STOCK DATA, TRADE DURATIONS, AND LIMIT ORDER BOOK INFORMATION

Ola Simonsen
Abstract


Paper [1] considers conditional duration models in which durations are in continuous time but measured in grouped or discretized form. This feature of recorded durations in combination with a frequently traded stock is expected to negatively influence the performance of conventional estimators for intraday duration models. A few estimators that account for the discreteness are discussed and compared in a Monte Carlo experiment. An EM-algorithm accounting for the discrete data performs better than those which do not. Empirically, the incorporation of level variables for past trading is rejected in favour of change variables. This enables an interpretation in terms of news effects. No evidence of asymmetric responses to news about prices and spreads is found.

Paper [2] considers an extension of the univariate autoregressive conditional duration model to which durations from a second stock are added. The model is empirically used to study duration dependence in four traded stocks, Nordea, Föreningssparbanken, Handelsbanken and SEB A on the Stockholm Stock Exchange. The stocks are all active in the banking sector. It is found that including durations from a second stock may add explanatory power to the univariate model. We also find that spread changes have significant effect for all series.

Paper [3] empirically tests whether an open limit order book contains information about future short-run stock price movements. To account for the discrete nature of price changes, the integer-valued autoregressive model of order one is utilized. A model transformation has an advantage over conventional count data approaches since it handles negative integer-valued price changes. The empirical results reveal that measures capturing offered quantities of a share at the best bid- and ask-price reveal more information about future short-run price movements than measures capturing the quantities offered at prices below and above. Imbalance and changes in offered quantities at prices below and above the best bid- and ask-price do, however, have a small and significant effect on future price changes. The results also indicate that the value of order book information is short-term.

Paper [4] This paper studies the impact of news announcements on trade durations in stocks on the Stockholm Stock Exchange. The news are categorized into four groups and the impact on the time between transactions is studied. Times before, during and after the news release are considered. Econometrically, the impact is studied within an autoregressive conditional duration model using intradaily data for six stocks. The empirical results reveal that news reduces the duration lengths before, during and after news releases as expected by the theoretical literature on durations and information flow.

Key words: Finance, Maximum likelihood, Estimation, ACD, News, Multivariate, Intraday, Market microstructure, Granger causality, Time series, INAR, Stock price, Open limit order book.
Introduction

The role for markets is to determine the prices of goods. The question is then, how is this done? The price is determined through an agreement between a buyer and a seller of a quantity of a good. In daily life this may be when a buyer and a seller contact each other directly, for example, when buying a car or a household good. Often there are a number of persons involved in buying or selling goods. The price is then the price at which the quantity for sale is equal to the quantity individuals want to buy. In the economic literature this is the intersection of the demand and supply curves of a specific good. Simply, if more individuals are interested in buying than selling a good the price will increase and if more people are interested in selling than buying a good the price will decrease.

When buying and selling financial assets the buyers and sellers, denoted traders in the finance literature, may either contact each other directly through a specialist or trade through a computerized system. A common wish among traders is to trade at a market that is liquid. This may be defined as the ability to quickly trade a quantity of shares with low price impact. To accomplish this efficiently several trading mechanisms have emerged. The three main solutions are order driven markets, Walrasian auction and price driven markets. The Stockholm Stock Exchange (SSE) is an order driven market. At the SSE traders enter their orders directly into a computerized system and orders are executed when they can be matched. Since the market is computerized small as well as large investors have direct access to the market. The access may be, e.g., through trading platforms on the internet. A second market mechanism used at financial markets is the Walrasian auction. It is one of the oldest market types and is used, e.g., at the London gold market. Finally, a third alternative market mechanism is the price driving market. In that market traders trade through a market maker. The market maker has to provide liquidity to the market, i.e. always standing ready to buy and sell assets. A leading example of such a market is the New York Stock Exchange (NYSE).

The common feature of the different trading mechanisms is that they are governed
by rules. For example, the order driven market at the SSE is a continuous market. Traders may enter their orders at any time during the opening hours and, hence, orders are matched and executed continuously over time. The implication of this is that trades occur irregularly and may be clustered over time. In the finance literature the irregularly spaced time between transactions has received attention as it is thought to reflect the information flow to the market. In order driven markets traders enter their orders to an order book. The order book is monitored by a computer and the content is visible to traders. Accordingly, traders’ order placements are observed and may influence other market participants. Another common feature of the SSE is the existence of automated trading or algorithm trading among institutional and private traders. Automated trading is performed by a computer, i.e. the trading decisions is transferred from a human trader to a computer. Strategies based on variables such as liquidity, time and price may trigger buy and sell orders. The purpose of the automated trading may be, for example, to lower transaction costs or to make profit out of trading strategies. Institutional investors may have research divisions devoted to automated trading or use software from companies providing sophisticated trading solutions. Also, private investors may have access to automated trading through online brokerage firms. The use of automated trading may have implications for the market characteristics at the SSE. One possible complication may be that reactions after events are faster and more powerful than without automated trading. The reason for this may be that the automated trading triggers a chain reaction. For example, an event resulting in a price movement may initiate a number of trades performed by computers.

In this thesis high frequency data are utilized. High frequency data or intraday data is financial data where every transaction or event is recorded. The recording frequency is often on a second scale in contrary to, e.g., daily or weekly. Such high frequency data is characterized by a lack of synchronization, i.e. event occurs irregularly over time. Recorded events in high frequency data may be, e.g., how market participants place their orders or every transaction with associated price, spread and volume. Consequently, even short time series of intraday data contain huge amounts of data. They give much finer information of events at the market and how market
participants are acting. The availability of high frequency data has had an important impact on research of how the market mechanisms are working and may be modelled. For example, economic questions like the role of information to stock market participants may be studied and modelled. The availability of high frequency data has also led to the development of new econometric tools. For example, the feature of irregularly spaced times between events has led to new econometric tools, e.g., for the modelling of the time between transactions. For an overview of high frequency data, market microstructure and econometrics, see, e.g., Bauwens and Giot (2001) and Tsay (2002, ch. 5).


**Trade durations**

Trade durations correspond to the time between two consecutive transactions and a transaction refers to a trade between a buyer and a seller of a volume of stocks at a given price. Trade durations have played a central roll in market microstructure research in recent decades. The time elapsed between transactions is thought to carry information about the information flow to market participants. Relevant information may be, e.g., related to the valuation of the stock. Consequently, trade durations may be important for understanding the price process. The idea to study durations originates from the information based model of Glosten and Milgrom (1985). In this model traders are either informed or uninformed, with information varying with regard to the value of the asset they trade. Uninformed traders mainly trade for liquidity reasons, while informed traders act on their superior information on the value of the asset. Diamond and Verrecchia (1987) show that short durations correspond to the presence of bad news. Easley and O’Hara (1992) extend the model of Glosten and Milgrom (1985) by highlighting the importance of time to distinguish between informed
and uninformed traders. They stress that new information to market participants leads to increased trade intensity, i.e., shorter durations between transactions. This corresponds to an increased number of informed agents trying to exploit their new information. The information may either be public or private. The public information may, e.g., be news announcements from news agencies or press releases. The part denoted private information may be not publicly released information, e.g., analysis performed by influential investment banks.

News may not only be important for the particular stock it concerns it may also be important for related stocks. For example, private information, discovered by increased trade intensity in a stock may influence trading in related stocks, e.g., in the same industry. If this is true we may find duration dependence between stocks. This feature has rendered attention in the finance literature. Recent studies of dependence between stock transaction series are, e.g., Bauwens and Hautsch (2004) and Spierdijk et al. (2002) who found dependence between transaction series.

Paper [2] in this thesis examines the dependence between stocks in the banking sector at the Stockholm Stock Exchange. The result shows that there is Granger causality (Granger, 1969) between the stocks in the banking sector at the Stockholm Stock Exchange, i.e., increased trade intensity in one of the stocks from the banking sector influence trading in other stocks in the same sector.

News agencies distribute price driving information from stock market companies. Companies are forced to release their price driving information publicly as the use of not publicly released price driving information is prohibited in most countries, including Sweden. Hence, news releases from news agencies may shorten durations. Several studies have focused on the impact of news releases on price and volatility rather than durations, e.g., Berry and Howe (1994), Ederington and Lee (1993), Mitchell and Mulherin (1994), Bollerslev et al. (2000), Bauwens et al. (2005) and Kalev et al. (2004). Their results show that news influence price and volatility. The potential link between price and trade durations has also been studied by, e.g., Grammig and Wellner (2002), Dufour and Engle (2000) and Engle (2000) who study the interdependence between intradaily prices, price volatility and trade durations. Dufour and Engle (2000) find
that as the time between transactions become shorter the speed of price adjustment increases suggesting that an active market with short durations demonstrates presence of informed traders. Accordingly, publicly released news announcements that contains price driving information may not only reduce duration lengths but may also affect prices. Paper [4] in this thesis deals with the impact of public news releases on durations, i.e. how trade durations are affected by news releases from news agencies.

**Autoregressive conditional duration model**

For econometric modelling of the time between transactions standard econometric tools may not be appropriate as transactions are irregularly spaced over time. An influential suggestion of how to model the irregularly spaced times is due to Engle and Russell (1998). Their autoregressive conditional duration (ACD) model explains the length of the next duration by conditioning on the length of past durations and explanatory variables. Several extensions and applications of the original ACD model have been presented by, e.g., Engle and Lunde (2003) and Bauwens and Giot (2001).

A duration arises as the time between two consecutive transactions at \( t_{i-1} \) and \( t_i \), i.e. \( d_i = t_i - t_{i-1} \). The conditional expectation of a duration \( d_i \) is specified as \( E(d_i | d_{i-1}, ..., d_1; x) = \theta_i \). Here \( d_i \) is conditioned on past durations and other explanatory variables \( x \), and \( \theta_i \) is specified in such a way that \( \epsilon_i = d_i / \theta_i \) is independent and identically distributed. The \( \theta_i \) may be parameterized as

\[
\theta_i = \omega + \sum_{j=1}^{p} \alpha_j d_{i-j} + \sum_{j=1}^{q} \beta_j \theta_{i-j} + \pi' x_{i-1}.
\]

Here, \( \theta_i \) is a function of \( p \) lagged durations and \( q \) lags of conditional durations. This is called an ACD\((p,q,x)\), where \( p \) and \( q \) are the orders of the lags in the mean function and \( x_i \) is a vector of explanatory variables such as volume, spread and price.

Engle and Russell (1998) popularized the quasi maximum likelihood (QML) estimator building on the exponential distribution for the estimation of the unknown parameters of the ACD model. The estimator maximizes the log-likelihood function
\[
\ln \ell = - \sum \left[ \ln \theta_i + \frac{d_i}{\theta_i} \right].
\]

The estimator is widely used for estimating these types of models and, hence, an increasing amount of empirical results builds on estimates from this estimator. The estimator is consistent if the model for the conditional mean is correctly specified and the duration is in the exponential family of distributions. A potential problem with the estimator is with the empirical duration data often used in the estimation. The data are discretized with a large fraction of short and zero durations, e.g., more than one transaction occurring during one second. The discreteness of the recorded transaction data is due to the second scale and data are hence integer-valued. As the distribution underlying the QML estimator is continuous the discreteness may negatively affect the performance of the estimator. Kulldorff (1961, ch. 2) shows that inconsistency of the maximum likelihood (ML) estimator may arise from, e.g., using mid-interval values to represent the interval when data are discretized or grouped. In paper [1] the problem is highlighted and estimators are proposed to account for the problem.

One of the prominent features of the results from the ACD model, e.g., from papers [1],[2] and [4], is the quick response of, for example, price changes and changes of traded volumes. The response takes place within seconds after events occur and fades out quickly thereafter. The presence of automated trading at the Stockholm Stock Exchange may be one explanation for this pattern. Trading rules programmed to react to, e.g., price moves of more than a predetermined number of ticks, may influence the characteristics of the trading behavior.

**Limit order book**

In order driven markets no market maker is involved in the trading process. Traders directly enter their orders in an order book monitored by a computer. The order book is visible to traders, private as well as institutional traders. Traders may enter two types of orders, either a market order or a limit order. Limit orders is placed in queue in the order book and is executed when they can be matched. The priority of limit
order execution is determined by price and time of order placement, i.e. the stocks with the lowest limit sell price are sold first and the stocks with the highest limit buy price are bought first. Market orders are, on the other hand, executed immediately to the best bid or ask price. For example, the publicly visible limit order book for a stock listed on the SSE shows the first five levels on the bid and ask side, respectively. This is illustrated in Figure 1 where $P_i^d$ and $P_i^s$ are the prices on the bid- (demand) and ask- (supply) side of an arbitrary order book for the levels $i = 1, 2...5$. The bid- and ask-volumes contained at level $i$ are denoted $Q_i^d$ and $Q_i^s$.

Figure 1: Illustration of the limit orderbook.

In the theoretical literature of limit order books the assumption is that informed traders always use market orders instead of using limit orders, e.g., Glosten (1994), Rock (1996) and Seppi (1997). Accordingly, there should be limited information of observing offered quantities of a share at other prices than the best bid- or ask-price, i.e. at the high and low end of the order book. However, recent papers of Cao et al. (2004) and Bloomfield et al. (2005) find evidence of information also in the order book beyond the first levels. Cao et al. (2004) introduce measures capturing the shape of the order book, e.g., if the order book is asymmetric, i.e. if there is more value on the bid- (ask-) side relative to the ask- (bid-) side. The measure is used to study if the asymmetry of the order book is informative regarding future price movements. The findings are that the order book contains information regarding future price movements.
using intradaily data with 5 minutes aggregation level. Further, Foucault (1999) argues that an increase in asset volatility increases the proportion of limit order traders and the limit order traders have to post higher ask prices and lower bid prices, i.e. market depth increase.

In paper [3] of this thesis the purpose is to empirically study the information contained in the open limit order book about future short-run stock price movements. To assess the information contained in the order book the paper presents new measures and extensions of existent measures. A new measure capturing the market depth discussed in Foucault (1999) is presented. We find the appropriate aggregation level to be 1 minute in contrary to Cao et al. (2004), who use 5 and 10 minutes aggregation levels.

**Summary of the papers**

**Paper [I] Discretized Time and Conditional Duration Modelling for Stock Transaction Data**

The paper considers conditional duration models in which durations are in continuous time but measured in grouped or discretized form. This feature of recorded durations in combination with a frequently traded stock is expected to negatively influence the performance of conventional estimators for intraday duration models. Estimators that account for the discreteness are discussed and compared in a Monte Carlo experiment, e.g., grouped maximum likelihood and EM-algorithm. In the small Monte Carlo study the EM-algorithm that accounts for the discrete nature of the data both in the outcome and the lagged explanatory variables comes out as the best estimator of the compared ones.

In the empirical part of the paper the differences between estimators are generally quite small and the EM-algorithm and ML estimators based on discrete data are not too different from ML based on grouped data and Weibull and Burr models. When it comes to the effects of explanatory variables the study provides support for the use of changes rather than levels to reflect news. There is throughout a significant and
positive effect of news about prices and a negative effect of a change in the spread. The spread effect is, however, not significant. A higher volume has an insignificant but prolonging effect in most cases. We could not find statistically significant support for an asymmetric response to news about spreads nor about prices. A contributory cause of the quick response after changes in the explanatory variables may be the computerized trading at the stock market. The log-likelihood function value of the Burr is larger than for other models but the models are not nested. In addition, the serial correlation properties of the exponential and Weibull models speak in favor of these two models. A generalized gamma was also employed and provided a better fit to the data than both the exponential and Weibull models.


This paper considers an extension of the univariate autoregressive conditional duration model to which durations from a second stock are added. By including durations from a second stock dependence between duration series is captured in the model. For example, Figure 2 illustrates the dependence from stock 2 to stock 1. Completed durations are added from stock 2, $d^2_{N(1)}$, to stock 1, and the completed durations from stock 2 are given weights dependent on the size of $\tau_1$, i.e. depending on how far away in time the completed durations are.

The model is empirically used to study duration dependence in four traded stocks, Nordea, Föreningssparbanken, Handelsbanken and SEB A on the Stockholm Stock Exchange. The stocks are all in the banking sector. In the empirical part we find Granger causality between all the stocks. This result indicates that there may be duration dependence between stocks active in the banking sector on the Stockholm Stock Exchange. The dependence may be caused by new information revealed by the trade intensity. In view of this empirical result the suggested model extension is able to capture dependence between duration series and to provide an improvement of the econometric specification of the model.
The purpose of this paper is to empirically study the information contained in the open limit order book about future short-run stock price movements. Specifically, attention is paid to whether changes or asymmetries in the order book concerning offered quantities of a share at prices below the best bid price (low end of the order book) and above the best ask price (high end of the order book) are informative. To assess the information contained in the order book the paper presents new measures as well as extensions to existing measures summarizing order book movements.

To account for the discrete nature of price changes, the integer-valued autoregressive model (e.g., McKenzie, 1985, 1986, Al-Osh and Alzaid, 1987) is utilized to model the discrete nature of high frequency stock price data.

The empirical results reveal that measures capturing offered quantities of a share at the best bid- and ask-price reveal more information about future short-run price movements than measures capturing the quantities offered at prices below and above.
Introduction and Summary

However, imbalance and changes in offered quantities at prices below and above the best bid- and ask-price do have a small and significant effect on future price changes. The results also indicate that the value of order book information is short-term. This can be compared to Cao et al. (2004) who found informational value of the higher levels of the order book at an aggregation level of 5 and 10 minutes. The effects in this paper are most apparent at a low aggregation level (1 minute) while estimation results for higher aggregation levels (2, 5 and 10 minutes) show mostly insignificantly results.


This paper examines empirically the short-run impact of public news announcements on trade durations in stocks traded on the Stockholm Stock Exchange in Sweden. Econometrically, the impact is studied within an autoregressive conditional duration model using intraday data for six stocks. The news are categorized into four groups and added as explanatory variables to the model. The news categories are Company/Sector, EU macro, Swedish macro and US macro news. The news categories are added to the model through a dummy variable structure that captures the impact before, during and after the news.

The empirical results reveal that news reduces the duration lengths before, during and after news releases. For example, Company/Sector related news shortens durations by 20 to 40 percent. The cause of reduced durations before the actual news releases may originate from anticipated news releases, while the significant result of the impact during and after the news release may be due to both anticipated and unanticipated news events. The result supports the predictions in the theoretical literature of shorter durations in connection to news.
References


