, a process for arrival times might capture the moments under which the market is intense. Thus, under the ACD framework, it is hypothesized that traders associate short durations with the existence of some kind of information whereas they associate long durations with the absence of information.

Following Clark (1973) and Engle (1996), the relationship between arrival times and movements in prices is investigated using a GARCH framework. Under the GARCH framework, it is hypothesized that long durations are associated with low volatility in the market. By including duration measures in the GARCH model for price the hypothesis could be tested. In sum, using intraday data for the four most traded stocks in the Helsinki Exchange at the sample time, interesting results emerged on the analysis of the persistence pattern.

It follows that the ACD model captures quite well the persistence pattern of the four most traded stocks in the Helsinki Exchanges, suggesting that intraday durations have potential to forecast the time of the next transaction. Extending the ACD model by including market microstructure effects, it comes out that trading volumes are negatively related to durations, suggesting that short durations are associated with higher trading volumes. Furthermore, defining prices as a GARCH process, it comes out that arrival times have implications on the observed market volatility. Precisely, long durations are associated with low volatility in the market. The main result leads to the conclusion that market participants do learn from watching trading time and transaction size as predicted by Easley and O'Hara (1992) and found previously by Engle (1996).

REFERENCES

- Bauwens, L. and Veredas, D. (1999): "The Stochastic Conditional Duration model: a latent factor model for the analysis of financial durations", Working paper, Center for operations research and economics (CORE), Louvain-la-neuve, Belgique.
- Bauwens, L., Giot, P., Grammig, J and Veredas, D. (2000): "A comparison of financial duration models via density forecasts", CORE, Discussion paper.
- Brock, W. A. and Kleidon, A.W. (1992): "Periodic Market Closure and Trading Volume: A Model of Intraday Bids and Asks", *Journal of Economic Dynamics and Control*, 16, 451-489.
- Cox, D. R., (1972): "A Regression Models and Life Tables (with discussion)", *J. R. Stat. Soc. B.* 187-220.
- Diamond, D.W. and Verrecchia, R. E., (1987): " Constraints on Short-selling and Asset Price Adjustments to Private Information", *Journal of Financial Economics*, 18, 277-311.
- Dufour, A. and Engle, R.F., (2000): "Time and the Price Impact of a Trade", *The Journal of Finance*, Vol. LV, 2467-2498.
- Easley, D. and O'Hara, M., (1992): "Time and the Process of Security Price Adjustment", *Journal of Finance*, 47, 905-927.
- Engle, R.F. and Russell, J. R., (1998): "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data", *Econometrica*, Vol. 66, 1127-1162.
- Engle, R.F., (1996): "The Econometrics of Ultra High Frequency Data", Discussion Paper, University of California.
- Favero, C. A., Pesaran, M. H. and Sunil, S. (1994): "A duration model of irreversible oil investment: theory and empirical evidence", *Journal of applied econometrics*, 9, 95-112.
- Gerhard, F. and Hautsch, N., (2001): "Volatility Estimation on the basis of Price intensities" *Journal of Empirical Finance*, 9, 57-89.
- Ghysels, E., and Jasiak, J., (1997): "GARCH for Irregularly Spaced Financial Data: The ACD-GARCH model", CIRANO
- Goodhart, A.E.C. and O'Hara, M. (1997): "High frequency data in financial markets: Issues and applications", *Journal of Empirical Finance*, 4, 73-114.

Heckman, J.J. and Walker, J. R., (1990): “The Relationship between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data”, *Econometrica*, Vol. 58, 1411-1441.

Hedvall, K. (1994): “ Essays on the market microstructure of the Helsinki stock exchange”, Dissertation, No 56, *Ekonomi och Samhälle*, Swedish School of Economics and Business Administration.

Jaggia, S. (1991): “Specification tests based on the heterogeneous generalized gamma model of duration: with an application to Kennan’s strike data, *Journal of applied econometrics*, 6, 169-180.

Lancaster, T., (1979): “Econometric Methods for Duration of Unemployment”, *Econometrica*, Vol. 47, 939-956.

Lee, S., and Hansen, B., (1994): “Asymptotic Theory for the GARCH (1,1) Quasi-Maximum Likelihood Estimator”, *Econometric Theory*, 10, 29-52.

Meddelanden från Svenska handelshögskolan

Swedish School of Economics and Business Administration. Working Papers.

2002

466. Sääksjärvi, Maria: Consumer Adoption of Technological Innovations.
467. Solitander, Nikodemus & Kovács, Gyöngyi: Corporate Environmental Management. A Blueprint for Categorising Environmental Policies.
468. von Nandelstadh, Alexander: Analysts' Accuracy of Estimation and the Relative Trading Volume.
469. Jern, Benny: Performance of Finnish Mutual Funds: A Decomposition Approach.
470. Pasternack, Daniel: Factors Driving Stock Option Grants - Empirical Evidence from Finland
471. Rosenberg, Matts: Does Uncertainty Affect Investment and Labor Demand?
472. Karakozova, Olga: An Econometric Analysis and Forecast of Office Property Rents in the Helsinki Area.
473. Felixson, Karl: The Expiration Day Effect of Index Options and Index Futures on the Underlying Shares.
474. Virtanen, Henrik: Samarbete mellan konkurrenter- en undersökning av österbottniska sm-företag.
475. Liljebloom, Eva & Pasternack, Daniel: Share Repurchases, Dividends, and Executive Options; Empirical Evidence from Finland.
476. Pasternack, Daniel & Rosenberg, Matts: The Impact of Stock Option Incentives on Investment and Firm Value.
477. Jern, Benny: Performance Persistence of International Stock Funds.
478. Ekholm, Anders & Pasternack, Daniel: Overconfidence and investor size.
479. Juselius, Mikael: Price Inflation, Unemployment and Devaluations: The Finnish Experience of the 1990s.
480. Aba Al-Khail, Mohammed & Berglund, Tom: The Impact of the EMU on International Portfolio Investments.
481. Aba Al-Khail, Mohammed: International Portfolio Investments and the Informational Value of Trade.
482. Aba Al-Khail, Mohammed: The Impact of FDI on International Portfolio Investments.
483. Snellman, Kenneth: Incentives and Substitute Personal Activities.
484. Kovács, Gyöngyi: Digital Asset Management in Marketing Communication Logistics.
485. Ben Sita, Bernard: An Empirical Study of the Mixture of Time and Movements in Prices.