# Modelling High Frequency Financial Time Series with Trading Information

by

Shinn-Juh Lin

Graduate Program in Economics

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
August 1998

0-612-31160-0

ABSTRACT

The purpose of this thesis is to study the modelling of high frequency financial time series when extra information is available from trading activities.

The first chapter starts with a literature review on the available econometric models that have been proposed for modelling financial time series, in particular, the returns series. These include models that assume returns being drawn from some time homogeneous distributions, and models that take into account the time dependency of returns. We concentrate on models that incorporate exogenous trading information. Since information is not practically observable, informational proxies have to be used, instead. Typically, informational proxies are constructed from observed trading variables such as trading volume, the number of trades, the number of price changes, the number of quote changes, and the number of executed order imbalance. The goal of this chapter is to identify suitable models and informational proxies that would be potentially useful for modelling high frequency data.

Applied research in high frequency financial data has a very short history. Furthermore, unlike daily or weekly financial time series, the high frequency 1-minute data studied in this thesis has several distinguished features that deserve our attention. Therefore, we devote the second chapter to a detailed examination of the 1-minute data. To facilitate the analysis, we select two representative stocks, IBM and INTEL, from the New York stock exchange, and British Telecom from the London Stock Exchange. We find, from the constructed one-minute data, that neither the returns series nor the trading variables are independently, identically, and normally distributed. There also exists a large significant negative first-order autocorrelation in the one-minute data. Since significant negative first-order autocorrelation typically does not exist in any returns series with frequency lower than one day, it differentiates high frequency data that we use from the lower frequency data. In the following chapters, we will try to model this interesting stylized fact. In addition to trading variables, the British Telecom data from the London Stock Exchange also contain extra buy/sell information on every single transaction. In the preliminary analysis, we find that different types of buy/sell trades do not arrive independently and could be potentially

helpful in explaining the returns autocorrelation.

The third chapter begins the model building of this thesis. Since information does not arrive in equally spaced intervals, we propose that a mixed jump-diffusion process is suitable for modelling both the returns dynamics and the information arrival. In this chapter, we detail the modelling, the estimation procedures, and the construction of some test statistics. The estimation and testing results of this model using those three aforementioned data sets indicate that the number of trades typically perform better than any other trading variable as the informational proxy. In order to accommodate the non-homogeneous nature of the information variables, we then extend our model by allowing the information arrival intensity to follow a stochastic process. When this doubly stochastic model is employed, we find that volatility persistence is much reduced compared to what is typically observed in the GARCH-type of models. The extended model is also capable of capturing the arrival of the number of trades.

Although the models proposed in the third chapter are able to capture the arrival of the number of trades and pin down the volatility persistence, it fails to describe the other stylized fact, significant negative first-order autocorrelation. In the final chapter, we extend our original homogeneous mixed jump-diffusion model by incorporating the extra buy/sell information existing in the British Telecom data. With this new setup, the significant negative first-order autocorrelation is on average reduced by 40%. In addition to explaining the autocorrelation, we also use buy/sell signals in exploring market asymmetry in a threshold autoreregresive (TAR) framework. Specifically, we develop a qualitative threshold model with conditional heteroskedasticity, where both the conditional mean and conditional variance are regime-dependent. This model fits high-frequency data better than the benchmark GARCH model. It also generates smaller volatility persistence. This finding is in accordance with the empirical evidence that high GARCH measure of volatility persistence may arise as a result of mis-specifying existing structural changes. Furthermore, our model provides much better in-sample and out-of-sample prediction on both returns and volatility.

ACKNOWLEDGEMENTS

This thesis would NEVER be completed without the excellent guidance, constant encouragement, and unlimited support that my supervisor, Professor John Knight, has offered me. His approachability and sense of humor have also turned my volatile academic life into an enjoyable one.

My special thanks also go to Professor Jian Yang. He spent much of his time in reading the whole draft of my thesis, and provided me with lots of valuable insights and suggestions. I would like to thank Professor Reg Kulperger for teaching me advanced probability theory and sparking my interest in modelling financial data. Professor Victor Aguirregabiria also deserves my thanks for providing many useful comments on this thesis.

I would also like to express my sincere gratitude to the Department of Economics at Birkbeck College, the University of London, as a major part of this thesis was conducted there. I would especially like to thank Dr. Stephen Satchell for his profound inspiration and discussion in Financial Economics. His invitation for handicap races at Regent's Park also contributed a lot to lift my momentum in getting through the Ph.D. program. I like to thank Dr. Mike Orszag for providing me with lots of valuable suggestions in numerical computation and for allowing me access to his very useful computing resources. Dr. Oliver Hansch also deserves my special thanks in sharing with me his invaluable expertise in the operation of the London Stock Exchange.

I should also mention some of my friends to whom I am deeply grateful for their help and support. These people are Yvonne Adams, Melissa Broadfoot, Jane McAndrew, and Song Zhang at UWO; and I-Hui Cheng, Ali Choudhary, Fabio Comelli, Uma Moorthy, and Marco Polenghi at Birkbeck College.

Finally, I would like to dedicate this thesis to my wife Autumn, who has given me the greatest support and two adorable kids, Philip and Jessica. Without her, this thesis would not have been possible.

*To Autumn, my wife*

# TABLE OF CONTENTS

# LIST OF TABLES

# Chapter 1

# Financial Time Series Models and Trading Information

## 1.1 Introduction

One of the long-lasting debates in financial econometrics is whether a sequence of speculative asset returns are independently, identically, and normally distributed. Such an issue is important at least for the following two reasons. First of all, in many financial studies, economic hypotheses are tested under the assumption that the daily stock returns are independently, identically, and normally distributed. Secondly, the distribution of speculative price returns has implications for pricing derivative securities, whose pricing formula rely heavily on the variance of the underlying returns distribution. It is obvious that any departure from the *i.i.d.* normality assumption may lead to dubious statistical inferences and results. Hence, how to best model asset returns has always been an important issue in financial econometrics.

The earliest parametric distributional model for stock returns is proposed by Bachelier (1900). By assuming the sequence of prices have independent increments with expected increment equal to zero, Bachelier constructs a random walk model of stock price changes. According to this model, price changes then follow a normal distribution with zero mean and variance proportional to the time difference. In other words, they follow a standard Brownian motion. However, empirical evidence suggests that daily stock returns typically exhibit higher kurtosis and fatter tails than what would be expected out of a normal distribution. Furthermore, it is well known that, in financial markets, large price changes tend to bunch together. Similarly, small price changes tend to be followed by small price changes. This so called volatility-persistence phenomenon is an indication of time dependency of the returns series. In other words, no matter how uncorrelated, a sequence of financial returns is rarely an independent series. Consequently, an independent homoskedastic Gaussian process is not adequate for describing financial asset dynamics.

Models proposed by researchers in responding to the incapability of the *i.i.d.* normal distribution in fitting stock returns can be broadly categorized into two groups. The first group contains models that assume returns being drawn from some time homogeneous distributions, while the second group contains models that take into account the time dependency. Examples of the first group include Mandelbrot's (1963) stable distribution, Praetz's $t$-distribution (1972), Clark's (1973) log-normal normal distribution, Kon's (1984) finite mixture of normal distributions, and Merton's (1976) mixed jump-diffusion model. This group of models originate from the stochastic nature of information arrival, and information asymmetry among market participants. Essentially, they are derived either by compounding a normal distribution with a variance parameter drawn from some other time-invariant distribution, or by compounding several normal distributions. Although unconditional distributions derived from these models exhibit fatter tails and higher kurtosis than those of a normal distribution, they do not take into account the time dependency of financial time series, and hence are unable to capture the volatility-persistence phenomenon. Therefore, we will not pursue any further discussion on these types of model. We include them here for the sake of completeness.

To capture the serial dependency of stock returns, several candidate models are available. The mixed jump-diffusion model provides a natural framework for us to study the time dependency of financial time series with exogenous trading information. Imagine that we can approximate the information arrival by some observable trading activities, the time dependency is built into the model via the dependency of trading activities. Other available models that are designed to capture the time dependency of financial time series include Engle's (1982) autoregressive conditional heteroskedasticity (ARCH) model, Bollerslev's (1986) generalized autoregressive conditional heteroskedasticity (GARCH) model, Hamilton's (1989) regime-switching model, and Tong's (1983) threshold model. These models enable us to study nonlinearity that exists in the data.

In this chapter, we survey these available models for modelling stock returns. In particular, since market participants react to informational available to them, it is natural to treat equity returns as results of the influx of new information into the market, and of the re-evaluation of existing information. Consequently, we concentrate on models that enable us to incorporate the impact of information. However, information is not practically observable, some proxies have to be used, instead. To construct these proxies, there are public news releases which contain information available to the general public. There is also private information that is ultimately reflected in the trading activities. Due to the data frequency that we are examining in this thesis, we will focus on reviewing private informational proxies that are adopted in the empirical literature. The candidates include trading volume, the number of trades, average trading volume, the number of quote arrivals, the number of price changes, and the executed order imbalance. It is well known that the operation of financial markets is far from that of a Walrasian competitive market, and tends to be in a sequence of disequilibrium. This is especially true in short time horizons. Hence, despite their imprecise role as informational proxies, trading activities are likely to contain information about the disequilibrium dynamics of asset returns.

The goal of this chapter is to identify suitable models and informational proxies that would be potentially useful for modelling high frequency financial data. We start surveying available returns distribution models in the next section. In Section 1.3, we review trading variables that have been proposed as proxies for private information. We provide some concluding remarks in the final section.

## 1.2 Available Returns Distribution Models

Stock returns typically consist of quite a few discontinuities. The mixed-jump diffusion model of Merton (1976) is able to capture any abnormal informational shocks and thus display discontinuous sample path. Basically, it is a mixture of a continuous normal compounding and a Poisson jump process to allow instantaneous stock price increase or decrease. It exhibits certain martingale properties (see Harrison and Pliska (1981)), and is thus consistent with the efficient market hypothesis.

The mixed jump-diffusion model could be described as the following

$$X(t) = \ln(\frac{P_t}{P_{t-1}}) = (\alpha - \frac{1}{2}\sigma^2) + \sigma(z(t) - z(t-1)) + \sum_{i=1}^{\Delta N(t)} Q_i \tag{1.1}$$

where $\Delta N(t)$ is the Poisson counting process with intensity parameter $\lambda$; $Q_i$ measures the jump size and is distributed $N(\mu_Q, \sigma_Q^2)$; $\alpha$ is the instantaneous conditional expectation; and $\sigma^2$ is the instantaneous conditional variance; $z(t)$ is the standard Brownian motion. The unconditional density function of $X(t)$ could be calculated as

$$pdf(X(t)) = e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \phi(\mu + \mu_Q \cdot j, \sigma^2 + \sigma_Q^2 \cdot j) \tag{1.2}$$

where $\mu = \alpha - \sigma^2/2$, and $\phi(a, b)$ is the normal density with mean $a$, and variance $b$. All moments of this unconditional distribution exist. The unconditional distribution of $X(t)$ is leptokurtic if $\lambda > 0$, and is skewed if $\mu_Q \neq 0$. If the random variable is standardized by setting $\mu = \mu_Q = 0$, and $\sigma^2 + \lambda\sigma_Q^2 = 1$, the unconditional distribution of $X(t)$ is more peaked around the center and exhibits longer tails than a standard normal distribution. To capture the serial dependency of stock returns, Oldfield, Rogalski and Jarrow (1977) employ the mixed jump-diffusion model and allow for serial correlation between jumps, which implies serial correlation between each transaction. Their formulation sheds light on how one could incorporate the informational variables through the mixed jump-diffusion model. Indeed, the mixed jump-diffusion model provides a natural framework for us to study the time dependency of financial time series with exogenous trading information. Imagine that we can approximate the information arrival in a mixed jump-diffusion model by some observable trading activities, the time dependency is built into the model via the dependency of trading activities. Furthermore, we can include buy/sell trading information into the model, or formulate the Poisson arrival intensity parameter by some other stochastic process that depends on its own history. We shall explore these possibilities in Chapter 3 and Chapter 4.

Other available models that are designed to capture the time-dependency of financial time series include Engle's (1982) autoregressive conditional heteroskedasticity (ARCH) model, Bollerslev's (1986) generalized autoregressive conditional heteroskedasticity (GARCH) model, Hamilton's (1989) regime-switching model, and Tong's (1983) threshold model. The GARCH model formulates the current conditional variance as a linear function of past squared innovations and historical conditional variances. This model can be specified as the following

$$\begin{aligned} X_t &= \varepsilon_t \qquad (1.3) \\ \text{where } \varepsilon_t &= \nu_t h_t^{1/2}, \;\; \nu_t \sim iid\,(0,1), \\ \text{and } h_t &= \alpha + \beta_1 \varepsilon_{t-1}^2 + .... + \beta_p \varepsilon_{t-p}^2 + \gamma_1 h_{t-1} + .... + \gamma_q h_{t-q} \end{aligned}$$

This model and its variants are specifically designed to capture the commonly observed volatility clustering phenomenon of which the persistent level is measured by $\sum_{i=1}^{p} \beta_i + \sum_{j=1}^{q} \gamma_j$. They are also capable of describing other observed stylized facts of financial data, such as fat tails, leverage effects, mean reversion, market asymmetry,..., etc. Therefore, this type of model is very popular in financial modelling.[1] However, the GARCH-type model has several shortcomings. First of all, due to its deterministic formulation of the conditional variance, a GARCH model is unable to display stochastic volatility which is another observed stylized fact of financial data. Secondly, a high GARCH measure of volatility persistence may be due to the failure to take into account of structural shifts in the model (Lamoureux and Lastrapes (1990b)). Furthermore, the prediction power of a generic GARCH is very poor when structural changes do exist (Hamilton and Susmel (1994)). Thirdly, the generic GARCH model does not consider market asymmetry.

Time series models with nonlinearity include a very broad class of models which could

[1] For detail reviews of these types of model and their applications in financial analyses, see Bera and Higgins (1995), Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994) and Diebold and Lopez (1995).

be estimated either parametrically or non-parametrically. Tong (1990) and Teräsvirta, Tjstheim, and Granger (1994), both provide a comprehensive coverage of the most popular models in this class. Among the numerous parametric models, the regime-switching (R-S) model of Hamilton (1989) has received extensive attention in the econometric literature. By assuming the variable of interest, $X_t$, follows an AR(1) process, we can write down a simple regime-switching model as follows:

$$X_t = \alpha_t X_{t-1} + \varepsilon_t, \; t = 0, 1, 2, \ldots \tag{1.4}$$

where $\{\alpha_t\}$ is a Markov chain which is irreducible and aperiodic and has a finite state space consisting of $k$ states (regimes), $s_1, \ldots, s_k$. $\{\varepsilon_t\}$ is a sequence of *i.i.d.* random variables with zero mean and constant variance $\sigma^2$. To describe the transition between regimes, we need a transition probability matrix $\mathrm{P}(p_{ij})$, where $p_{ij} = P(\alpha_t = s_j \mid \alpha_{t-1} = s_i)$, $i,\, j = 1, \ldots, k$. This type of model has the advantage of allowing the interaction between the data and the Markov chain to endogenously determine the state of the world. It has been applied successfully to determine structural breaks in long horizons, such as applications in estimating business cycles (Hamilton (1989), Durland and McCurdy (1994), Filardo (1994, 1998)), in fitting weekly and monthly Treasury bill excess returns (Cai (1994), Gray (1996)), and in fitting weekly stock returns (Hamilton and Susmel (1994)). However, a clear-cut major structural break is not likely to happen in a shorter horizon such as an intra-day period. This renders the application of the R-S model to intra-day analysis problematic. Furthermore, the estimation of the R-S type of model is usually computationally intensive and is quite sensitive to the specification of the model, especially the specification of the transition probability function of the Markov process.

Threshold autoregressive (TAR) models of Tong (1983) provide an alternative in the parametric class of nonlinear models. A time series $X_t$ is a TAR model if it has the following functional form

$$X_t = \phi_0^i + \phi_1^i X_{t-1} + \ldots + \phi_p^i X_{t-p} + \varepsilon_t^i, \text{ if } \theta_{t-d} \in L_i,\; i = 1, 2, \ldots, k \tag{1.5}$$

where $L_i$ form a mutually exclusive partition of the real line in the sense that $\cup_{i=1}^k L_i = R$, and $L_i \cap L_j = \phi$, if $i \neq j$; $p$ is the $AR$ order; $\theta$ is the decision rule variable; $d$ is the delay parameter; $k$ is the threshold parameter; and $\{\varepsilon_t^i\}$ is a sequence of *i.i.d.* random variables with zero mean and variance $\sigma_i^2$, $\{\varepsilon_t^i\}$ and $\{\varepsilon_t^j\}$ are independent whenever $i \neq j$. This type of model exogenously sorts returns into different regimes according to some threshold, usually a function of the lagged dependent variable ($\theta_{t-d} = X_{t-d}$) and is thus termed self-exciting TAR (SETAR) model. Estimating TAR models are relatively easier than estimating the R-S model. TAR models are also capable of capturing time–irreversibility, asymmetric limit cycle and jump phenomenon. The major criticism of TAR models comes from the fact that researchers rarely know which state of the world they are currently in, which hinders the application of the TAR-type of model in economic time series. However, we think that the criticism of the TAR model is not always justified, especially when we have exogenously observable information to determine the threshold.

To sum up, extensions of the mixed-jump diffusion model and the TAR model both allow us to easily accommodate trading information into the model and are capable of capturing time dependency, market asymmetry, and nonlinearity existing in financial data. In Chapter 3 and Chapter 4, we shall extend these two models and apply them in studying high frequency stock return data.

## 1.3 Informational Proxies

Informational proxies adopted in the literature can be broadly categorized into two types, public information and private information. Examples of public information are scheduled macroeconomic news releases, and announcements of survey statistics. The impact of these publicly released news on interest rates, foreign exchange rates, foreign exchange rate futures, and stock returns has been examined by DeGennaro and Shrieves (1997), Ederington and Lee (1993), Goodhart *et al.* (1993), Hakkio and Pearce (1985), Harvey and Huang (1991), and Payne (1997). This literature shows that intraday volatility is significantly reduced when announcement effects of public information are included in the estimation.[2]

Unlike public information, private information by its very nature is not directly observable. However, we do know that private information invariably affects trading decisions.

[2]See Payne (1997) for a survey of a list of public information variables examined in this literature, and the associated announcement effects of these variables.

The observed trading activities are therefore the most suitable and popular proxies for private information, see Admati and Pfleiderer (1988), Glosten and Milgrom (1985), and Kyle (1985) for theoretical arguments.[3] Empirically, there is also a long history of using trading variables to explain returns/volatility dynamics at lower frequencies. The candidates include, trading volume in Clark (1973), the number of trades, and average trading volume in Jones, Kaul, and Lipson (1994), the number of quote arrivals in Bollerslev and Domowitz (1993), and the executed order imbalance in Locke and Sayers, (1993). In this thesis, we will only examine private information. Our reasons are twofold. First of all, public news announcements do not arrive at the same frequency as the high frequency transactional data we study in this thesis. Secondly, trading activities ultimately incorporate both public and private information.

Trading volume is, by far, the most often used informational proxy in the empirical study of stock returns. Ever since Clark (1973), trading volume has been used to test the mixture of distributions hypothesis, see Harris (1987); to test the price-volume relationship, see Gallant *et al.* (1992);[4] and to examine stock returns volatility, see Andersen (1996) and Lamoureux *et al.* (1994).

However, in a recent paper, Jones, Kaul, and Lipson (1994) show that trading volume has no informational content beyond that contained in the number of trades. The use of number of trades as the informational proxy dates back to Osborne (1959), who modified Bachelier's random walk model by incorporating a diffusion process into the evolution of stock prices, with an instantaneous variance dependent on the number of trades sampled from a uniform distribution. The uniform distribution assumption on the number of trades is however dubious, because transaction time intervals are certainly not uniformly distributed, see Oldfield *et al.* (1977). Recently, several researchers have revitalized the use of the number of trades and have given empirical support for using them as alternative informational proxies.[5] Marsh and Rock (1986) find that the net number of trades (number of seller-initiated minus buyer-initiated trades) explains as much as does the net volume.

[3]For example, market makers with asymmetric information will adjust their inventories before the news is announced.

[4]See also Karpoff (1987) for a survey of previous studies on the price - trading volume relationship.

[5]This is probably due to the availability of data. As pointed out by Jones, Kaul, and Lipson

Geman and Ané (1996) demonstrate that the moments of the time change needed to induce returns normality match the moments of the number of trades for the S&P 500 one minute returns. Madan and Chang (1997) propose a variance gamma stock price process and confirm that normality is attained in the trade-based measure of time. All of this new evidence indicates that the number of trades could be a better instrument for the non-quantifiable information than trading volume.

In addition to trading volume and number of trades, average trading volume (trading volume divided by the number of trades) is also used in the empirical analysis of stock returns. Jones, Kaul, and Lipson (1994) actually use average trading volume, instead of total trading volume, in their comparison of the explanatory power of different information proxies. Their justification comes from the observation that both the number of trades and the average trading volume are highly correlated with the total trading volume; however, there is little correlation between the number of trades and average trading volume. In other words, the number of trades and the average trading volume seem to contain different information.

Quote changes are the number of times the valid market quoted prices changed throughout the day for a certain security. Although Bollerslev and Domowitz (1993) find that market activity, as measured by the number of quote arrivals, has no statistically significant effect on returns volatility, Smaby (1995) and Takezawa (1995) suggest that the number of quotes is positively and significantly related to the intra-day volatility of foreign exchange rates in their recent studies.

An equivalent measure of the number of quote changes is the number of price changes for the trade data set. Again, this variable has not often been used as a proxy for information arrival, possibly due its unavailability. Both the number of quote changes and the number of price changes seem to be intuitively good instruments for the discrete price-jumps often observed in equity markets. In addition to the aforementioned information variables, Locke and Sayers (1993) have also examined the impact of executed order imbalance in reducing

---

(1994), although trading volume for the NASDAQ securities have been available for many years, historical data on the number of transactions were not available until recent years.

volatility persistence.

All of these variables are observable and have empirical implication of a random rate of flow of information. In this thesis, we only examine the performance of three information variables, namely trading volume and the number trades, and the number of price changes.

## 1.4 Concluding Remarks

In this chapter, we survey available models for modelling the returns series and concentrate on those models that enable us to incorporate information variables and to examine their usefulness in explaining the returns dynamics. It appears that the mixed jump-diffusion model provides a natural framework for us to study the time dependency of financial time series with exogenous trading information. The threshold autoregressive (TAR) model provides an alternative to study the nonlinearity existing in financial time series. It is capable of capturing structural changes and it is easier to implement than the regime-switching models. We shall rely on the mixed jump-diffusion model and the TAR model in our modelling.

Regarding the information contained in the trading activities, trading volume, the number trades, and the number of price changes seem to be the most popular proxies for private information. These are the three informational variables that we examine in this thesis.

# Chapter 2

# High Frequency One-Minute Data

## 2.1 Introduction

Applied research employing high frequency data has been an attractive and challenging topic in the area of Financial Econometrics since the ever-increasing availability of good quality tick-by-tick data. This can be seen by the amount of research devoted to the two international seminars on high-frequency data organized by the Olsen and Associates. It has a rather short history, although related earlier papers exist, see, for example, Roll (1984). Due to the availability of data sets, researchers are able to uncover more interesting features of asset dynamics at intraday frequency in recent years. Goodhart and O'Hara (1997) review a large literature which contains the availability of databases, statistical properties, problems and difficulties associated with high frequency data. Unlike daily or weekly financial time series, the high frequency data studied in this thesis has several distinguished features that deserve our attention. Therefore, we devote this chapter to the detail construction of the 1-minute data, and examine their statistical properties.

To facilitate the analysis, we select two representative stocks, IBM and INTEL, from the New York stock exchange, and British Telecom from the London Stock Exchange. We find that there exists a large significant negative first-order autocorrelation in the one-minute data. Since significant negative first-order autocorrelation typically does not exist in any returns series with frequency lower than one day, it differentiates high frequency data that we use from the lower frequency data. In the following chapters, we will try to model this stylized fact. In addition to trading variables, the British Telecom data from the

London Stock Exchange also contain extra buy/sell information on every single transaction. Based on a preliminary analysis, we find that different types of buy/sell trade do not arrive independently and could be potentially useful in modelling the returns series.

## 2.2 Data Construction

Data used in this paper come from two sources. The US data are extracted from the January 1994 Trade and Quote (TAQ) database, and the British data are extracted from the August 1994 London Stock Exchange Transaction and Quotation Database (LSETQD).[1]

Among the numerous stocks in TAQ, we only choose two frequently traded stocks, namely IBM and INTEL. IBM has been employed in several intraday trading analyses, for example Engle (1996) and Engle and Russell (1994). INTEL had the highest total trading volume among all available stocks. In January 1994, there were 21 trading days, which are treated separately in the following study. Only those transactions that occurred between 9:30 a.m. to 4:00 p.m. are extracted, since the exchanges where IBM and INTEL were mostly traded, NYSE and NASDAQ respectively, were open during that period.[2] We ended up with 13,095 observations of IBM stock traded on NYSE, and 72,831 observations of INTEL traded on NSADAQ. Variables recorded for each observation include: a time stamp, a traded price and the associated trading volume (shares).

British Telecom (BT) is one of the major LSE securities that constitutes the FTSE 100 index. The data set includes trade-by-trade information of BT from 8:30a.m. to 4:30 p.m., during the period of August 1994. It contains 22 trading dates, and 18,116 observations. In addition to the types of information contained in the US data, the LSETQD data set also contains buyer's and seller's identifiers being one of the following: market makers (M), inter-dealer brokers (I), exchange members who are not registered as market makers (P),

[1] The TAQ database is produced monthly by New York Stock Exchange (NYSE). This database contains virtually every trade and every quote of every stock traded on major American stock exchanges, like NYSE, AMEX, NASDAQ,..., etc.

[2] Although, there are some transactions that occurred before 9:30 a.m. and some occurred after 4:00 p.m., the percentage of these exceptions are quite small. Therefore, we decided to delete these observations.

brokers (A), and direct customers (N).[3] Because some of the prices may simply reflect large measurement errors, we have filtered out trades with anomalous prices using the following rule: any trades with prices that deviate from the adjacent price by 7 pence, roughly 4.5 times the average bid-ask spread, were deleted. After applying this filter, we ended up with 18,078 observations.[4] In this subset, we categorize trade types into the following buyer-seller combinations:

1. MM: inter market makers trades;
2. MI & IM: trades between market makers and inter-dealer brokers;
3. MP & PM: trades between market makers and non-market maker exchange members;
4. MA & AM: trades between market makers and brokers;
5. MN & NM: trades between market makers and direct customers.

where for each $XY$ combination, $X, Y \in \{M, I, P, A, N\}$, $X$ denotes the buyer, and $Y$ denotes the seller. For example, MA refers to the case where a market maker buys shares from a broker.

Since the function of inter-dealer brokers is just to facilitate anonymous trades between market makers, a pair of MI and IM trades should actually be treated as one MM trade. Otherwise, we will double-count the frequencies and quantities of transactions. After converting 467 MI & IM trades to 216 MM trades, our sample size is reduced to 17,827.[5] Table 2.1 presents the frequencies for the various types of trades and the associated average trading quantity and value. We show, in Table 2.2, that about 95% of the trades came from brokers and direct customers dealing with market makers, we refer to these as *external trades*. The rest of the trades are defined as *internal trades*. On average, the quantity and

[3]Inter-dealer brokers exist to facilitate anonymous trading between market makers. This helps market makers to unwind their excess inventory, especially when they are hit by trades with large volume. See Hansch (1997) for a good description of the market structure of the LSE. See also Board and Sutcliffe (1995) for the impact of the market structure.

[4]Since there are only two observations that involve transactions between two brokers, we deleted them from the rest of the analysis as well.

[5]See Appendix A for a detailed description of the conversion scheme that we use.

value for the external trades are much smaller than those of internal trades.

Table 2.1: Distribution of Trade Types

| Types | Frequency | Quantity | Value |
|---|---|---|---|
| MM | 438 | 138493 | 53001357 |
| MA | 13472 | 859 | 326426 |
| MN | 305 | 37744 | 14349387 |
| MP | 14 | 81571 | 30975893 |
| AM | 3131 | 6417 | 2451340 |
| NM | 431 | 86438 | 33121842 |
| PM | 36 | 42697 | 16315273 |

MM: Inter-Market Maker trades;
MA: Brokers SELL to Market Makers;
MN: Customers SELL to Market Makers;
AM: Brokers BUY from Market Makers;
NM: Customers BUY from Market Makers;
MP :Non-MM Exchange Member SELL to Market Maker;
PM : Non-MM Exchange Member BUY from Market Maker.

Table 2.2: Internal vs. External Trades

| | Frequency | Quantity | Value |
|---|---|---|---|
| Internal | 488 | 129793 | 49663129 |
| External | 17339 | 4639 | 1772007 |

Internal trades: MM, MP, PM;
External trades: MA, AM MN, NM.

To facilitate our analysis in fixed time intervals, we sample these tick-by-tick data every minute. Our sampling procedure, adopted by Locke and Sayers (1993, p.17), is described as follows,

1. Select the first recorded trade as the observation for each minute.

2. Retain the previous trade information for those following minutes with no trades.

This yields roughly 390 observations per day for IBM and INTEL, and roughly 490 for BT. While sampling the 1-minute data, not only have we extracted the price series, we have also calculated total trading volume (TVol), total number of trades (N), and total number of price changes (NPC) for each one-minute interval.

## 2.3 Descriptive Analyses

We first examine the returns processes. For each stock and each day, we calculate various moments and test statistics of the returns data, these are reported in Tables 2.5, 2.6, and 2.7. The stock return concept used in this paper is the one-minute log-return, defined as the difference of the logarithm prices of two consecutive minutes. By examining these tables, we notice that most return processes are highly kurtotic and non-normally distributed. We compute the Bera-Jarque normality test statistics, which is asymptotically distributed as $\chi^2(2)$. This test is rejected in most cases for IBM and BT, and is rejected in every case for INTEL. Also, most returns processes are clearly not *i.i.d.*, judging from the BDS test statistics proposed by Brock, Dechert and Scheinkman (1986). Another salient feature of the returns series is that most of them have significant large negative first-order autocorrelation, as reported in Table 2.3. On average, the first-order autocorrelation is -0.484 for INTEL, -0.269 for IBM, and -0.428 for BT. Summary statistics of trading frequency are reported in Table 2.4. We do this to see if autocorrelation in the data is caused by non-synchronicity, see Lo and McKinlay (1990). We note, from Table 2.4, that INTEL has the lowest rate of no-trades whilst having the largest correlation as reported earlier. Therefore, Lo and McKinlay's nonsynchronous trading hypothesis does not seem to be able to explain the observed significant autocorrelation.

We provide summary statistics of the three trading variables in Tables 2.8 to 2.14. We observe that all three trading variables are positively skewed, highly kurtotic, and non-normal. Similar to return processes, we reject that these trading variables are *i.i.d.* - distributed in most cases. Both IBM and BT have smaller turnover rates than INTEL, and share a very similar distribution.

## 2.4 A Preliminary Analysis on the Buy/Sell Patterns

In this section, we analyse the buy/sell pattern of the BT data. Since market makers in the UK are allowed to delay reports of large trades (above 3 × Normal Market Size) for up

Table 2.3: First Lag Autocorrelation

| Date | INTEL | IBM | BT | Date |
|---|---|---|---|---|
| 01/03/94 | -0.485 | -0.399 | -0.429 | 08/01/94 |
| 01/04/94 | -0.464 | -0.302 | -0.346 | 08/02/94 |
| 01/05/94 | -0.538 | -0.255 | -0.482 | 08/03/94 |
| 01/06/94 | -0.452 | -0.283 | -0.437 | 08/04/94 |
| 01/07/94 | -0.476 | -0.272 | -0.467 | 08/05/94 |
| 01/10/94 | -0.491 | -0.237 | -0.406 | 08/08/94 |
| 01/11/94 | -0.462 | -0.271 | -0.397 | 08/09/94 |
| 01/12/94 | -0.467 | -0.316 | -0.445 | 08/10/94 |
| 01/13/94 | -0.467 | -0.236 | -0.424 | 08/11/94 |
| 01/14/94 | -0.527 | -0.292 | -0.292 | 08/12/94 |
| 01/17/94 | -0.477 | -0.227 | -0.479 | 08/15/94 |
| 01/18/94 | -0.577 | -0.248 | -0.406 | 08/16/94 |
| 01/19/94 | -0.497 | -0.287 | -0.414 | 08/17/94 |
| 01/20/94 | -0.514 | -0.309 | -0.359 | 08/18/94 |
| 01/21/94 | -0.465 | -0.275 | -0.474 | 08/19/94 |
| 01/24/94 | -0.309 | -0.554 | -0.565 | 08/22/94 |
| 01/25/94 | -0.524 | 0.030 | -0.358 | 08/23/94 |
| 01/26/94 | -0.474 | -0.105 | -0.387 | 08/24/94 |
| 01/27/94 | -0.481 | -0.203 | -0.480 | 08/25/94 |
| 01/28/94 | -0.536 | -0.330 | -0.460 | 08/26/94 |
| 01/31/94 | -0.488 | -0.274 | -0.458 | 08/30/94 |
| | | | -0.451 | 08/31/94 |
| Average | -0.484 | -0.269 | -0.428 | |

to 90 minutes to offset their inventory risk,[6] we suspect that very little information would be contained in the internal trades (MM, MP, and PM) and decide to exclude those trades from the rest of the analysis. For the rest of the transactions (MN, NM, MA, and AM), we construct two-way contingency tables to help us understand the information contained in this extra variable. States listed along the first column refer to what happens in period $t$, while those listed along the first row refer to what happens in period $t + 1$. Table 2.15 describes the observed frequencies of different types of transitions. The transitional probabilities are presented in Table 2.16. Since we are interested in whether different types of transactions arrive independently, we construct expected arrival frequencies in Table 2.17 under the assumption that different types of trades occur independently. To examine this null hypothesis, we calculate the contribution of each cell and test the hypothesis with a

[6]See Board and Sutcliffe (1995).

Table 2.4: Summary Statistics of the Trading Frequency

| | IBM | INTEL | BT |
|---|---|---|---|
| Minimum | 0.000 | 0.000 | 0.000 |
| 1st Quartile | 0.000 | 3.000 | 1.000 |
| Median | 1.000 | 6.000 | 1.000 |
| Mean | 1.596 | 8.689 | 1.609 |
| 3rd Quartile | 2.000 | 10.000 | 2.000 |
| Maximum | 14.000 | 170.000 | 18.000 |
| No Trade (%) | 28.32 | 1.98 | 22.93 |

$\chi^2$ statistics in Table 2.18. The null hypothesis is clearly rejected. We next reduce the types of trades to buyer-initiated trades (by aggregating NM and AM) and seller-initiated trades (by aggregating MN and MA), and report the results in Tables 2.19, 2.20, 2.21, and 2.22. Under this new categorization, we still reject the independence hypothesis. The above results indicate that buy/sell transitions do contain information.

## 2.5 Concluding Remarks

We devote this chapter to a detailed examination of the 1-minute data, and study their statistical properties. We find, from our constructed one-minute data, that neither the returns series nor the trading variables are independently, identically, and normally distributed. Furthermore, there also exists a large significant negative first-order autocorrelation in the one-minute data. Since significant negative first-order autocorrelation typically does not exist in any returns series with frequency lower than one day, it differentiates high frequency data that we use from the lower frequency data. In the following chapters, we will try to model this interesting stylized fact. In addition to trading variables, the British Telecom data from the London Stock Exchange also contain extra buy/sell information on every single transaction. In the preliminary analysis, we find that different types of buy/sell trades do not arrive independently and could be potentially helpful in explaining the returns autocorrelation.

Table 2.5: Descriptive Statistics of Log-Returns (INTEL)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 389 | -4.1632E-005 | 8.6799E-006 | -0.074 | 0.261 | 1.5 | 8.47 |
| 01/04/94 | 389 | 1.0284E-004 | 7.9931E-006 | 0.118 | 1.326 | 29.4 | 7.55 |
| 01/05/94 | 389 | 4.0012E-005 | 7.6488E-006 | 0.121 | 0.997 | 17.1 | 8.99 |
| 01/06/94 | 389 | -4.0012E-005 | 8.5810E-006 | 0.096 | 0.656 | 7.6 | 7.89 |
| 01/07/94 | 389 | 9.9267E-005 | 8.6340E-006 | 0.030 | 1.665 | 45.0 | 6.58 |
| 01/10/94 | 389 | 2.9048E-005 | 8.6570E-006 | -0.267 | 2.862 | 137.4 | 7.81 |
| 01/11/94 | 389 | -4.8095E-006 | 8.5162E-006 | 0.116 | 0.442 | 4.0 | 6.84 |
| 01/12/94 | 389 | 0.0000E+000 | 9.9269E-006 | -0.180 | 4.449 | 322.9 | 7.96 |
| 01/13/94 | 389 | 0.0000E+000 | 6.0198E-006 | -0.000 | -0.435 | 3.1 | 7.62 |
| 01/14/94 | 389 | -9.3650E-006 | 9.1142E-006 | 0.018 | 0.890 | 12.9 | 6.88 |
| 01/17/94 | 389 | -8.6177E-005 | 9.8798E-006 | 0.196 | 4.878 | 388.2 | 8.44 |
| 01/18/94 | 389 | 3.3667E-005 | 1.4521E-005 | 0.281 | 8.987 | 1314.1 | 8.82 |
| 01/19/94 | 389 | -1.0221E-005 | 2.0207E-005 | -0.354 | 17.402 | 4916.5 | 8.75 |
| 01/20/94 | 389 | 4.0644E-005 | 1.1315E-005 | 0.023 | 2.334 | 88.3 | 6.91 |
| 01/21/94 | 389 | 2.9834E-005 | 7.1557E-006 | -0.033 | 1.116 | 20.3 | 10.69 |
| 01/24/94 | 389 | 1.9775E-005 | 1.3870E-005 | 0.018 | 27.735 | 12467.9 | 7.93 |
| 01/25/94 | 389 | -9.8683E-006 | 1.1452E-005 | 0.036 | 36.546 | 21647.6 | 8.04 |
| 01/26/94 | 389 | -5.0554E-006 | 9.2477E-006 | -0.027 | 0.599 | 5.9 | 8.90 |
| 01/27/94 | 389 | 1.0061E-005 | 8.1759E-006 | -0.078 | 0.243 | 1.4 | 7.74 |
| 01/28/94 | 389 | 3.9857E-005 | 8.5613E-006 | 0.041 | -0.001 | 0.1 | 9.43 |
| 01/31/94 | 389 | 0.0000E+000 | 7.1059E-006 | 0.023 | -0.373 | 2.3 | 8.88 |

• N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis; BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99. BDS = Brock-Dechert-Scheinkman $\sim N(0,1)_{asy}$, Embedding Dimension = 3, Epsilon = Standard Deviation/Spread

Table 2.6: Descriptive Statistics of Log-Returns (IBM)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 388 | 3.3764E-005 | 1.3221E-006 | 0.229 | 2.210 | 82.4 | 9.65 |
| 01/04/94 | 387 | 2.7636E-005 | 8.6563E-007 | -0.123 | 3.260 | 172.3 | 6.01 |
| 01/05/94 | 387 | 2.7229E-005 | 7.3054E-007 | 0.331 | 4.317 | 307.6 | 3.67 |
| 01/06/94 | 385 | -5.4684E-005 | 1.0495E-006 | -0.206 | 1.859 | 58.2 | 4.06 |
| 01/07/94 | 388 | 2.1935E-005 | 8.2120E-007 | 0.061 | 2.590 | 108.7 | 3.99 |
| 01/10/94 | 388 | 2.1796E-005 | 8.8587E-007 | 0.048 | 2.146 | 74.6 | 1.20 |
| 01/11/94 | 389 | -2.1832E-005 | 8.6346E-007 | -0.053 | 2.299 | 85.9 | 3.63 |
| 01/12/94 | 387 | -2.2085E-005 | 9.0234E-007 | -0.048 | 2.133 | 73.5 | 3.39 |
| 01/13/94 | 389 | 2.1972E-005 | 7.9455E-007 | -0.143 | 3.809 | 236.5 | 4.56 |
| 01/14/94 | 389 | -1.0962E-005 | 7.1724E-007 | 0.197 | 4.554 | 338.6 | 5.89 |
| 01/17/94 | 388 | -5.0047E-005 | 1.2484E-006 | -0.600 | 3.788 | 255.3 | 4.28 |
| 01/18/94 | 389 | -1.6857E-005 | 1.1145E-006 | -0.164 | 2.000 | 66.6 | 7.29 |
| 01/19/94 | 388 | -4.0048E-005 | 9.7589E-007 | -0.084 | 2.075 | 70.0 | 2.35 |
| 01/20/94 | 389 | -2.8917E-005 | 1.1908E-006 | -0.169 | 1.856 | 57.7 | 4.29 |
| 01/21/94 | 386 | 5.8679E-006 | 9.1534E-007 | 0.016 | 2.644 | 112.4 | 2.90 |
| 01/24/94 | 387 | 1.2993E-004 | 2.7850E-005 | -0.146 | 25.620 | 10585.2 | 9.93 |
| 01/25/94 | 388 | -2.1982E-005 | 4.3778E-006 | -2.478 | 20.380 | 7112.0 | 2.97 |
| 01/26/94 | 382 | -1.0802E-004 | 1.6862E-006 | -0.352 | 2.498 | 107.2 | 3.09 |
| 01/27/94 | 389 | 1.6968E-005 | 1.1024E-006 | -0.119 | 3.206 | 167.5 | 3.44 |
| 01/28/94 | 388 | -2.2266E-005 | 9.9185E-007 | -0.039 | 1.768 | 50.6 | 5.75 |
| 01/31/94 | 389 | -5.0462E-005 | 1.0457E-006 | -0.077 | 1.606 | 42.2 | 2.31 |

* N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis;
BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99.
BDS = Brock-Dechert-Scheinkman $\sim N(0,1)_{asy}$, Embedding Dimension = 3,
Epsilon = Standard Deviation/Spread

Table 2.7: Descriptive Statistics of Log-Returns (BT)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 08/01/94 | 483 | 5.6184E-006 | 4.9268E-006 | -0.038 | 1.610 | 52.3 | 8.99 |
| 08/02/94 | 495 | 2.1492E-005 | 4.1006E-006 | -0.037 | 1.948 | 78.4 | 5.39 |
| 08/03/94 | 487 | 0.0000E+000 | 1.1116E-005 | 0.033 | 1.643 | 54.9 | 11.62 |
| 08/04/94 | 484 | -2.4744E-005 | 4.0706E-006 | 0.406 | 4.045 | 343.2 | 6.79 |
| 08/05/94 | 486 | 3.8359E-005 | 5.8192E-006 | 0.029 | 1.343 | 36.6 | 8.58 |
| 08/08/94 | 482 | 1.6315E-005 | 5.2781E-006 | -0.177 | 4.345 | 381.8 | 7.60 |
| 08/09/94 | 489 | -2.1470E-005 | 3.9849E-006 | -0.046 | 2.478 | 125.2 | 7.69 |
| 08/10/94 | 491 | 2.6745E-006 | 5.7969E-006 | 0.037 | 1.803 | 66.6 | 7.92 |
| 08/11/94 | 486 | -1.6309E-005 | 4.9622E-006 | -0.033 | 1.394 | 39.4 | 7.16 |
| 08/12/94 | 485 | 0.0000E+000 | 3.8430E-006 | 0.006 | 3.305 | 220.7 | 3.93 |
| 08/15/94 | 484 | -1.9143E-005 | 5.0618E-006 | 0.059 | 1.013 | 21.0 | 7.99 |
| 08/16/94 | 486 | 8.2140E-006 | 6.0304E-006 | 0.058 | 0.766 | 12.2 | 5.99 |
| 08/17/94 | 501 | 0.0000E+000 | 4.4222E-006 | 0.109 | 0.826 | 15.2 | 5.29 |
| 08/18/94 | 488 | 1.8810E-005 | 7.0930E-006 | 0.049 | 8.748 | 1556.1 | 9.02 |
| 08/19/94 | 490 | 1.6006E-005 | 3.5818E-006 | 0.110 | 2.705 | 150.4 | 8.41 |
| 08/22/94 | 492 | -3.4608E-005 | 5.9356E-006 | 0.016 | 2.747 | 154.8 | 10.48 |
| 08/23/94 | 487 | -5.4328E-006 | 6.9500E-006 | -0.098 | 0.888 | 16.8 | 5.02 |
| 08/24/94 | 489 | 4.2829E-005 | 5.1671E-006 | 0.021 | 0.840 | 14.4 | 5.76 |
| 08/25/94 | 500 | -1.5365E-005 | 3.9066E-006 | 0.131 | 2.513 | 133.0 | 8.23 |
| 08/26/94 | 490 | 3.4210E-005 | 4.3316E-006 | 0.160 | 1.540 | 50.5 | 9.14 |
| 08/30/94 | 492 | -1.5575E-005 | 5.7453E-006 | 0.003 | 1.038 | 22.1 | 9.20 |
| 08/31/94 | 485 | 5.2936E-006 | 3.8824E-006 | -0.041 | 2.379 | 114.5 | 7.34 |

- N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis; BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99. BDS = Brock-Dechert-Scheinkman $\sim \mathcal{N}(0,1)_{asy}$, Embedding Dimension = 3, Epsilon = Standard Deviation/Spread

Table 2.8: Descriptive Statistics of the Total Trading Volume (INTEL)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 390 | 7257 | 1.223E+008 | 3.644 | 21.873 | 8637.6 | 7.13 |
| 01/04/94 | 390 | 9778 | 2.395E+008 | 3.648 | 16.515 | 5297.2 | 5.19 |
| 01/05/94 | 390 | 11397 | 3.374E+008 | 3.812 | 18.175 | 6312.7 | 3.33 |
| 01/06/94 | 390 | 7621 | 1.069E+008 | 2.665 | 9.259 | 1854.9 | 6.79 |
| 01/07/94 | 390 | 15337 | 5.906E+008 | 3.342 | 13.695 | 3774.0 | 9.36 |
| 01/10/94 | 390 | 10946 | 2.030E+008 | 2.912 | 11.483 | 2693.7 | 4.90 |
| 01/11/94 | 390 | 7537 | 1.382E+008 | 3.436 | 14.815 | 4334.1 | 5.15 |
| 01/12/94 | 390 | 15224 | 5.142E+008 | 2.838 | 9.350 | 1944.0 | 10.87 |
| 01/13/94 | 390 | 8333 | 1.412E+008 | 2.433 | 6.476 | 1066.4 | 7.55 |
| 01/14/94 | 390 | 9849 | 1.944E+008 | 3.253 | 14.595 | 4149.4 | 7.88 |
| 01/17/94 | 390 | 12826 | 3.947E+008 | 4.077 | 25.435 | 11593.6 | 4.04 |
| 01/18/94 | 390 | 18317 | 9.682E+008 | 6.341 | 58.890 | 58968.3 | 3.60 |
| 01/19/94 | 390 | 62260 | 3.883E+009 | 1.855 | 4.394 | 537.4 | 11.87 |
| 01/20/94 | 390 | 20868 | 1.624E+009 | 10.006 | 146.424 | 354906.8 | 8.64 |
| 01/21/94 | 390 | 22353 | 1.183E+009 | 3.627 | 19.061 | 6759.0 | 3.13 |
| 01/24/94 | 390 | 14056 | 5.656E+008 | 3.912 | 20.940 | 8120.2 | 5.41 |
| 01/25/94 | 390 | 10155 | 3.237E+008 | 4.593 | 28.431 | 14506.7 | 3.25 |
| 01/26/94 | 390 | 18585 | 9.252E+008 | 3.815 | 20.081 | 7499.1 | 7.06 |
| 01/27/94 | 390 | 14278 | 3.513E+008 | 2.237 | 5.746 | 861.6 | 5.627 |
| 01/28/94 | 390 | 9490 | 2.760E+008 | 3.104 | 11.694 | 2848.2 | 8.86 |
| 01/31/94 | 390 | 8823 | 3.594E+008 | 9.540 | 134.074 | 298024.3 | 5.95 |

* N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis; BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99. BDS = Brock-Dechert-Scheinkman $\sim N(0,1)_{asy}$, Embedding Dimension = 3, Epsilon = Standard Deviation/Spread

Table 2.9: Descriptive Statistics of the Total Trading Volume (IBM)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 389 | 2781.0 | 3.858E+007 | 6.289 | 54.822 | 51278.1 | 1.47 |
| 01/04/94 | 388 | 3450.3 | 6.309E+007 | 7.724 | 93.777 | 146029.0 | 3.65 |
| 01/05/94 | 388 | 5356.4 | 1.526E+008 | 6.238 | 52.767 | 47530.5 | 5.52 |
| 01/06/94 | 386 | 4816.8 | 2.682E+008 | 13.836 | 230.536 | 867097.6 | 3.24 |
| 01/07/94 | 389 | 3085.9 | 5.445E+007 | 7.091 | 69.460 | 81459.2 | 2.60 |
| 01/10/94 | 389 | 3654.2 | 5.468E+007 | 3.889 | 18.455 | 6500.7 | 0.51 |
| 01/11/94 | 390 | 3145.6 | 9.841E+007 | 7.440 | 64.934 | 72113.6 | 1.52 |
| 01/12/94 | 388 | 3162.6 | 5.333E+007 | 4.968 | 31.639 | 17778.7 | 2.66 |
| 01/13/94 | 390 | 3854.6 | 6.716E+007 | 4.136 | 24.657 | 10991.3 | 6.29 |
| 01/14/94 | 390 | 3101.8 | 5.128E+007 | 6.687 | 66.937 | 75715.5 | 2.21 |
| 01/17/94 | 389 | 2930.3 | 4.088E+007 | 3.894 | 19.024 | 6848.8 | 4.06 |
| 01/18/94 | 390 | 4350.8 | 7.373E+007 | 3.593 | 16.041 | 5020.3 | 4.46 |
| 01/19/94 | 389 | 3928.5 | 7.836E+007 | 4.496 | 24.729 | 11222.5 | 1.38 |
| 01/20/94 | 390 | 4535.1 | 7.614E+007 | 3.632 | 19.634 | 7121.5 | 0.52 |
| 01/21/94 | 387 | 5741.1 | 2.401E+008 | 9.786 | 134.918 | 299699.6 | 1.45 |
| 01/24/94 | 388 | 8797.4 | 2.418E+008 | 6.477 | 71.144 | 84540.7 | 3.85 |
| 01/25/94 | 389 | 22047.8 | 1.172E+009 | 5.040 | 43.368 | 32130.5 | 4.85 |
| 01/26/94 | 383 | 8929.5 | 5.049E+008 | 12.836 | 209.362 | 710007.0 | 0.72 |
| 01/27/94 | 390 | 4600.8 | 9.495E+007 | 6.040 | 59.584 | 60062.0 | 4.28 |
| 01/28/94 | 389 | 2800.3 | 5.344E+007 | 7.468 | 82.443 | 113782.7 | 1.40 |
| 01/31/94 | 390 | 3863.3 | 5.466E+007 | 4.067 | 23.864 | 10328.7 | 1.04 |

* N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis; BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99. BDS = Brock-Dechert-Scheinkman $\sim N(0,1)_{asy}$, Embedding Dimension = 3, Epsilon = Standard Deviation/Spread

Table 2.10: Descriptive Statistics of the Number of Trades (INTEL)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 390 | 5.6 | 31.1 | 3.427 | 18.411 | 6271.7 | 13.59 |
| 01/04/94 | 390 | 6.1 | 34.6 | 3.565 | 19.955 | 7296.8 | 13.30 |
| 01/05/94 | 390 | 7.2 | 62.6 | 5.135 | 38.847 | 26237.0 | 11.53 |
| 01/06/94 | 390 | 5.9 | 31.1 | 2.963 | 16.141 | 4804.5 | 13.44 |
| 01/07/94 | 390 | 9.8 | 174.6 | 6.211 | 61.325 | 63620.0 | 15.04 |
| 01/10/94 | 390 | 8.5 | 59.2 | 4.565 | 31.809 | 17796.9 | 7.96 |
| 01/11/94 | 390 | 6.1 | 30.9 | 3.255 | 16.362 | 5039.2 | 9.27 |
| 01/12/94 | 390 | 9.8 | 131.9 | 3.911 | 18.829 | 6755.5 | 15.67 |
| 01/13/94 | 390 | 5.4 | 19.9 | 2.148 | 6.437 | 973.3 | 12.97 |
| 01/14/94 | 390 | 7.4 | 60.1 | 4.589 | 32.046 | 18056.4 | 12.36 |
| 01/17/94 | 390 | 6.8 | 36.6 | 2.553 | 9.529 | 1898.9 | 13.49 |
| 01/18/94 | 390 | 9.2 | 50.5 | 2.199 | 8.540 | 1499.5 | 14.66 |
| 01/19/94 | 390 | 33.9 | 571.6 | 1.976 | 5.341 | 717.3 | 23.57 |
| 01/20/94 | 390 | 13.1 | 157.0 | 3.828 | 21.546 | 8496.3 | 18.09 |
| 01/21/94 | 390 | 9.5 | 69.6 | 4.554 | 34.008 | 20142.5 | 13.22 |
| 01/24/94 | 390 | 7.4 | 43.0 | 3.058 | 15.990 | 4762.7 | 13.27 |
| 01/25/94 | 390 | 5.6 | 25.3 | 2.868 | 11.437 | 2660.3 | 10.71 |
| 01/26/94 | 390 | 7.4 | 40.1 | 2.244 | 6.740 | 1065.4 | 12.84 |
| 01/27/94 | 390 | 6.6 | 31.0 | 2.177 | 7.164 | 1142.0 | 14.25 |
| 01/28/94 | 390 | 5.5 | 36.3 | 4.436 | 27.983 | 14003.7 | 13.97 |
| 01/31/94 | 390 | 5.8 | 55.8 | 8.253 | 97.587 | 159180.7 | 9.92 |

* N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis; BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99. BDS = Brock-Dechert-Scheinkman $\sim N(0,1)_{asy}$, Embedding Dimension = 3, Epsilon = Standard Deviation/Spread

Table 2.11: Descriptive Statistics of the Number of Trades (IBM)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 389 | 1.4 | 1.8 | 1.115 | 1.320 | 108.8 | 8.21 |
| 01/04/94 | 388 | 1.4 | 2.1 | 1.315 | 1.953 | 173.5 | 5.18 |
| 01/05/94 | 388 | 1.6 | 2.3 | 1.835 | 8.170 | 1296.7 | 4.01 |
| 01/06/94 | 386 | 1.6 | 2.3 | 1.040 | 0.815 | 80.2 | 6.67 |
| 01/07/94 | 389 | 1.3 | 1.5 | 0.918 | 0.419 | 57.5 | 2.21 |
| 01/10/94 | 389 | 1.5 | 2.0 | 1.020 | 0.979 | 82.9 | 2.31 |
| 01/11/94 | 390 | 1.2 | 1.6 | 1.159 | 1.553 | 126.5 | 2.31 |
| 01/12/94 | 388 | 1.3 | 1.7 | 1.267 | 2.049 | 171.7 | 5.02 |
| 01/13/94 | 390 | 1.3 | 1.8 | 1.381 | 2.008 | 189.4 | 6.84 |
| 01/14/94 | 390 | 1.2 | 1.7 | 1.243 | 1.542 | 139.1 | 5.48 |
| 01/17/94 | 389 | 1.3 | 2.0 | 1.290 | 1.367 | 138.2 | 4.12 |
| 01/18/94 | 390 | 1.6 | 2.2 | 1.060 | 1.144 | 94.2 | 4.84 |
| 01/19/94 | 389 | 1.4 | 1.9 | 1.386 | 2.781 | 250.0 | 1.36 |
| 01/20/94 | 390 | 1.5 | 1.9 | 1.360 | 2.535 | 224.6 | 4.51 |
| 01/21/94 | 387 | 1.5 | 2.2 | 1.292 | 2.246 | 189.1 | 5.03 |
| 01/24/94 | 388 | 2.6 | 4.9 | 1.329 | 2.465 | 212.4 | 7.41 |
| 01/25/94 | 389 | 3.7 | 6.8 | 0.774 | 0.306 | 40.4 | 6.94 |
| 01/26/94 | 383 | 2.1 | 3.4 | 1.131 | 1.407 | 113.2 | 5.15 |
| 01/27/94 | 390 | 1.4 | 1.9 | 1.166 | 1.184 | 111.1 | 3.16 |
| 01/28/94 | 389 | 1.3 | 1.9 | 1.407 | 2.504 | 230.0 | 5.42 |
| 01/31/94 | 390 | 1.5 | 1.6 | 1.035 | 1.605 | 111.5 | 0.40 |

• N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis;
BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99.
BDS = Brock-Dechert-Scheinkman $\sim N(0,1)_{asy}$, Embedding Dimension = 3,
Epsilon = Standard Deviation/Spread

Table 2.12: Descriptive Statistics of the Number of Trades (BT)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 08/01/94 | 484 | 1.9 | 2.4 | 1.101 | 2.110 | 187.5 | 2.14 |
| 08/02/94 | 496 | 1.7 | 2.1 | 0.926 | 0.696 | 80.9 | 1.03 |
| 08/03/94 | 488 | 1.9 | 2.4 | 1.260 | 2.633 | 270.1 | 0.31 |
| 08/04/94 | 485 | 1.5 | 1.8 | 1.274 | 2.342 | 242.1 | 3.20 |
| 08/05/94 | 487 | 1.5 | 1.7 | 0.968 | 1.151 | 103.0 | 0.08 |
| 08/08/94 | 483 | 1.9 | 3.1 | 2.452 | 14.974 | 4996.1 | 3.11 |
| 08/09/94 | 490 | 1.6 | 1.8 | 0.819 | 0.460 | 59.0 | 0.72 |
| 08/10/94 | 492 | 1.3 | 1.6 | 1.318 | 2.790 | 302.0 | 0.26 |
| 08/11/94 | 487 | 1.4 | 1.6 | 1.059 | 1.456 | 134.0 | 1.29 |
| 08/12/94 | 486 | 1.3 | 1.5 | 1.162 | 1.884 | 181.2 | 1.88 |
| 08/15/94 | 485 | 1.6 | 1.8 | 0.823 | 0.449 | 58.8 | 0.99 |
| 08/16/94 | 487 | 1.4 | 1.6 | 0.807 | 0.192 | 53.5 | 4.21 |
| 08/17/94 | 502 | 1.7 | 2.1 | 0.793 | 0.319 | 54.7 | 1.06 |
| 08/18/94 | 489 | 1.7 | 2.0 | 0.818 | 0.556 | 60.8 | 1.01 |
| 08/19/94 | 491 | 1.6 | 1.9 | 0.999 | 0.997 | 102.0 | 1.48 |
| 08/22/94 | 493 | 1.7 | 1.9 | 0.866 | 0.921 | 79.0 | 1.38 |
| 08/23/94 | 488 | 1.3 | 1.4 | 0.884 | 0.468 | 68.0 | 1.45 |
| 08/24/94 | 490 | 1.6 | 1.7 | 0.867 | 1.446 | 104.0 | 1.74 |
| 08/25/94 | 501 | 2.0 | 2.2 | 1.114 | 2.438 | 227.7 | 0.35 |
| 08/26/94 | 491 | 1.8 | 2.3 | 1.110 | 1.685 | 159.0 | 2.16 |
| 08/30/94 | 493 | 2.1 | 2.5 | 0.983 | 1.378 | 118.4 | 1.25 |
| 08/31/94 | 486 | 1.7 | 1.7 | 0.694 | -0.074 | 39.2 | 0.27 |

* N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis;
BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99.
BDS = Brock-Dechert-Scheinkman $\sim \mathcal{N}(0,1)_{\text{asy}}$, Embedding Dimension = 3,
Epsilon = Standard Deviation/Spread

Table 2.13: Descriptive Statistics of the Number of Price Changes (INTEL)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 390 | 3.0103 | 7.2081 | 1.8741 | 6.1987 | 852.7 | 10.42 |
| 01/04/94 | 390 | 3.1231 | 6.8691 | 1.7950 | 4.3085 | 511.1 | 4.76 |
| 01/05/94 | 390 | 3.6564 | 10.7762 | 2.1332 | 6.8053 | 1048.4 | 9.89 |
| 01/06/94 | 390 | 3.2026 | 8.6504 | 1.8694 | 5.1254 | 654.0 | 9.50 |
| 01/07/94 | 390 | 4.6513 | 18.9835 | 1.8770 | 5.0057 | 636.2 | 1.62 |
| 01/10/94 | 390 | 4.4718 | 10.1110 | 1.5547 | 3.7506 | 385.7 | 6.31 |
| 01/11/94 | 390 | 3.2103 | 9.3696 | 2.7147 | 11.4268 | 2600.8 | 8.37 |
| 01/12/94 | 390 | 5.3821 | 38.4012 | 3.4377 | 14.5498 | 4208.2 | 16.20 |
| 01/13/94 | 390 | 3.0923 | 7.5647 | 2.1491 | 7.0446 | 1106.6 | 10.29 |
| 01/14/94 | 390 | 4.0282 | 12.1560 | 2.4341 | 10.0782 | 2035.6 | 8.65 |
| 01/17/94 | 390 | 3.5103 | 10.7441 | 2.4117 | 9.8556 | 1956.5 | 8.80 |
| 01/18/94 | 390 | 4.8179 | 12.9668 | 1.8517 | 5.6909 | 749.1 | 10.62 |
| 01/19/94 | 390 | 17.4667 | 144.3575 | 1.9079 | 5.2997 | 693.0 | 21.77 |
| 01/20/94 | 390 | 6.7846 | 60.7915 | 5.4213 | 44.4658 | 34040.0 | 19.55 |
| 01/21/94 | 390 | 4.9538 | 19.9824 | 3.6961 | 22.7886 | 9326.9 | 9.77 |
| 01/24/94 | 390 | 4.1231 | 11.3833 | 1.9455 | 6.7445 | 985.2 | 9.57 |
| 01/25/94 | 390 | 2.9077 | 6.3771 | 1.9183 | 5.5525 | 740.2 | 5.99 |
| 01/26/94 | 390 | 3.9333 | 12.4480 | 2.1248 | 6.5732 | 995.6 | 9.12 |
| 01/27/94 | 390 | 3.3846 | 7.7334 | 1.4063 | 2.7355 | 250.1 | 7.20 |
| 01/28/94 | 390 | 2.7769 | 6.1481 | 2.0716 | 7.7779 | 1262.0 | 10.79 |
| 01/31/94 | 390 | 2.9000 | 7.5298 | 3.5401 | 23.4108 | 9720.7 | 4.14 |

• N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis; BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99. BDS = Brock-Dechert-Scheinkman $\sim N(0,1)_{asy}$, Embedding Dimension = 3, Epsilon = Standard Deviation/Spread

Table 2.14: Descriptive Statistics of the Number of Price Changes (IBM)

| Date | N | Mean | Variance | K3 | K4 | BJ | BDS |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 389 | 0.3907 | 0.5531 | 2.0608 | 3.9658 | 530.3 | 9.06 |
| 01/04/94 | 388 | 0.2526 | 0.3391 | 2.8183 | 9.4730 | 1964.4 | 4.97 |
| 01/05/94 | 388 | 0.2088 | 0.2896 | 3.1204 | 11.6976 | 2841.8 | 3.79 |
| 01/06/94 | 386 | 0.3264 | 0.3867 | 2.1803 | 5.5918 | 808.7 | 3.85 |
| 01/07/94 | 389 | 0.2468 | 0.3101 | 2.4469 | 6.0589 | 983.2 | 4.70 |
| 01/10/94 | 389 | 0.3033 | 0.3614 | 2.0460 | 3.8552 | 512.3 | 3.99 |
| 01/11/94 | 390 | 0.2795 | 0.3613 | 2.5780 | 7.8688 | 1438.2 | 3.88 |
| 01/12/94 | 388 | 0.2526 | 0.2926 | 2.1648 | 4.2024 | 588.6 | 2.72 |
| 01/13/94 | 390 | 0.2000 | 0.2272 | 2.5245 | 6.5558 | 1112.7 | 4.17 |
| 01/14/94 | 390 | 0.1897 | 0.2364 | 2.8549 | 8.8073 | 1790.3 | 5.30 |
| 01/17/94 | 389 | 0.3188 | 0.3878 | 2.0971 | 4.3028 | 585.2 | 4.08 |
| 01/18/94 | 390 | 0.3487 | 0.4950 | 2.4629 | 6.9755 | 1185.0 | 6.49 |
| 01/19/94 | 389 | 0.2725 | 0.3585 | 2.4906 | 6.9718 | 1190.0 | 0.43 |
| 01/20/94 | 390 | 0.3308 | 0.3710 | 1.9398 | 3.7109 | 468.4 | 5.30 |
| 01/21/94 | 387 | 0.2248 | 0.2887 | 2.8445 | 9.9777 | 2127.2 | 2.40 |
| 01/24/94 | 388 | 0.6005 | 0.9589 | 2.2807 | 6.8562 | 1096.3 | 4.29 |
| 01/25/94 | 389 | 0.8483 | 1.1445 | 1.5243 | 2.3735 | 242.0 | 8.07 |
| 01/26/94 | 383 | 0.4099 | 0.4938 | 1.7337 | 2.4843 | 290.4 | 4.12 |
| 01/27/94 | 390 | 0.2410 | 0.2862 | 2.6665 | 9.2613 | 1856.0 | 2.52 |
| 01/28/94 | 389 | 0.2853 | 0.3282 | 2.0529 | 3.8841 | 517.8 | 5.77 |
| 01/31/94 | 390 | 0.3154 | 0.4016 | 2.0610 | 3.7419 | 503.6 | 2.39 |

* N : Number of Observations; K3 : Coefficient of Skewness; K4 : Coefficient of Kurtosis;
BJ : Bera-Jarque Normality Test Statistics $\sim \chi^2(2)$ with 5% critical value = 5.99.
BDS = Brock-Dechert-Scheinkman $\sim N(0,1)_{asy}$, Embedding Dimension = 3,
Epsilon = Standard Deviation/Spread

Table 2.15: Observed Frequency of Different Type of Trades

| | MN | MA | NM | AM | TOTAL |
|---|---|---|---|---|---|
| MN | 10657 | 209 | 2297 | 308 | 13471 |
| MA | 198 | 33 | 61 | 13 | 305 |
| NM | 2326 | 47 | 683 | 75 | 3131 |
| AM | 290 | 16 | 90 | 35 | 431 |
| TOTAL | 13471 | 305 | 3131 | 431 | 17338 |

Table 2.16: Observed Transitional Probability of Different Type of Trades

| | MN | MA | NM | AM |
|---|---|---|---|---|
| MN | 0.791 | 0.685 | 0.734 | 0.715 |
| MA | 0.015 | 0.108 | 0.019 | 0.030 |
| NM | 0.173 | 0.154 | 0.218 | 0.174 |
| AM | 0.022 | 0.052 | 0.029 | 0.081 |

Table 2.17: Expected Frequency of Different Type of Trades

| | MN | MA | NM | AM | TOTAL |
|---|---|---|---|---|---|
| MN | 10467 | 237 | 2433 | 335 | 13471 |
| MA | 237 | 5 | 55 | 8 | 305 |
| NM | 2433 | 55 | 565 | 78 | 3131 |
| AM | 335 | 8 | 78 | 11 | 431 |
| TOTAL | 13471 | 305 | 3131 | 431 | 17338 |

Table 2.18: Chi-square Contribution of Different Type of Trades

| | MN | MA | NM | AM |
|---|---|---|---|---|
| MN | 3.5 | 3.3 | 7.6 | 2.2 |
| MA | 6.4 | 142.3 | 0.6 | 3.9 |
| NM | 4.7 | 1.2 | 24.5 | 0.1 |
| AM | 6.0 | 9.3 | 1.9 | 55.0 |
| $\chi^2$ statistics = 272.5 with 12 degrees of freedom | | | | |

SELL: MA or MN; i.e. Broker or customers SELL to Market Maker;
BUY: AM or NM; i.e. Broker or customers BUY from Market Maker.
The 99% critical value of a $\chi^2$ statistics with 2 degrees of freedom equals 9.21034;
States listed in the first column refer to transaction in period $t$;
States listed in the first row refer to transaction in period $t+1$.

Table 2.19: Observed Frequency of Different Type of Buy/Sell Trades

| | SELL | BUY | TOTAL |
|---|---|---|---|
| SELL | 11097 | 2697 | 13776 |
| BUY | 2697 | 883 | 3562 |
| TOTAL | 13776 | 3562 | 17338 |

Table 2.20: Observed Transitional Probability of Different Type of Buy/Sell Trades

| | SELL | BUY |
|---|---|---|
| SELL | 0.806 | 0.194 |
| BUY | 0.752 | 0.248 |

Table 2.21: Expected Frequency of Different Type of Buy/Sell Trades

| | SELL | BUY | TOTAL |
|---|---|---|---|
| SELL | 10946 | 2830 | 13776 |
| BUY | 2830 | 732 | 3562 |
| TOTAL | 13776 | 3562 | 17338 |

Table 2.22: Chi-Square Contribution of Different Type of Buy/Sell Trades

| | SELL | BUY |
|---|---|---|
| SELL | 2.1 | 8.1 |
| BUY | 8.1 | 31.2 |
| $\chi^2$ statistics = 49.51 with 2 degrees of freedom | | |

**SELL: MA or MN; i.e. Broker or customers SELL to Market Maker;**
**BUY: AM or NM; i.e. Broker or customers BUY from Market Maker.**
**The 99% critical value of a $\chi^2$ statistics with 2 degrees of freedom equals 9.21034;**
**States listed in the first column refer to transaction in period $t$;**
**States listed in the first row refer to transaction in period $t + 1$.**

# CHAPTER 3

# MODELLING INTRA-DAY EQUITY PRICES AND VOLATILITY USING INFORMATION ARRIVALS - A COMPARATIVE STUDY OF DIFFERENT CHOICES OF INFORMATIONAL PROXIES

## 3.1 INTRODUCTION

The purpose of this chapter is to present a model for intra-day asset prices and volatility generation processes. In particular, we consider alternative choices of conditioning variables, i.e. exogenous variables, to help us in modelling. Although our methodology is general, we restrict ourselves to two U.S. stocks, IBM and INTEL and we take tick by tick data for January 1994; these stocks were chosen on the basis of their high liquidity. It might be argued that this is insufficient information to carry out our analysis, our response is that our use of data here is illustrative and that a full analysis involving many stocks and longer time periods could be carried out by researchers following the methodology presented here. Our Data comes from the New York Stock Exchange (NYSE) Trade and Quote (TAQ) database. This database contains virtually every trade and quote of every stock traded on major American stock exchanges.

In Section 3.2, we present our initial models, investigate their statistical properties, and identify certain problems. We then discuss estimation techniques and provide estimation results in the same section. We find that our information variables do not satisfy the requirement of being independently and identically distributed. To deal with this problem, we present an extended model based on doubly stochastic processes in Section 3.3. Surprisingly, these models are straightforward to estimate for all information variables except volume. We find that volume does not appear to be a suitable variable for measuring information flow, whilst the number of trades or the number of price changes seem to work very well. Finally, and importantly in our opinion, we find no evidence of volatility persistence in our model, although GARCH models measured on the same data show strong evidence of persistence. This indicates, to us at least, that the claimed persistence of volatility may be an artifact of the choice of model and does not reflect a market opportunity or a forecastable feature of the data. We conclude this chapter in Section 3.4.

## 3.2 A Homogeneous Mixed Jump-Diffusion Model

The first model we examine is a homogeneous mixed jump-diffusion model. The rationale of using such a model is that it reveals systematic discontinuities, which coincide with occasional jumps in financial time series. This model was first examined by Satchell and Yoon (1993) to study the influence of the number of transactions on the conditional mean and conditional variance of daily returns of five British stocks. By solving the standard jump-diffusion stochastic differential equation describing the evolution of asset prices, Satchell and Yoon show that log-returns are conditionally normal with mean and variance being linear functions of the number information arrival.

We define our price generating equation as follows

$$P(t) = P(0)\exp[(\alpha - \frac{1}{2}\sigma^2)t + \sigma(z(t) - z(0)) + \sum_{i=1}^{N(t)} Q_i] \tag{3.1}$$

where $P(t)$ denotes the price of an asset at time $t$, $\alpha$ and $\sigma$ are parameters, $z(t)$ is a standard Brownian motion, $N(t)$ is a homogeneous Poisson process with parameter $\lambda$ (we shall relax the assumption of homogeneity later in the chapter), $Q$ is a normal random variate with mean $\mu_Q$ and variance $\sigma_Q^2$ in the interval $(t, t + \Delta t]$. In what follows, we assume that there is some observed variable which measures the number of information arrival, $N(t)$.

As discussed in Chapter 2, such a variable could be trading volume, the number of trades, or the number of price changes,$\cdots$, etc.

It is straightforward to derive the probability density function (*pdf*) of logarithmic returns, $X(t) = \ln(P(t)/P(t-1))$, since

$$X(t) = (\alpha - \frac{1}{2}\sigma^2) + \sigma(z(t) - z(t-1)) + \sum_{i=1}^{\Delta N(t)} Q_i \tag{3.2}$$

We see that $X(t)$ is independent and identically distributed (*i.i.d.*) and

$$pdf(X(t)) = e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \phi(\mu + \mu_Q \cdot j, \sigma^2 + \sigma_Q^2 \cdot j) \tag{3.3}$$

where $\mu = \alpha - \sigma^2/2$, and $\phi(a,b)$ is the normal density with mean $a$, and variance $b$.

We next compute $pdf(X(t)|\Delta N(t))$. Simple manipulations with equation (3.2) show that

$$pdf(X(t) \mid \Delta N(t)) = \phi(\mu + \mu_Q \Delta N(t), \sigma^2 + \sigma_Q^2 \Delta N(t)) \tag{3.4}$$

It follows, in regression notation, that

$$X(t) = \mu + \mu_Q \Delta N(t) + \sqrt{\sigma^2 + \sigma_Q^2 \Delta N(t)}\,\varepsilon(t) \tag{3.5}$$

where $\varepsilon(t)$ is $N(0,1)$ and is independent of $\Delta N(t)$. Equation (3.5) may also be interpreted as a linear regression model with linear heteroskedasticity in $\Delta N(t)$, as such it is an example of the heteroskedasticity models popular in econometrics, see Judge *et al.* (1985, p.419) for a survey. As long as $\mu_Q$ and $\sigma_Q^2$ are found to be significant, our assumption that $\Delta N(t)$ influences the conditional mean and variance of returns is not rejected.

To test if $\Delta N(t)$ influences the rate of returns, an appropriate test would be the joint hypothesis, that is, $\mu_Q$=0 and $\sigma_Q^2$=0. We now devote some arguments to testing our various hypotheses. We shall consider the different hypotheses in turn. Let the three tests be

$$\begin{aligned} H_{10} &: \mu_Q = 0 & vs.\ H_{1A} : \mu_Q \neq 0 \\ H_{20} &: \sigma_Q^2 = 0 & vs.\ H_{2A} : \sigma_Q^2 \neq 0 \\ H_{30} &: \mu_Q = 0 \ and\ \sigma_Q^2 = 0 & vs.\ H_{3A} : \mu_Q \neq 0 \text{ or } \sigma_Q^2 \neq 0 \end{aligned} \tag{3.6}$$

A test that $\mu_Q$=0 implies that the number of trades does not influence the expected rate of returns, whilst it increases the volatility of the asset, an assertion investigated by Lamoureux and Lastrapes (1990a). A test that $\sigma_Q^2$=0 implies that jump magnitudes are of constant size, albeit unknown to the econometricians. In this case, each arrival of new information has the same kind of effect on stock prices, i.e., each transaction generates the same amount of trading volume and consequently the same impact on prices. Similarly, the joint hypothesis implies that the $\Delta N(t)$ is completely independent of price changes. Since all these hypotheses are interesting, it is worthwhile estimating and testing our model. We note two of the tests above have the difficulty that the point $\sigma_Q^2$=0 lies on the boundary of the parameter space, so that the asymptotic distribution of the one-sided test will be non-standard, i.e., not $\chi^2(1)$. For this reason the Lagrange Multiplier ($LM$) test would be preferred to Wald or Likelihood Ratio ($LR$) tests, since it is well-known that the $LM$ test retains its $\chi^2(1)$ distribution under $H_0$ even for boundary points. Here, computational ease is required at the cost of potential loss of power. It is straightforward to derive the $LM$ test, see Breusch and Pagan (1980). The derivation of the score statistics for the three hypotheses is shown in Appendix B. We present the results as a theorem.

**Theorem 3.2.1** *The LM tests for our hypotheses given in equation (3.6) are*

$$\begin{aligned}
LM_1 &= \left(\sum \frac{(x_t-\hat{\mu})\Delta N_t}{\hat{h}_t}\right)^2 \left(\sum \frac{\Delta N_t^2}{\hat{h}_t} - (\sum \frac{\Delta N_t}{\hat{h}_t})^2 (\sum \frac{1}{\hat{h}_t})^{-1}\right)^{-1} \\
\textit{where } \hat{h}_t &= \hat{\sigma}^2 + \hat{\sigma}_Q^2 \Delta N_t \\
LM_2 &= LM_2(\hat{e}_t) = \frac{(\sum(\hat{e}_t^2-\hat{\sigma}^2)\Delta N_t)^2}{2\hat{\sigma}^4\sum(\Delta N_t - \Delta\overline{N})^2} \quad \textit{where } \hat{e}_t = x_t - \hat{\mu} - \hat{\mu}_Q \Delta N_t \\
LM_3 &= LM_2(\tilde{e}_t) + \frac{(\sum \tilde{e}_t \Delta N_t)^2}{\hat{\sigma}^2 \sum(\Delta N_t - \Delta\overline{N})^2} \quad \textit{where } \tilde{e}_t = x_t - \hat{\mu}
\end{aligned} \tag{3.7}$$

*where $LM_i$ is the test appropriate for $H_{0i}$.*

It should be noted that our test procedures are asymptotically $\chi^2(1)$ for test statistics $LM_1, LM_2$ and $\chi^2(2)$ for test statistics $LM_3$. This is true despite the fact that the alternative hypothesis involves $\sigma_Q^2$ being positive, so that one may wish to use this information explicitly. This has been done in general by Rogers (1986) in which he proposes a test procedure based on the Kuhn-Tucker test of Gourieroux, Holly and Monfort (1982). This should lead to a more powerful test but involves substantially more computation; we shall not investigate this point any further.

We now turn to the estimation of equation (3.5), which can be estimated by the following iterative feasible generalized least squares procedure,

1. First, we regress $X(t)$ on $\Delta N(t)$ by ordinary least squares (OLS), and calculate the residuals, say $\hat{\epsilon}$'s, from the resulting OLS estimates $\hat{\mu}$, $\hat{\mu}_Q$. In other words, $\hat{\epsilon} = \hat{\epsilon}(t) = X(t) - \hat{\mu} - \Delta N(t)\hat{\mu}_Q$.

2. Regress $\hat{\epsilon}^2$ on $\Delta N(t)$ by nonlinear least squares (NLS) to obtain $\hat{\sigma}^2$, and $\hat{\sigma}_Q^2$.[1]

3. Apply generalized least squares on equation (3.5) after dividing both sides of the equation by $\sqrt{\hat{\sigma}^2 + \Delta N(t)\hat{\sigma}_Q^2}$. This will produce another set of estimates, $\hat{\mu}$ and $\hat{\mu}_Q$. Based on these estimates, calculate the new squared residual $\hat{\epsilon}^2 = [X(t) - \hat{\mu} - \Delta N(t)\hat{\mu}_Q]^2$ and iterate on step 2 and step 3.

Estimates derived from the above procedure will converge to maximum likelihood estimates by a familiar linearized maximum likelihood argument. Usually, only the first three iterations are required to produce convergent estimates.

Initially, we use all three trading variables, Tvol, N, and NPC, as informational proxies. However, we find that whenever Tvol is employed, the above iterative procedure has trouble converging. This is probably due to the high skewness typically observed for the data on trading volume. The nonconvergence problem also indicates that Tvol is not a suitable informational proxy in our model, which is consistent with recent empirical finding on the informational role of trading volume, as described in Chapter 2. Therefore, we exclude Tvol as one of the informational proxies in the following analysis.

The estimation results are reported in Tables 3.3 to 3.6. Judging from the $t$-statistics, neither the number of trades nor the number of price changes significantly influence the mean and the variance of returns of INTEL. In contrast, both number of trades and number of price changes have significant impact on the variance of returns of IBM. This may

[1]To avoid getting negative variance estimates, we use NLS to obtain $\hat{\sigma}$, and $\hat{\sigma}_Q$. The standard errors of $\hat{\sigma}^2$, and $\hat{\sigma}_Q^2$ are then derived by $\delta$-method.

be related to the data in Table 2.4 where the higher numbers of trades and price changes for INTEL relative to IBM mean that their impact is less important. Technically, it is as if the Poisson process may be converging to Brownian motion again.

Table 3.1: LM Test Statistics for Equation

| | INTEL | | | IBM | | |
|---|---|---|---|---|---|---|
| Date | LM1 | LM2 | LM3 | LM1 | LM2 | LM3 |
| 01/03/94 | 2.532 | 0.1 | 295879 | 0.607 | 69.5 | 390701 |
| 01/04/94 | 0.108 | 0.6 | 14125 | 0.022 | 41.4 | 420148 |
| 01/05/94 | 3.067 | 0.3 | 402553 | 0.038 | 26.6 | 236021 |
| 01/06/94 | 0.104 | 1.2 | 13134 | 0.243 | 24.8 | 215395 |
| 01/07/94 | 2.418 | 4.0 | 309104 | 1.007 | 55.1 | 2286418 |
| 01/10/94 | 0.156 | 0.4 | 18769 | 0.000 | 32.0 | 196540 |
| 01/11/94 | 0.030 | 0.0 | 3366 | 0.003 | 44.4 | 49963 |
| 01/12/94 | 0.603 | 5.7 | 67172 | 1.188 | 34.1 | 1571143 |
| 01/13/94 | 0.005 | 7.6 | 10195 | 2.092 | 41.9 | 14511512 |
| 01/14/94 | 0.002 | 0.2 | 178 | 0.273 | 48.5 | 295116 |
| 01/17/94 | 0.051 | 1.1 | 12681 | 1.863 | 51.5 | 3808499 |
| 01/18/94 | 0.387 | 15.1 | 20392 | 0.010 | 40.9 | 427964 |
| 01/19/94 | 0.022 | 20.1 | 6482 | 1.975 | 18.4 | 3463029 |
| 01/20/94 | 2.083 | 27.2 | 107936 | 1.016 | 25.4 | 1136782 |
| 01/21/94 | 0.720 | 1.2 | 110418 | 0.330 | 27.5 | 330128 |
| 01/24/94 | 0.105 | 9.6 | 9392 | 0.328 | 77.9 | 24908 |
| 01/25/94 | 0.180 | 0.7 | 16356 | 0.675 | 29.4 | 96355 |
| 01/26/94 | 0.077 | 2.3 | 9226 | 1.734 | 84.3 | 2195317 |
| 01/27/94 | 0.096 | 0.0 | 11746 | 1.785 | 21.8 | 7945026 |
| 01/28/94 | 2.073 | 0.7 | 251009 | 0.247 | 29.9 | 2203044 |
| 01/31/94 | 0.437 | 1.3 | 65353 | 0.840 | 14.1 | 2574018 |

To further investigate the effect of the number of trades, we conduct the *LM* test procedures detailed in Theorem 3.2.1. The results are reported in Table 3.1. In general, we cannot reject the null hypothesis $\mu_Q = 0$ for either INTEL or IBM. The null hypothesis $\sigma_Q^2 = 0$ cannot be rejected for most INTEL cases, while it is rejected for all IBM cases. These findings based on $LM_1$ and $LM_2$ are consistent with the findings based on the previously reported *t*-statistics. However, the $LM_3$ test is significant in all cases. This provides evidence that return processes are indeed influenced by the trading processes. The insignificant results are possibly due to modelling mis-specification. One way of examining the modelling mis-specification is to examine the independence of the fitted residuals from the model.

Since $\epsilon_t$ in equation (3.5) is assumed to be *i.i.d.*, we test this assumption by running the BDS test on the standardized estimated residuals, namely

$$\widehat{\epsilon}(t) = \frac{X(t) - \widehat{\mu} - \widehat{\mu}_Q \cdot \Delta N(t)}{\sqrt{\widehat{\sigma}^2 + \widehat{\sigma}_Q^2 \cdot \Delta N(t)}} \tag{3.8}$$

From those results reported in Table 3.2, the *i.i.d.* assumption on $\epsilon_t$ is clearly rejected for most cases. This leads us to the doubly stochastic modelling of information arrivals in the next section, since rejection of the *i.i.d.* process is likely to be a consequence of more complex arrival processes than the ones we have modelled.

Table 3.2: BDS Test on the Estimated Residuals

| Date | INTEL | IBM |
|---|---|---|
| 01/03/94 | 8.617 | 9.1082 |
| 01/04/94 | 7.546 | 5.9476 |
| 01/05/94 | 9.205 | 3.6334 |
| 01/06/94 | 7.886 | 4.0592 |
| 01/07/94 | 6.410 | 3.9121 |
| 01/10/94 | 7.809 | 1.1909 |
| 01/11/94 | 6.837 | 3.5701 |
| 01/12/94 | 7.725 | 3.3836 |
| 01/13/94 | 7.647 | 4.5410 |
| 01/14/94 | 6.877 | 5.8569 |
| 01/17/94 | 8.231 | 4.0395 |
| 01/18/94 | 7.615 | 7.2038 |
| 01/19/94 | 6.974 | 2.3438 |
| 01/20/94 | 6.255 | 4.2690 |
| 01/21/94 | 10.590 | 2.8906 |
| 01/24/94 | 7.959 | 9.7016 |
| 01/25/94 | 7.923 | 0.5750 |
| 01/26/94 | 8.903 | 2.4349 |
| 01/27/94 | 7.744 | 3.3970 |
| 01/28/94 | 10.022 | 5.7085 |
| 01/31/94 | 8.953 | 2.3108 |

Standardized Residual: $\frac{X(t)-\widehat{\mu}-\widehat{\mu}_Q \cdot \Delta N(t)}{\sqrt{\widehat{\sigma}^2+\widehat{\sigma}_Q^2 \cdot \Delta N(t)}}$

BDS : Brock-Dechert-Scheinkman$\sim N(0,1)_{asy}$,

Embedding Dimension = 3,

Epsilon : Standard Deviation/Spread

## 3.3 Doubly Stochastic Process and the Marginal Distribution of Returns

We have shown that information arrivals are not consistent with the (homogeneous) Poisson process with a fixed arrival intensity parameter. In order to accommodate the non-homogeneous nature of the information data, we now introduce a more complex (non-homogeneous) Poisson process which allows $\lambda(t)$ to vary. By an appropriate choice of $\lambda(t)$, we can model the marginal distribution of $\ln(P(t)/P(t-1))$. There are many candidates for the process of $\lambda(t)$. Here, $\lambda(t)$ is defined in the form of a stochastic volatility model as

$$\lambda(t) = \alpha v^2(t-1) + \beta Var(X(t-1)|I(t-2)) \tag{3.9}$$

where $v(t)$ is $N(0,1)$ unconditional in $N(t)$, in fact, $v(t) = \Delta z(t) = z(t) - z(t-1)$, and $I(t)$ contains information up to the end of the minute.[2] Heuristically, the expected number of jumps depends upon the previous volatility and the deviation from fundamental $\Delta z^2(t)$, see equation (3.2).

It follows from equation (3.2) that

$$\begin{aligned} E[X(t)|I(t-1)] &= \mu + \mu_Q \lambda(t) & (3.10) \\ Var(X(t)|I(t-1)) &= \sigma^2 + (\mu_Q^2 + \sigma_Q^2)\lambda(t) & (3.11) \end{aligned}$$

since $\lambda(t)$ is known given $I(t-1)$. To simplify our model we shall assume that $X(t)$ and $\Delta N(t)$ are (weakly) stationary. Under the assumption of weak stationarity, equation (3.9) becomes

$$\lambda(t) = \alpha v^2(t-1) + \beta\sigma^2 + \theta\lambda(t-1) = \frac{\beta\sigma^2}{1-\theta} + \alpha\sum_{j=0}^{\infty}\theta^j v^2(t-j-1) \tag{3.12}$$

where $\theta{=}\beta(\mu_Q^2+\sigma_Q^2), 0{<}\theta{<}1$. We can calculate the mean and the variance of $\lambda(t)$, detailed in Appendix C, as follows;

[2] A GARCH-type model would involve interpreting $v^2(t-1)$ as $Var(X(t-1)|I(t-2)) \times \Delta z^2(t)$ for $z(t) \sim i.i.d.\ N(0,1)$. This complicates the model without adding to its explanatory power. We shall refer to this as GARCH-type effects, although the model in equation (3.9) is closer to a stochastic volatility model.

$$E[\lambda(t)] = \frac{\beta\sigma^2 + \alpha}{1-\theta} \tag{3.13}$$

$$Var(\lambda(t)) = \frac{2\alpha^2}{1-\theta^2} \tag{3.14}$$

using the fact that $v^2(t)$ has a $\chi^2(1)$ distribution.

Given our model, the information variable is not purely exogenous any more: its intensity is dependent upon the past history of prices. This framework is attractive because it allows a feedback effect through the variables. It can explain certain phenomena in financial time series such as volatility clustering where large price changes tend to bunch together. This non-homogeneous Poisson process also resolves the restrictive aspect of the homogeneous Poisson distribution which implies that mean and variance are equal. From the moment generating function of $\Delta N(t)$ derived in Appendix C,[3]

$$m_{\Delta N}(s) = \exp\left(\frac{\beta\sigma^2}{1-\theta}(\exp(s)-1)\right) \cdot \prod_{t=1}^{\infty}\left(1-2\alpha\theta^j(\exp(s)-1)\right)^{-1/2} \tag{3.15}$$

we can derive the mean and the variance of $\Delta N(t)$,

$$E[\Delta N(t)] = \frac{\beta\sigma^2 + \alpha}{1-\theta} \tag{3.16}$$

$$Var(\Delta N(t)) = \frac{\beta\sigma^2 + \alpha}{1-\theta} + \frac{2\alpha^2}{1-\theta} \tag{3.17}$$

Note that the mean and the variance are not equal in the presence of stochastic $\lambda(t)$. Moreover, the serial correlation of $\Delta N(t)$ can be shown to be

$$\begin{aligned} Corr(\Delta N(t), \Delta N(t-s)) &= \frac{Cov(\theta^s\lambda(t-s), \Delta N(t-s))}{Var(\Delta N(t))} \\ &= \frac{\theta^s Var(\lambda(t-s))}{Var(\Delta N(t))} \\ &= \begin{cases} \frac{2\alpha^2\theta^s}{2\alpha^2+(\beta\sigma^2+\alpha)(1+\theta)}, & \text{if } s = 1, 2, \ldots \\ 1, & \text{if } s = 0 \end{cases} \end{aligned} \tag{3.18}$$

Then from equation (3.2), the moment conditions of $X(t)$, detailed in Appendix D, are obtained as follows,

$$E[X(t)] = \mu + \frac{\mu_Q(\beta\sigma^2 + \alpha)}{1-\theta}$$

[3]We have assumed that $\Delta N(t)$ is stationary, an alternative expression can be calculated if we start from a fixed starting point.

$$Var[X(t)] = \sigma^2 + \frac{(\mu_Q^2+\sigma_Q^2)(\beta\sigma^2+\alpha)}{1-\theta} + \frac{2\alpha^2\mu_Q^2}{1-\theta^2} \tag{3.19}$$

$$Corr(X(t), X(t-s)) = \begin{cases} \frac{2\alpha^2\mu_Q^2\theta^s}{\sigma^2(1-\theta^2)+(\mu_Q^2+\sigma_Q^2)(1+\theta)(\beta\sigma^2+\alpha)+2\alpha^2\mu_Q^2}, & \text{if } s=1,2,\ldots \\ 1, & \text{if } s=0 \end{cases}$$

Given that $\Delta N(t)$ is observable, the joint likelihood function of $X(t)$ and $\Delta N(t)$ can be written as

$$\begin{aligned} L &= \prod_{t=1}^{T} pdf(X(t), \Delta N(t)|I(t-1)) \qquad (3.20) \\ &= \prod_{t=1}^{T} pdf(X(t)|\Delta N(t), I(t-1)) \cdot pdf(\Delta N(t)|I(t-1)) \\ &= \prod_{t=1}^{T} \phi(\mu+\mu_Q\Delta N(t), \sigma^2+\sigma_Q^2\Delta N(t)) \frac{\exp(-\lambda(t))(\lambda(t))^{\Delta N(t)}}{\Delta N(t)!} \end{aligned}$$

Thus, the log-likelihood function becomes

$$\begin{aligned} \ln L = const. - \Sigma\ln(\sigma^2+\sigma_Q^2\Delta N(t)) - \frac{1}{2}\Sigma\frac{(X(t)-\mu-\mu_Q\Delta N(t))^2}{\sigma^2+\sigma_Q^2\Delta N(t)} \\ -\Sigma\lambda(t) + \Sigma\Delta N(t)\ln(\lambda(t)) \qquad (3.21) \\ \text{where } \lambda(t) = \beta\sigma^2 + \alpha\left(\frac{X(t-1)-\mu-\mu_Q\Delta N(t-1)}{\sqrt{\sigma^2+\sigma_Q^2\Delta N(t-1)}}\right)^2 + \beta(\mu_Q^2+\sigma_Q^2)\lambda(t-1) \end{aligned}$$

Although the joint density is tractable, the marginal densities of $X(t)$ and $\Delta N(t)$ cannot be derived explicitly. This is one of the characteristics of a mixture of distributions which contain several random variables.

We estimate equation (3.21) and report the results of in Tables 3.7, 3.8, 3.9, and 3.10. In terms of the significance of coefficients, $\widehat{\mu}_Q$ and $\widehat{\sigma}_Q^2$, in the variance equation, we obtain similar results as those obtained from estimating the homogeneous mixed jump-diffusion model in equation (3.5). Namely, $\widehat{\sigma}_Q^2$ is highly significant for all IBM cases, while $\widehat{\mu}_Q$ is insignificant in most cases. In addition, $\widehat{\beta}$ is highly significant for all cases. From equation (3.9), we know that $\widehat{\beta}$ measure the sensitivity of information arrival intensity $\lambda(t)$ to previous period's realized volatility. This is a strong evidence for the existence of a stochastic arrival intensity process.

Another way of examining the performance of the doubly stochastic Poisson model is to compare the sample moments (mean and variance) of trading variables with those implied by the model in equation (3.16). The results are reported in Tables 3.11, 3.12, 3.13, and 3.14. An interesting result shown in these tables is that the implied expected values, $\widehat{E}(\Delta N_t)$, of the trading variables match their sample counterparts, $E(\Delta N_t)$, quite well. To give a quick measure of how close they are, we calculate a $\chi^2$ test statistics $TCF$ as

$$TCF = \sum_{t=1}^{21} CF_t = \sum_{t=1}^{21} \frac{\left[E(\Delta N_t) - \widehat{E}(\Delta N_t)\right]^2}{\widehat{E}(\Delta N_t)} \sim \chi^2(20) \tag{3.22}$$

All of the $TCF$'s are well inside the critical region under conventional significance levels. By comparing the magnitudes of $TCF$s, we also find that, for both INTEL and IBM, the number of trades seem to be a slightly better proxy of information arrival. In addition, for IBM, implied variances, $\widehat{V}(\Delta N_t)$, of the trading variables are also very close to their sample counterparts, $V(\Delta N_t)$.

Also reported in Tables 23 and Tables 24 are the $\widehat{\theta}$ values, which represent the degree of dependence of $\lambda(t)$ on $\lambda(t-1)$ from equation (3.12). This is an equivalent measure of volatility persistence in GARCH models. On average, when the number of trades is used as the informational proxy, $\widehat{\theta} \simeq 0.05$ for INTEL, and $\widehat{\theta} \simeq 0.64$ for IBM. Similar results are obtained when the number of price changes is used as the informational proxy. In that case, $\widehat{\theta} \simeq 0.1$ for INTEL, and $\widehat{\theta} \simeq 0.64$ for IBM on average. In other words, the volatility persistence implied by our model is much smaller than those implied by the GARCH-type models.

To compare the persistence of GARCH-type models versus that of informational volatility models as in this paper, we fit a GARCH(1,1) model on the same data sets. We compare the $\widehat{\theta}$ values listed in Tables 3.11, 3.12, 3.13, and 3.14 with value of $\widehat{\alpha}+\widehat{\beta}$ reported in Table 3.15. We recall that Tables 3.11 and 3.12, describe the models for the number of trades, whilst Tables 3.13 and 3.14 describe the models for the number of price changes. In Tables 3.11, for INTEL, there are no values of $\widehat{\theta}$ greater than 0.5, and there are only three values greater than 0.1. Similarly, in Table 3.13, there are only 6 values of $\widehat{\theta}$ greater than 0.1. However, for the GARCH(1,1) model for INTEL in Table 3.15, there are seven values of $\widehat{\alpha}+\widehat{\beta}$ greater than 0.9, and most others greater than 0.5. Likewise, for IBM, there are no values of $\widehat{\theta}$ greater than 0.8 in either Table 3.12 or Table 3.14. But, for the GARCH(1,1) model for IBM in Table 3.15, there are eight values of $\widehat{\alpha}+\widehat{\beta}$ greater than 0.9. This indicates, to us at least, that the claimed persistence of volatility may be an artifact of the choice of model and does not reflect a market opportunity or a forecastable feature of the data.

However, the failure of this model is that autocorrelation of returns must be positive, as given by equation (3.19). This is enough to invalidate it. We discuss this further in the next chapter.

## 3.4 Conclusions

This chapter has had three objectives. They are (i), to compare three different proxies for informational variables in high frequency equity data; (ii), to model dynamic processes using doubly stochastic Poisson models; (iii), to investigate intra-day volatility persistence. We find that the number of trades and the number of price changes seem to be the best choices for informational variables, volume being decidedly inferior. Secondly, we find that our model does seem to be estimable without undue difficulty and finally we find that persistence in volatility is much reduced when our model is used rather than GARCH(1,1). The use of informational variables seems to substantially eliminate much of the persistence. Since persistence in volatility is a stylized fact that seems somewhat flawed in terms of theoretical explanations, the results lead toward better modelling and understanding of intra-day volatility.

Table 3.3: Estimation Results of Equation (3.5) (INTEL) Using the Number of Trades as the Conditioning Variable

| Date | $\widehat{\mu}$ | $\widehat{\mu}_Q$ | $\widehat{\sigma}^2$ | $\widehat{\sigma}_Q^2$ |
|---|---|---|---|---|
| 01/03/94 | 1.996E-004 | -4.329E-005 | 8.838E-006 | 1.312E-015 |
| | (0.94474) | (-1.61103) | (9.55341) | (-1.428E-007) |
| 01/04/94 | 1.527E-004 | -8.231E-006 | 8.429E-006 | 3.862E-014 |
| | (0.73996) | (-0.33619) | (7.97083) | (-4.753E-007) |
| 01/05/94 | -1.843E-004 | 3.119E-005 | 7.841E-006 | 5.282E-017 |
| | (-0.97418) | (1.75996) | (8.67204) | ( 4.500E-008) |
| 01/06/94 | 1.299E-005 | -8.968E-006 | 9.263E-006 | 4.198E-014 |
| | (0.05988) | (-0.33559) | (9.00200) | (-4.769E-007) |
| 01/07/94 | -8.050E-005 | 1.840E-005 | 9.466E-006 | 1.247E-014 |
| | (-0.43540) | (1.63519) | (9.17861) | (-6.909E-007) |
| 01/10/94 | -3.718E-005 | 7.850E-006 | 9.052E-006 | 5.319E-014 |
| | (-0.16752) | (0.40331) | (6.32805) | (-4.952E-007) |
| 01/11/94 | -3.314E-005 | 4.730E-006 | 8.637E-006 | 9.960E-014 |
| | (-0.15011) | (0.17307) | (8.63161) | (-4.698E-007) |
| 01/12/94 | 1.141E-004 | -1.173E-005 | 1.137E-005 | 1.406E-014 |
| | (0.54093) | (-0.82905) | (6.75941) | (-4.008E-007) |
| 01/13/94 | 1.168E-005 | -2.199E-006 | 4.492E-006 | 2.846E-007 |
| | (0.05941) | (-0.06648) | (7.59214) | (3.30385) |
| 01/14/94 | -3.387E-006 | -8.203E-007 | 9.382E-006 | 4.411E-016 |
| | (-0.01597) | (-0.04081) | (8.67428) | (-9.808E-008) |
| 01/17/94 | -4.191E-005 | -6.560E-006 | 8.997E-006 | 1.266E-007 |
| | (-0.17292) | (-0.22794) | (4.58957) | (0.58414) |
| 01/18/94 | 2.183E-004 | -2.009E-005 | 9.172E-006 | 5.782E-007 |
| | (0.70048) | (-0.62542) | (2.33811) | (1.71119) |
| 01/19/94 | 4.652E-005 | -1.674E-006 | 1.093E-005 | 2.722E-007 |
| | (0.12095) | (-0.14970) | (1.40840) | (1.45586) |
| 01/20/94 | -3.573E-004 | 3.071E-005 | 5.678E-006 | 4.363E-007 |
| | (-1.32397) | (1.47524) | (3.16975) | (4.36392) |
| 01/21/94 | -1.071E-004 | 1.453E-005 | 7.759E-006 | 6.499E-014 |
| | (-0.52230) | (0.89053) | (8.10009) | (-8.193E-007) |
| 01/24/94 | -5.805E-005 | 1.052E-005 | 1.735E-005 | 1.295E-018 |
| | (-0.20290) | (0.36238) | (3.00957) | (-9.974E-010) |
| 01/25/94 | 7.789E-005 | -1.587E-005 | 1.063E-005 | 1.432E-007 |
| | (0.29767) | (-0.42287) | (1.97415) | (0.19728) |
| 01/26/94 | 4.803E-005 | -7.185E-006 | 1.042E-005 | 1.688E-015 |
| | (0.20184) | (-0.29312) | (8.95570) | (-1.573E-007) |
| 01/27/94 | -4.314E-005 | 8.078E-006 | 8.167E-006 | 3.522E-017 |
| | (-0.19172) | (0.30938) | (8.52901) | (-2.953E-008) |
| 01/28/94 | -1.591E-004 | 3.636E-005 | 8.955E-006 | 4.182E-015 |
| | (-0.79349) | (1.47203) | (10.87673) | (-2.776E-007) |
| 01/31/94 | -7.144E-005 | 1.234E-005 | 7.523E-006 | 5.033E-020 |
| | (-0.41732) | (0.68072) | (12.95443) | (-1.828E-009) |

• Numbers in the parentheses are t-values

Table 3.4: Estimation Results of Equation (3.5) (IBM) Using the Number of Trades as the Conditioning Variable

| Date | $\widehat{\mu}$ | $\widehat{\mu}_Q$ | $\widehat{\sigma}^2$ | $\widehat{\sigma}^2_Q$ |
|---|---|---|---|---|
| 01/03/94 | 8.476E-005 | -3.769E-005 | 5.061E-007 | 6.011E-007 |
| | (1.49222) | (-0.80970) | (2.70837) | (6.11218) |
| 01/04/94 | 3.548E-005 | -5.719E-006 | 4.606E-007 | 2.940E-007 |
| | (0.69125) | (-0.15700) | (3.39504) | (4.30002) |
| 01/05/94 | 1.729E-005 | 6.371E-006 | 4.399E-007 | 1.849E-007 |
| | (0.32467) | (0.20052) | (3.35640) | (3.07845) |
| 01/06/94 | -2.508E-005 | -1.887E-005 | 6.229E-007 | 2.701E-007 |
| | (-0.39776) | (-0.50228) | (4.20342) | (3.96198) |
| 01/07/94 | -3.140E-005 | 4.066E-005 | 3.467E-007 | 3.590E-007 |
| | (-0.65941) | (1.05446) | (2.75879) | (5.12474) |
| 01/10/94 | 2.205E-005 | -1.628E-007 | 4.955E-007 | 2.531E-007 |
| | (0.38751) | (-0.00458) | (3.78897) | (4.05828) |
| 01/11/94 | -1.922E-005 | -2.098E-006 | 4.518E-007 | 3.284E-007 |
| | (-0.36991) | (-0.05135) | (3.65432) | (4.70252) |
| 01/12/94 | 3.584E-005 | -4.622E-005 | 5.312E-007 | 2.934E-007 |
| | (0.64646) | (-1.12057) | (4.20136) | (4.16692) |
| 01/13/94 | -5.142E-005 | 5.839E-005 | 4.264E-007 | 2.805E-007 |
| | (-1.06162) | (1.57593) | (3.35118) | (4.08225) |
| 01/14/94 | 1.316E-005 | -1.963E-005 | 3.785E-007 | 2.742E-007 |
| | (0.28891) | (-0.54274) | (3.03904) | (3.95972) |
| 01/17/94 | 3.241E-005 | -6.373E-005 | 6.459E-007 | 4.567E-007 |
| | (0.54283) | (-1.40802) | (3.25467) | (4.41821) |
| 01/18/94 | -1.056E-005 | -3.887E-006 | 5.499E-007 | 3.466E-007 |
| | (-0.16881) | (-0.10026) | (3.40469) | (4.72099) |
| 01/19/94 | -1.180E-004 | 5.772E-005 | 6.743E-007 | 2.176E-007 |
| | (-1.87331) | (1.42191) | (4.92592) | (3.05282) |
| 01/20/94 | 3.810E-005 | -4.573E-005 | 7.201E-007 | 3.179E-007 |
| | (0.54601) | (-1.02940) | (4.26880) | (3.77777) |
| 01/21/94 | 3.717E-005 | -2.098E-005 | 5.734E-007 | 2.275E-007 |
| | (0.63357) | (-0.58568) | (4.13782) | (3.47190) |
| 01/24/94 | -6.196E-005 | 7.478E-005 | 7.402E-006 | 7.937E-006 |
| | (-0.18828) | (0.48924) | (0.65732) | (2.38591) |
| 01/25/94 | 1.102E-004 | -3.546E-005 | 1.902E-006 | 6.626E-007 |
| | (0.71898) | (-0.83679) | (1.04523) | (1.65493) |
| 01/26/94 | -4.424E-006 | -4.935E-005 | 3.848E-007 | 6.162E-007 |
| | (-0.07222) | (-1.32686) | (1.53389) | (6.87969) |
| 01/27/94 | -6.900E-005 | 6.118E-005 | 6.882E-007 | 2.824E-007 |
| | (-1.04948) | (1.43605) | (3.88852) | (3.12824) |
| 01/28/94 | -5.002E-005 | 2.189E-005 | 6.070E-007 | 2.991E-007 |
| | (-0.86796) | (0.52766) | (4.69518) | (4.30378) |
| 01/31/94 | -1.133E-004 | 4.083E-005 | 6.943E-007 | 2.235E-007 |
| | (-1.54928) | (0.94092) | (4.47511) | (2.87036) |

* Numbers in the parentheses are t-values

Table 3.5: Estimation Results of Equation (3.5) (INTEL) Using the Number of Price Changes as the Conditioning Variable

| Date | $\widehat{\mu}$ | $\widehat{\mu}_Q$ | $\widehat{\sigma}^2$ | $\widehat{\sigma}_Q^2$ |
|---|---|---|---|---|
| 01/03/94 | 2.797E-004 | -1.073E-004 | 8.908E-006 | 1.1962E-014 |
| | (1.24956) | (-1.92245) | (9.05379) | (-1.741E-007) |
| 01/04/94 | 1.290E-004 | -8.417E-006 | 7.850E-006 | 3.918E-008 |
| | (0.57753) | (-0.15131) | (6.86882) | (0.13952) |
| 01/05/94 | -6.153E-005 | 2.780E-005 | 7.208E-006 | 1.126E-007 |
| | (-0.29259) | (0.62172) | (7.19134) | (0.55261) |
| 01/06/94 | -1.284E-005 | -8.498E-006 | 8.232E-006 | 1.021E-007 |
| | (-0.05847) | (-0.16307) | (7.89459) | (0.42606) |
| 01/07/94 | -2.415E-004 | 7.327E-005 | 9.623E-006 | 2.476E-014 |
| | (-1.11331) | (2.15278) | (7.95842) | (-2.841E-007) |
| 01/10/94 | -1.618E-004 | 4.273E-005 | 7.125E-006 | 3.356E-007 |
| | (-0.63430) | (0.84874) | (4.30540) | (1.11198) |
| 01/11/94 | -2.863E-005 | 7.484E-006 | 7.878E-006 | 1.9320E-007 |
| | (-0.13154) | (0.14092) | (8.07765) | (0.86884) |
| 01/12/94 | 1.663E-005 | -3.113E-006 | 1.083E-005 | 4.441E-014 |
| | (0.07846) | (-0.11960) | (6.42574) | (-2.945E-007) |
| 01/13/94 | 6.764E-005 | -2.202E-005 | 4.874E-006 | 3.655E-007 |
| | (0.36191) | (-0.42621) | (8.63130) | (2.65741) |
| 01/14/94 | 1.385E-004 | -3.688E-005 | 8.525E-006 | 1.380E-007 |
| | (0.58783) | (-0.79119) | (7.17790) | (0.61662) |
| 01/17/94 | -1.547E-004 | 1.960E-005 | 7.308E-006 | 7.280E-007 |
| | (-0.68114) | (0.34487) | (3.83022) | (1.82811) |
| 01/18/94 | 1.832E-004 | -3.109E-005 | 9.065E-006 | 1.130E-006 |
| | (0.58729) | (-0.50424) | (2.24456) | (1.68045) |
| 01/19/94 | 7.124E-005 | -4.669E-006 | 1.097E-005 | 5.260E-007 |
| | (0.18416) | (-0.21381) | (1.39101) | (1.41326) |
| 01/20/94 | -8.547E-005 | 1.868E-005 | 5.818E-006 | 8.168E-007 |
| | (-0.34831) | (0.51756) | (3.74100) | (5.40690) |
| 01/21/94 | -1.682E-004 | 4.006E-005 | 6.814E-006 | 6.004E-008 |
| | (-0.81873) | (1.24356) | (7.25026) | (0.42637) |
| 01/24/94 | -9.627E-005 | 2.821E-005 | 1.778E-005 | 1.684E-021 |
| | (-0.32268) | (0.50274) | (2.95583) | (-1.912E-011) |
| 01/25/94 | 2.472E-004 | -8.903E-005 | 8.935E-006 | 8.328E-007 |
| | (0.96579) | (-1.16360) | (1.64408) | (0.58457) |
| 01/26/94 | 5.779E-005 | -1.603E-005 | 9.595E-006 | 2.082E-016 |
| | (0.25008) | (-0.36543) | (8.53751) | (-3.283E-008) |
| 01/27/94 | -4.747E-006 | 4.377E-006 | 6.801E-006 | 4.003E-007 |
| | (-0.02126) | (0.07827) | (7.01254) | (1.80849) |
| 01/28/94 | -8.705E-005 | 4.584E-005 | 7.879E-006 | 2.305E-007 |
| | (-0.39123) | (0.72296) | (8.61021) | (0.93629) |
| 01/31/94 | -1.228E-004 | 4.249E-005 | 7.325E-006 | 5.385E-014 |
| | (-0.62444) | (0.85998) | (10.95842) | (-3.701E-007) |

* Numbers in the parentheses are t-values

Table 3.6: Estimation Results of Equation (3.5) (IBM) Using the Number of Price Changes as the Conditioning Variable

| Date | $\widehat{\mu}$ | $\widehat{\mu}_Q$ | $\widehat{\sigma}^2$ | $\widehat{\sigma}_Q^2$ |
|---|---|---|---|---|
| 01/03/94 | 6.026E-005 | -6.762E-005 | 6.172E-007 | 1.769E-006 |
| | (1.46349) | (-0.61591) | (4.55751) | (10.97768) |
| 01/04/94 | 7.588E-005 | -1.905E-004 | 3.694E-007 | 1.906E-006 |
| | (2.48022) | (-1.41183) | (4.16398) | (13.63471) |
| 01/05/94 | 1.571E-005 | 5.502E-005 | 3.686E-007 | 1.727E-006 |
| | (0.50900) | (0.37617) | (4.30642) | (11.64407) |
| 01/06/94 | -5.648E-005 | 5.477E-006 | 4.852E-007 | 1.715E-006 |
| | (-1.50418) | (0.04517) | (4.81681) | (11.95872) |
| 01/07/94 | -5.255E-006 | 1.099E-004 | 3.225E-007 | 2.008E-006 |
| | (-0.18449) | (0.79623) | (4.26582) | (16.17123) |
| 01/10/94 | 3.789E-005 | -5.292E-005 | 4.159E-007 | 1.529E-006 |
| | (1.10772) | (-0.45048) | (4.77369) | (11.81812) |
| 01/11/94 | -4.200E-005 | 7.198E-005 | 3.318E-007 | 1.885E-006 |
| | (-1.40281) | (0.55665) | (4.33180) | (16.30588) |
| 01/12/94 | 4.232E-005 | -2.543E-004 | 3.934E-007 | 1.919E-006 |
| | (1.28267) | (-1.79089) | (4.67616) | (13.61706) |
| 01/13/94 | -1.270E-005 | 1.729E-004 | 2.664E-007 | 2.609E-006 |
| | (-0.48989) | (0.99633) | (3.24443) | (16.41958) |
| 01/14/94 | 3.216E-005 | -2.267E-004 | 3.079E-007 | 2.081E-006 |
| | (1.20046) | (-1.44434) | (3.63448) | (12.81442) |
| 01/17/94 | -3.231E-005 | -5.548E-005 | 5.335E-007 | 2.228E-006 |
| | (-0.79160) | (-0.38794) | (3.53979) | (10.34310) |
| 01/18/94 | -6.470E-005 | 1.368E-004 | 4.756E-007 | 1.836E-006 |
| | (-1.69856) | (1.11706) | (4.60249) | (13.95372) |
| 01/19/94 | -6.733E-005 | 9.985E-005 | 5.116E-007 | 1.666E-006 |
| | (-1.77032) | (0.75106) | (5.50270) | (11.79130) |
| 01/20/94 | -4.766E-005 | 5.651E-005 | 5.806E-007 | 1.827E-006 |
| | (-1.15499) | (0.44705) | (4.92443) | (10.74088) |
| 01/21/94 | 2.324E-005 | -7.708E-005 | 4.426E-007 | 2.074E-006 |
| | (0.69293) | (-0.50643) | (4.99179) | (13.61826) |
| 01/24/94 | 3.607E-005 | 1.559E-004 | 7.168E-016 | 5.060E-005 |
| | (0.25796 ) | ( 4.134E-006 ) | (-1.554E-009) | (7.13034) |
| 01/25/94 | 4.234E-005 | -7.563E-005 | 1.933E-006 | 2.864E-006 |
| | (0.44948) | (-0.64119) | (1.47035) | (2.97583) |
| 01/26/94 | -2.175E-005 | -2.099E-004 | 7.664E-007 | 2.215E-006 |
| | (-0.42620) | (-1.58869) | (3.99335) | (9.39011) |
| 01/27/94 | -1.460E-005 | 1.306E-004 | 4.411E-007 | 2.689E-006 |
| | (-0.45798) | (0.84308) | (3.88487) | (13.89358) |
| 01/28/94 | -2.892E-005 | 2.324E-005 | 3.725E-007 | 2.156E-006 |
| | (-0.88786) | (0.16630) | (4.49665) | (16.65284) |
| 01/31/94 | -1.956E-005 | -9.774E-005 | 4.688E-007 | 1.827E-006 |
| | (-0.54239) | (-0.78937) | (5.08685) | (14.03249) |

• Numbers in the parentheses are t-values

Table 3.7: MLE Estimation Results on Equation (3.21) (INTEL) when the Number of Trades is Used as the Conditioning Variable

| Date | $\widehat{\mu}$ | $\widehat{\mu}_Q$ | $\widehat{\sigma}^2$ | $\widehat{\sigma}_Q^2$ | $\widehat{\alpha}$ | $\widehat{\beta}$ |
|---|---|---|---|---|---|---|
| 01/03/94 | 5.64E-004 | -1.38E-004 | 9.67E-006 | 8.25E-014 | 0.154 | 583748 |
| | (2.395) | (-5.281) | (11.709) | (2.0E-006) | (4.563) | (12.059) |
| 01/04/94 | -3.26E-004 | 1.19E-004 | 9.30E-006 | 1.62E-015 | 0.359 | 649453 |
| | (-1.291) | (3.964) | (11.379) | ( -1.2E-006) | (5.699) | (14.303) |
| 01/05/94 | -2.32E-004 | 3.24E-005 | 7.45E-006 | 1.79E-008 | -0.107 | 958372 |
| | (-1.159) | (1.311) | (10.703) | (0.188) | (-2.258) | (16.172) |
| 01/06/94 | 1.52E-005 | -9.44E-006 | 8.59E-006 | 2.66E-018 | 0.008 | 685890 |
| | (0.064) | (-0.283) | (9.188) | (6.6E-009) | (0.190) | (15.300) |
| 01/07/94 | 4.95E-005 | 9.96E-006 | 8.59E-006 | 1.39E-016 | -0.206 | 1159475 |
| | (0.293) | (0.732) | (16.557) | ( -8.5E-009) | (-6.326) | (15.847) |
| 01/10/94 | -3.56E-005 | 9.63E-006 | 8.59E-006 | 3.45E-015 | 0.039 | 971904 |
| | (-0.131) | (0.354) | (9.329) | ( -2.4E-007) | (0.870) | (20.818) |
| 01/11/94 | -8.07E-005 | 5.88E-006 | 8.53E-006 | 3.2E-019 | 0.238 | 676035 |
| | (-0.371) | (0.26277) | (12.731) | ( -3.7E-009) | (5.949) | (15.062) |
| 01/12/94 | 8.05E-005 | -7.58E-006 | 9.92E-006 | 7.54E-020 | -0.197 | 990418 |
| | (0.458) | (-0.706) | (22.246) | ( -5.8E-009) | (-7.774) | (24.090) |
| 01/13/94 | 7.77E-004 | -2.03E-004 | 7.56E-006 | 1.47E-014 | 0.608 | 637174 |
| | (3.486) | (-7.779) | (9.801) | ( -1.1E-007) | (7.303) | (10.853) |
| 01/14/94 | -6.02E-006 | 5.01E-006 | 9.00E-006 | 9.45E-009 | -0.160 | 815408 |
| | (-0.025) | (0.178) | (12.640) | (0.162) | (-2.914) | (16.443) |
| 01/17/94 | -5.46E-005 | -6.31E-006 | 8.29E-006 | 2.32E-007 | -0.035 | 687441 |
| | (-0.204) | (-0.172) | (10.074) | (1.803) | (-1.135) | (22.495) |
| 01/18/94 | 3.44E-004 | -4.281E-005 | 1.20E-005 | 2.43E-007 | 0.199 | 628407 |
| | (0.917) | (-1.39052) | (13.412) | (3.624) | (4.468) | (26.134) |
| 01/19/94 | -0.004 | -0.001 | 4.13E-011 | 6.32E-005 | 26.025 | 1465 |
| | (-6.721) | (-54.772) | ( -5.8E-007) | (4.790) | (4.917) | (4.777) |
| 01/20/94 | -0.003 | 6.86E-004 | 2.10E-005 | 2.56E-006 | 3.124 | 162954 |
| | (-17.544) | (48.216) | (13.694) | (14.230) | (16.743) | (16.825) |
| 01/21/94 | -8.05E-005 | 1.26E-005 | 7.13E-006 | 3.87E-017 | -0.090 | 1326696 |
| | (-0.358) | (0.597) | (11.562) | ( -4.5E-008) | (-1.461) | (15.532) |
| 01/24/94 | -6.51E-005 | 1.11E-005 | 1.38E-005 | 2.90E-017 | -0.068 | 537886 |
| | (-0.228) | (0.315) | (30.146) | ( -3.0E-008) | (-4.019) | (33.728) |
| 01/25/94 | 1.04E-004 | -2.29E-005 | 1.14E-005 | 5.41E-016 | 0.043 | 477442 |
| | (0.336) | (-0.494) | (10.800) | ( -5.4E-008) | (1.369) | (34.725) |
| 01/26/94 | 1.49E-004 | -2.36E-005 | 8.94E-006 | 3.21E-008 | -0.257 | 829013 |
| | (0.757) | (-1.056) | (12.334) | (0.562) | (-5.787) | (15.641) |
| 01/27/94 | -1.20E-004 | 2.97E-005 | 8.18E-006 | 3.99E-017 | 0.256 | 767316 |
| | (-0.541) | (1.404) | (11.245) | ( -5.1E-008) | (4.763) | (14.353) |
| 01/28/94 | -0.001 | 3.37E-004 | 1.20E-005 | 6.75E-014 | 1.125 | 342264 |
| | (-5.749) | (16.455) | (14.095) | ( -2.2E-006) | (14.852) | (13.736) |
| 01/31/94 | -4.45E-005 | 6.29E-006 | 6.92E-006 | 3.36E-008 | -0.416 | 864783 |
| | (-0.316) | (0.380) | (10.954) | (0.436) | (-6.621) | (12.163) |

* Numbers in the parentheses are t-values

Table 3.8: MLE Estimation Results on Equation (3.21) (IBM) when the Number of Trades is Used as the Conditioning Variable

| Date | $\hat{\mu}$ | $\hat{\mu}_Q$ | $\hat{\sigma}^2$ | $\hat{\sigma}_Q^2$ | $\hat{\alpha}$ | $\hat{\beta}$ |
|---|---|---|---|---|---|---|
| 01/03/94 | 9.26E-005 | -6.55E-005 | 3.71E-007 | 6.80E-007 | 0.063 | 990164 |
| | ( 1.334 ) | ( -1.277 ) | ( 13.681 ) | ( 12.816 ) | ( 3.450 ) | ( 18.043) |
| 01/04/94 | 5.40E-005 | -2.77E-005 | 2.47E-007 | 4.46E-007 | 0.040 | 1541930 |
| | ( 0.643 ) | ( -0.507 ) | ( 11.556 ) | ( 13.800 ) | ( 2.413 ) | ( 19.553 ) |
| 01/05/94 | 3.00E-005 | -1.35E-006 | 2.67E-007 | 3.08E-007 | 0.055 | 2009605 |
| | ( 0.541 ) | ( -0.035 ) | ( 16.672 ) | ( 12.957 ) | ( 3.189 ) | ( 21.223 ) |
| 01/06/94 | -7.38E-006 | -2.65E-005 | 3.28E-007 | 5.15E-007 | 0.064 | 1325644 |
| | ( -0.129 ) | ( -0.590 ) | ( 16.292 ) | ( 11.283 ) | ( 3.891 ) | ( 17.039 ) |
| 01/07/94 | -3.43E-005 | 3.56E-005 | 1.06E-007 | 5.67E-007 | -0.008 | 1547819 |
| | ( -0.346 ) | ( 0.498 ) | ( 10.055 ) | ( 16.582 ) | ( -1.153 ) | ( 18.828 ) |
| 01/10/94 | 2.13E-006 | 8.56E-006 | 2.28E-007 | 4.50E-007 | 0.033 | 1627898 |
| | ( 0.037 ) | ( 0.191 ) | ( 17.481 ) | ( 13.483 ) | ( 2.013 ) | ( 18.406 ) |
| 01/11/94 | -2.17E-005 | 1.66E-006 | 2.69E-007 | 5.30E-007 | -0.028 | 1371322 |
| | ( -0.450 ) | ( 0.034 ) | ( 20.217 ) | ( 10.820 ) | ( -2.562 ) | ( 15.386 ) |
| 01/12/94 | 5.00E-005 | -6.16E-005 | 2.41E-007 | 6.08E-007 | 0.030 | 1213033 |
| | ( 0.925 ) | ( -1.140 ) | ( 17.946 ) | ( 11.543 ) | ( 1.963 ) | ( 15.455 ) |
| 01/13/94 | -3.88E-005 | 4.46E-005 | 2.76E-007 | 3.99E-007 | 0.011 | 1586913 |
| | ( -0.775 ) | ( 1.020 ) | ( 19.927 ) | ( 13.099 ) | ( 0.772 ) | ( 19.828 ) |
| 01/14/94 | 2.10E-005 | -4.30E-005 | 2.11E-007 | 4.44E-007 | 0.045 | 1559873 |
| | ( 0.523 ) | ( -1.012 ) | ( 24.277 ) | ( 13.235 ) | ( 3.486 ) | ( 18.371 ) |
| 01/17/94 | 2.19E-005 | -3.17E-005 | 4.26E-007 | 6.56E-007 | 0.037 | 979631 |
| | ( 0.371 ) | ( -0.584 ) | ( 16.900 ) | ( 12.956 ) | ( 2.637 ) | ( 20.980 ) |
| 01/18/94 | -3.08E-005 | 1.03E-005 | 3.70E-007 | 4.99E-007 | 0.089 | 1303136 |
| | ( -0.468 ) | ( 0.222 ) | ( 24.074 ) | ( 9.527 ) | ( 4.399 ) | ( 14.605 ) |
| 01/19/94 | -1.16E-004 | 5.75E-005 | 5.05E-007 | 3.65E-007 | -0.032 | 1376933 |
| | ( -1.677 ) | ( 1.140 ) | ( 13.603 ) | ( 7.357 ) | ( -1.852 ) | ( 15.900 ) |
| 01/20/94 | 4.06E-005 | -5.17E-005 | 4.21E-007 | 5.53E-007 | 0.039 | 1153405 |
| | ( 0.597 ) | ( -0.981 ) | ( 14.042 ) | ( 9.478 ) | ( 1.912 ) | ( 15.159 ) |
| 01/21/94 | 3.56E-005 | -2.25E-005 | 3.33E-007 | 4.28E-007 | 0.021 | 1514227 |
| | ( 0.637 ) | ( -0.507 ) | ( 17.762 ) | ( 10.393 ) | ( 1.211 ) | ( 16.688 ) |
| 01/24/94 | -3.71E-006 | 5.24E-005 | 1.86E-005 | 3.17E-006 | -0.007 | 96231 |
| | ( -0.011 ) | ( 0.493 ) | ( 32.624 ) | ( -9.914 ) | ( -0.807 ) | ( 35.528 ) |
| 01/25/94 | 1.34E-004 | -4.70E-005 | 2.02E-006 | 6.17E-007 | 0.118 | 833443 |
| | ( 0.724 ) | ( -0.994 ) | ( 10.245 ) | ( 9.010 ) | ( 4.901 ) | ( 37.122 ) |
| 01/26/94 | -4.01E-005 | 4.44E-006 | 5.31E-007 | 4.91E-007 | 0.230 | 1215796 |
| | ( -0.501 ) | ( 0.118 ) | ( 12.227 ) | ( 11.135 ) | ( 6.678 ) | ( 15.76848 ) |
| 01/27/94 | -4.89E-005 | 4.30E-005 | 4.52E-007 | 4.95E-007 | 0.035 | 1191080 |
| | ( -0.670 ) | ( 0.779 ) | ( 19.765 ) | ( -10.157 ) | ( 1.972 ) | ( 17.873 ) |
| 01/28/94 | -3.14E-005 | 1.67E-005 | 3.08E-007 | 5.79E-007 | 0.022 | 1188794 |
| | ( -0.622 ) | ( 0.326 ) | ( 19.500 ) | ( 10.141 ) | ( 1.523 ) | ( 15.041 ) |
| 01/31/94 | -8.91E-005 | 2.35E-005 | 4.22E-007 | 4.40E-007 | 0.037 | 1367254 |
| | ( -1.334 ) | ( 0.474 ) | ( 13.929 ) | ( 8.577 ) | ( 1.534 ) | ( 15.040 ) |

* Numbers in the parentheses are t-values

Table 3.9: MLE Estimation Results on Equation (3.21) (INTEL) when the Number of Price Changes is Used as the Conditioning Variable

| Date | $\hat{\mu}$ | $\hat{\mu}_Q$ | $\hat{\sigma}^2$ | $\hat{\sigma}_Q^2$ | $\hat{\alpha}$ | $\hat{\beta}$ |
|---|---|---|---|---|---|---|
| 01/03/94 | 4.39E-005 | 2.28E-005 | 9.12E-006 | 3.64E-014 | 0.266 | 321238 |
| | (0.185) | (0.422) | (12.326) | (-2.35E-007) | (5.073) | (11.988) |
| 01/04/94 | 1.19E-004 | -5.33E-006 | 7.29E-006 | 2.27E-007 | -0.020 | 389596 |
| | (0.478) | (-0.075) | (7.883) | (0.896) | (-0.387) | (15.557) |
| 01/05/94 | -5.53E-005 | 2.62E-005 | 6.97E-006 | 1.91E-007 | -0.002 | 477996 |
| | (-0.226) | (0.466) | (8.877) | (-0.878) | (-0.041) | (15.798) |
| 01/06/94 | 8.37E-004 | -3.98E-004 | 1.18E-005 | 2.96E-014 | 0.526 | 231624 |
| | (3.026) | (-7.727) | (10.949) | (-3.75E-007) | (8.413) | (12.133) |
| 01/07/94 | -9.62E-004 | 4.28E-004 | 1.26E-005 | 6.85E-015 | 0.752 | 306435 |
| | (-3.918) | (11.341) | (13.034) | (-2.64E-007) | (9.465) | (12.311) |
| 01/10/94 | -2.42E-004 | 6.93E-005 | 6.25E-006 | 5.4E-007 | 0.086 | 505050 |
| | (-0.814) | (1.147) | (5.639) | (-2.360) | (1.877) | (19.265) |
| 01/11/94 | 2.89E-005 | 2.22E-005 | 1.02E-005 | 5.09E-013 | 0.185 | 374330 |
| | (0.112) | (0.459) | (9.540) | ( 3.70E-006) | (3.328) | (11.865) |
| 01/12/94 | 2.70E-005 | -5.20E-006 | 9.92E-006 | 1.19E-018 | -0.074 | 540704 |
| | (0.131) | (-0.184) | (20.182) | ( 9.42E-009) | (-3.100) | (23.680) |
| 01/13/94 | 5.897E-004 | -2.37E-004 | 7.08E-006 | 4.10E-018 | 0.384 | 392595 |
| | (2.891) | (-5.978) | (9.530) | (-1.09E-008) | (5.932) | (10.159) |
| 01/14/94 | 1.296E-004 | -3.70E-005 | 9.06E-006 | 1.17E-018 | 0.050 | 433865 |
| | (0.524) | (-0.821) | (10.110) | ( 3.78E-009) | (0.975) | (15.376) |
| 01/17/94 | -1.81E-004 | 2.37E-005 | 6.50E-006 | 9.23E-007 | -0.007 | 359775 |
| | (-0.726) | (0.354) | (9.025) | (-5.750) | (-0.275) | (20.147) |
| 01/18/94 | 1.49E-004 | -2.42E-005 | 1.14E-005 | 5.80E-007 | 0.053 | 334170 |
| | (0.410) | (-0.406) | (12.090) | (3.344) | (1.158) | (27.265) |
| 01/19/94 | -0.00366 | -0.00174 | 1.28E-012 | 5.98E-005 | 12.524 | 1339 |
| | (-8.867) | (-49.338) | (-6.8E-007) | (-4.461) | (4.549) | (4.300) |
| 01/20/94 | 0.00229 | -8.21E-004 | 2.22E-005 | 1.85E-006 | 2.011 | 120470 |
| | (9.122) | (-27.059) | (13.866) | (9.510) | (19.491) | (17.977) |
| 01/21/94 | -1.67E-004 | 4.01E-005 | 6.86E-006 | 5.19E-008 | 0.002 | 689308 |
| | (-0.756) | (1.189) | (9.402) | (0.369) | (0.034) | (16.462) |
| 01/24/94 | -9.70E-005 | 2.79E-005 | 1.38E-005 | 1.29E-021 | -0.033 | 299702 |
| | (-0.290) | (0.388) | (15.391) | (-5.29E-011) | (-3.595) | (23.610) |
| 01/25/94 | 1.88E-004 | -7.05E-005 | 7.51E-006 | 1.35E-006 | 0.019 | 251587 |
| | (0.583) | (-0.647) | (13.289) | (6.948) | (1.268) | (40.919) |
| 01/26/94 | 0.001 | -3.92E-004 | 1.19E-005 | 6.05E-014 | 0.462 | 297088 |
| | (3.558) | (-7.349) | (12.596) | (-3.72E-007) | (6.509) | (12.237) |
| 01/27/94 | 8.13E-006 | 2.84E-006 | 6.86E-006 | 3.88E-007 | 0.050 | 407225 |
| | (0.034) | (0.049) | (6.983) | (-1.655) | (1.095) | (13.391) |
| 01/28/94 | -5.68E-004 | 2.37E-004 | 9.18E-006 | 7.87E-018 | 0.190 | 299799 |
| | (-2.084) | (3.463) | (7.977) | (-4.52E-009) | (3.219) | (11.426) |
| 01/31/94 | -1.44E-004 | 3.01E-005 | 6.92E-006 | 6.09E-008 | -0.241 | 436765 |
| | (-0.732) | (0.530) | (9.902) | (-0.288) | (-4.050) | (11.775) |

* Numbers in the parentheses are t-values

Table 3.10: MLE Estimation Results on Equation (3.21) (IBM) when the Number of Price Changes is Used as the Conditioning Variable

| Date | $\widehat{\mu}$ | $\widehat{\mu}_Q$ | $\widehat{\sigma}^2$ | $\widehat{\sigma}^2_Q$ | $\widehat{\alpha}$ | $\widehat{\beta}$ |
|---|---|---|---|---|---|---|
| 01/03/94 | 5.89E-005 | -1.29E-004 | 3.64E-007 | 2.60E-006 | 0.037 | 255342 |
| | ( 1.566 ) | ( -0.885 ) | ( 27.271 ) | ( 5.297 ) | ( 3.638) | ( 7.332 ) |
| 01/04/94 | 7.76E-005 | -2.06E-004 | 2.48E-007 | 2.58E-006 | 0.033 | 243157 |
| | ( 2.234 ) | ( -1.236 ) | ( 28.674 ) | ( 4.376 ) | ( 4.026 ) | ( 6.077 ) |
| 01/05/94 | 1.58E-005 | 4.90E-005 | 2.84E-007 | 2.14E-006 | 0.004 | 276309 |
| | ( 0.525 ) | ( 0.282 ) | ( 34.075 ) | ( 4.627 ) | ( 0.674 ) | ( 7.614 ) |
| 01/06/94 | -5.52E-005 | 2.05E-005 | 3.32E-007 | 2.37E-006 | 0.019 | 279091 |
| | ( -1.545 ) | ( 0.135 ) | ( 28.126 ) | ( 4.258 ) | ( 2.148 ) | ( 6.254 ) |
| 01/07/94 | -8.86E-006 | 1.06E-004 | 2.20E-007 | 2.56E-006 | 0.017 | 268972 |
| | ( -0.332 ) | ( 0.627 ) | ( 38.656 ) | ( 3.159 ) | ( 3.057 ) | ( 4.379 ) |
| 01/10/94 | 3.79E-005 | -4.75E-005 | 2.93E-007 | 2.08E-006 | 0.014 | 312964 |
| | ( 1.169 ) | ( -0.324 ) | ( 31.357 ) | ( 3.901 ) | ( 1.820 ) | ( 5.829 ) |
| 01/11/94 | -4.31E-005 | 8.50E-005 | 2.34E-007 | 2.56E-006 | 0.019 | 276424 |
| | ( -1.486 ) | ( 0.512 ) | ( 35.148 ) | ( 3.149 ) | ( 2.870 ) | ( 4.296 ) |
| 01/12/94 | 4.09E-005 | -2.47E-004 | 3.00E-007 | 2.46E-006 | 0.008 | 261973 |
| | ( 1.263 ) | ( -1.390 ) | ( 31.646 ) | ( 3.238 ) | ( 1.157 ) | ( 4.865 ) |
| 01/13/94 | -1.53E-005 | 1.85E-004 | 1.95E-007 | 3.35E-006 | 0.017 | 212449 |
| | ( -0.622 ) | ( 0.808 ) | ( 42.182 ) | ( 3.317 ) | ( 2.675 ) | ( 4.366 ) |
| 01/14/94 | 2.67E-005 | -2.40E-004 | 2.02E-007 | 3.07E-006 | 0.026 | 209445 |
| | ( 1.032 ) | ( -1.081 ) | ( 39.583 ) | ( 4.297 ) | ( 3.144 ) | ( 5.669 ) |
| 01/17/94 | -3.02E-005 | -1.76E-005 | 4.57E-007 | 2.50E-006 | 0.016 | 240100 |
| | ( -0.651 ) | ( -0.111 ) | ( 42.806 ) | ( 3.055 ) | ( 1.627 ) | ( 5.001 ) |
| 01/18/94 | -7.62E-005 | 1.50E-004 | 3.93E-007 | 2.07E-006 | 0.045 | 270105 |
| | ( -1.814 ) | ( 1.167 ) | ( 30.615 ) | ( 3.885 ) | ( 4.419 ) | ( 6.247 ) |
| 01/19/94 | -6.49E-005 | 7.66E-005 | 3.91E-007 | 2.43E-006 | 0.014 | 245865 |
| | ( -1.752 ) | ( 0.440 ) | ( 28.120 ) | ( 2.995 ) | ( 1.710 ) | ( 4.816 ) |
| 01/20/94 | -4.87E-005 | 6.29E-005 | 3.97E-007 | 2.58E-006 | 0.009 | 257071 |
| | ( -1.262 ) | ( 0.393 ) | ( 27.684 ) | ( 4.155 ) | ( 0.983 ) | ( 6.145 ) |
| 01/21/94 | 2.22E-005 | -6.17E-005 | 3.09E-007 | 2.90E-006 | 0.007 | 227475 |
| | ( 0.702 ) | ( -0.307 ) | ( 34.496 ) | ( 2.813 ) | ( 0.973 ) | ( 4.277 ) |
| 01/24/94 | 4.54E-005 | 1.38E-004 | 7.67E-006 | 2.50E-005 | 7.57E-004 | 26339 |
| | ( 0.249 ) | ( 0.437 ) | ( -73.583 ) | ( -19.148 ) | ( 0.224 ) | ( 25.795 ) |
| 01/25/94 | 7.70E-005 | -1.33E-004 | 1.20E-006 | 4.15E-006 | 0.073 | 164220 |
| | ( 0.952 ) | ( -0.899 ) | ( 16.352 ) | ( 22.946 ) | ( 3.990) | ( 26.642 ) |
| 01/26/94 | -5.69E-006 | -2.4E-004 | 6.46E-007 | 2.47E-006 | 0.069 | 203008 |
| | ( -0.113 ) | ( -1.650 ) | ( 23.548 ) | ( 7.077 ) | ( 3.822 ) | ( 10.399 ) |
| 01/27/94 | -1.36E-005 | 1.34E-004 | 2.56E-007 | 3.46E-006 | 0.018 | 203408 |
| | ( -0.434 ) | ( 0.686 ) | ( 45.409 ) | ( 4.042 ) | ( 2.893 ) | ( 5.453 ) |
| 01/28/94 | -2.40E-005 | 7.15E-005 | 2.90E-007 | 2.35E-006 | 0.032 | 261737 |
| | ( -0.746 ) | ( 0.489 ) | ( 32.247 ) | ( 3.631 ) | ( 3.452 ) | ( 5.395 ) |
| 01/31/94 | -1.84E-005 | -1.06E-004 | 3.26E-007 | 2.50E-006 | 0.008 | 274554 |
| | ( -0.551 ) | ( -0.664 ) | ( 31.510 ) | ( 3.188 ) | ( 1.146 ) | ( 4.530 ) |

* Numbers in the parentheses are t-values

Table 3.11: Moments and Persistence Level of Information Arrival (N of INTEL)

| Date | E(N) | V(N) | $\widehat{V}$(N) | $\widehat{E}$ | $\widehat{V}(\lambda)$ | $\widehat{\theta}$ | $CF$ |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 5.600 | 31.058 | 5.914 | 5.866 | 4.77E-002 | 1.11E-002 | 1.21E-002 |
| 01/04/94 | 6.079 | 34.567 | 6.716 | 6.459 | 2.57E-001 | 9.13E-003 | 2.24E-002 |
| 01/05/94 | 7.223 | 62.621 | 7.189 | 7.166 | 2.29E-002 | 1.81E-002 | 4.53E-004 |
| 01/06/94 | 5.928 | 31.090 | 5.896 | 5.896 | 1.14E-004 | 6.12E-005 | 1.74E-004 |
| 01/07/94 | 9.779 | 174.604 | 9.834 | 9.749 | 8.49E-002 | 1.15E-004 | 9.23E-005 |
| 01/10/94 | 8.467 | 59.211 | 8.387 | 8.384 | 3.11E-003 | 9.02E-005 | 8.22E-004 |
| 01/11/94 | 6.051 | 30.861 | 6.116 | 6.002 | 1.13E-001 | 2.34E-005 | 4.00E-004 |
| 01/12/94 | 9.831 | 131.946 | 9.709 | 9.631 | 7.73E-002 | 5.69E-005 | 4.15E-003 |
| 01/13/94 | 5.362 | 19.908 | 6.314 | 5.574 | 7.40E-001 | 2.63E-002 | 8.06E-003 |
| 01/14/94 | 7.359 | 60.112 | 7.286 | 7.235 | 5.11E-002 | 7.73E-003 | 2.13E-003 |
| 01/17/94 | 6.787 | 36.626 | 6.746 | 6.744 | 2.53E-003 | 1.60E-001 | 2.74E-004 |
| 01/18/94 | 9.213 | 50.461 | 9.258 | 9.177 | 8.13E-002 | 1.54E-001 | 1.41E-004 |
| 01/19/94 | 33.931 | 571.561 | 1395.703 | 28.760 | 1.37E+003 | 9.51E-002 | 9.30E-001 |
| 01/20/94 | 13.064 | 156.955 | 38.734 | 12.924 | 2.58E+001 | 4.94E-001 | 1.52E-003 |
| 01/21/94 | 9.456 | 69.565 | 9.386 | 9.370 | 1.61E-002 | 2.12E-004 | 7.89E-004 |
| 01/24/94 | 7.438 | 42.987 | 7.386 | 7.376 | 9.12E-003 | 6.59E-005 | 5.21E-004 |
| 01/25/94 | 5.579 | 25.283 | 5.503 | 5.499 | 3.69E-003 | 2.51E-004 | 1.16E-003 |
| 01/26/94 | 7.423 | 40.132 | 7.486 | 7.354 | 1.32E-001 | 2.71E-002 | 6.47E-004 |
| 01/27/94 | 6.597 | 30.951 | 6.668 | 6.537 | 1.31E-001 | 6.78E-004 | 5.51E-004 |
| 01/28/94 | 5.500 | 36.256 | 7.991 | 5.458 | 2.53E+000 | 3.89E-002 | 3.23E-004 |
| 01/31/94 | 5.808 | 55.806 | 6.079 | 5.732 | 3.47E-001 | 2.91E-002 | 1.01E-003 |
| | | | | | | $TCF$ | 9.87E-001 |

E=Mean; V=Variance; $\widehat{E}$=Estimated Mean; $\widehat{E} = \widehat{E}(N) = \widehat{E}(\lambda)$; $\widehat{V}$=Estimated Variance; $\widehat{\theta} = \beta(\widehat{\mu}_Q^2 + \widehat{\sigma}_Q^2)$; $CF$ =Criterion Function=$\frac{[E(N)-\widehat{E}(N)]^2}{\widehat{E}(N)}$; $TCF = \sum_{t=1}^{21} CF_t$;

Table 3.12: Moments and Persistence Level of Information Arrival (N of IBM)

| Date | E(N) | V(N) | $\widehat{V}$(N) | $\widehat{E}$ | $\widehat{V}(\lambda)$ | $\widehat{\theta}$ | $CF$ |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 1.352 | 1.780 | 1.351 | 1.336 | 1.49E-002 | 0.678 | 1.92E-004 |
| 01/04/94 | 1.371 | 2.053 | 1.353 | 1.347 | 6.01E-003 | 0.688 | 4.28E-004 |
| 01/05/94 | 1.562 | 2.324 | 1.564 | 1.554 | 9.83E-003 | 0.620 | 4.12E-005 |
| 01/06/94 | 1.568 | 2.256 | 1.588 | 1.573 | 1.52E-002 | 0.683 | 1.59E-005 |
| 01/07/94 | 1.316 | 1.505 | 1.311 | 1.311 | 5.06E-004 | 0.880 | 1.91E-005 |
| 01/10/94 | 1.537 | 2.048 | 1.516 | 1.512 | 4.70E-003 | 0.732 | 4.13E-004 |
| 01/11/94 | 1.246 | 1.579 | 1.248 | 1.244 | 3.25E-003 | 0.726 | 3.22E-006 |
| 01/12/94 | 1.257 | 1.655 | 1.258 | 1.253 | 4.08E-003 | 0.743 | 1.28E-005 |
| 01/13/94 | 1.256 | 1.847 | 1.235 | 1.234 | 4.01E-004 | 0.637 | 3.92E-004 |
| 01/14/94 | 1.228 | 1.724 | 1.235 | 1.228 | 7.70E-003 | 0.696 | 0.00E+000 |
| 01/17/94 | 1.296 | 2.013 | 1.277 | 1.272 | 4.58E-003 | 0.643 | 4.53E-004 |
| 01/18/94 | 1.618 | 2.216 | 1.661 | 1.634 | 2.73E-002 | 0.650 | 1.57E-004 |
| 01/19/94 | 1.352 | 1.863 | 1.349 | 1.346 | 2.74E-003 | 0.508 | 2.68E-005 |
| 01/20/94 | 1.467 | 1.869 | 1.463 | 1.458 | 5.16E-003 | 0.640 | 5.56E-005 |
| 01/21/94 | 1.504 | 2.271 | 1.496 | 1.495 | 1.46E-003 | 0.649 | 5.42E-005 |
| 01/24/94 | 2.568 | 4.885 | 2.562 | 2.562 | 1.18E-004 | 0.305 | 1.41E-005 |
| 01/25/94 | 3.728 | 6.750 | 3.753 | 3.715 | 3.82E-002 | 0.516 | 4.55E-005 |
| 01/26/94 | 2.129 | 3.539 | 2.257 | 2.121 | 1.37E-001 | 0.597 | 3.02E-005 |
| 01/27/94 | 1.403 | 1.871 | 1.409 | 1.405 | 3.76E-003 | 0.592 | 2.85E-006 |
| 01/28/94 | 1.270 | 1.852 | 1.251 | 1.249 | 1.89E-003 | 0.689 | 3.53E-004 |
| 01/31/94 | 1.538 | 1.596 | 1.550 | 1.545 | 4.28E-003 | 0.603 | 3.17E-005 |
| | | | | | | $TCF$ | 2.74E-003 |

E=Mean; V=Variance; $\widehat{E}$=Estimated Mean; $\widehat{E} = \widehat{E}(N) = \widehat{E}(\lambda)$; $\widehat{V}$=Estimated Variance; $\widehat{\theta} = \beta(\widehat{\mu}_Q^2 + \widehat{\sigma}_Q^2)$; $CF$ =Criterion Function=$\frac{[E(N)-\widehat{E}(N)]^2}{\widehat{E}(N)}$; $TCF = \sum_{t=1}^{21} CF_t$

Table 3.13: Moments and Persistence Level of Information Arrival (NPC of INTEL)

| Date | E(NPC) | V(NPC) | $\widehat{V}$(NPC) | $\widehat{E}$ | $\widehat{V}(\lambda)$ | $\widehat{\theta}$ | $CF$ |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 3.010 | 7.208 | 3.338 | 3.197 | 1.42E-001 | 1.67E-004 | 1.09E-002 |
| 01/04/94 | 3.123 | 6.869 | 3.094 | 3.093 | 8.26E-004 | 8.84E-002 | 2.91E-004 |
| 01/05/94 | 3.656 | 10.776 | 3.666 | 3.666 | 1.05E-005 | 9.18E-002 | 2.73E-005 |
| 01/06/94 | 3.203 | 8.650 | 3.945 | 3.391 | 5.54E-001 | 3.67E-002 | 1.04E-002 |
| 01/07/94 | 4.651 | 18.983 | 6.022 | 4.888 | 1.13E+000 | 5.62E-002 | 1.15E-002 |
| 01/10/94 | 4.472 | 10.111 | 4.490 | 4.474 | 1.60E-002 | 2.75E-001 | 8.94E-007 |
| 01/11/94 | 3.210 | 9.370 | 4.088 | 4.019 | 6.87E-002 | 1.84E-004 | 1.63E-001 |
| 01/12/94 | 5.382 | 38.401 | 5.302 | 5.291 | 1.10E-002 | 1.46E-005 | 1.57E-003 |
| 01/13/94 | 3.092 | 7.565 | 3.528 | 3.233 | 2.95E-001 | 2.21E-002 | 6.15E-003 |
| 01/14/94 | 4.028 | 12.156 | 3.989 | 3.984 | 5.08E-003 | 5.93E-004 | 4.86E-004 |
| 01/17/94 | 3.510 | 10.744 | 3.492 | 3.492 | 1.19E-004 | 3.32E-001 | 9.28E-005 |
| 01/18/94 | 4.818 | 12.967 | 4.809 | 4.803 | 5.88E-003 | 1.94E-001 | 4.68E-005 |
| 01/19/94 | 17.467 | 144.357 | 329.588 | 13.673 | 3.16E+002 | 8.41E-002 | 1.05E+000 |
| 01/20/94 | 6.785 | 60.792 | 15.638 | 6.729 | 8.91E+000 | 3.04E-001 | 4.66E-004 |
| 01/21/94 | 4.954 | 19.982 | 4.915 | 4.915 | 7.15E-006 | 3.69E-002 | 3.09E-004 |
| 01/24/94 | 4.123 | 11.383 | 4.118 | 4.116 | 2.12E-003 | 2.33E-004 | 1.19E-005 |
| 01/25/94 | 2.908 | 6.377 | 2.890 | 2.889 | 7.85E-004 | 3.40E-001 | 1.25E-004 |
| 01/26/94 | 3.933 | 12.448 | 4.617 | 4.189 | 4.28E-001 | 4.56E-002 | 1.56E-002 |
| 01/27/94 | 3.385 | 7.733 | 3.383 | 3.378 | 5.06E-003 | 1.58E-001 | 1.45E-005 |
| 01/28/94 | 2.777 | 6.148 | 3.065 | 2.993 | 7.23E-002 | 1.68E-002 | 1.56E-002 |
| 01/31/94 | 2.900 | 7.530 | 2.973 | 2.857 | 1.16E-001 | 2.70E-002 | 6.47E-004 |
| | | | | | | $TCF$ | 1.29E+000 |

E=Mean; V=Variance; $\widehat{E}$=Estimated Mean; $\widehat{E} = \widehat{E}(NPC) = \widehat{E}(\lambda)$; $\widehat{V}$=Estimated Variance; $\widehat{\theta} = \beta(\widehat{\mu}_Q^2 + \widehat{\sigma}_Q^2)$; $CF$ =Criterion Function=$\frac{\left[E(N)-\widehat{E}(N)\right]^2}{\widehat{E}(N)}$; $TCF = \sum_{t=1}^{21} CF_t$;

Table 3.14: Moments and Persistence Level of Information Arrival (NPC of IBM)

| Date | E(NPC) | V(NPC) | $\widehat{V}$(NPC) | $\widehat{E}$ | $\widehat{V}(\lambda)$ | $\widehat{\theta}$ | $CF$ |
|---|---|---|---|---|---|---|---|
| 01/03/94 | 0.391 | 0.553 | 0.395 | 0.390 | 5.04E-003 | 0.666 | 2.56E-006 |
| 01/04/94 | 0.253 | 0.339 | 0.265 | 0.261 | 3.80E-003 | 0.641 | 2.45E-004 |
| 01/05/94 | 0.209 | 0.290 | 0.202 | 0.202 | 5.75E-005 | 0.590 | 2.43E-004 |
| 01/06/94 | 0.328 | 0.387 | 0.333 | 0.332 | 1.34E-003 | 0.662 | 4.82E-005 |
| 01/07/94 | 0.247 | 0.310 | 0.249 | 0.248 | 1.15E-003 | 0.692 | 4.03E-006 |
| 01/10/94 | 0.303 | 0.361 | 0.302 | 0.301 | 6.54E-004 | 0.650 | 1.33E-005 |
| 01/11/94 | 0.279 | 0.361 | 0.290 | 0.288 | 1.46E-003 | 0.710 | 2.81E-004 |
| 01/12/94 | 0.252 | 0.292 | 0.256 | 0.256 | 2.26E-004 | 0.662 | 6.25E-005 |
| 01/13/94 | 0.200 | 0.227 | 0.208 | 0.207 | 1.17E-003 | 0.719 | 2.37E-004 |
| 01/14/94 | 0.190 | 0.236 | 0.201 | 0.198 | 2.42E-003 | 0.653 | 3.23E-004 |
| 01/17/94 | 0.319 | 0.388 | 0.314 | 0.313 | 7.58E-004 | 0.599 | 1.15E-004 |
| 01/18/94 | 0.349 | 0.495 | 0.356 | 0.349 | 6.08E-003 | 0.566 | 0.00E+000 |
| 01/19/94 | 0.272 | 0.359 | 0.277 | 0.276 | 6.33E-004 | 0.600 | 5.80E-005 |
| 01/20/94 | 0.331 | 0.371 | 0.335 | 0.334 | 2.98E-004 | 0.667 | 2.69E-005 |
| 01/21/94 | 0.234 | 0.309 | 0.225 | 0.225 | 1.50E-004 | 0.658 | 3.60E-004 |
| 01/24/94 | 0.607 | 0.971 | 0.595 | 0.595 | 2.02E-006 | 0.659 | 2.42E-004 |
| 01/25/94 | 0.848 | 1.144 | 0.874 | 0.854 | 2.01E-002 | 0.686 | 4.22E-005 |
| 01/26/94 | 0.420 | 0.523 | 0.424 | 0.411 | 1.30E-002 | 0.512 | 1.98E-004 |
| 01/27/94 | 0.241 | 0.286 | 0.239 | 0.238 | 1.23E-003 | 0.707 | 3.78E-005 |
| 01/28/94 | 0.285 | 0.328 | 0.283 | 0.280 | 3.28E-003 | 0.614 | 8.93E-005 |
| 01/31/94 | 0.315 | 0.402 | 0.313 | 0.313 | 2.36E-004 | 0.688 | 1.28E-005 |
| | | | | | | $TCF$ | 2.64E-003 |

E=Mean; V=Variance; $\widehat{E}$=Estimated Mean; $\widehat{E} = \widehat{E}(NPC) = \widehat{E}(\lambda)$; $\widehat{V}$=Estimated Variance; $\widehat{\theta} = \beta(\widehat{\mu}_Q^2 + \widehat{\sigma}_Q^2)$; $CF$ =Criterion Function=$\frac{\left[E(N)-\widehat{E}(N)\right]^2}{\widehat{E}(N)}$; $TCF = \sum_{t=1}^{21} CF_t$;

Table 3.15: Estimation Results of a GARCH(1,1) Model

| | INTEL | | | IBM | | |
|---|---|---|---|---|---|---|
| Date | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\alpha}+\widehat{\beta}$ | $\widehat{\alpha}$ | $\widehat{\beta}$ | $\widehat{\alpha}+\widehat{\beta}$ |
| 01/03/94 | 0.0750 | 0.8504 | 0.9255 | 0.3032 | 0.3064 | 0.6096 |
| | (2.289) | ( 13.271) | | (3.786) | ( 2.080) | |
| 01/04/94 | 0.1345 | 0.5361 | 0.6706 | 0.1848 | 0.5668 | 0.7516 |
| | (2.123) | ( 2.115) | | (3.559) | ( 7.342) | |
| 01/05/94 | 0.1176 | 0.7069 | 0.8245 | 0.0756 | 0.7296 | 0.8051 |
| | (2.537) | ( 6.235) | | (2.225) | ( 4.646) | |
| 01/06/94 | 0.0643 | 0.8674 | 0.9317 | 0.0628 | 0.9134 | 0.9761 |
| | (2.155) | ( 15.254) | | (3.004) | (28.020) | |
| 01/07/94 | 0.0326 | 0.9475 | 0.9801 | 0.0579 | 0.8950 | 0.9529 |
| | (2.190) | ( 40.441) | | (3.379) | (32.166) | |
| 01/10/94 | 0.1140 | 0.3627 | 0.4767 | 0.0158 | 0.8985 | 0.9142 |
| | (2.126) | ( 1.722) | | (0.882) | ( 8.227) | |
| 01/11/94 | 0.0536 | 0.8071 | 0.8607 | 0.1090 | 0.7308 | 0.8398 |
| | (1.642) | ( 7.651) | | (3.091) | ( 8.371) | |
| 01/12/94 | 0.2070 | 0.5692 | 0.7761 | 0.0856 | 0.6787 | 0.7644 |
| | (2.615) | ( 2.935) | | (2.305) | ( 5.012) | |
| 01/13/94 | 0.0062 | 0.8897 | 0.8959 | 0.1506 | 0.6368 | 0.7874 |
| | (0.210) | ( 3.470) | | (2.948) | ( 4.988) | |
| 01/14/94 | 0.0790 | 0.7777 | 0.8567 | 0.0731 | 0.8973 | 0.9704 |
| | (1.721) | ( 5.739) | | (3.696) | (30.635) | |
| 01/17/94 | 0.2700 | 0.2186 | 0.4885 | 0.1716 | 0.5462 | 0.7178 |
| | (3.109) | ( 1.353) | | (3.215) | ( 4.869) | |
| 01/18/94 | 0.1046 | 0.5887 | 0.6932 | 0.1119 | 0.8736 | 0.9855 |
| | (2.597) | ( 4.144) | | (4.733) | (35.903) | |
| 01/19/94 | 0.4192 | 0.5548 | 0.9740 | 0.0650 | 0.0000 | 0.0650 |
| | (4.337) | ( 7.962) | | (1.456) | ( 0.000) | |
| 01/20/94 | 0.0865 | 0.8616 | 0.9481 | 0.1147 | 0.4616 | 0.5762 |
| | (3.228) | ( 20.637) | | (2.297) | ( 2.914) | |
| 01/21/94 | 0.0679 | 0.7176 | 0.7855 | 0.0858 | 0.0654 | 0.1512 |
| | (2.010) | ( 6.288) | | (2.177) | ( 0.090) | |
| 01/24/94 | 0.2779 | 0.2249 | 0.5028 | 0.6866 | 0.3134 | 1.0000 |
| | (2.426) | ( 1.385) | | (8.520) | ( 7.587) | |
| 01/25/94 | 0.4140 | 0.1855 | 0.5995 | 0.1218 | 0.8451 | 0.9669 |
| | (3.627) | ( 0.689) | | (3.180) | (17.185) | |
| 01/26/94 | 0.0980 | 0.8135 | 0.9115 | 0.1146 | 0.8103 | 0.9249 |
| | (2.897) | ( 13.318) | | (3.352) | (14.984) | |
| 01/27/94 | 0.0587 | 0.6096 | 0.6683 | 0.1547 | 0.7330 | 0.8877 |
| | (1.248) | ( 2.215) | | (3.813) | (11.422) | |
| 01/28/94 | 0.0778 | 0.8468 | 0.9246 | 0.1005 | 0.7660 | 0.8665 |
| | (2.555) | ( 13.144) | | (3.414) | (11.561) | |
| 01/31/94 | 0.0275 | 0.7464 | 0.7738 | 0.0439 | 0.5490 | 0.5928 |
| | (0.770) | ( 2.438) | | (1.204) | ( 2.080) | |

GARCH(1,1) model: $r_t = \varepsilon_t$, $\varepsilon_t = \nu_t \cdot h_t^{1/2}$, $\nu_t \sim N(0,1)$
$h_t = a_0 + \alpha\varepsilon_{t-1}^2 + \beta\ h_{t-1}$; Numbers in the parentheses are t-values

# Chapter 4

# Modelling Volatility, Correlation, and Market Asymmetry Using Trading Information and Buy/Sell Signals

## 4.1 Introduction

Although the models proposed in the third chapter are able to capture the arrival of the number of trades and pin down the volatility persistence, it fails to describe the other stylized fact, significant negative first-order autocorrelation. In the preliminary analysis in Chapter 2, we show evidence that different types of buy/sell trades do not arrive independently. Whereas buy or sell information may influence the price, returns will be influenced by whether buy followed sell or any of the three other possibilities. In this chapter, we explore the usefulness of the extra buy/sell information in explaining the returns autocorrelation.

A straightforward way of incorporating buy/sell information, is to decompose transactions into the following four types: (1) *BS*: a buyer-initiated transaction followed by a seller-initiated transaction; (2) *SB*: a seller-initiated transaction followed by a buyer-initiated transaction; (3) *BB*: a buyer-initiated transaction followed by another buyer-initiated transaction; (4) *SS*: a seller-initiated transaction followed by another seller-initiated transaction.

According to this decomposition, we modify our homogeneous mixed jump-diffusion model in Chapter 3 correspondingly. We show that this model allows us to examine the impact of buy/sell information within a simple regression framework. Most importantly, the significant negative first-order autocorrelation is on average reduced by 40% within this new framework. Furthermore, once we have captured this feature, conditional volatility is reduced.

In addition to explaining the autocorrelation, we also use buy/sell signals in exploring market asymmetry in a threshold autoreregresive (TAR) framework. TAR models are capable of capturing time–irreversibility, asymmetric limit cycle and jump phenomenon. The major criticism of TAR models comes from the fact that researchers rarely know which state of the world they are currently in, which hinders the application of the TAR models in economic time series. In this chapter, we demonstrate that the TAR model is a promising alternative for analyzing high frequency time series. We argue that the criticism of the TAR model is not always justified, especially when we have exogenously observable information to determine the threshold. Specifically, we use the sign of net buy/sell trading volume from the preceding period as the switching rule and develop a qualitative threshold autoregressive model with conditional heteroskedasticity, where both the conditional mean and conditional variance are regime-dependent. We show this model has a lower volatility persistence and outperforms a benchmark GARCH(1,1) specification in terms of data fitting. We also compare their in-sample prediction and the one-step-ahead out-of-sample forecasting performance. Surprisingly, our QTAR-GARCH model consistently outperforms the GARCH(1,1) model throughout the 22 examined trading days. Although the issue of fitting the trading process as in our extended mixed jump-diffusion model is not discussed here, our QTAR-GARCH model is endowed with superior predicting power and thus provides a good alternative to the GARCH models in studying nonlinearity existing in financial data.

In the next section, we start with a short survey on empirical evidence that show significant negative autocorrelation in high frequency returns data, followed by some proposed theoretical explanations. We also provide general conditions under which large negative autocorrelation in returns could occur. We detail the construction of our extended mixed jump-diffusion model with buy/sell information and the estimation results in Section 4.3. In Section 4.4, we present our QTAR-GARCH model and the estimation results. In the final section, we discuss extensions of our model, and conclude the paper.

## 4.2 AUTOCORRELATION

As demonstrated in Chapter 2, one salient feature of the intraday individual stock returns series is that most of them have extremely high negative first-order autocorrelation. Most literature on the intraday analysis of exchange rates and stock indices note the presence of significant negative but small, first-order autocorrelation, see Andersen and Bollerslev (1997), Goodhart and Figliuoli (1991), and Goodhart and O'Hara (1997). Earlier research by Roll (1983) suggests that bid/ask bounce can cause negative serial correlation in transactional price changes, although Roll is unable to demonstrate this point by using daily data from AMEX and NYSE listed stocks. Zhou (1996) reports significant large negative first-order autocorrelation for tick-by-tick DM/US$ and JPY/US$ exchange rates. He models the logarithm of exchange rates by a Brownian motion with noise, and shows that, as the sampling time span keeps shrinking, the market noise is going to dominate the variation that comes from the diffusion term. The sample autocorrelation of such return series is thus -50%. Other explanations for the negative first-order autocorrelation include: nonsynchronous trading of Lo and McKinlay (1990), and brokers' inventory considerations of Goodhart and O'Hara (1997). It would seem prudent at this point to address the question of what sort of assumptions are required on the price process to imply negative autocorrelation in returns. These are presented in the following Lemma.

**Lemma 4.2.1** *If $\ln P_t$ follows a stationary process with a negative first-order autocorrelation, the associated returns process $X_t = \ln P_t - \ln P_{t-1}$ must have negative first-order autocorrelation.*

**Proof.**

Consider the autocovariance between $X_t$ and $X_{t-1}$

$$\begin{aligned} Cov(X_t, X_{t-1}) &= Cov(\ln P_t - \ln P_{t-1}, \ln P_{t-1} - \ln P_{t-2}) \\ &= Cov(\ln P_t, \ln P_{t-1}) - Var(\ln P_{t-1}) \end{aligned}$$

$$
\begin{aligned}
& -Cov(\ln P_t, \ln P_{t-2}) + Cov(\ln P_{t-1}, \ln P_{t-2}) \\
= & -\gamma_0 + 2\gamma_1 - \gamma_2 \\
= & -\gamma_0(1 - 2\rho_1 + \rho_2) \qquad (4.1)
\end{aligned}
$$

where $\gamma_j = Cov(\ln P_t, \ln P_{t-j})$ and $\rho_j = \gamma_j/\gamma_0$.

For $Cov(X_t, X_{t-1})$ to be negative requires $1 - 2\rho_1 + \rho_2 > 0$. Thus examining the autocorrelation matrix for a stationary process, we have, for some $3 \times 3$ matrix $\Omega$

$$
\Omega = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix} \qquad (4.2)
$$

and since $\Omega$ is positive definite, we note, from the Hawkins-Simons condition for positive definiteness,

(i) $1 - \rho_1^2 > 0$.

(ii) $|\Omega| = (1 - \rho_2)(1 + \rho_2 - 2\rho_1^2) > 0$.

(ii) implies $1 + \rho_2 > 2\rho_1^2$ and thus

$$
1 - 2\rho_1 + \rho_2 > 2\rho_1(\rho_1 - 1) > 0, \quad \text{if } \rho_1 < 0 \qquad (4.3)
$$

Thus any stationary process with $\rho_1 < 0$ will give negative first order autocorrelation in first differences. ■

The result in the above Lemma cannot be extended to cover general stationary processes with positive first-order autocorrelation.[1] However, we note that $Cov(X_t, X_{t-1})$ will be non-positive for all AR(1) and MA(1) processes as we now show

AR(1): $x_t = \phi x_{t-1} + \varepsilon_t$ with $\rho_1 = \phi$, $\rho_2 = \phi^2$, gives $1 - 2\rho_1 + \rho_2 = (1 - \phi)^2 > 0$.

MA(1): $x_t = \varepsilon_t + \theta\varepsilon_{t-1}$with $\rho_1 = \theta/(1 + \theta^2)$, $\rho_2 = 0$, gives $1 - 2\rho_1 + \rho_2 = \frac{(1-\theta)^2}{1+\theta^2} \geq 0$.

[1] A straightforward counter-example can be constructed from a MA(2) process. Let $P_t = v_t + \phi_1 v_{t-1} + \phi_2 v_{t-2}$, where $v_t \sim i.i.d.(0, 1)$. In this case, $\rho_1 = (\phi_1 + \phi_1\phi_2)/(1 + \phi_1^2 + \phi_2^2)$. If $\phi_1 = 0.8$ and $\phi_2 = 0.1$, $\rho_1 \simeq 0.533 > 0$. Thus the result does not follow for this MA(2) process.

## 4.3 An Extended mixed jump-diffusion Model

In this section, we extend our initial models analyzed in Sections 3.2 and 3.3 by incorporating buy/sell information to help us explain the significant negative autocorrelation in the data.

Our motivation comes form the fact that the initial models in Chapter 3 were unable to capture the significant first-order autocorrelation in the data. We find that the decrease in the autocorrelation, after incorporating the number of trades in the model, is less than 5% for all three stocks. Although our model in the previous section captures the information arrival very well, it also predicts, in equation (3.19), that the autocorrelation in returns is positive, which contradicts the empirical facts presented in Chapter 2. To overcome this difficulty, we note, in Tables 2.15 to 2.22, that different types of transactions do not arrive randomly. This evidence indicates that buy/sell transitions do contain information and would help in explaining the significant negative autocorrelation.

Notice that, by definition, the sum of buyer-initiated and seller-initiated transactions is equal to the total number of external transactions. Whereas buy or sell information may influence the price, returns will be influenced by whether buy followed sell or any of the three other possibilities. A straightforward way of incorporating buy/sell information, is to decompose transactions into the following four types: (1) *BS*: a buyer-initiated transaction followed by a seller-initiated transaction; (2) *SB*: a seller-initiated transaction followed by a buyer-initiated transaction; (3) *BB*: a buyer-initiated transaction followed by another buyer-initiated transaction; (4) *SS*: a seller-initiated transaction followed by another seller-initiated transaction. According to this decomposition, we modify our initial model correspondingly. First of all, the price generation process described in equation (3.1) becomes

$$P(t) = P(0)\exp[(\alpha - \frac{1}{2}\sigma^2)t + \sigma(z(t) - z(0)) + \sum_{i=1}^{4}\sum_{j=1}^{N_i(t)} Q_{ij}] \tag{4.4}$$

where $Q_{ij} \sim ind\ N(\mu_i, \sigma_i^2)$. Assuming $\Delta N_i(t)$, $i = 1, \ldots, 4$, are jointly Poisson distributed, we can rewrite the returns process described in equation (3.2) as

$$X(t) = (\alpha - \frac{1}{2}\sigma^2) + \sigma(z(t) - z(t-1)) + \sum_{i=1}^{4}\sum_{j=1}^{\Delta N_i(t)} Q_{ij} \tag{4.5}$$

For simplicity, we also set $\alpha - \frac{1}{2}\sigma^2 = 0$, and $\sigma_i^2 = \sigma_Q^2$, $i = 1, \ldots, 4$. Consequently, the conditional density function of $X(t)$ becomes

$$pdf(X(t) \mid \Delta N_i(t)) = \phi(\sum_{i=1}^{4} \mu_i \Delta N_i(t), \sigma^2 + \sigma_Q^2 \Delta N(t)) \tag{4.6}$$

where $\Delta N(t) = \sum_{i=1}^{4} \Delta N_i(t)$. We set $\Delta N_1 = \Delta N_{BS}, \Delta N_2 = \Delta N_{SB}, \Delta N_3 = \Delta N_{BB}, \Delta N_4 = \Delta N_{SS}$, and likewise for the mean parameters $\mu_i$'s. This allows us to examine the impact of buy/sell information within a simple regression framework stated as follows;

$$X(t) = \mu_{BS}\Delta N_{BS} + \mu_{SB}\Delta N_{SB} + \mu_{BB}\Delta N_{BB} + \mu_{SS}\Delta N_{SS} + \sqrt{\sigma^2 + \sigma_Q^2 \Delta N(t)}\varepsilon(t) \tag{4.7}$$

We estimate this new model by the same iterative feasible generalized least squares procedure described in last section. The results are presented in Table 4.1. We find that the four new explanatory variables are all highly significant under conventional significance levels. We also examine the autocorrelation of the fitted residuals. As shown in Table 4.2, we observe that, on average, the first-order autocorrelation drops by 40 percent after incorporating buy/sell information in the conditional mean. The buy/sell information also help in explaining intraday volatility, this could be seen by comparing the estimated conditional variance term $\widehat{\sigma}^2 + \widehat{\sigma}_Q^2 \Delta N(t)$ in equation (3.5) and equation (4.7). We estimate equation (3.5) using BT data and present the estimation results in Table 4.3. Comparing Table 4.1 with Table 4.3, we find that both $\widehat{\sigma}^2$ and $\widehat{\sigma}_Q^2$ are reduced for every day once buy/sell information is incorporated in the estimation. Therefore, the conditional variance $\widehat{\sigma}^2 + \widehat{\sigma}_Q^2 \Delta N(t)$ is also reduced.

It might be possible to extend our second model a little further by postulating a joint *pdf* for $\Delta N_i(t)$, $i = 1, \ldots, 4$, and deriving a relationship similar to equation (3.21). This would require modelling some joint Poison process and finding the joint *pdf* of $\Delta N_i(t)$ and $\Delta N(t)$. One model that could be used is the continuous time time-homogeneous Markov chain defined on a finite state space $S^*$ of dimension 4, where $S^* = \{BS, SB, BB, SS\}$. We would need to specify a $4 \times 4$ generator matrix $\Lambda = [(\lambda_{ij})]$ where $\lambda_{ij} \geq 0$, $i, j \in S^*$, and where $\lambda_{ij}$ represents an intensity or infinitesimal probability of moving from one state to the rest. The probability that the system is in state $j$ at time $t$, given it was in an initial state $i$ at time 0, is given by $P_{ij}(t)$. This probability is the $i, j^{th}$ element of $P(t)$ and can be computed from

$$P(t) = \exp(\Lambda t) \tag{4.8}$$

which is interpreted as a matrix equation, see Chiang (1968), especially chapter 7. Over a given time interval, it is possible, but complicated, to compute the joint *pdf* of $\Delta N_i(t)$, the number of times the system is in each of the four states. This is a fascinating extension but would require considerable further analysis.

## 4.4 The QTAR-GARCH Model

The autoregressive conditional heteroskedastic (ARCH) model of Engle (1982) and the generalized ARCH (GARCH) model of Bollerslev (1987) are very popular in financial modelling, and are capable of capturing the commonly observed volatility clustering phenomenon. A GARCH model can be specified as the following

$$\begin{aligned} X_t &= \varepsilon_t \quad (4.9)\\ \text{where } \varepsilon_t &= \nu_t h_t^{1/2},\ \nu_t \sim iid\,(0,1),\\ \text{and } h_t &= \alpha + \beta_1 \varepsilon_{t-1}^2 + .... + \beta_p \varepsilon_{t-p}^2 + \gamma_1 h_{t-1} + .... + \gamma_q h_{t-q} \end{aligned}$$

However, this generic GARCH model has several shortcomings. First of all, a high GARCH measure of volatility persistence may be due to the failure to take into account of structural shifts in the model (Lamoureux and Lastrapes (1990b)). Furthermore, the prediction power of a generic GARCH is very poor when structural changes do exist (Hamilton and Susmel (1994)). Secondly, the generic GARCH model does not consider market asymmetry. In this section, we explore the usefulness of buy/sell signals in capturing market asymmetry. We shall develop a qualitative threshold framework with conditional heteroskedasticity, where both the conditional mean and conditional variance are regime-dependent. With additional information obtained from trading activities, we expect our model to have a better predicting power than the GARCH model.

A time series $X_t$ is a TAR model if it has the following functional form

$$X_t = \phi_0^i + \phi_1^i X_{t-1} + \ldots + \phi_p^i X_{t-p} + \varepsilon_t^i, \text{ if } \theta_{t-d} \in L_i,\ i = 1, 2, \ldots, k \tag{4.10}$$

where $L_i$ form a mutually exclusive partition of the real line in the sense that $\cup_{i=1}^{k} L_i = R$, and $L_i \cap L_j = \phi$, if $i \neq j$; $p$ is the $AR$ order; $\theta$ is the decision rule variable; $d$ is the delay parameter; $k$ is the threshold parameter; and $\{\varepsilon_t^i\}$ is a sequence of $i.i.d.$ random variables with zero mean and variance $\sigma_i^2$, $\{\varepsilon_t^i\}$ and $\{\varepsilon_t^j\}$ are independent whenever $i \neq j$. When $\theta_{t-d} = X_{t-d}$, i.e. decision rules depend on some past value of the dependent variable, the model in equation (4.10) becomes a self-exciting TAR (SETAR) model of Tong (1983). In the application, $p$ is chosen by the best linear $AR$ model, while the choice of $d$ and $k$ are dependent on the data according to some threshold linearity tests.

Obviously, the switching rule does not have to depend on the lagged dependent variable. In this paper, we propose to use the signs of net buy/sell trading volume as the decision rule variable. While constructing the 1-minute returns series, we also calculate the total trading volume initiated by buyers and that by sellers in each one-minute interval.[2] We construct a dummy series ($D_t$) according to the relative magnitudes of buyer-initiated trading volume ($BV_{t-1}$) and seller-initiated trading volume ($SV_{t-1}$) from the preceding period. Specifically, we assign

$$D_t = \begin{cases} 1, & \text{if } (BV_{t-1} - SV_{t-1}) > 0 \\ 0, & \text{if } (BV_{t-1} - SV_{t-1}) < 0 \end{cases} \tag{4.11}$$

where the subscript, $t \in N$, indicates time. Subsequently, we name periods associated with $D_t = 1$ as buyer-dominating periods, and those associated with $D_t = 0$ as seller-dominating periods. Note that $BV$ and $SV$ are the accumulated trading volume in each period and thus reflect market activities better than the price series.

It remains to justify why $D_t$ is a suitable threshold indicator. Our explanations are as follows. First of all, as demonstrated in Chapter 2, buyer-initiated trades and seller-initiated trades do not arrive independently, and possess different levels of persistence. Clearly, buy/sell transitions do contain valuable information. Secondly, due to the risk that market makers bear, we believe that returns should perform asymmetrically when a market is dominated by buy-contracts and when a market is dominated by sell-contracts. Thirdly, as pointed out by Ross (1987), the operation of financial markets is far from that

[2]A trade with the market maker as the buyer is classified as a seller-initiated trade, while a trade with the market maker as the seller is classified as a buyer-initiated trade.

of a Walrasian competitive market, and tends to be in a sequence of disequilibrium. This is especially true in short time horizons. Hence, excess demand is likely to reveal information about the disequilibrium asset dynamics. In other words, net buy/sell trading volume from the preceding period should at least reveal as much information as lagged returns. Furthermore, due to the re-sampling scheme commonly used in constructing fixed-interval high frequency returns data, the selected representative price in each interval and hence returns may not fully reflect current market activities and information contained therein.

We present our model as follows

$$\begin{aligned} X_t &= D_t\alpha_{01} + (1-D_t)\alpha_{02} + [D_t\alpha_{11} + (1-D_t)\alpha_{12}]\,X_{t-1} + \varepsilon_t \qquad (4.12) \\ &\quad \text{where } \varepsilon_t = \nu_t h_t^{1/2},\ \nu_t \sim N(0,1), \\ &\quad \text{and } h_t = D_t\beta_{01} + (1-D_t)\beta_{02} + \beta_1\varepsilon_{t-1}^2 + \beta_2 h_{t-1} \end{aligned}$$

where the lagged return $X_{t-1}$ is specified in the conditional mean equation to capture the significant first-order autocorrelation in the 1-minute data. Since the dummy, $D_t$, is a qualitative variable, we call this model a qualitative threshold autoregressive model with GARCH(1,1) conditional heteroskedasticity (QTAR-GARCH). Note that only the intercept terms in the conditional variance equation are regime dependent. We specify our model as such to avoid numerical problems and over-parameterization as suggested in Hamilton and Susmel (1994).

We estimate this model by quasi maximum likelihood (QML) and present the results in Table 4.4. The table shows at least two major findings. First of all, nonlinearity not only exists in the conditional variance but also exists in the conditional mean of intraday high frequency returns data. Secondly, buyer-dominating regimes are associated with negative returns ($\widehat{\alpha}_{01}$), higher serial correlation ($\widehat{\alpha}_{11}$) and higher volatility ($\widehat{\beta}_{01}$), while seller-dominating regimes are associated with the opposite. These results accord quite well with the hypothesis that financial markets constantly stay in disequilibrium states in short time horizons. In addition, in Table 4.5, we find that the volatility persistence measure $\beta_1 + \beta_2$ has been reduced in our model by more than 20% in 10 out of 22 days, while only slightly increases in 2 days. This finding is in accordance with the empirical evidence that high GARCH measure of volatility persistence may arise as a result of mis-specifying existing structural changes.

We compare our QTAR-GARCH model with a linear regression model (OLS, hereafter) and a GARCH(1,1) models. In Table 4.6, we report their log-likelihood functions, and Akaike Information Criterion (AIC). The QTAR-GARCH model clearly dominates the other two models in both criteria throughout the 22 trading dates. To examine their relative forecasting performance, we first compare their in-sample prediction on the conditional mean and conditional variance.. For the conditional mean, the criterion we use is the mean squared error (MSE) between the observed and fitted returns

$$MSE = \frac{1}{\text{\# of obs}} \sum_{t=1}^{\text{\# of obs}} (X_t - \widehat{X}_t)^2 \tag{4.13}$$

For the conditional variance, we use the mean absolute error (MAE) between the fitted squared residuals and the fitted conditional variance

$$MAE = \frac{1}{\text{\# of obs}} \sum_{t=1}^{\text{\# of obs}} |\widehat{\varepsilon}_t^2 - \widehat{h}_t| \tag{4.14}$$

In the OLS case, $\widehat{h}_t = \widehat{\sigma}^2$. The results are reported in Table 4.7. For almost every case, our QTAR-GARCH model has the smallest values of MSE and MAE. To strengthen our confidence on the QTAR-GARCH model, we perform one-step ahead out-of-sample prediction on all three models. We reserve the first 400 observations (about 80% of the sample sizes) for estimating the parameters. Based on these parameter estimates, we generate one-step ahead prediction iteratively, i.e. on the $401^{th}$ observation, then the $402^{th}$,..., etc. . The performance criteria we use for the conditional mean and variance are similar to the in-sample prediction measures. For the conditional mean,

$$MSE = \frac{1}{(\text{\# of obs} - 400)} \sum_{t=401}^{\text{\# of obs}} (X_t - \widehat{X}_t)^2 \tag{4.15}$$

For the conditional variance,

$$MAE = \frac{1}{(\text{\# of obs} - 400)} \sum_{t=401}^{\text{\# of obs}} |\widehat{\varepsilon}_t^2 - \widehat{h}_t| \tag{4.16}$$

We report the results in Table 4.8. It is interesting to note that, in most cases, the QTAR-GARCH model consistently outperforms the other two models in predicting both the conditional mean and the conditional variance. All these results are quite robust throughout all 22 days of data that we examine, and hence provide strong support for the potential usefulness of the qualitative threshold model in analyzing the dynamics of high frequency equity returns.

## 4.5 Conclusions

In this chapter, we investigate the information content of trading activities. In particular, we demonstrate the usefulness of the extra buy/sell information in explaining the returns autocorrelation and market asymmetry. With additional information on buy/sell transitions from the UK stocks, we have pinned down a significant amount of negative autocorrelation in returns. We propose two models, of which the extended mixed jump-diffusion model is better at reducing the conditional volatility, while the QTAR-GARCH model is endowed with good predicting power. Both of these two models outperform the popular GARCH(1,1) model. The conclusion of our analysis is that buy/sell transitions do contain information. They augment the information arrival models we use for the US data to explain the negative autocorrelation in returns.

A potentially interesting extension of our mixed jump-diffusion model is to derive the joint or marginal distribution of stock returns with buy/sell information. We have not done this due to the complexity introduced by the assumption that the four buy/sell variables are jointly Poisson distributed. Regarding our QTAR-GARCH model, a useful extension is to introduce a smooth transition between the buyer-dominating and seller-dominating regimes. Teräsvirta (1994) suggests a logistic formulation of the transition. However, Schlittgen (1997), from a simulation on Teräsvirta's model, finds that it is almost impossible to estimate the parameters due to the convergence problem. We leave these problems for future research.

Table 4.1: Estimation Results of Equation (4.7) using Buy/Sell Information

| Date | $\widehat{\mu}_{BS}$ | $\widehat{\mu}_{SB}$ | $\widehat{\mu}_{BB}$ | $\widehat{\mu}_{SS}$ | $\widehat{\sigma}^2$ | $\widehat{\sigma}^2_Q$ |
|---|---|---|---|---|---|---|
| 08/01/94 | 0.00208 | -0.00182 | 0.00162 | -1.71E-004 | 3.55E-006 | 1.06E-007 |
| | ( 10.308 ) | ( -8.489 ) | ( 4.352 ) | ( -3.560 ) | ( 6.740 ) | ( 0.500 ) |
| 08/02/94 | 0.00198 | -0.00124 | 5.41E-004 | -2.31E-004 | 2.69E-006 | 3.75E-007 |
| | ( 9.275 ) | ( -5.746 ) | ( 2.122 ) | ( -3.839 ) | ( 5.979 ) | ( 1.777 ) |
| 08/03/94 | 0.00325 | -0.00248 | 0.00193 | -2.67E-004 | 7.19E-006 | 6.49E-007 |
| | ( 10.836 ) | ( -7.942 ) | ( 3.611 ) | ( -3.331 ) | ( 5.947 ) | ( 1.275 ) |
| 08/04/94 | 0.00174 | -0.00153 | 0.00143 | -2.50E-004 | 2.52E-006 | 5.67E-007 |
| | ( 8.085 ) | ( -6.990 ) | ( 4.601 ) | ( -3.711 ) | ( 4.838 ) | ( 2.129 ) |
| 08/05/94 | 0.00279 | -0.00190 | 0.00119 | -3.33E-004 | 3.04E-006 | 9.22E-007 |
| | ( 11.204) | ( -7.323 ) | ( 2.780 ) | ( -4.452 ) | ( 5.369 ) | ( 3.177 ) |
| 08/08/94 | 0.00237 | -0.00136 | 0.00123 | -2.37E-004 | 3.53E-006 | 3.43E-007 |
| | ( 9.735 ) | ( -5.404) | ( 3.534 ) | ( -4.427 ) | ( 4.669 ) | ( 1.173 ) |
| 08/09/94 | 0.00187 | -0.00163 | 0.00171 | -1.87E-004 | 2.78E-006 | 1.63E-007 |
| | ( 9.548 ) | ( -8.004 ) | ( 5.718 ) | ( -3.483 ) | ( 6.122 ) | ( 0.745 ) |
| 08/10/94 | 0.00281 | -0.00206 | 7.40E-004 | -3.02E-004 | 3.32E-006 | 5.84E-007 |
| | ( 12.454 ) | ( -8.829 ) | ( 3.216 ) | ( -3.835 ) | ( 6.261 ) | ( 2.021 ) |
| 08/11/94 | 0.00258 | -0.00133 | 0.00162 | -3.63E-004 | 2.08E-006 | 1.03E-006 |
| | ( 12.204 ) | ( -6.200 ) | ( 4.219 ) | ( -4.876 ) | ( 4.469 ) | ( 4.173 ) |
| 08/12/94 | 0.00179 | -0.00122 | 0.00120 | -1.87E-004 | 2.93E-006 | 2.54E-007 |
| | ( 7.618 ) | ( -5.137 ) | ( 3.803 ) | ( -2.876 ) | ( 5.823 ) | ( 0.877 ) |
| 08/15/94 | 0.00215 | -0.00134 | 4.60E-004 | -3.48E-004 | 2.66E-006 | 7.67E-007 |
| | ( 10.646 ) | ( -6.596 ) | ( 1.841 ) | ( -4.644 ) | ( 5.586 ) | ( 3.370 ) |
| 08/16/94 | 0.00210 | -0.00151 | 8.34E-004 | -4.36E-004 | 3.15E-006 | 1.04E-006 |
| | ( 9.282 ) | ( -6.623 ) | ( 3.350 ) | ( -4.833 ) | ( 6.54117 ) | ( 4.110 ) |
| 08/17/94 | 0.00180 | -0.00148 | 7.4E-004 | -2.94E-004 | 2.96E-006 | 2.41E-007 |
| | ( 10.372 ) | ( -8.274 ) | ( 4.252 ) | ( -4.316 ) | ( 7.349 ) | ( 1.305 ) |
| 08/18/94 | 0.00283 | -0.00211 | 0.00154 | -3.56E-004 | 3.98E-006 | 6.41E-007 |
| | ( 12.349 ) | ( -8.587 ) | ( 4.480 ) | ( -4.978 ) | ( 3.193 ) | ( 1.103 ) |
| 08/19/94 | 0.00256 | -0.00124 | 5.39E-004 | -1.94E-004 | 2.18E-006 | 2.02E-007 |
| | ( 13.858 ) | ( -6.295 ) | ( 1.667 ) | ( -4.105 ) | ( 5.560 ) | ( 1.093 ) |
| 08/22/94 | 0.00285 | -0.00193 | 3.74E-004 | -3.4E-004 | 2.72E-006 | 9.18E-007 |
| | ( 12.846 ) | ( -8.226 ) | ( 0.889 ) | ( -4.810 ) | ( 4.462 ) | ( 3.201 ) |
| 08/23/94 | 0.00234 | -0.00187 | 0.00141 | -5.10E-004 | 3.69E-006 | 1.32E-006 |
| | ( 9.461 ) | ( -7.146 ) | ( 5.117 ) | ( -4.515 ) | ( 6.461 ) | ( 3.923 ) |
| 08/24/94 | 0.00177 | -0.00128 | 8.79E-004 | -2.66E-004 | 3.73E-006 | 3.70E-007 |
| | ( 8.618 ) | ( -5.939 ) | ( 2.810 ) | ( -3.470 ) | ( 7.352 ) | ( 1.475 ) |
| 08/25/94 | 0.00193 | -0.00160 | 0.00128 | -1.68E-004 | 2.27E-006 | 1.18E-007 |
| | ( 12.322 ) | ( -9.762 ) | ( 5.317 ) | ( -3.729 ) | ( 6.554 ) | ( 0.804 ) |
| 08/26/94 | 0.00190 | -0.00165 | 9.26E-004 | -2.01E-004 | 2.38E-006 | 6.81E-007 |
| | ( 9.034 ) | ( -7.282 ) | ( 2.994 ) | ( -3.180 ) | ( 5.251 ) | ( 3.424 ) |
| 08/30/94 | 0.00239 | -0.00201 | 4.76E-004 | -1.46E-004 | 4.10E-006 | 2.73E-007 |
| | ( 9.909 ) | ( -8.236 ) | ( 1.628 ) | ( -2.559 ) | ( 6.817 ) | ( 1.186 ) |
| 08/31/94 | 0.00206 | -0.00152 | 7.54E-004 | -1.91E-004 | 2.60E-006 | 3.21E-007 |
| | ( 9.802 ) | ( -6.932 ) | ( 2.464 ) | ( -3.440 ) | ( 5.444 ) | ( 1.399 ) |

* Numbers in the parentheses are t-values

Table 4.2: First-Order Autocorrelation (BT)

| Date | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| 08/01/94 | -0.429 | -0.424 | -0.261 | -39.2% |
| 08/02/94 | -0.346 | -0.339 | -0.243 | -29.8% |
| 08/03/94 | -0.482 | -0.474 | -0.249 | -48.3% |
| 08/04/94 | -0.437 | -0.430 | -0.277 | -36.6% |
| 08/05/94 | -0.467 | -0.458 | -0.368 | -21.2% |
| 08/08/94 | -0.406 | -0.399 | -0.293 | -27.8% |
| 08/09/94 | -0.397 | -0.390 | -0.248 | -37.5% |
| 08/10/94 | -0.445 | -0.430 | -0.247 | -44.5% |
| 08/11/94 | -0.424 | -0.415 | -0.229 | -46.0% |
| 08/12/94 | -0.292 | -0.289 | -0.220 | -24.7% |
| 08/15/94 | -0.479 | -0.448 | -0.248 | -48.2% |
| 08/16/94 | -0.406 | -0.381 | -0.233 | -42.6% |
| 08/17/94 | -0.414 | -0.417 | -0.263 | -36.5% |
| 08/18/94 | -0.359 | -0.348 | -0.170 | -52.6% |
| 08/19/94 | -0.474 | -0.467 | -0.213 | -55.1% |
| 08/22/94 | -0.565 | -0.553 | -0.383 | -32.2% |
| 08/23/94 | -0.358 | -0.355 | -0.226 | -36.9% |
| 08/24/94 | -0.387 | -0.388 | -0.294 | -24.0% |
| 08/25/94 | -0.480 | -0.459 | -0.221 | -54.0% |
| 08/26/94 | -0.460 | -0.434 | -0.275 | -40.2% |
| 08/30/94 | -0.458 | -0.455 | -0.225 | -50.9% |
| 08/31/94 | -0.451 | -0.444 | -0.270 | -40.1% |
| Average | -0.428 | -0.418 | -0.257 | -39.5% |

(1) refers to autocorrelation of raw returns

(2) refers to autocorrelation of fitted Residuals of Equation. (3.5)

(3) refers to autocorrelation of fitted Residuals with Buy/Sell Identities

(4) refers to percentage decrease in autocorrelation, (4)=[(1)-(3)]/(1)

Table 4.3: Estimation Results of Equation (BT) Using the Number of Trades as the Conditioning Variable

| Date | $\widehat{\mu}$ | $\widehat{\mu}_Q$ | $\widehat{\sigma}^2$ | $\widehat{\sigma}_Q^2$ |
|---|---|---|---|---|
| 08/01/94 | 2.37E-004 | -1.19E-004 | 4.62E-006 | 1.36E-007 |
| | ( 1.472) | (-1.782) | ( 6.827) | ( 0.500) |
| 08/02/94 | 1.58E-004 | -8.42E-005 | 3.15E-006 | 5.84E-007 |
| | ( 1.215) | (-1.221) | ( 5.620) | ( 2.241) |
| 08/03/94 | 1.47E-004 | -8.09E-005 | 9.06E-006 | 1.12E-006 |
| | ( 0.651) | (-0.767) | ( 6.087) | ( 1.784) |
| 08/04/94 | 1.45E-004 | -1.19E-004 | 3.60E-006 | 3.11E-007 |
| | ( 1.107) | (-1.644) | ( 5.426) | ( 0.919) |
| 08/05/94 | 3.05E-004 | -1.84E-004 | 3.94E-006 | 1.25E-006 |
| | ( 2.096) | (-2.054) | ( 5.596) | ( 3.469) |
| 08/08/94 | 1.26E-004 | -5.815E-005 | 4.26E-006 | 5.30E-007 |
| | ( 0.832) | (-0.867) | ( 4.891) | ( 1.573) |
| 08/09/94 | 6.68E-005 | -5.59E-005 | 3.43E-006 | 3.44E-007 |
| | ( 0.499) | (-0.805) | ( 5.871) | ( 1.221) |
| 08/10/94 | 1.72E-005 | -1.10E-005 | 4.35E-006 | 1.09E-006 |
| | ( 0.119) | (-0.119) | ( 5.981) | ( 2.737) |
| 08/11/94 | -1.37E-004 | 8.72E-005 | 3.18E-006 | 1.26E-006 |
| | ( -1.051) | (1.016) | ( 5.295) | ( 3.935) |
| 08/12/94 | 1.07E-005 | -8.51E-006 | 3.38E-006 | 3.59E-007 |
| | ( 0.085) | (-0.109) | ( 5.841) | ( 1.077) |
| 08/15/94 | 1.33E-005 | -2.02E-005 | 3.23E-006 | 1.14E-006 |
| | ( 0.095) | (-0.252) | ( 5.276) | ( 3.878) |
| 08/16/94 | 2.02E-004 | -1.38E-004 | 3.82E-006 | 1.55E-006 |
| | ( 1.415) | (-1.482) | ( 5.789) | ( 4.454) |
| 08/17/94 | -5.49E-005 | 3.32E-005 | 3.57E-006 | 5.09E-007 |
| | ( -0.402) | (0.479) | ( 7.061) | ( 2.190) |
| 08/18/94 | 2.80E-004 | -1.58E-004 | 5.60E-006 | 8.60E-007 |
| | ( 1.571) | (-1.713) | ( 3.401) | ( 1.120) |
| 08/19/94 | -1.04E-004 | 7.45E-005 | 2.72E-006 | 5.21E-007 |
| | ( -0.841) | (1.119) | ( 5.073) | ( 2.054) |
| 08/22/94 | 1.55E-004 | -1.16E-004 | 3.87E-006 | 1.25E-006 |
| | ( 1.014) | (-1.353) | ( 4.210) | ( 2.886) |
| 08/23/94 | 8.12E-005 | -7.02E-005 | 4.94E-006 | 1.62E-006 |
| | ( 0.524) | (-0.644) | ( 6.438) | ( 3.598) |
| 08/24/94 | 1.67E-004 | -7.96E-005 | 4.57E-006 | 3.71E-007 |
| | ( 1.062) | (-0.965) | ( 7.435) | ( 1.225) |
| 08/25/94 | -4.20E-005 | 1.41E-005 | 3.00E-006 | 4.75E-007 |
| | ( -0.302) | (0.216) | ( 4.913) | ( 1.837) |
| 08/26/94 | 2.65E-004 | -1.33E-004 | 3.11E-006 | 6.81E-007 |
| | ( 1.982) | (-1.955) | ( 5.611) | ( 2.803) |
| 08/30/94 | 6.09E-005 | -3.72E-005 | 4.72E-006 | 4.93E-007 |
| | ( 0.357) | (-0.522) | ( 6.428) | ( 1.744) |
| 08/31/94 | 1.81E-004 | -1.08E-004 | 3.20E-006 | 4.05E-007 |
| | ( 1.318) | (-1.506) | ( 5.383) | ( 1.417) |

* Numbers in the parentheses are t-values

Table 4.4: Estimation Results of a QTAR-GARCH(1, 1) Model

| Date | $\widehat{\alpha}_{01}$ | $\widehat{\alpha}_{02}$ | $\widehat{\alpha}_{11}$ | $\widehat{\alpha}_{12}$ | $\widehat{\beta}_{01}$ | $\widehat{\beta}_{02}$ | $\widehat{\beta}_{1}$ | $\widehat{\beta}_{2}$ |
|---|---|---|---|---|---|---|---|---|
| 08/01/94 | -0.0871 | 0.0435 | -0.4960 | -0.1010 | 0.0467 | 0.0216 | 0.0776 | 0.0000 |
| | (-4.364) | (5.139) | (-6.152) | (-1.872) | (8.561) | (19.384) | (1.771) | ( 0.000) |
| 08/02/94 | -0.0786 | 0.0351 | -0.2684 | -0.1602 | 0.0209 | 0.0148 | 0.0885 | 0.3323 |
| | (-4.985) | (3.941) | (-3.388) | (-2.489) | (2.779) | ( 2.764) | (2.150) | ( 1.584) |
| 08/03/94 | -0.1401 | 0.0472 | -0.3928 | -0.2596 | 0.0025 | 0.0010 | 0.0773 | 0.9039 |
| | (-5.908) | (3.889) | (-5.439) | (-4.541) | (1.094) | ( 1.316) | (3.024) | (32.445) |
| 08/04/94 | -0.1065 | 0.0394 | -0.3969 | -0.0856 | 0.0194 | 0.0171 | 0.3760 | 0.0000 |
| | (-7.784) | (5.037) | (-4.560) | (-1.110) | (4.525) | ( 5.702) | (3.988) | ( 0.000) |
| 08/05/94 | -0.0721 | 0.0339 | -0.4420 | -0.0998 | 0.0204 | 0.0007 | 0.0000 | 0.8565 |
| | (-2.960) | (4.037) | (-5.955) | (-1.518) | (4.478) | ( 2.194) | (0.000) | (32.673) |
| 08/08/94 | -0.0987 | 0.0361 | -0.4667 | -0.0671 | 0.0708 | 0.0236 | 0.0000 | 0.0000 |
| | (-3.178) | (4.215) | (-4.485) | (-1.440) | (2.945) | ( 4.874) | (0.000) | ( 0.000) |
| 08/09/94 | -0.0694 | 0.0262 | -0.4302 | -0.1404 | 0.0168 | 0.0085 | 0.1259 | 0.4625 |
| | (-3.627) | (3.470) | (-4.840) | (-2.178) | (2.872) | ( 2.662) | (2.676) | ( 2.906) |
| 08/10/94 | -0.0884 | 0.0507 | -0.4237 | -0.1164 | 0.0423 | 0.0240 | 0.0877 | 0.1701 |
| | (-4.049) | (5.063) | (-5.568) | (-1.820) | (4.308) | ( 4.953) | (1.753) | ( 1.237) |
| 08/11/94 | -0.0695 | 0.0431 | -0.5100 | -0.2174 | 0.0176 | 0.0093 | 0.1290 | 0.5478 |
| | (-2.904) | (4.721) | (-5.289) | (-3.737) | (2.699) | ( 2.634) | (2.852) | ( 4.119) |
| 08/12/94 | -0.0787 | 0.0314 | -0.2204 | -0.0436 | 0.0153 | 0.0043 | 0.0640 | 0.6775 |
| | (-3.983) | (4.169) | (-2.421) | (-0.637) | (2.924) | ( 3.072) | (2.704) | ( 8.614) |
| 08/15/94 | -0.0624 | 0.0443 | -0.4577 | -0.2119 | 0.0028 | 0.0005 | 0.0604 | 0.9001 |
| | (-4.259) | (4.766) | (-7.121) | (-3.398) | (2.102) | ( 1.147) | (3.661) | (39.930) |
| 08/16/94 | -0.0989 | 0.0547 | -0.2790 | -0.2268 | 0.0009 | 0.0016 | 0.0806 | 0.8897 |
| | (-6.541) | (4.732) | (-4.514) | (-3.429) | (0.816) | ( 1.522) | (3.498) | (29.120) |
| 08/17/94 | -0.0495 | 0.0417 | -0.3624 | -0.1368 | 0.0099 | 0.0040 | 0.0377 | 0.7567 |
| | (-3.339) | (4.274) | (-5.339) | (-2.045) | (2.318) | ( 1.657) | (1.555) | ( 7.736) |
| 08/18/94 | -0.0395 | 0.0343 | -0.4320 | -0.2625 | 0.0612 | 0.0229 | 0.2813 | 0.0740 |
| | (-1.788) | (3.455) | (-5.268) | (-4.155) | (5.679) | ( 7.458) | (3.557) | ( 1.028) |
| 08/19/94 | -0.0395 | 0.0242 | -0.5828 | -0.1988 | 0.0022 | 0.0000 | 0.0411 | 0.9415 |
| | (-2.264) | (3.791) | (-8.077) | (-3.259) | (2.129) | ( 0.313) | (3.884) | (86.707) |
| 08/22/94 | -0.0655 | 0.0212 | -0.5740 | -0.3329 | 0.0014 | 0.0028 | 0.0647 | 0.8711 |
| | (-3.474) | (2.097) | (-7.882) | (-5.832) | (0.665) | ( 1.853) | (2.604) | (16.512) |
| 08/23/94 | -0.1012 | 0.0532 | -0.1922 | -0.1993 | 0.0250 | 0.0114 | 0.0727 | 0.6013 |
| | (-4.766) | (4.552) | (-2.336) | (-3.147) | (1.162) | ( 0.925) | (1.552) | ( 1.804) |
| 08/24/94 | -0.0578 | 0.0502 | -0.3841 | -0.1790 | 0.0031 | 0.0030 | 0.0616 | 0.8565 |
| | (-3.921) | (4.647) | (-6.016) | (-2.736) | (1.132) | ( 1.390) | (2.293) | (10.627) |
| 08/25/94 | -0.0362 | 0.0200 | -0.4988 | -0.3580 | 0.0023 | 0.0004 | 0.0482 | 0.9147 |
| | (-2.560) | (2.501) | (-7.725) | (-5.991) | (2.291) | ( 1.150) | (2.657) | (35.129) |
| 08/26/94 | -0.0597 | 0.0204 | -0.4579 | -0.2744 | 0.0002 | 0.0002 | 0.0774 | 0.9155 |
| | (-4.731) | (3.239) | (-6.922) | (-4.581) | (0.248) | ( 1.470) | (4.512) | (60.265) |
| 08/30/94 | -0.0945 | 0.0343 | -0.4523 | -0.2120 | 0.0408 | 0.0281 | 0.1168 | 0.0925 |
| | (-4.585) | (3.455) | (-6.140) | (-3.481) | (4.104) | ( 4.245) | (2.334) | ( 0.541) |
| 08/31/94 | -0.0590 | 0.0296 | -0.5560 | -0.1004 | 0.0316 | 0.0204 | 0.1073 | 0.0000 |
| | (-3.164) | (3.760) | (-6.987) | (-1.678) | (5.499) | ( 7.042) | (1.764) | ( 0.000) |

† GARCH(1,1)-QTARCH model: $X_t = D \cdot (\alpha_{01} + \alpha_{11} X_{t-1}) + (1-D) \cdot (\alpha_{02} + \alpha_{12} X_{t-1}) + \varepsilon_t$
$\varepsilon_t = \nu_t \cdot h_t^{1/2}$, $\nu_t \sim N(0,1)$, and $h_t = D \cdot \beta_{01} + (1-D) \cdot \beta_{02} + \beta_1 \varepsilon_{t-1}^2 + \beta_2\, h_{t-1}$

‡ Numbers in the parentheses are t-values

Table 4.5: Reduction in Volatility Persistence

| Date | GARCH $\widehat{\beta}_1$ | GARCH $\widehat{\beta}_2$ | GARCH $\widehat{\beta}_1+\widehat{\beta}_2$ | QTAR-G $\widehat{\beta}_1$ | QTAR-G $\widehat{\beta}_2$ | QTAR-G $\widehat{\beta}_1+\widehat{\beta}_2$ |
|---|---|---|---|---|---|---|
| 08/01/94 | 0.10413 | 0.40712 | 0.5112 | 0.077581 | 0.000005 | 0.0776 |
| | ( 2.62 ) | ( 2.38 ) | | ( 1.771 ) | ( 0.000 ) | |
| 08/02/94 | 0.04908 | 0.63830 | 0.6874 | 0.088548 | 0.332278 | 0.4208 |
| | ( 1.67 ) | ( 1.69 ) | | ( 2.150 ) | ( 1.584 ) | |
| 08/03/94 | 0.05364 | 0.93537 | 0.9890 | 0.077256 | 0.903916 | 0.9812 |
| | ( 3.43 ) | ( 45.88 ) | | ( 3.024 ) | ( 32.445 ) | |
| 08/04/94 | 0.20482 | 0.19522 | 0.4000 | 0.375963 | 0.000012 | 0.3760 |
| | ( 3.68 ) | ( 1.74 ) | | ( 3.988 ) | ( 0.000 ) | |
| 08/05/94 | 0.08818 | 0.86671 | 0.9549 | 0.000006 | 0.856465 | 0.8565 |
| | ( 3.56 ) | ( 24.35 ) | | ( 0.000 ) | ( 32.673 ) | |
| 08/08/94 | 0.10247 | 0.85511 | 0.9576 | 0.000000 | 0.000003 | 0.0000 |
| | ( 5.15 ) | ( 39.35 ) | | ( 0.000 ) | ( 0.000 ) | |
| 08/09/94 | 0.16016 | 0.45165 | 0.6118 | 0.125854 | 0.462483 | 0.5884 |
| | ( 3.09 ) | ( 2.92 ) | | ( 2.676 ) | ( 2.906 ) | |
| 08/10/94 | 0.10573 | 0.34554 | 0.4512 | 0.087694 | 0.170137 | 0.2578 |
| | ( 2.50 ) | ( 1.78 ) | | ( 1.753 ) | ( 1.237 ) | |
| 08/11/94 | 0.10039 | 0.74491 | 0.8453 | 0.129005 | 0.547811 | 0.6768 |
| | ( 1.79 ) | ( 3.74 ) | | ( 2.852 ) | ( 4.119 ) | |
| 08/12/94 | 0.05764 | 0.87059 | 0.9282 | 0.064047 | 0.677458 | 0.7415 |
| | ( 3.12 ) | ( 17.92 ) | | ( 2.704 ) | ( 8.614 ) | |
| 08/15/94 | 0.05846 | 0.91441 | 0.9729 | 0.060364 | 0.900052 | 0.9605 |
| | ( 3.40 ) | ( 41.94 ) | | ( 3.661 ) | ( 39.930 ) | |
| 08/16/94 | 0.07028 | 0.90186 | 0.9722 | 0.080590 | 0.889724 | 0.9703 |
| | ( 3.75 ) | ( 37.43 ) | | ( 3.498 ) | ( 29.120 ) | |
| 08/17/94 | 0.05345 | 0.84302 | 0.8964 | 0.037698 | 0.756657 | 0.7944 |
| | ( 2.43 ) | ( 13.25 ) | | ( 1.555 ) | ( 7.736 ) | |
| 08/18/94 | 0.08345 | 0.87956 | 0.9631 | 0.281280 | 0.074017 | 0.3553 |
| | ( 3.96 ) | ( 26.77 ) | | ( 3.557 ) | ( 1.028 ) | |
| 08/19/94 | 0.06421 | 0.93035 | 0.9945 | 0.041087 | 0.941493 | 0.9826 |
| | ( 4.40 ) | ( 63.49 ) | | ( 3.884 ) | ( 86.707 ) | |
| 08/22/94 | 0.07690 | 0.85154 | 0.9284 | 0.064743 | 0.871089 | 0.9358 |
| | ( 3.07 ) | ( 17.72 ) | | ( 2.604 ) | ( 16.512 ) | |
| 08/23/94 | 0.06042 | 0.84648 | 0.9069 | 0.072709 | 0.601325 | 0.6740 |
| | ( 2.03 ) | ( 10.15 ) | | ( 1.552 ) | ( 1.804 ) | |
| 08/24/94 | 0.04055 | 0.90845 | 0.9490 | 0.061577 | 0.856492 | 0.9181 |
| | ( 2.53 ) | ( 20.25 ) | | ( 2.293 ) | ( 10.627 ) | |
| 08/25/94 | 0.06673 | 0.90326 | 0.9700 | 0.048153 | 0.914736 | 0.9629 |
| | ( 3.70 ) | ( 30.64 ) | | ( 2.657 ) | ( 35.129 ) | |
| 08/26/94 | 0.08854 | 0.90340 | 0.9919 | 0.077443 | 0.915530 | 0.9929 |
| | ( 4.59 ) | ( 54.40 ) | | ( 4.512 ) | ( 60.265 ) | |
| 08/30/94 | 0.09881 | 0.29734 | 0.3961 | 0.116779 | 0.092472 | 0.2093 |
| | ( 2.08 ) | ( 0.93 ) | | ( 2.334 ) | ( 0.541 ) | |
| 08/31/94 | 0.13620 | 0.10042 | 0.2366 | 0.107287 | 0.000032 | 0.1073 |
| | ( 2.79 ) | ( 0.53 ) | | ( 1.764 ) | ( 0.000 ) | |

QTAR-G indicates QTAR-GARCH(1,1) model.

Table 4.6: Comparison of Log-likelihood and AIC

| | Log-L | | | AIC | | |
|---|---|---|---|---|---|---|
| Date | OLS | GARCH | QTAR-G | OLS | GARCH | QTAR-G |
| 08/01/94 | 117 | 123 | 170 | 114 | 118 | 162 |
| 08/02/94 | 152 | 155 | 179 | 149 | 150 | 171 |
| 08/03/94 | -74 | -40 | -12 | -77 | -45 | -20 |
| 08/04/94 | 155 | 177 | 219 | 152 | 172 | 211 |
| 08/05/94 | 64 | 80 | 121 | 61 | 75 | 113 |
| 08/08/94 | 93 | 131 | 159 | 90 | 126 | 151 |
| 08/09/94 | 173 | 190 | 217 | 170 | 185 | 209 |
| 08/10/94 | 70 | 76 | 110 | 67 | 71 | 102 |
| 08/11/94 | 101 | 115 | 140 | 98 | 110 | 132 |
| 08/12/94 | 156 | 174 | 207 | 153 | 169 | 199 |
| 08/15/94 | 117 | 128 | 155 | 114 | 123 | 147 |
| 08/16/94 | 47 | 59 | 90 | 44 | 54 | 82 |
| 08/17/94 | 141 | 147 | 168 | 138 | 142 | 160 |
| 08/18/94 | 0 | 57 | 66 | -3 | 52 | 58 |
| 08/19/94 | 193 | 233 | 256 | 190 | 228 | 248 |
| 08/22/94 | 90 | 101 | 116 | 87 | 96 | 108 |
| 08/23/94 | 16 | 22 | 51 | 13 | 17 | 43 |
| 08/24/94 | 93 | 100 | 120 | 90 | 95 | 112 |
| 08/25/94 | 184 | 204 | 214 | 181 | 199 | 206 |
| 08/26/94 | 147 | 209 | 229 | 144 | 204 | 221 |
| 08/30/94 | 75 | 79 | 103 | 72 | 74 | 95 |
| 08/31/94 | 171 | 179 | 207 | 168 | 174 | 199 |
| Average | 104 | 123 | 149 | 101 | 118 | 141 |

[a] QTAR-G indicates QTAR-GARCH(1,1) model.
[b] Log-L: Log-Likelihood function.
[c] AIC: Akaike Information Criterion
AIC= Log-Likelihood function - number of parameters.

Table 4.7: Comparison of In-Sample Prediction

| | A | | | B | | |
|---|---|---|---|---|---|---|
| Date | OLS | GARCH | QTAR-G | OLS | GARCH | QTAR-G |
| 08/01/94 | 0.03600 | 0.03606 | 0.03117 | 0.04788 | 0.04534 | 0.03926 |
| 08/02/94 | 0.03145 | 0.03146 | 0.02908 | 0.04142 | 0.04106 | 0.03479 |
| 08/03/94 | 0.07939 | 0.07946 | 0.07160 | 0.09757 | 0.09142 | 0.08364 |
| 08/04/94 | 0.03072 | 0.03075 | 0.02811 | 0.04252 | 0.03915 | 0.03385 |
| 08/05/94 | 0.04488 | 0.04503 | 0.03965 | 0.06038 | 0.05721 | 0.05192 |
| 08/08/94 | 0.03966 | 0.03977 | 0.03422 | 0.05559 | 0.05132 | 0.04651 |
| 08/09/94 | 0.02879 | 0.02905 | 0.02615 | 0.03845 | 0.03542 | 0.03265 |
| 08/10/94 | 0.04396 | 0.04399 | 0.03934 | 0.05735 | 0.05550 | 0.04817 |
| 08/11/94 | 0.03854 | 0.03859 | 0.03484 | 0.04769 | 0.04387 | 0.04050 |
| 08/12/94 | 0.03067 | 0.03072 | 0.02796 | 0.04646 | 0.04576 | 0.03834 |
| 08/15/94 | 0.03608 | 0.03611 | 0.03254 | 0.04379 | 0.04200 | 0.03701 |
| 08/16/94 | 0.04828 | 0.04833 | 0.04283 | 0.05763 | 0.05596 | 0.04567 |
| 08/17/94 | 0.03311 | 0.03312 | 0.03073 | 0.03942 | 0.03905 | 0.03559 |
| 08/18/94 | 0.05856 | 0.05868 | 0.05812 | 0.07402 | 0.06852 | 0.06629 |
| 08/19/94 | 0.02629 | 0.02631 | 0.02339 | 0.03639 | 0.03386 | 0.03224 |
| 08/22/94 | 0.04051 | 0.04059 | 0.03786 | 0.05117 | 0.04866 | 0.04785 |
| 08/23/94 | 0.05477 | 0.05478 | 0.04975 | 0.07007 | 0.06925 | 0.05567 |
| 08/24/94 | 0.03981 | 0.03986 | 0.03707 | 0.04898 | 0.04771 | 0.04137 |
| 08/25/94 | 0.02797 | 0.02799 | 0.02684 | 0.03442 | 0.03187 | 0.03077 |
| 08/26/94 | 0.03200 | 0.03214 | 0.02972 | 0.04128 | 0.03686 | 0.03518 |
| 08/30/94 | 0.04302 | 0.04306 | 0.03970 | 0.05492 | 0.05334 | 0.04960 |
| 08/31/94 | 0.02880 | 0.02885 | 0.02584 | 0.03812 | 0.03558 | 0.03416 |
| Average | 0.03969 | 0.03976 | 0.03620 | 0.05116 | 0.04858 | 0.04368 |

[a] QTAR-G indicates QTAR-GARCH(1,1) model;

[b] Panel **A** reports the mean squared error ($MSE$) between the observed and fitted returns $MSE = \frac{1}{\#\text{ of obs}} \sum_{t=1}^{\#\text{ of obs}} (X_t - \widehat{X}_t)^2$;

[c] Panel **B** reports the mean absolute error ($MAE$) between the fitted squared residuals and the fitted conditional variance.
$MAE = \frac{1}{\#\text{ of obs}} \sum_{t=1}^{\#\text{ of obs}} |\widehat{\epsilon}_t^2 - \widehat{h}_t|$. In the OLS case, $\widehat{h}_t = \widehat{\sigma}^2$.

Table 4.8: Comparison of One-Step ahead Out-of-Sample Prediction

| | A | | | B | | |
|---|---|---|---|---|---|---|
| Date | OLS | GARCH | QTAR-G | OLS | GARCH | QTAR-G |
| 08/01/94 | 0.05233 | 0.05212 | 0.04894 | 0.05887 | 0.05681 | 0.05217 |
| 08/02/94 | 0.03524 | 0.03512 | 0.03143 | 0.04250 | 0.04260 | 0.03493 |
| 08/03/94 | 0.02339 | 0.02335 | 0.02199 | 0.07854 | 0.04306 | 0.03627 |
| 08/04/94 | 0.03380 | 0.03381 | 0.03119 | 0.04598 | 0.04184 | 0.03591 |
| 08/05/94 | 0.02149 | 0.02163 | 0.01863 | 0.05388 | 0.04056 | 0.03789 |
| 08/08/94 | 0.04145 | 0.04109 | 0.03744 | 0.05264 | 0.05151 | 0.04981 |
| 08/09/94 | 0.02336 | 0.02341 | 0.01969 | 0.03654 | 0.03232 | 0.02681 |
| 08/10/94 | 0.05816 | 0.05772 | 0.05474 | 0.06590 | 0.06404 | 0.05730 |
| 08/11/94 | 0.05387 | 0.05389 | 0.04649 | 0.05852 | 0.05821 | 0.05131 |
| 08/12/94 | 0.01105 | 0.01099 | 0.01043 | 0.03428 | 0.02829 | 0.02281 |
| 08/15/94 | 0.01971 | 0.01961 | 0.01762 | 0.03207 | 0.02591 | 0.02243 |
| 08/16/94 | 0.03230 | 0.03242 | 0.02909 | 0.05161 | 0.04494 | 0.03660 |
| 08/17/94 | 0.03215 | 0.03220 | 0.02877 | 0.03218 | 0.03215 | 0.03078 |
| 08/18/94 | 0.02159 | 0.02147 | 0.02126 | 0.04999 | 0.03275 | 0.03737 |
| 08/19/94 | 0.04947 | 0.04906 | 0.04565 | 0.05244 | 0.05671 | 0.04678 |
| 08/22/94 | 0.04287 | 0.04297 | 0.03863 | 0.05126 | 0.04852 | 0.04613 |
| 08/23/94 | 0.04437 | 0.04455 | 0.04448 | 0.06656 | 0.06372 | 0.05362 |
| 08/24/94 | 0.02333 | 0.02358 | 0.02606 | 0.04382 | 0.03930 | 0.03441 |
| 08/25/94 | 0.03108 | 0.03102 | 0.02883 | 0.03406 | 0.03203 | 0.03055 |
| 08/26/94 | 0.01558 | 0.01562 | 0.01406 | 0.03140 | 0.02046 | 0.01970 |
| 08/30/94 | 0.03314 | 0.03315 | 0.03304 | 0.04739 | 0.04602 | 0.04408 |
| 08/31/94 | 0.02633 | 0.02646 | 0.02482 | 0.03492 | 0.03272 | 0.03285 |
| Average | 0.03300 | 0.03297 | 0.03060 | 0.04797 | 0.04248 | 0.03821 |

[a] We reserve the first 400 observations to obtain parameter estimates. The one-step ahead out-of-sample prediction is then based these estimates.

[b] QTAR-G indicates QTAR-GARCH(1,1) model;

[c] Panel **A** reports the mean squared error ($MSE$) between the observed and fitted returns $MSE = \frac{1}{(\text{\# of obs}-400)}\sum_{t=401}^{\text{\# of obs}}(X_t - \widehat{X}_t)^2$;

[d] Panel **B** reports the mean absolute error ($MAE$) between the fitted squared residuals and the fitted conditional variance.
$MAE = \frac{1}{(\text{\# of obs}-400)}\sum_{t=401}^{\text{\# of obs}}|\widehat{\varepsilon}_t^2 - \widehat{h}_t|$. In the OLS case, $\widehat{h}_t = \widehat{\sigma}^2$.

# REFERENCES


[1] Admati, A. R., and P. Pfleiderer, "A Theory of Intraday Trading Pattern: Volume and Price Variability," *Review of Financial Studies*, 1 (1988), 3-40.

[2] Andersen, T. G., "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility," *Journal of Finance*, 51 (1996), 169-204.

[3] Andersen, T. G., and T. Bollerslev, "Intraday Periodicity and Volatility Persistence in Financial Market," *Journal of Empirical Finance*, 4 (1997), 115-158.

[4] Bachelier, L., *Theorie de la Speculation*, 1900, Paris: Gauthier-Villars.

[5] Bera, A. K. and M. L. Higgins, On ARCH Models: Properties, Estimating and Testing, in L. Exley, D. A. R. George, C. J. Roberts and S. Sawyer (eds.), *Surveys in Econometrics*, Basil Blackwell: Oxford, 1995.

[6] Board, J., and C. Sutcliffe, "The Effects of Trade Transparency in the London Stock Exchange: A Summary," Financial Markets Group, London School of Economics, Special Paper No. 67, January 1995.

[7] Bollerslev, T., "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31 (1986), 307-327.

[8] Bollerslev, T., R. Y. Chou and K. F. Kroner, "ARCH Modeling in Finance, A Review of the Theory and Empirical Evidence", *J of Econometrics*, 52 (1992), 5-59.

[9] Bollerslev, T., and I. Domowitz, "Trading Patterns and Prices in the Interbank Foreign Exchange Market," *Journal of Finance*, 48 (1993), 1421-1443.

[10] Bollerslev, T., R. Engle, and D. Nelson, ARCH Models, in R. F. Engle and D. MacFadden (eds.), *Handbook of Econometrics*, vol IV, North-Holland , Amsterdam, 1994.

[11] Breusch, T. and A. Pagan, "The Lagrange Multiplier Test and its Applications to Model Specification in Economics," *Review of Economic Studies*, 47 (1980), 239-253.

[12] Brock, W., W. Dechert, and J. Sheinkman, "A Test for Independence Based on the Correlation Dimension," mimeo, University of Wisconsin, Madison, 1986.

[13] Cai, J., "A Markov Model of Switching-Regime ARCH," *Journal of Business and Economic Statistics*, 12 (1994), 309-316.

[14] Chiang, C. L., *Introduction to Stochastic Processes in Biostatistics*, John Wiely & Sons Ltd, 1968.

[15] Clark, P., "A Subordinated Stochastic Process Model With Finite Variance for Speculative Prices," *Econometrica*, 41 (1973), 135-155.

[16] DeGennaro, R. P., and R. E. Shrieves, "Public Information Releases, Private Information Arrival and Volatility in the Foreign Exchange," *Journal of Empirical Finance*, 4 (1997), 195-315.

[17] Diebold, F., and J. A. Lopez, Modelling Volatility Dynamics, in K. Hoover (ed.), *Macroeconometrics: Developments, Tensions and Prospects*, 1995, 427-466.

[18] Durland, J., and T. McCurdy, "Duration-Dependent Transitions in a Markov Model of U.S. GNP Growth," *Journal of Business and Economic Statistics*, 12 (1994), 279-288.

[19] Ederington, L., and J. Lee, "How Market Process Information: News Releases and Volatility," *Journal of Finance*, 48 (1993), 1161-1191.

[20] Engle, R. F., "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50 (1982), 987-1007.

[21] Engle, R. F., "The Econometrics of Ultra-High Frequency Data," NBER 5816, 1996.

[22] Engle, R. F., and J. R. Russell, "Forecasting Transaction Rates: The Autoregressive Conditional Duration Model," UCSD 94-27, 1994.

[23] Filardo, A., "Business Cycle Phases and their Transitional Dynamics," *Journal of Business and Economic Statistics*, 12 (1994), 299-308.

[24] Filardo, A., and S. Gordon, "Business Cycle Durations," *Journal of Econometrics*, 85 (1998), 99-123.

[25] Gallant, A., P. Rossi, and G. Tauchen, "Stock Prices and Volume," *Review of Financial Studies*, 5 (1992), 198-242.

[26] Geman, H., and T. Ané, "Stochastic Subordination," *Risk*, 9 (1996), 145-149.

[27] Glosten, L. R., and P. R. Milgrom, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14 (1985), 71-100.

[28] Goodhart, C. A. E., and L. Figliuoli, "Every Minute Counts in Financial Markets," *Journal of International Money and Finance*, 10 (1991), 23-52.

[29] Goodhart, C. A. E., S. G. Hall, S. G. B. Henry, and B. Pesaran, "News Effects in a High-Frequency Model of the Sterling-Dollar Exchange Rate," *Journal of Applied Econometrics*, 8 (1993), 1-13.

[30] Goodhart, C. A. E., and M. O'Hara, "High Frequency Data in Financial Markets: Issues and Applications," *Journal of Empirical Finance*, 4 (1997), 73-114.

[31] Gourieroux, C., A. Holly, and A. Monfort, "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear with Inequality Constraints on the Regression Parameters," *Econometrica*, 50 (1982), 63-80.

[32] Gray, S., "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," *Journal of Financial Economics*, 42 (1996), 27-62.

[33] Hakkio, C., and D. Pearce, "The Reaction of Exchange Rates to Economic News," *Economic Inquiry*, 23 (1985), 621-635.

[34] Hamilton, J., "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57 (1989), 357-384.

[35] Hamilton, J., and R. Susmel, "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64 (1994), 307-333.

[36] Hansch, O., "Market Structure of the London Stock Exchange," Appendix in Ph.D. Thesis, London Business School, 1997.

[37] Harris, L., "Transaction Data Tests of the Mixture of Distributions Hypothesis," *Journal of Financial and Quantitative Analysis*, 22 (1987), 127-141.

[38] Harrison, M., and Pliska, S. "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications*, 11 (1981), 215-260.

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

[40] Jones, C., G. Kaul, and M. Lipson, "Transactions, Volume and Volatility," *Review of Financial Studies*, 7 (1994), 631-651.

[41] Judge, G., C. Hill, W. Griffiths, H. Lutkepohl, and T-C Lee, *The Theory of Econometrics*, John Wiley and Sons, 1985.

[42] Karpoff, J., "The Relation Between Price Changes and Trading Volume: A Survey," *Journal of Financial and Quantitative Analysis*, 22 (1987), 109-126.

[43] Kon, S., "Model Stock Returns – A Comparison," *Journal of Finance*, 39 (1984), 147-165.

[44] Kyle, A. S., "Continuous Auctions and Insider Trading," *Econometrica*, 53 (1985), 1315-1335.

[45] Lamoureux, C., and W. Lastrapes, "Heteroskedasticity in Stock Return Data: Volume versus GARCH Effects," *Journal of Finance*, 45 (1990a), 7-38.

[46] Lamoureux, C., and W. Lastrapes, "Persistence in Variance, Structural Change, and the GARCH model," *Journal of Business and Economic Statistics*, 8 (1990b), 225-234.

[47] Lamoureux, C., and W. Lastrapes, "Endogenous Trading Volume and momentum in Stock-Return Volatility," *Journal of Business and Economic Statistics*, 12 (1994), 253-260.

[48] Lo, A., and A. G. McKinlay, "An Econometric Analysis of Nonsynchronous-Trading," *Journal of Econometrics*, 45 (1990), 181-212.

[49] Locke, P. R., and C. L. Sayers, (1993) "Intra-day Futures Price Volatility: Information Effects and Variance Persistence," *Journal of Applied Econometrics*, 8 (1993), 15-30.

[50] Madan, D., E. Chang, "The Variance Gamma Option Pricing Model," mimeo, 1997.

[51] Marsh, T., and K. Rock, "The Transactional Process and Rational Stock Price Dynamics," University of California at Berkeley, 1986.

[52] Mandelbrot, B., "The Variations of Certain Speculative Prices," *Journal of Business*, 36 (1963), 394-419.

[53] Merton, R., "Option Pricing when Underlying Stock Returns Are Discontinuous," *Journal of Financial Economics*, 3 (1976), 125-144.

[54] Oldfield, G., Rogalski, R., and Jarrow, R., "An Autoregressive Jump Process for Common Stock Returns," *Journal of Financial Economics*, 5 (1977), 389-418.

[55] Osborne, M., "Browning Motion and the Stock Market," *Operations Research*, 7 (1959), 145-173.

[56] Payne, R., "Announcement Effects and Seasonality in the Intra-day Foreign Exchange Market," Discussion Paper No. 0238, Financial Markets Group, London School of Economics, 1997.

[57] Praetz, P., "The Distribution of Share Price Changes," *Journal of Business*, 45 (1972), 49-55.

[58] Rogers, A., "Modified Lagrange Multiplier Tests for Problems with One-Sided Alternatives," *Journal of Econometrics*, 31 (1986), 341-361.

[59] Roll, R., "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *Journal of Financial Economics*, 39 (1984), 1127-1140.

[60] Ross, S., "The Interrelations of Finance and Economics: Theoretical Perspectives," *American Economic Review*, 77 (1987), 29-34.

[61] Satchell, S., Y. Yoon, "Some Statistics for Testing the Influence of the number of Transactions on the distributions of Returns," University of Cambridge DAE working paper No. 9307, 1993.

[62] Smaby, T., "An Examination of the Intraday Behavior of the Yen/Dollar Exchange Rate: The Relationship between Trading Activity and Returns Volatility," *Journal of Economics and Finance*, 19 (1995), 39-50.

[63] Schlittgen, R., "Fitting of Threshold Models for Time Series," mimeo, University of Hamburg, 1997.

[64] Takezawa, N., "Note on Intraday Foreign Exchange Volatility and the Informational Role of Quote Arrivals," *Economics Letters*, 48 (1995), 399-404.

[65] Teräsvirta, T., "Specification, Estimation, and Evolution of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89 (1994), 208-217.

[66] Teräsvirta, T., D. Tjstheim, and C. Granger, "Aspects of Modelling Nonlinear Time Series," in R. Engle and D. McFadden (eds.), *Handbook of Econometrics*, Vol. IV, Elsevier, Amsterdam, 1994.

[67] Tong, H., *Threshold Models in Nonlinear Time Series Analysis*, Springer-Verlag, New York, 1983.

[68] Tong, H., *Nonlinear Time Series Analysis: A Dynamical System Approach*, Oxford University Press, New Tork, 1990.

[69] Zhou, B., "High-Frequency Data and Volatility in Foreign-Exchange Rates," *Journal of Business and Economic Statistics*, 14 (1996), 45-52.

# Appendix A

# Converting IM-MI Trades to MM Trades

In this appendix, we describe our scheme of re-coding pairs of IM-MI trades into an observations of MM trade. The following table shows an extracted example of a sequence of IM-MI trades.

Table A.1: An Example of a Sequence of IM-MI Trades

| Seq. No* | Time | Buyer | Seller | Qty. | Value | Ask | Bid | Price |
|---|---|---|---|---|---|---|---|---|
| 13787 | 11:01 | M | I | 30000 | 11540700 | 385 | 384 | 384.69 |
| 13788 | 11:01 | I | M | 30000 | 11535000 | 385 | 384 | 384.50 |
| 13829 | 11:17 | I | M | 100000 | 38450000 | 385 | 384 | 384.50 |
| 13831 | 11:18 | M | I | 100000 | 38469000 | 385 | 384 | 384.69 |
| 13850 | 11:26 | M | I | 10000 | 3840000 | 385 | 383 | 384.00 |
| 13851 | 11:26 | I | M | 10000 | 3838100 | 385 | 383 | 383.81 |
| 13956 | 12:26 | M | I | 100000 | 38350000 | 385 | 383 | 383.50 |
| 13957 | 12:27 | I | M | 100000 | 38331000 | 385 | 383 | 383.31 |
| 13986 | 12:48 | M | I | 25000 | 9604750 | 385 | 383 | 384.19 |
| 13987 | 12:48 | I | M | 100000 | 38400000 | 385 | 383 | 384.00 |
| 13988 | 12:49 | I | M | 25000 | 9600000 | 385 | 383 | 384.00 |
| 13989 | 12:49 | M | I | 100000 | 38419000 | 385 | 383 | 384.19 |
| 14159 | 14:35 | M | I | 10000 | 3856900 | 386 | 385 | 385.69 |
| 14162 | 14:35 | I | M | 10000 | 3855000 | 386 | 385 | 385.50 |

*Seq. No corresponds to the sequence that each observation occurs in the original data set. Note that all the other non-IM-MI trades are taken out of this table for illustration purpose.

In the simplest cases, if an IM trade is followed by an MI trade, or vice versa, with the same traded quantity, we treat this pair of IM and MI trades as an MM trade. The 13,787$^{th}$ observation and the 13,788$^{th}$ observation in the table constitute such an example. However, complication occurs in the following three situations:

1. Pairs of IM & MI trades may not be located right besides each other along the sequence of recorded observations. For example, there is a non-IM-MI trade occurs between the 13,829$^{th}$ and the 13831$^{th}$ observation. However, by matching the quantity, we would still be able that say that these two observations belong to the same MM trade. In this situation, we take the first recorded time as the time for the newly re-coded observation, and also maintain its sequence in the dataset.

2. Since the inter-dealer brokers charge commission when facilitating trades between market makers, there would be two different prices for each pair in the table, as is invariably the case in the table. The price we choose is the "cleaner price", which is a multiple of 25 pence.

3. Trades between market makers and inter-dealer brokers may involve more than three parties: a market maker as a buyer, a market maker as a seller and an inter-dealer. For example, the 13,986$^{th}$, 13,987$^{th}$, 13,988$^{th}$ and 13,989$^{th}$observations. Again, we can treat this sequence of trades as the same block of MM trade by matching the involved quantities and buy-sell prices. The previous two criteria also apply in this case.

# Appendix B

# LM tests

Our equation can be written as

$$X(t) = \mu + \mu_Q \Delta N(t) + \varepsilon(t) \tag{B.1}$$

where $\varepsilon(t)$, conditional on $N(t)$, is $N(0, \sigma^2 + \sigma_Q^2 \Delta N(t))$. Let $Y$ be a $(T \times 2)$ matrix $Y' = \begin{bmatrix} 1, & 1, & \cdots & 1 \\ \Delta N_1, & \Delta N_2, & \cdots, & \Delta N_T \end{bmatrix}$, $\gamma' = [\mu, \mu_Q]$, $\beta' = [\sigma^2, \sigma_Q^2]$. Then the likelihood function $L$ can be written as

$$L = L(\gamma, \beta) = -\frac{1}{2}\Sigma \ln h_t - \frac{1}{2}\Sigma e_t^2 / h_t \tag{B.2}$$

where $h_t = \beta' Y_t'$, $Y_t$ is the $t$-th row of $Y$, $e_t = X_t - Y_t \gamma$, and the constant term is omitted. It follows that the first derivatives of the log-likelihood function are

$$\begin{aligned} \frac{\partial L}{\partial \gamma} &= \Sigma \frac{e_t Y_t}{h_t} \\ \frac{\partial L}{\partial \beta} &= -\frac{1}{2}\Sigma \frac{Y_t}{h_t} + \frac{1}{2}\Sigma \frac{e_t^2 Y_t}{h_t^2} \end{aligned} \tag{B.3}$$

The second derivatives follow immediately,

$$\begin{aligned} \frac{\partial^2 L}{\partial \gamma \partial \gamma'} &= \Sigma \frac{Y_t' Y_t}{h_t} \\ \frac{\partial^2 L}{\partial \beta \partial \beta'} &= \frac{1}{2}\Sigma \frac{Y_t' Y_t}{h_t^2} - \Sigma \frac{e_t^2 Y_t' Y_t}{h_t^3} \\ \frac{\partial^2 L}{\partial \gamma \partial \beta'} &= -\Sigma \frac{e_t Y_t' Y_t}{h_t^2} \end{aligned} \tag{B.4}$$

Given the above information, the Fisher's information matrix can be constructed as

$$-\frac{\partial^2 L}{\partial \theta \partial \theta'} = \begin{bmatrix} \Sigma \frac{Y_t' Y_t}{h_t} & \Sigma \frac{e_t Y_t' Y_t}{h_t^2} \\ \Sigma \frac{e_t Y_t' Y_t}{h_t^2} & \Sigma \frac{e_t Y_t' Y_t}{h_t^3} - \frac{1}{2}\Sigma \frac{Y_t' Y_t}{h_t^2} \end{bmatrix} \tag{B.5}$$

We now calculate $\phi_{H_j} = E_{H_j}(-\frac{\partial^2 L}{\partial\theta\partial\theta'})$, $j$=1, 2, 3 where $\theta' = [\gamma', \beta']$, that is,

$$\begin{aligned}
\phi_{H_1} &= \begin{bmatrix} \Sigma\frac{Y_t'Y_t}{\hat{h}_t} & 0 \\ 0 & \frac{1}{2}\Sigma\frac{Y_t'Y_t}{\hat{h}_t^2} \end{bmatrix} \\
\phi_{H_2} &= \begin{bmatrix} \frac{1}{\sigma^2}(Y'Y) & 0 \\ 0 & \frac{1}{2\sigma^4}(Y'Y) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{bmatrix} \otimes (Y'Y) \\
\phi_{H_3} &= \phi_{H_2}
\end{aligned} \tag{B.6}$$

The score test for the hypothesis $H_j$, $j$=1, 2, 3, is of the form $D'_{H_j}\phi^{-1}_{H_j}D_{H_j}$ which will be asymptotically distributed as $\chi^2(1)$ for $j$=1, 2 and $\chi^2(2)$ for $j$=3. In turn, from (A.3)

$$D_{H_1} = \begin{bmatrix} 0 \\ \Sigma\frac{(x_t-\hat{u})\Delta N_t}{\hat{h}_t} \\ 0 \\ 0 \end{bmatrix} \tag{B.7}$$

where $\hat{h}_t$ is equal to $\hat{\beta}'Y_t$ (evaluated under $H_0$).

$$D_{H_2} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{2\hat{\sigma}^4}\Sigma(\hat{e}_t^2 - \hat{\sigma}^2)\Delta N_t \end{bmatrix} \tag{B.8}$$

where $\hat{e}_t = X_t - \hat{\mu} - \hat{\mu}_0\Delta N_t$ and $\hat{\sigma}^2 = \Sigma\hat{e}^2/T$, and

$$D_{H_3} = \begin{bmatrix} 0 \\ \frac{1}{\tilde{\sigma}^2}\Sigma\tilde{e}_t^2\Delta N_t \\ 0 \\ \frac{1}{2\tilde{\sigma}^4}\Sigma(\tilde{e}_t^2 - \tilde{\sigma}^2)\Delta N_t \end{bmatrix} \tag{B.9}$$

where $\tilde{e}_t = X_t - \hat{\mu}$ and $\tilde{\sigma}^2 = \Sigma\tilde{e}_t^2/T$.

We can compute $LM_j = D'_{H_j}\phi^{-1}_{H_j}D_{H_j}$ for $j = 1, 2, 3$ explicitly as follows:

$$\begin{aligned}
LM_1 &= \left(\sum\frac{(x_t-\hat{\mu})\Delta N_t}{\hat{h}_t}\right)^2\left(\sum\frac{\Delta N_t^2}{\hat{h}_t} - (\sum\frac{\Delta N_t}{\hat{h}_t})^2(\sum\frac{1}{\hat{h}_t})^{-1}\right)^{-1} \\
&\quad \text{where } \hat{h}_t = \hat{\sigma}^2 + \hat{\sigma}_Q^2\Delta N_t \\
LM_2 &= LM_2(\hat{e}_t) = \frac{(\sum(\hat{e}_t^2 - \hat{\sigma}^2)\Delta N_t)^2}{2\hat{\sigma}^4\sum(\Delta N_t - \Delta\overline{N})^2} \quad \text{where } \hat{e}_t = x_t - \hat{\mu} - \hat{\mu}_Q\Delta N_t \\
LM_3 &= LM_2(\tilde{e}_t) + \frac{1}{\tilde{\sigma}^2}\frac{(\sum\tilde{e}_t\Delta N_t)^2}{\sum(\Delta N_t - \Delta\overline{N})^2} \quad \text{where } \tilde{e}_t = x_t - \hat{\mu}
\end{aligned} \tag{B.10}$$

# Appendix C

# Moment Generating Function and Moments of $\Delta N(t)$

$$\begin{aligned} \Delta N(t)|\lambda(t) &\sim \text{Poisson } (\lambda(t)) \\ E\left[\exp(s \cdot \Delta N(t))|\lambda(t)\right] &= mgf \text{ of a Poisson Process} \end{aligned} \tag{C.1}$$

A Poisson distributed random variable $x$ with density function $f(x)$

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \ x = 0, 1, 2, \ldots \tag{C.2}$$

has the following moment generating function

$$\begin{aligned} E[e^{sx}] &= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^s)^x}{x!} \\ &= e^{-\lambda} \exp(\lambda e^s) \\ &= \exp(-\lambda(1 - e^s)) \\ &= \exp(\lambda(e^s - 1)) \end{aligned} \tag{C.3}$$

Therefore, $E\left[\exp(s \cdot \Delta N(t))|\lambda(t)\right] = \exp(\lambda(t)(e^s - 1))$.

Now,

$$\lambda(t) = \frac{\beta\sigma^2}{1-\theta} + \alpha \sum_{j=0}^{\infty} \theta^j v^2_{(t-j-1)} \tag{C.4}$$

which implies

$$E[\exp(q\lambda(t))] = \exp(q\frac{\beta\sigma^2}{1-\theta}) E[\exp(q\alpha \sum_{j=0}^{\infty} \theta^j v^2_{(t-j-1)})] \tag{C.5}$$

Since $v^2_{(t-j-1)} \sim i.i.d.\ \chi^2(1)$,

$$\begin{aligned} E[\exp(q\alpha \sum_{j=0}^{\infty} \theta^j v^2_{(t-j-1)})] &= \prod_{t=1}^{\infty} E[\exp(q\alpha\theta^j \chi^2(1))] \\ &= \prod_{t=1}^{\infty} (1 - 2q\alpha\theta^j)^{-1/2} \end{aligned} \tag{C.6}$$

Thus

$$E\left[\exp(s\cdot\Delta N(t))\right] = E[\exp(\lambda(t)(e^s-1))] \tag{C.7}$$
$$= \exp(\frac{\beta\sigma^2}{1-\theta}(e^s-1))\cdot\prod_{t=1}^{\infty}\left[1-2\alpha\theta^j(e^s-1)\right]^{-1/2}$$

Based on this moment generating function, we can derive associated moments of $\Delta N(t)$ as the following,

$$E\left[\Delta N(t)|\lambda(t)\right] = \lambda(t) \tag{C.8}$$
$$E(\Delta N(t)) = E\left[\lambda(t)\right]$$
$$Var(\Delta N(t)) = E\left[Var\left(\Delta N(t)|\lambda(t)\right)\right] + Var\left[E\left(\Delta N(t)|\lambda(t)\right)\right]$$
$$= E\left[\lambda(t)\right] + Var\left[\lambda(t)\right]$$

where

$$E\left[\lambda(t)\right] = \frac{\beta\sigma^2+\alpha}{1-\theta} \tag{C.9}$$

Define

$$M(q) = \ln E\left[\exp\left(q\lambda(t)\right)\right] \tag{C.10}$$
$$= q\left(\frac{\beta\sigma^2}{1-\theta}\right) - \frac{1}{2}\sum\ln(1-2q\alpha\theta^j)$$

$$\frac{\partial M(q)}{\partial q} = \frac{\beta\sigma^2}{1-\theta} - \frac{1}{2}\sum_{j=0}^{\infty}\frac{-2\alpha\theta^j}{(1-2q\alpha\theta^j)} \tag{C.11}$$

It follows that

$$Var\left[\lambda(t)\right] = \frac{\partial}{\partial q}\left[\alpha\sum\frac{\theta^j}{(1-2q\alpha\theta^j)}\right]_{q=0} \tag{C.12}$$
$$= \alpha\sum\theta^j(-1)(1-2q\alpha\theta^j)^{-2}(-2\alpha\theta^j)|_{q=0}$$
$$= \frac{2\alpha^2}{1-\theta^2}$$

Therefore,

$$E\left[\Delta N(t)\right] = \frac{\beta\sigma^2+\alpha}{1-\theta} \tag{C.13}$$
$$Var\left[\Delta N(t)\right] = \frac{\beta\sigma^2+\alpha}{1-\theta} + \frac{2\alpha^2}{1-\theta^2}$$

# Appendix D

# Moment Conditions of $\Delta N(t)$, and $X(t)$

From our model,

$$\begin{aligned} Cov(\Delta N(t), \Delta N(t-1)) &= E_{I(t-1)}\left[Cov\left(\Delta N(t), \Delta N(t-1) \,|I(t-1)\right)\right] \\ &\quad + Cov_{I(t-1)}\left(E\left[\Delta N(t)|I(t-1)\right], E\left[\Delta N(t-1)|I(t-1)\right]\right) \\ &= Cov_{I(t-1)}\left(\alpha v^2(t-1) + \beta\sigma^2 + \theta\lambda(t-1), \lambda(t-1)\right) \\ &= \theta Cov_{I(t-1)}\left(\lambda(t-1), \lambda(t-1)\right) = \theta Var_{I(t-1)}\left(\lambda(t-1)\right) \\ &= \frac{2\theta\alpha^2}{1-\theta^2} \qquad \text{(D.1)} \end{aligned}$$

Then we derive the following recursively,

$$Cov(\Delta N(t), \Delta N(t-s)) = \frac{2\theta^s\alpha^2}{1-\theta^2} \qquad \text{(D.2)}$$

Finally, it follows that

$$\begin{aligned} Corr(\Delta N(t), \Delta N(t-s)) &= \frac{Cov(\Delta N(t), \Delta N(t-s))}{Var(\Delta N(t))} \qquad \text{(D.3)} \\ &= \begin{cases} \frac{2\alpha^2\theta^s}{2\alpha^2 + (\beta\sigma^2+\alpha)(1+\theta)}, & \text{if } s \geq 1 \\ 1, & \text{if } s = 0 \end{cases} \end{aligned}$$

For the moments of $X(t)$,

$$X(t) = \mu + \sigma(z(t) - z(t-1)) + \sum_{i=1}^{\Delta N(t)} Q_i \qquad \text{(D.4)}$$

it follows that

$$\begin{aligned} E[X(t)] &= \mu + \mu_Q E[\lambda(t)] = \mu + \frac{\mu_Q(\beta\sigma^2 + \alpha)}{1-\theta} \qquad \text{(D.5)} \\ Var[X(t)] &= \sigma^2 + (\mu_Q^2 + \sigma_Q^2)\frac{\beta\sigma^2 + \alpha}{1-\theta} + \frac{2\alpha^2\mu_Q^2}{1-\theta^2} \end{aligned}$$

The correlation between $X(t)$ and $X(t-s)$ can be recursively derived by using $Cov(X(t), X(t-1))$. It is straightforward to show that

$$\begin{aligned} Cov(X(t), X(t-1)) &= Cov\left(\sigma v(t) + \sum_{i=1}^{\Delta N(t)} Q_i, \sigma v(t-1) + \sum_{i=1}^{\Delta N(t-1)} Q_i\right) \\ &= \sigma Cov\left(\sum_{i=1}^{\Delta N(t)} Q_i, v(t-1)\right) + Cov\left(\sum_{i=1}^{\Delta N(t)} Q_i, \sum_{i=1}^{\Delta N(t-1)} Q_i\right) \\ &= E_{I(t-1)}\left[Cov\left(\sum_{i=1}^{\Delta N(t)} Q_i|I(t-1), \sum_{i=1}^{\Delta N(t-1)} Q_i|I(t-1)\right)\right] \\ &\quad + Cov_{I(t-1)}\left(E\left[\sum_{i=1}^{\Delta N(t)} Q_i|I(t-1)\right], E\left[\sum_{i=1}^{\Delta N(t-1)} Q_i|I(t-1)\right]\right) \\ &= \mu_Q^2 Cov_{I(t-1)}\left(\Delta N(t), \Delta N(t-1)\right) = \mu_Q^2 Cov\left(\lambda(t), \lambda(t-1)\right) \\ &= \mu_Q^2 \frac{2\alpha^2\theta}{1-\theta^2} \end{aligned} \tag{D.6}$$

Then we derive the following recursively,

$$Cov(X(t), X(t-s)) = \frac{2\mu_Q^2\alpha^2\theta^s}{1-\theta^2} \tag{D.7}$$

Finally, it follows that

$$\begin{aligned} Corr(X(t), X(t-s)) &= \frac{Cov(X(t), X(t-s))}{Var(X(t))} \\ &= \begin{cases} \frac{2\alpha^2\mu_Q^2\theta^s}{\sigma^2(1-\theta^2)+(\mu_Q^2+\sigma_Q^2)(1+\theta)(\beta\sigma^2+\alpha)+2\alpha^2\mu_Q^2}, & \text{if } s \geq 1 \\ 1, & \text{if } s = 0 \end{cases} \end{aligned} \tag{D.8}$$

# IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989