The Microstructure of a Dealership Market:

An Empirical Investigation of the London Stock Exchange

Hung-Neng Lai

A thesis submitted for the degree
Doctor of Philosophy

January 1999

The London School of Economics and Political Science
University of London

To my parents

ABSTRACT

This thesis investigates the trading on the London Stock Exchange, a multiple dealership market. It consists of four chapters. Chapter one motivates the study and summarises the major findings of the thesis.

Chapter two examines the individual quoting behaviour of market makers. Quote revisions are often first made by one of the few price leaders in response to market information, and the rest follow suit. Price leaders tend to quote one side of the yellow strip to attract unbalanced orders; price followers often straddle the strip, which results in smaller but more balanced order flows. Moreover, the negative correlation between effective spreads and orders flows is detected in all but very small trades.

The third chapter attempts to solve an apparent puzzle on the Exchange: market makers appear to charge different costs from different market participants. The chapter uses the theories of market microstructure to explain why the costs are different, and examines the extent to which the difference is related to the collusion of market makers or to the trading mechanisms. The evidence suggests the negotiation power of trading parties may play an important role in determining the costs of trades.

The last chapter presents a new approach to estimate bid-ask spreads in multiple dealership markets. The traditional approach of estimating spreads is not applicable in those markets because observed prices cannot be ordered sequentially. An alternative method is proposed for which data sequentiality is not needed. The model is put into state space form to use the Kalman filter to estimate the fundamental price and the spread. The new method is implemented to the data of three liquid stocks on the Exchange.

Acknowledgement

The completion of this thesis would never come about without the guidance of Professor David Webb, my supervisor in my fourth year of Ph.D. study. I am deeply indebted to his tolerance and judgement. Moreover, I am extremely lucky to be supervised by Dr. John Board in the following years of study. I would like to thank him for, among many other things, the patience to answer my endless questions. The last chapter of the thesis originates from a joint research project with Professor Siem Koopman. I have learned a lot from working with him, including how to write computer programs to implement econometrics, and how to make a point more forcefully.

The thesis has greatly benefited from the insightful comments by many people. In particular, I thank Philip Chan, George Cheng, Andrew Ellul, Libon Fung, Lawrence Glosten, Jie-Haun Lee, Anne Fremault Vila, Stephen Wells, and Ingrid Werner for numerous suggestions and criticisms. The data used in the thesis is obtained from the London Stock Exchange, and the help from Nigel Fawn, Christine Jackson, and especially Stephen Wells is very much appreciated. The financial and administrative assistance from the Department of Accounting of Finance is very helpful for my study; I would like to express my gratitude to the convenor Professor Michael Power and the administrative team.

During my six and half years' stay in London, I have been overwhelmed by the encouragement of the warmest kind from many friends. With immense appreciation, my thanks go to Clara Chen, Hsuan-Ju Chiu, Wenyi Chu, Mahmood Delkhasteh, Vicky Hu, Charleen Huang, James Hueng, Michio Iijima, Dorota Lubańska, Chibuike Ugochukwu Uche and Carol Yang.

Finally, I would like to thank my sister for doing everything for me without any complaint, and my parents for giving me the freedom to do whatever I want to do with wholehearted support.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

OVERVIEW

This thesis attempts to explain how market makers post bid and ask quotes in dealership markets, and how the transaction prices are determined. Because the investigations are based on the data collected from the London Stock Exchange, it is necessary to give a brief description about how the Exchange is operated. Three main topics of this thesis will be introduced in Section 1.2, Section 1.3 and Section 1.4 respectively. Section 1.5 explains some of the notions used in this thesis.

1.1 The operation

The investigations of this thesis are based on the data collected between January and June 1996, when the London Stock Exchange (LSE) was still a typical multiple dealership market.[1] According to the Exchange (1997), there are 299 member firms of the Exchange in 1996. Member firms are free to register as market makers. During 1996, thirty-three market makers post bid and ask prices of more than 2,600 stocks on Stock Exchange Automated Quotation system (SEAQ). The registered market makers are obliged to maintain two-way quotes on SEAQ between 8:30 and 16:30 every trading day. The middle of SEAQ screen of each security shows the current highest bid and lowest ask prices, and up to four names of market makers

[1] Although an electronic order book was introduced for the trading of FTSE-100 stocks in October 1997, the order book only captures 30% of the trading volumes of those stocks, and the rest of the trades are still executed by dealers (Board and Wells 1998; Martinson 1998).

who are posting the best quotes on each side. As the best-quote information is highlighted with yellow colour, the best quotes are often referred to "the yellow strip". Furthermore, those market makers who post the best quotes are "on the yellow strip" despite the fact that some of their names may not be literally on the yellow strip when more than four market makers quote the best bid or ask prices. Those market makers who do not quote the best bid nor the best ask are said to "straddle the yellow strip" (Reiss and Werner 1996). Another synonym of yellow strip is "touch". Market makers honour the trades up to a certain number of shares, which is often equal to the Normal Market Size (NMS), a measure which attempts to capture the equivalence of 2.5% of daily trading volumes; see Appendix A.1. If the order is smaller than 10% of the NMS, the market participants may execute the trade through a computer network called Small-order Automated Execution Facility (SAEF), otherwise they will negotiate the price with the market makers through telephones. If the trade size is smaller than or equal to that honoured in the SEAQ screen, the market maker is allowed to trade at the price better than what is currently displayed in the screen. If the trade size is bigger than the quote size, then the market maker is free to quote any price on the phone.

The SEAQ screen provides the relevant information for the customers of market makers to decide whether to trade and with whom to trade, but the member firms of the Exchange are free to choose anybody to trade with. Therefore, those market makers who do not post the best quotes may obtain preferenced orders. Non-member firms can even trade with one another without the involvement of any market maker, but such trades do not occur frequently. For those investors who are not members of the Exchange, they can trade with market makers through a broker who is a member firm. The brokers are essentially "broker-dealers". They can trade with market makers on their own accounts, or they can act on behalf of their customers. Moreover, market makers are allowed to trade with their private clients, who are almost exclusively institutional investors.

The inter-dealer trading on the Exchange is conducted in two ways. First, market makers can trade with one another in SEAQ in the same way as trading with any other member firms, in which case the trades are called IMM trades (Board and Sutcliffe 1995). Second, they can trade with one another in the inter-dealer broker (IDB) market, in which the trades are called IDB trades. IDB market is one to which only market makers have exclusive access. Four IDBs operate separate electronic bulletin systems to allow market to place limit orders, and the computer screens only display the best buy and sell orders. Trades are not executed electronically. If another market makers wish to take the orders, they may call the IDB to execute the trades. More negotiation may take place if one of the market makers wish to trade more shares (Reiss and Werner 1997). Unlike SEAQ, trading in the IDB market is anonymous: only the IDB knows who place and who take the limit orders. On the other hand, both SEAQ and IDB trades are under the same post-trade publication rules. Trades are reported to the Exchange within three minutes of executions. The Exchange publishes small trades immediately, and it publishes the trades one hour after the executions if the trade sizes are more than six times NMS. If the trade sizes are more than seventy-five times NMS, then the trading parties may request the Exchange to delay the publication up to five days or until 90% of the trades are off-loaded.

1.2 Changing the quotes

In theory, investors do not pay much attention to the quotes by individual market makers. They regard the best quote, the lowest ask and highest bid price, as the market price. Investors are expected to reward those market makers who are able to execute the trades of any sizes at any time with negligible costs. Individual market makers may post only the highest bid, only the lowest ask, both highest bid and lowest ask, or neither of them. Those who post the lowest ask attract all of the buy orders, and those who post the highest bid attract all of the sell orders. In practice, market makers attract preferenced order flows. There are explicit contracts or implicit agreements between market makers and some of their customers, by which the latter are obliged to route certain amount of orders to the former, who in return execute the orders with the best bid or ask prices.

The objective of Chapter 2 is to explain the quoting behaviour of market makers on the LSE: why some market makers quote the best prices while the others avoid being on the yellow strip. Before explaining the quoting behaviour, it is necessary to explore the consequence of being on or off the yellow strip. After all, if it does not make any difference for the market maker to post the best quote, then any talk about the quoting behaviour is meaningless. Most of the theoretical models in the literature take it for granted that posting the best quotes captures the whole order flows. However, owing to preferenced order flows and some other reasons, early studies on the LSE do not provide strong evidence of the link between posting the best quotes and attracting order flows (Board et al. 1996; Hansch et al. 1998). By examining the link more closely, Chapter 2 first argues that a lot of market makers spend most of the time on straddling the yellow strip, but the number of trades executed by straddling the strip is not proportional to the time spent on the strip. Second, the imbalance of buy and sell orders may obscure the link between order flow and quote status. After adjusting the overall order imbalance, it becomes evident that market makers attract more buy than sell orders by quoting only the best ask, and they attract more sell than buy orders by quoting only the best bid. In contrast, those market makers who quote both the best ask and best bid are not expected to receive unbalanced order flows, and they do receive balance orders. Third, when market makers quote the best bid or best ask, the order flows they receive depend on how many market makers are posting the best quote at the same time. Since "a market maker quotes both best bid and best ask" often means "everybody else quotes both best bid and ask", the order flows market makers receive while on both sides of yellow strip are not different from when they straddle the strip. Fourth, posting the best quote attracts orders bigger than the quote size shown in the SEAQ screen, which implies the presence on the yellow strip has signalling effects. Finally, not only the quote spreads but also the effective spreads matter: the order flows are negatively related to the effective spreads for the medium and large trades. Based on those findings, Chapter 2 concludes that market makers do attract additional order flows by quoting the yellow strip prices, and how substantial the orders are depends on the number of market makers on the strip.

This chapter then explores the patterns of quote change and investigates the events before the quote change. Cross-sectional regressions are conducted. The dependent variable is the change in the mid-quote of individual market makers. The independent variables are the number of trades, the volumes of trades, and the inventory levels before the quote change, as well as the preceding quote change. The evidence from Section 2.5 suggests the majority of the market makers are price followers while a small number of market makers update the quote rather aggressively. There are two different types of quote revisions. The first type of the revision is to quote one side of the yellow strip aggressively in order to attract unbalanced order flows. The quotes are often made by price leaders in response to public information. The upward revisions are often associated with the arrival of buy trades, and the downward revisions are often with sell trades. The second type of quote revision is to deliberately straddle the yellow strip or to quote both sides of the strip. The revision is often made by a price follower soon after the quote change of price leaders. The benefit of straddling the yellow strip is to avoid receiving unbalanced order flows; the drawback is the order flows received tend to be less than standing on the strip.

1.3 Charging for the liquidity

Chapter 3 focuses on the comparison of the costs of trading, and on the choice of the trading platforms. A large number of trades in the LSE are carried out by negotiation among traders. During negotiation, the trading parties may gather more information about the securities than merely the price and the quantity of the trade. When a member firm approaches the market maker to initiate the trade, the market maker will try to detect the motivation behind the trade from the negotiation and from the past experience of dealing with the firm. On the other hand, the trade data collected from the LSE consists of not only the prices and quantities of the trades, but also the identities of the trading parties and trading conditions. Although it is not possible to identify the end investors behind brokers, the data provides sufficient clues to understand how the market makers differentiate their customers. The data reveals that market makers charge different customers for different costs. Specifically, the small investors appear to pay more than the private clients of the market makers. Member firms of the Exchange who act as principals have favourable prices despite of the fact that they pay more when they trade on behalf of the investors outside the Exchange. The trading platform affects the costs of trading, too. Market makers trade with one another frequently in both SEAQ and IDB market, but the costs of trading in IDB market are much lower than those in SEAQ. Similar findings have been documented in Board and Sutcliffe (1995), Reiss and Werner (1996, 1997) and Hansch et al. (1999).

Can the widely observed differences in costs be explained by the theories of market microstructure? This chapter attempts to answer this question by running a cross-sectional regression. The dependent variable is half of the effective spread, and the independent variables include the variables documented in the literature which may determine the costs of the trades, as well as the dummy variables representing the different types of customers. The analysis shows the variables suggested in the literature successfully explain part of the effective spreads, but they cannot explain most of the costs difference. It is possible, however, that different trades contains different information, which reflect different costs of trades and which the regression model fails to detect. Therefore, another regression model is used to analyse the effects of different trades on the quote change. The evidence reveals that market makers change quotes after customer trades, which imply those trades have certain inventory or information implications. IMM trades also tend to trigger the quote change which implies those trades contain information. On the contrary, the posters of IDB trades change their quotes to the opposite direction of the one predicted by the literature.

The success of IDB market to provide a low-cost trading platform coincides with the inability of IDB market to provide liquidity timely. Chapter 3 further finds that four IDBs altogether assist merely less than six deals per stock per day. Market makers place limit orders only when they can afford waiting for unwinding their inventories. The posting in IDB market clearly signals the trades only have inventory implication, so the costs are low. In contrast, market makers are very sensitive to IMM trades. If a market maker cannot wait for trading with customers or resort to the less-liquid IDB market, it indeed signals something unusual to the counter party, who has to charge the trade with higher costs to compensate the potential adverse selection problem. Another interpretation of the expensive IMM trades is that it is one market maker exercises price discrimination against the other, since the latter is unable to find the liquidity elsewhere.

The cost structure of the agency trades, client trades and principal trades can be explained by the preferenced orders and the different bargaining powers. Most of the end customers of agency trades are individual investors, who do not trade very frequently nor very heavily. They have little if any control over broker-dealers, so most of their orders are preferenced and are traded at the touch. The private clients of market makers trade regularly, and they are able to negotiate with market makers to get better deals. Finally, member firms of the Exchange trade with market makers most frequently. Their bargaining powers are further enhanced by the fact that they can decide to which market makers the customer orders are routed. Therefore, they are able to obtain the best deals from market makers.

1.4 Measuring the spread

The price of securities in financial markets consists of two main components: the fundamental price and the spread. Both components cannot be observed. There are two standard approaches to estimate the spread from the trade prices. One is to use the autocovariance structure of the differenced price and quote data (Roll 1984; Stoll 1989). The other is to regress the change in prices on some other variables such as change in trade directions or change in quotes (Glosten and Harris 1988; Huang and Stoll 1997). From the theoretical point of view, it implies that the parties have observed the past trades and have absorbed at least part of the information from the trades before conducting the current trade. From the estimation's point of view, it follows that the prices must be ordered sequentially over time: at each time period only one price is observed.

Unfortunately, the prices reported to the LSE are not sequential. Several negotiations among trading parties can occur at same time without the full knowledge of the activities in the market, so it is not theoretically appealing to assume the trading parties have some knowledge of the other trades taking place at the same time, and to impose a trading sequence on the data. Furthermore, this thesis uses the trade data constructed from the settlement records, which are time-stamped in minutes. For the liquid stocks, several trades are reported to take place at the same minute on a regular basis, and there is no way to define a trading sequence from the data set.

Chapter 4 develops a new statistical approach to model bid-ask spreads for non-sequential trade markets. The model decomposes the trade price into an underlying fundamental price and the half bid-ask spread. The fundamental price is assumed to follow a random walk plus the information effect from market volumes of customer trades. The spread is determined by the size of the trade and the time of the day, and both effects are non-linear and are modelled by regression splines. Each time interval has a different number of trades executed, so the dimension of the variables in the model is not constant. The problem of non-sequentiality is solved by putting the model into state space form and to estimate the model by maximising likelihood using the Kalman filter. The updating recursions of the Kalman filter do not require the dimension of the observational vector to be constant, and the missing observations can be dealt with in a straightforward manner. The model is applied to three heavily traded stocks on the LSE. The analyses have been successful. It is shown the model is capable of identifying the fundamental price and the spread in a straightforward manner. Compared with an alternative approach of using the mid-touch as the fundamental price, the proposed model explains more of the observed trade prices, and the residuals of the model exhibit much less autocorrelation than those from the model using mid-touch.

1.5 Notations

This thesis uses the term "spread" extensively. "Spread" and "bid-ask spread", denoted by s, are used generically to refer to the unobservable component of either the transaction price or the quote offered by market makers. This thesis considers four different types of bid-ask spreads:

1. The *quote spread* is defined as the difference between bid and ask prices of a stock quoted by a specific market maker. Corresponding to the quote spread, the *mid-quote* is defined as the average of the bid and ask prices of a stock quoted by a specific market maker.

2. The *touch spread* is the lowest ask, a_t, minus the highest bid, b_t, across all market makers of a specific stock. Synonym for the touch spread include the inside spread and the touch. The definition of the touch is the same as the "Spread" in Board and Sutcliffe (1995) and the "PCT" in Chan, Christie, and Schultz (1995). Corresponding to the touch spread, the mid-touch m_t is the average of the best bid and ask prices.[2]

3. The *effective spread* is defined as

$$2s = \begin{cases} 2 \times (p - m_t)/m_t & \text{for a buy} \\ 2 \times (m_t - p)/m_t & \text{for a sell} \end{cases} . \tag{1.1}$$

where p is the transaction price of a specific trade and m_t is the mid-touch at time t, when the trade is executed. The definition of the

[2] m_t is denoted by q_t in Huang and Stoll (1996) and by M_t in Huang and Stoll (1997). Occasionally, m_t is also used as the mid-quote in this thesis accompanied by clear specification.

effective spread used in this thesis is similar to the one in Hansch et al. (1999), who term it the Execution Quality (Exqual).

The direction of a customer trade is identified from the prespective of the counter party of market maker. In Chapter 2, no attempt is made to determine the initiator of an inter-dealer trade, and the direction of an inter-dealer trade is identified in the same way as that of a customer trade. In Chapter 3, especially in Section 3.5, where it is necessary to identify the initiators of inter-dealer trades, the directions of the trades are those of initiators.

This definition of the effective spread is different from the one in Huang and Stoll (1996), who define the effective spread as the absolute value of the difference between the price and the mid-touch. The difference results from the fact that the identities of trading parties are given in TDS data, while such information is often unavailable in the transaction data of the exchanges in the US.

4. Finally, the *Kalman-filter spread* is twice the half spread $d_{t,i} \times s$ estimated from the model proposed in (4.2) of Chapter 4. $d_{t,i}$ is the trade indicator which takes the value 1 when the trade is a buy and -1 when it is a sell. It has another notation d_b in Chapter 3 because the time of the trade is unimportant in that chapter. Note that the definition of trade indicator is the same as Q_t in Huang and Stoll (1997) and x_t in Madhavan et al. (1997).

Chapter 2

POSTING QUOTES IN MULTIPLE DEALERSHIP MARKETS

2.1 Introduction

How market makers determine the quotes is an intriguing question, and it is especially true when they are in a multiple-dealer environment. In theory, market makers observe the public information in the markets. They extract information from the customers with whom they trade. They observe the actions taken by another market makers. They need a lot of information to decide how to revise the quotes. The customers, on the other hand, do not pay much attention to the quotes by individual market makers. They use the best quote, the lowest ask and highest bid price, as the market price. Individual market makers may post only the lowest ask, only the highest bid, both lowest ask and highest bid, or neither of them. Two interesting questions immediately emerge. First, what makes market makers quote different prices? Second, what are the consequences if they do not quote the best bid or ask?

Both questions, especially the second one, have important policy implications. The spirit of a competitive dealership market is to allow market makers competing with one another to offer the best services to the customers. Customers are expected to reward those market makers who are able to execute the trades of any sizes at any time with negligible costs. Therefore, *ceteris paribus*, the smaller bid-ask spread a market maker offers, the more orders he should receive. From the regulator's point of view, it is essential to make sure that at least some of the market makers commit themselves to competing in prices, that the best bid-ask spread is satisfactorily narrow, that the market price reflects available information, and that the market attracts investors.

The main theme of this chapter is to explain the quoting behaviour of market makers on the London Stock Exchange, a multiple dealership market. Much of the emphasis will be on the consequence of on and off the yellow strip. Market makers are willing to post competitive bid and ask prices if posting the best quotes makes a difference in attracting business. Most of the theoretical models in the literature take it for granted that posting the best quotes captures the whole order flow. In reality, owing to preferenced order flows and some other reasons, early studies on the London Stock Exchange do not provide strong evidence of the link between posting the best quotes and attracting order flows (Board et al. 1996; Hansch et al. 1998). Board et al. (1997) even challenge the commitment of market makers to price discovery. However, this chapter will show that posting the best quote does attract business, the scale of which depends on how many market makers are posting the best quote at the same time. In addition, there are two new findings. First, posting the best quotes attracts orders bigger than the quote size shown in the computer screen, which implies that quotes have signalling effects. Second, not only the quote spreads but also the effective spreads matter: the order flows are negatively related to the effective spreads for the medium and large trades.

This chapter then explores the patterns of quote changes and provides descriptive statistics. Contrary to what models of competitive dealership markets suggest, market makers often deliberately post quotes that are away from the yellow strip. On the other hand, some market makers actively participate in the market and quote aggressively. The findings are consistent with Board et al. (1997) that there are different degrees of the commitments of market makers to price discovery. The cross-sectional regression analysis reveals that those who do not post the best quote tend to revise the quote soon after another market makers move to the yellow strip, and that quote changes made by price leaders often reflect trade information. The inventory consideration demonstrated by Hansch et al. (1998) is minor and is often over-shadowed by information effects. It appears market makers do not often resort to quote revision in response to the change in inventory.

The rest of this chapter is organised as follows. The next section briefly reviews the literature on the quote revision and quote status. The third section establishes the link between order flow and quote status. Section 2.4 further strengthens the link by demonstrating the signalling effect of quote status on attracting big trades. In addition, a new method is proposed to examine the relationship between trades and effective spreads. Section 2.5 highlights the quoting behaviour of market makers. Section 2.6 investigates what trigger the change of quote and discusses the implications. Section 2.7 concludes the chapter.

2.2 Quote revisions: a brief survey

2.2.1 Theoretical arguments

The decision to post the bid and ask price is potentially extremely complex, and most of the theoretical contributions in the literature do not directly analyse how market makers determine the quotes in a multiple dealership environment. Instead, a lot of assumptions are made to simplify the question. Stoll (1978b), Glosten and Milgrom (1985), Easley and O'Hara (1987) and Glosten (1989) discuss how a single dealer determines the bid and ask prices. In Kyle (1985) and Grossman and Miller (1988), market makers determine a single price, and the bid-ask spread is not even mentioned. Theories of competitive dealership markets are abundant. Diamond and Verrecchia (1991) study the effect of the disclosure of information on the cost of capital in multiple dealership markets. Madhavan (1992) compares the dealership and auction market. Dennert (1993) explores the relationship between the spread and the number of dealers in the market. The interactions among market makers are often introduced by including an inter-dealer market, such as the models by Vogler (1997), Saporta (1997) and Naik et al. (1999). Most of the results from their models rely on the assumptions of homogeneous market makers and the zero-profit condition. The justifications of zero profit come from the exchange regulations, Bertrand competition, or the strategic behaviours of market makers. The assumption of homogeneous market makers results in identical bid and ask prices for all of the market makers, which are not often observed in the real world. The exceptions are Ho and Stoll (1983) and Kandel and Marx (1997). Ho and Stoll assume dealers have different reservation prices because of different inventory levels or different opinions about the true price of the security. Kendal and Marx examine the decision of a market maker to quote according to the prevalent best bid and ask.

Two main issues dominate the discussion of quoting behaviour: inventory and information. The inventory consideration drives up the quote when the inventory level is high, and it moves down the quotes when the level is low. As a result, market makers prevent from keeping too much or too little inventory. Even whey they do face undesirable inventory level, they have acquired adequate compensation. The information consideration of quote changes is arguably part of inventory management — market makers want to buy more of desired securities and to get rid of unwanted ones. The difference is that inventory is a firm-specific consideration, whereas information of the securities concerns the whole market.

Quite a few researchers have shown that there are strong information effects on changing the quotes, for example, Hasbrouck (1988, 1991). Evidence of intraday returns such as in Glosten and Harris (1988) and Madhavan and Smidt (1991) may be regarded as supporting information effects because quotes are not different from transaction prices very often. Most of the empirical work employs single dealership models, even though the data are collected from multiple dealership markets (Snell and Tonks 1995, 1998). Moreover, very few attempts are made to explain the diversity of quotes posted by market makers.

2.2.2 Empirical evidence in London

Some work has been done to understand the behaviour of market makers. Reiss and Werner (1996) find the quoting behaviour of market makers in

SEAQ during 1991 is different from what they imagine competitive market makers should be. The average market maker changes his quotes by fewer than seven times a day. The quoted spread of a stock is the same for almost all of the market makers who quote the stock. When they change the quotes, they raise or lower the bid and ask prices by the same amount. Market makers do not often post the best bid and ask price simultaneously. They may post the best bid, they may post the best ask, or they may post neither, in which case they are said to "straddle the yellow strip". Different quotes by different market makers are widely observed, which appears to be in favour of the Ho-Stoll/Kendal-Marx types of models with heterogeneous market makers. These properties of quotes are not limited to only on the London Stock Exchange. For example, Goodhart and Figliuoli (1991) do not find foreign exchange dealers quote the same bid and ask prices. Chan, Christie, and Schultz (1995) find that market makers often quote one side of the inside quote in NASDAQ; almost nobody quotes both best ask and best bid at the same time.

Some efforts have been made to examine the change of quotes. Hansch et al. (1998) conduct a comprehensive survey of the inventory management of market makers. They find some evidence supporting the effects of inventory on quote changes, but they do not consider information effects. Snell and Tonks (1995, 1998) propose models of monopolistic market makers and estimate the inventory and information effects of the change of mid-touch. They justify this approach by assuming market makers come from the same collusive group. This chapter does not hold the view that market makers collude. In fact, the quoting behaviours of market makers are diverse as shown in Board et al. (1997) and later in this chapter. Other work such as Hansch et al. (1999) does not find the evidence of supporting the collusion assumptions, either. The change in mid-touch is not at all uninteresting, but the scope of this chapter will be limited to examine the change of quotes by individual market makers.

Any discussion of quoting behaviour is meaningful only if the quote status affects future order flows. If how market makers quote has nothing to do with the order flows, the quoting behaviour is irrelevant — there is no obvious reason for market makers to stay on or off the yellow strip. By definition, there is always a highest bid and a lowest ask in the market, and there are always somebody posting the best quotes. In absence of the links between quote status and order flow, even if market makers seem to quote aggressively, it is still not the evidence to support the inventory or information hypothesis, and it is not the evidence to support that market makers compete with one another.

How strong the link between order flows and quotes is an important empirical question. Two early studies suggest that the link is less than moderate. Board et al. (1996) examine the relationship of the presence on the yellow strip and the small trades received. They find market makers on average execute 19.58% of the customer trades while they are on the yellow strip. It implies that nearly 80% of the orders are preferenced. When a dummy variable of the presence of yellow strip is included into a regression model of the spread, they find that being on the yellow strip does not make significant difference on the terms of trades. Hansch et al. (1998) compare the market shares of the market makers on the yellow strip with the market shares of those market makers of the whole sample. They conclude the market shares increase by 6.6% if market makers quote the best bid and 5.3% if on the best ask.

If the quote status indeed has little to do with attracting order flows, then market makers may not care about the quote status very much. This is not the case. The next section will take appropriate measure to show that the quote status is closely related to order flows indeed.

2.3 The link between order flow and quote status

2.3.1 Critiques

The finding of the link between order flow and quote status in the literature is, at most, weak. The finding by Board et al. (1996) demonstrates the scale of preferenced order flows of the liquid stocks appears to be large. The increase in the market shares by five to seven percents in Hansch et al. (1998) is difficult to interpret; in fact, it is not the evidence supporting or undermining the link between order flow and quote status. What follows is a simple example to illustrate the change in market shares is a function of many variables that may have little to do with the ability of market makers to obtain order flows. The average increase in market share is defined by Hansch et al. (1998) as

$$\frac{1}{T}\sum_{t=1}^{T}[(\frac{\begin{array}{c}\text{orders of firms}\\ \text{on yellow strip}\end{array}}{\text{total orders of market}})_t - (\frac{\sum_{t=1}^{T}\begin{array}{c}\text{orders of firms on}\\ \text{yellow strip at } t\end{array}}{\sum_{t=1}^{T}\text{total orders of market at } t})], \quad (2.1)$$

where t is the time when some market makes are on the yellow strip. The first term in the square bracket is the market shares of the firm on the yellow strip at time t, and the second term is the total market shares of those firms who are on the yellow strip. Consider the following example. Suppose there are M market makers who take turns to be on the yellow strip with equal probability, and there are m market makers on the yellow strip at any time. There is no preferenced order, market makers will receive no orders if they are off the yellow strip, and the market share of those market makers who are on the yellow strip is 100%. As a result, the total market shares of those who are on the yellow strip are always m/M whoever they are, and the increase in the market share is $(1 - m/M) \times 100\%$. For example, if all of the market makers are always on the yellow strip ($m = M$), then the increase in the market shares is zero. Even in this simple case, the increase in the market shares depends on the number of market makers who are on the yellow strip.

Following the previous example, if some of the market makers stay on the yellow strip longer than the others, then the change in market shares will be reduced. Suppose n out of the m market makers remain on the yellow strips all the time, and the rest $M-n$ market makers take turns on the yellow strip with equal probability. The market share of those who are always on the yellow strip is $n/m \times 100\%$, and of the rest of the market makers who are on the yellow strip is $(1-n/m) * (m-n)/(M-n) \times 100\%$. The increase in market share of being on the yellow strip is

$$(1-\frac{n}{m}-(1-\frac{n}{m})\frac{m-n}{M-n}) * 100\% = \frac{(m-n)(M-m)}{m(M-n)} \times 100\%,$$

which is a decreasing function of n. The more market makers who stay long on the yellow strip, the less the increase in the market share.

Moreover, if $x\%$ of the market share is preferenced, then the increase in the market share will be only limited in the non-preferenced order flow, that is, the increase in the market share is

$$\frac{(m-n)(M-m)}{m(M-n)} \times (100-x)\%.$$

It confirms the argument made by Hansch et al. (1998) that the greater the preferenced order flow, the smaller the increase in market share. If the preferenced order flows are 100%, then the increase in market shares while on the yellow strip is zero.[1]

To sum up, even in the simplest case, the increase in market shares depends on how many and how often market makers are in the yellow strip, and how often they are on the yellow strip. Although the increase in the market share is less than 7% in Hansch et al. (1998), it is not clear that such a small increase is the result of the preferenced order flows or the presence of few market makers who are always on the yellow strip. A more desirable way to detect the link between trades and quotes is to modify the approach by Board et al. (1996), which will be demonstrated below.

[1] Here the preferenced orders are defined as those which will be always routed to a certain market maker. The definition is different from, for example, Hansch et al. (1999), in which they define preferenced orders as those which are executed by market makers who are not on the yellow strip.

2.3.2 Data description

Transaction and quote records are retrieved from the CD-ROM *Transaction Data Service* provided by the London Stock Exchange. Appendix A.2 provides the details of the data. Both the transaction records and quote records contain the identification of market makers. Best quotes are constructed by the quotes records, so whether or not a market maker posts the best quotes is known for any given time. The information of quote status is then matched with the transaction data. Each trade can be classified according to whether the market maker executing the trade is posting the best quote. The sample includes most of the customer and inter-dealer trades of the stocks of which more than one market maker posts the quotes during all the trading days between February and March 1996. Crosses, put-throughs, basket trades and trades without the involvement of market makers are excluded. Crosses are trades by the same market maker as the buy and the sell party. Put-throughs are essentially agency crosses except under the book of market makers; their occurrence is not the result of the quote status of market makers. Basket trades are trades of a portfolio of ten stocks or more simultaneously and are arranged one day before the execution. Although market makers may take into account the effects of the basket trades when posting the quotes, the trades themselves are not attracted by the quotes. The number of trades is calculated from market makers' point of view, so a trade between a broker and a market maker is counted as one trade, and a trade between two market makers is counted as two. Other procedures of editing the data are in appendix A.4.

Summary statistics are shown in Table 2.1. There are more than 1.2 million trades in the sample, and the volumes of trades are more than 80 billion pounds. The direction of the trade is defined as the opposite direction of the market maker. Sells are much more than buys in the sample, but the difference between trading volumes are rather small. It implies the sample consists of relatively big buys and many small sells, which is also revealed in the average size of trade: the average size of buy is nearly eighty thousand pounds, and the average sell is less than sixty thousand. Some of the stocks are thinly traded; ten stocks are without buys and three stocks are without sells. The next row shows the statistics of a stock whose trading volume is the median volume of the sample stocks. As the FTSE-100 stocks dominate the trading on the Exchange, the number of trades and the trading volumes of the median stock are small. The next three rows are data from three different market makers. Market maker A has substantial market shares and makes a market in nearly two third of the stocks. As the market maker trades heavily with brokers who represent small investors, its average trade size is smaller than the average. The fact that the firm trades small stocks frequently also makes the trade size small. Market maker B is among the smallest market makers on the Exchange. It only quotes in twenty-one stocks and trades very rarely, but the average size of the trades is large. Note that Firm B is also a big investment bank, but it entered the market only after the "Big Bang", the creation of SEAQ in 1986. Market maker C is a different type of firm. It quotes half of the stocks on the Exchange and trades frequently with a rather small average trade size. Finally, the last few rows show the number of trades and volumes of the trades excluded from the analysis; those trades only constitute a small proportion of the sample.

2.3.3 Preliminaries

A naive way to examine the relationship between order flow and quote status is to compare the proportion of buy and sell orders when the market makers are on and off the yellow strip. It is illustrated in Table 2.2. Trades are classified according to whether market markers quote the best bid or the best ask, and whether the customers are buyers or sellers. For example, the first two columns show the number of buys of the sample and its percentage as the sum of buys and sells. 193,222 trades occurred when market makers are on the best bid, and 312,683 trades occurred when market makers are off the best bid. The next two columns show the number of sells and its percentage. It appears the sells are always more than the buys regardless the quote status of market makers. Define Δ^N as the percentage of the number of sells subtracted from the percentage of the number of buys. Δ^N ranges from -5.6% to -24.5%. Even if the market maker posts the best ask, the sells are still more than the buys despite the increase of the incoming buy orders.

The sample consists of a lot of small sell trades and some big buy trades. If the trading volumes instead of number of trades are examined, the conclusion may be altered. The right-hand side of the table presents the volumes of buys and sells in the on-bid or on-ask categories. Define Δ^V the percentage of sell volumes subtracted from the percentage of buy volumes. Δ^V is positive when market makers do not post the best bid or when they post the best ask, and it is negative when they are on the best bid or when the are not on the best ask. Therefore, those market makers who are on the best bid attract more sell volumes, and those on the best ask attract more buy volumes. The results from trading volumes are consistent with what is expected to be a "normal" situation: those who post the best ask (bid) attract more buy (sell) orders. However, two questions remain to be answered. First, the absolute value of Δ^V ranges between 6.4% and 11.9%. Is the difference an incentive big enough to let market makers decide to be on or off the yellow strip? Second, both the number of trades and the volumes of the off-bid or off-ask category are always greater than those in the on-bid and on-ask category. Does it mean that market makers attract more orders when they are not on the yellow strip?

The answer to the second question is that an average market maker of liquid stocks spends a great deal of time in straddling the yellow strip, where they can still attract order flows. When market makers straddle the yellow strip, trades are classified into off-bid and off-ask categories in Table 2.2. As a large number of market makers are off the yellow strip, the number of trades executed off the strip is large. Does it mean that straddling the yellow strip attracts more orders? Consider Table 2.3 which classifies the trades into three categories according to the quote status of market makers: on both sides of the yellow strip, on one side only, and straddling the strip. Most of the trades are executed when market makers are quoting one side of the yellow strip, about 30% of the trades are executed when market makers are straddling the strip, and less than 10% of the trades are executed when market makers quote both sides of the strip. The last two columns show the daily duration of each quote status by an average market maker. The percentage of daily duration closely matches the percentages of order flows in each quote status. However, it may not be correct to conclude that market makers will obtain more trades by straddling the yellow strip or by being on one side of the strip. Market makers straddle the yellow strip much more frequently in liquid stocks, where the majority of the trades take place. On the other hand, trades in less-liquid stocks are thin, but market makers are more likely to be on the yellow strip.[2] Hence, the numbers in the first panel of Table 2.3 are seriously distorted.

To clarify the distortion, the second panel classifies the trades by the number of market makers posting quotes during the day in which trades occur, so the stocks are essentially sorted according to the liquidity. The stocks have at most nineteen market makers. To save the space, the panel only shows the statistics of the trades with two, ten and nineteen market makers. When there are two market makers, both of them tend to quote the same bid and ask prices, so all of them are present on both sides of the yellow strip more than half of the day. As a result, "quoting both sides" appears to attract slightly more orders. When they are ten market makers, it is less common to quote both sides of the yellow strip, and "quoting one side" attracts most orders. When there are nineteen market makers, straddling becomes much more common. More trades go to the market makers who straddle the strip, but the majority of the trades still go to those who post the quote on one side. For each group in the panel, the percentage of order flows received when market makers are on one side of the yellow strip is no less than the percentage of the time when they are on the yellow strip. In contrast, the percentage of order flows received when market makers straddle the yellow strip is less than the percentage of the time when they straddle strip.

[2] For example, if there are only two market makers, then either one of them are on both sides of the yellow strip, or both of them are at least on one side of the strip.

The last panel examines the trades of three market makers. Firm A spends most of the time quoting one side of the yellow strip, which contribute 60% of the orders. However, more than 30% of the order flows arrive when it is straddling the yellow strip. Firm B spends most of the time straddling the strip, but most of the trades occur when it is on one side of the strip. Firm C rarely straddles the yellow strips, and most of the trades occur when it is on the yellow strip. To sum up, if there are only few market makers, they are often present on both sides of the yellow strip, and being on both sides attracts orders as much as being on one side. If there are many market makers, straddling the yellow strip is common, but being on one side of the strip attracts more orders than straddling the strip even after considering the duration of quote status.

It requires further investigation to answer the first question raised from the results of Table 2.2, that is, whether a difference between 6% to 12% in order flow is a big incentive to for market maker to post the best quotes. To answer the question briefly, the naive investigation of quote status and order flow as presented in Table 2.2 does not find a big difference between on and off the yellow strip for the following reasons:

1. It does not acknowledge the impact of preferenced order flows, for example, one should expect straddling the yellow strip brings roughly the same buys and sells.

2. It does not consider the unbalanced buys and sells of the market.

3. It does not consider the number of market makers on the yellow strip.

4. It does not consider the size of the trades, for example, big trades often go through lengthy negotiations, and the times of their of executions may have little to do with quote status.

The remaining of this section will discuss the first three issues in turn, and the fourth issue will be dealt with in the next section.

2.3.4 Four-way classification

If quote status has anything to do with the order flows, one may expect those market makers who post only the best ask receive more buy orders than sells, and those who post only the best bid receive more sell orders than buys. If market makers post on both sides of the yellow strip or straddle the strip, it is unclear that whether the market makers wish to receive more buy or more sell orders. In Table 2.2, the trade occurred when the market maker quotes both sides of the yellow strip is classified into both "on-bid" and "on-ask" categories. The trade occurred when the market maker straddles the yellow strip is classified into both "off-bid" and "off-ask" categories. Those trades inevitably obscure the difference of buys and sells received by those market makers who only quote one side of the yellow strip. Therefore, it is essential to examine the order flows by classifying the trades according to four different quote status of market makers, namely, on both sides of yellow strip (on-both), on best ask only (on-ask), on best bid only (on-bid) and straddling the strip (straddle). The four-way classification is shown in Table 2.4. The first panel is the summary of the sample. Similar to Table 2.2, Δ^N is defined as the percentage difference between buys and sells, and Δ^V is defined as the percentage difference between buy volumes and sell volumes. The Δ^N is still always negative in the four categories, but it is the smallest when market makers quote only the best bid, which is -27%. The Δ^N of on-ask category is -7%, which is the second biggest of the four categories. On the other hand, Δ^V equals 14% in the on-ask category, it equals -13 % in the on-bid category. For the rest of the two categories, the differences in volumes are small, only between -1% and -2%.

The second panel examines three subsets of the sample, that is, the groups of trades during the day in which two, ten and nineteen firms make the market. For the two-market-maker group, Δ^N is smallest in the "on-bid" category and biggest in the "on-ask" category. Although it is consistent with the expectation, the magnitude of Δ^N of the "on-bid" category is small. The magnitude of Δ^V is even smaller despite that it is negative for the on-bid category and positive for the on-ask category. As the number of market makers gets bigger, such as in the ten- and nineteen-market-maker groups, Δ^N becomes all negative. The Δ^V of the straddle category is the smallest in the ten-market-maker group, but its magnitude in the on-bid and on-ask category has increased. The magnitude of Δ^V of on-bid and on-ask category in the nineteen-market-maker group are the biggest of the three groups.

The last panel compares the three market makers. The Δ^N and Δ^V of Firm A are similar to those in the whole sample both in directions and in magnitude. Firm B attracts much more buy orders than sells when on the best ask, and much more sells than buys when on the best bid. However, its Δ^N and Δ^V of the straddle category are both negative. Firm C has completely different order patterns. Its Δ^N is positive and Δ^V is negative in all four categories. To sum up, being on bid only attracts more sell orders, being on ask only attracts less sell orders, and the orders are relatively balanced when market makers are on both sides of the yellow strip or are straddling the strip. Nevertheless, the results are less clear for the two-market-maker stocks and for Firm C who quote a lot of small stocks.

2.3.5 Adjusting for order imbalance

The four-way classification in Table 2.4 reveals that the Δ^V of the whole sample is consistent with the expectation. Although there is some evidence that both Δ^N and Δ^V of the on-ask category are often bigger than those of on-bid group, most of the Δ^N is negative. One of the main problems of using the percentages of buys and sells directly, as Δ^N and Δ^V are calculated above, is that the order flows of the market are unbalanced. Section 2.3.2 reveals that the sample consists of a lot of small sells and some large buys. Since the total buys and sells are unbalanced, there is no reason to believe the preferenced orders are balanced. Many preferenced sell orders pour into the market, and the orders are executed by market makers regardless of their quote status. The all-negative Δ^N just reflect the fact that the market is swamped with many sells. On the other hand, the difference of buy and sell volumes is relatively small, and presumably the preferenced buy volumes are not very different from the preferenced sell volumes. The balanced volumes between buys and sells appear to be the reason why Δ^V is more consistent with the expectation that posting the best bid (ask) attracts more sell (buy) trades.

The order flows are not balanced in the aggregate level, and they are even more unbalanced in some individual stocks. Table 2.1 has shown that ten stocks are without buys and three without sells. There are stocks whose sells are much more than buys or the other way round. In order to assess the relationship between order flow and quote status correctly, first exclude the stocks without either buy or sells. Secondly, Δ^N and Δ^V have to be adjusted. A simple way of adjustment is to subtract the means from them. Define

$$\begin{aligned}\text{raw } \Delta^N_{i,j,k,l} &= \text{percentage difference of the number of buy and sell} \\ &\quad \text{trades of market maker } i \text{ with quote status } j \text{ in stock} \\ &\quad k \text{ when there are } l \text{ market makers quoting the stock,} \\ \text{mean } \Delta^N_k &= \text{percentage difference of the number of buy and sell} \\ &\quad \text{trades of stock } k \text{ in the sample,}\end{aligned}$$

and

$$\text{adjusted } \Delta^N_{i,j,k,l} = \text{raw } \Delta^N_{i,j,k,l} - \text{mean } \Delta^N_k .$$

The adjusted $\Delta^V_{i,j,k,l}$ is defined in a similar way. Table 2.5 presents the means of $\Delta^N_{i,j,k,l}$ and their t-statistics. The first row is the means of Δ^N of all stocks by all market makers in different quote categories. The effect of the adjustment is substantial. The mean of Δ^N in the on-both category is -0.08% and not significantly different from zero. The mean of Δ^N in the on-ask category is 22.04%, and that in on-bid category is -15.21%. In other words, after controlling for the market imbalance between buy and sell orders, if a market maker quotes only the best bid, then it will attract more sell orders than buy orders by more than fifteen percentage points on average. Similarly, the market maker who quotes only the best ask will attract on average twenty-two percents more the buy orders than the sells. Finally, when the market maker straddles the yellow strip, it will attract 2.68% more of buy orders, which is significantly bigger than zero statistically, but may be of less economic significance. When Δ^N is averaged according to the number of market makers of the stocks, the results are very similar to the aggregate result for each group. Δ^N is generally small in the on-both category, is between 15% and 40% in the on-ask category, is between -0.9% and -21% in the on-bid category, and is also small in the straddle category. Moreover, almost all of the t-statistics of Δ^N in the on-ask and on-bid category are significantly bigger or smaller than zero at 0.01 level. The last panel examines the mean Δ^N of the three market makers. All of the three Δ^N's in the on-ask category are positive, and those in the on-bid category are all negative. The magnitude of the Δ^N of Firm B tend to be big, but they have to be interpreted with caution. For example, there are only two observations of the firm in the on-both category. Although the mean Δ^N is huge, it is not statistically significant.

Table 2.6 summarises the means of $\Delta^V_{i,j,k,l}$ and their t-statistics. The first row shows the means of Δ^V of all stocks by all market makers in different quote categories. The results are very similar to those in Table 2.5. The means of Δ^V in the on-both and straddle category are -0.08% and -4.76% respectively. The mean of Δ^V in the on-ask category is 15.13%, and that in on-bid category is -18.95%. Therefore, market makers attract more buy volumes when they quote only the best ask, and there are more sell volumes when they quote only at the best bid. The magnitudes of Δ^V in the on-both and straddle category are moderate despite statistically significant. When averaging Δ^V according to the number of market makers of the stocks, it ranges between 4.94% and 24.09% in the on-ask category, it ranges between -13.95% and -28.78% in the on-bid category, and the t-statistics are significantly different from zero at 0.01 level in all but one case. The market maker who straddles the yellow strip or who quotes both sides of the strip tends to receive more balanced order flows. There are much fewer Δ^V in the on-both or straddle category which are significantly different from zero. Even if they are statistically significant, they are almost always bigger than their counterparts in the on-bid category and smaller than in the on-ask category with only two exceptions. Finally, the last panel examines the average Δ^V of the three market makers. With the exception of Firm B in the on-both category, the Δ^V in the on-ask category is always bigger than that in on-both or straddle category, which in turns are bigger than that in the on-bid category.

2.3.6 Number of market makers on the yellow strip

Theoretical models of competitive dealership markets often assume that investors randomly select a market maker who are on the yellow strip to trade with (Vogler 1997). It implies the number of market makers who quote the best prices may affect the imbalance of order flows of those market makers. For example, if there are fifteen market makers quoting the best ask, the non-preferenced buy orders each market maker receives are presumably less than if only one market maker quotes the best ask. The more market makers on the yellow strip, the smaller order flow each market maker receives.

To investigate whether the number of market makers on the yellow strip indeed affects the order imbalance of those market makers, trades are classified by whether the market maker is on the yellow strip, and by how many market makers are on the yellow strip. Three categories are identified if the market maker is on the yellow strip: (a) there is only one market maker standing alone on the best bid (ask), (b) the number of market makers on the best bid (ask) is less than half of the total number of market makers who post the quote, and (c) the number of market makers on the best bid (ask) is equal to or more than half of the total number of market makers. Category (d) includes the trades by the market makers who are not on the yellow strip. Compared with the four-way classification used in Table 2.4, 2.5 and 2.6, if the best ask is examined, the trades in the previous "on-bid" and "straddle" category will fall into the category (d), the trades in the "on-both" category will be mostly in category (c),[3] and the trades in "on-ask" category and a small part of the "on-both" trades will be in either category (a), (b) or (c) according to the number of market makers on the best ask. Likewise, if the best bid is examined, the trades in the previous "on-ask" and "straddle" category will be in category (d), the "on-both" trades will be mostly in category (c), and the "on-bid" trades and a small proportion of "on-both" trades will

[3] The quote of a market maker is in "on-both" category often means the quotes of everybody else is in "on-both", too. See Section 2.5.

be in one of the category (a), (b), or (c) according to the number of market makers on the best bid. Define Δ^V like before:

$$\begin{aligned}
\text{raw } \Delta^V_{i,j,k,l} &= \text{percentage difference of the buy and sell volumes of} \\
&\quad \text{market maker } i \text{ in category } j \text{ in stock } k \text{ when there} \\
&\quad \text{are total } l \text{ market makers quoting the stock,} \\
\text{mean } \Delta^V_k &= \text{percentage difference of the buy and sell volumes of} \\
&\quad \text{stock } k \text{ in the sample,} \\
\text{and} & \\
\text{adjusted } \Delta^V_{i,j,k,l} &= \text{raw } \Delta^V_{i,j,k,l} - \text{mean } \Delta^V_k.
\end{aligned}$$

It should be noted that Δ^N can be defined in a similar way, but the results obtained by using Δ^N are not different from those by Δ^V. Table 2.7 shows the means of the adjusted Δ^V of the four categories on the ask side. The mean of the percentage difference between buy and sell volumes is 29.68% when the market maker is alone on the best ask, is 15.72% when the market maker is with than less than half of the market makers on the best ask, is 0.39% when the market maker is with more than half of the market makers, and is −12.69% when the market maker is not on the best ask. The table also provides the t-statistics of the difference between the adjacent Δ^V. The mean Δ^V in category (a) is significantly bigger than that in category (b), which is in turns significantly bigger than that in category (c), which is significantly bigger than that in (d). The number of market makers on the best ask appears to play an important role in determining the order imbalance: the fewer market makers on the yellow strip, the more unbalanced the orders each one receives.

The next panel groups the trades according to the number of market makers posting the quotes during the day. The mean of Δ^V ranges between 12.40% and 47.37% in category (a), between 6.78% and 25.27% in category (b), between −10.05% and 6.95 in category (c), and between −16.88% and −7.19% in category (d). With only one exception, Δ^V in category (a) is always greater than that in category (b), which is in turns greater than that in (c), which is greater than that in (d). The t-statistics are significant at

0.01 level in most cases. The last panel shows the Δ^V of the three chosen market makers. The pattern of Δ^V of Firm A is similar to what is observed in the aggregate level. For the remaining two firms, the results are less clear. Although Δ^V in category (a) is still bigger than that in (d), it does not monotonously decline from category (a) to (b) to (c) and to (d). Nor are most of the t-statistics significant.

Table 2.8 shows the means of the adjusted Δ^V of the four categories on the bid side. The results are again consistent with the hypothesis that the order imbalance of the market maker quoting the best bid decreases in the number of market makers on the best bid. The mean Δ^V of category (a), (b), (c) and (d) are respectively -28.44%, -22.75%, -11.29% and 6.31%. The differences in the adjacent Δ^V are all significantly smaller than zero at 0.01 level. When the trades are grouped by the number of market makers posting the quotes, Δ^V of category (a) is almost always smaller than that of category (b) with only one exception, Δ^V of category (b) is always smaller than that in category (c), which is in turns smaller than that in category (d). More than half of the t-statistics are significant. Finally, the rising pattern of Δ^V can be found in the trading by Firm A and Firm B, but it is not the case for Firm C. However, many of the t-statistics at the firm's level are not significant.

2.4 Quote status, order flows, and trade sizes

In a dealership market, if market makers do not post the best quotes, they can still receive some orders from the customers. The orders come from two main sources. The first source is preferenced order flows. Last section has shown that how preferenced orders increase the difficulty to detect the link between quote status and order flow. The second source of the orders is large trades. Market makers only honour the quotes up to a certain size. The price of a trade with the size bigger than the quote size has to be negotiated between the market maker and the customer. When a customer seeks to execute a large order, it is not necessary for her to trade with those market makers who are posting the best quotes. Market makers who do not post the best quotes may be willing to offer more competitive prices for the large trades than those who are on the yellow strip do. It is possible, however, that posting the best quotes signals the willingness of the market makers to execute the large trades with lower costs. The signalling effect will strengthen the link between quote and order flow, but the absence of the link between large trades and quotes, is by no means an indication of the lack of competition in the market.

2.4.1 Order imbalance and trade sizes

To know whether the link between order flow and quote status exists for any sizes of trades, trades are classified into ten groups based on the multiples of the most popular quotes sizes of the stocks in the sample (the Modal Quote Size, or MQS), which are often the same as the NMS.[4] Summary statistics are presented in Table 2.9. The majority of the trades are of less than 0.05 MQS multiples, and there are much more sells than buys in this group. There is not much difference between buys and sells when the size becomes bigger. Despite the number is large, the very small trades only contribute less than 3% of the volumes to the market. The volumes concentrate on the trades with the sizes between one and seventy-five times MQS.

Define Δ^N similar to before as the adjusted percentage difference between the numbers of buys and sells in one of the four-way category of a particular stock by a particular market maker within a particular size range. Define Δ^V as the percentage difference between buy and sell volumes of a stock of a market maker within a particular size range. Table 2.10 presents the means and t-statistics of Δ^N and Δ^V in different size groups. All the percentage differences are adjusted for the total market imbalance. The first panel shows the means of Δ^N. The familiar pattern emerges from most of the size group. The Δ^N in the on-ask category is almost always bigger than that in on-both or straddle category, which is bigger than that in on-bid category. Δ^N in

[4] *Transaction Data Service* does not provide NMS data.

the on-ask category is positive except for the group of trades bigger than 75 times MQS, and it appears to be bigger in the medium-size groups than in the small-size or large-size groups. The biggest Δ^N in the category is 45.72%, which occurs with the trades between 0.5 and 1 times MQS. On the other hand, Δ^N in the on-bid category is negative except for the two biggest groups. The smallest Δ^N in the on-bid category is -28.06%, which occurs with the trades between 0.5 and 1 times MQS, too. The case in Δ^V is very similar, as shown in the second panel. Δ^V in the on-ask category in each size group is often bigger than that in on-both or straddle category, which is in turns bigger than that in on-bid category. Δ^V of on-ask category is positive for the size groups between 0.1 and 6 times MQS, but it is negative elsewhere. Δ^V of on-bid category is negative except for the group with the size between 6 and 75 times MQS.

Overall, the most serious order imbalance in both on-bid and on-ask categories occurs around the size of one times MQS. When market makers post only one side of the best quotes, they are more likely to attract the orders of medium-size trades from one direction. Their ability of attracting trades bigger than one times MQS implies that being present on the yellow strip essentially advertises their willingness to trade with the public at any size. It is not surprising the means of Δ^N and Δ^V of the group with 75 times MQS or more show different signs in on-bid or on-ask category. The very large trades often experience lengthy negotiations. Some of them are pre-arranged (Franks and Schaefer 1995); some of the trades are matched by market makers between two or more customers. All of the particularities of very large trades imply that market makers are aware of the trades long before they are executed. Therefore, the quote status at the time of trade execution has little to do with the big trades. What is really interesting is the order imbalance of the very small trades is less severe than all but the very large trades. Order preferencing seems to be prevalent in small trades, and being on one side of the yellow strip does not attract many orders from one direction. In fact, the result is consistent with the early investigation of Board et al. (1996), in which they examine exclusively small trades and cannot detect the difference between on and off the yellow strip.

2.4.2 The real spread and order flows

Numerous researchers have shown that the touch spread is wider than the effective spread (Wells 1992; Board and Sutcliffe 1995; Reiss and Werner 1996). It implies that market makers often offer better prices on the phone than the yellow-strip prices. It may be argued that the effective spread offered on the phone is the "real" spread that the customers of market makers obtained, and the order flows depend on the real spread instead of the touch. Therefore, the "real-spread hypothesis" states that those market makers who offer a smaller spread (on the phone) receive more orders. Unfortunately, the telephone conversations between market makers and the customers are not available, and the source of real spread is the observed effective spread from transaction data.[5] When there is no trade, there is no way to find out what the real spreads are, so this hypothesis can only be examined when more than two market makers obtain the order flows at the same time.

To test the real-spread hypothesis, obtain the number of trades in different size categories by each market maker in a certain period of time, and calculate the average effective spread. The "certain period of time" is either a day or an hour. Trades are grouped according to the trade sizes because effective spreads are related to trade sizes (de Jong et al. 1995 and Chapter 4). Buys and sells are grouped separately because market makers may offer different spreads for the trades in different directions. For example, if a market maker wishes to attract more buys and less sells, then he should offer both low ask and low bid price relative with the mid-touch, so the effective spread of a buy is smaller than that of a sell. The effective spread is s defined in (1.1). An alternative of the effective spread is to measure the "gain" from the trade, i.e., the distance between the transaction price and the best bid or best ask price (Board and Sutcliffe 1995; Reiss and Werner 1997). However, Chapter 3 will argue that this measure is undesirable when the size of touch spread

[5] As is argued in Chapter 4, the spread is an unobservable component of the transaction price. The effective spread used in this section serves as a proxy for the real spread.

varies.

The real-spread hypothesis implies the number of trade is negatively related to s. Within each size-time-direction group, calculate the correlation coefficients of the number of trades and the effective spread. Table 2.11 reports the means of the correlation coefficients. The first panel presents the results by grouping the trades every day. The first column shows the mean coefficients between the number of trades and s. The mean coefficient is 0.04, which is positive and statistically significant, contrary to the real-spread hypothesis. When the coefficients are grouped according to the size, the coefficients of the smallest and the largest trade group are positive, but the coefficients are negative when the trade size is between 0.1 and 75 times MQS. The smallest coefficient is -0.12 with the group of trades between two and three times MQS. The second panel shows the coefficients when the trades are grouped every hour. The mean coefficient of the full sample violates the real-spread hypothesis. The violation of the hypothesis again comes from the small trades, most of which may be executed electronically and without negotiation. The coefficients of medium and big trades are consistent with the hypothesis.

Table 2.11 provides the evidence supporting the real-spread hypothesis for large trades but not small trades. The "real" quotes on the phone appear to play a role in attracting order flows. The main problem of measuring the real spread, however, is the lack of the data. First, telephone conversations are not available, so the spread has to be measured from the transaction data. Second, since the spread is measured from the transaction data, there is a spread only if there is a trade. It is not known that whether those market makers who do not receive order flows during a day or an hour offer better or worse quotes than those who do. It is even very rare for liquid stocks that two or more market makers execute the trades of the same stock within the same size group at the same time period in the same trade direction. The last two columns of Table 2.11 shows the number of trades and volumes of the groups. Compared with Table 2.9, the trades used in testing the real-spread hypothesis are only a small proportion of the full sample. Nevertheless, the bottom line is that the order flows of medium trades, and large trades to some extent, appear to be negatively related to the real spread offered by market makers on the phone.

2.5 Quoting behaviour

2.5.1 Preliminaries

Because the quote status is related to the number and the balance of the orders, market makers must decide how to quote thoughtfully. The quoting behaviour of market makers is examined in this section. All the stocks with quotes in the forty-two trading days between February and March 1996 are included. Apart from the 1832 stocks used in Section 2.3 and 2.4, another ten stocks without trade data are included. For each stock, there were at least two market makers maintaining the quotes for all of the forty-two trading days during the two-month period.

The mandatory quote period is between 8:30 and 16:30 every weekday. Most of the market makers start posting the quotes well before 8:20, and most of them withdraw the quotes before 16:40. Occasionally, some of the market makers may be late in posting the quotes, sometimes even after 9:00, and some of them may withdraw the quotes before 16:30. To avoid the case in which some of the market makers are inactive, only the period during which all of the market makers are present is considered. In other words, if the quote of any market maker is not available, then the rest of the quotes are not used, either. Whether or not the quotes are inside the mandatory quote period is irrelevant. An implicit assumption behind this sample selection criterion is that quotes outside the mandatory quote period are as informative as inside.

Table 2.12 provides summary statistics of the quote data. The stocks are classified according to the number of market makers present during the day. Market makers are free to provide or withdraw the quotes for any stock after notifying the Stock Exchange in the previous day, the number of market makers for any stock is stable but not necessarily constant, and the stocks may be included into more than one of the market-maker groups. The maximum number of market makers in the sample is nineteen, more than those in Reiss and Werner (1996) and Hansch et al. (1998), in which both are seventeen. Column three and four are respectively the medians of quote and inside spreads. The next column shows the median of quote sizes in pounds sterling, which is defined as

$$\text{quote size}=0.5 \times (\text{bid quote} \times \text{bid quantity}+\text{ask quote} \times \text{ask quantity}).$$

The next column provides the median number of quotes per market maker per day. If there are few market makers, then they often maintain the same quotes during the day, and the median number of daily quote is one. The more market makers, the more quote changes they make. However, even thought there are nineteen market makers, a market maker on average only posts eight quotes per day, or in other words, changes the quote once per hour on average. The quote spread, the touch spread, the quote size and the number of quotes per day in the sample are very similar to what have been reported in the literature (Board and Sutcliffe 1995; Reiss and Werner 1996; Hansch et al. 1999).

The last four columns of the table indicate the means of the number of market makers in one of the four quote status during the day: at best ask only, at best bid only, at both best bid and ask, and straddling the yellow strip. It is the major difference between the current sample and those in previous studies. The numbers of market makers who straddle the quotes are bigger than those found in Reiss and Werner (1996), and there are less market makers who are on the yellow strip. Reiss and Werner (1996) find the median of market makers on the best bid, ask and straddle are 4,4 and 1 respectively. Hansch et al. (1998) report that 70% of market makers quote one side, and there are typically two to four market makers at either side at any time. Board et al. (1996) also report that more than a half of market makers are on either side of the best quote. In contrast, straddling the yellow strip seems to be more common in the current sample.

Despite the apparent difference, this sample exhibits some other features similar to those in previous studies. Table 2.13 presents the common practices of quoting and trading behaviour of market makers. First, as documented in Board et al. (1997), market makers use a few common quote spreads, among which three, five, seven and ten pence are the most popular ones. Coincidentally, Goodhart and Figliuoli (1991) find five, seven and ten are commonly used values in foreign exchange markets. Second, although there is no formal tick rule, market makers appear to post the quotes in a certain units. Define the implicit tick of a stock as the minimum difference of the consecutive ask prices, then it turns out that most of the quotes are rounded in pence, similar to Board et al. (1997), where 97.5% of the quotes are rounded in pence. There are more trade prices that are rounded in half a penny, but the majority the trade prices are still in pence. Regarding the scale of change in quotes, it is very common for market makers to move the mid-quote by one or two pence. Finally, more than 99% of the quote changes are to move the mid-quotes upwards or downwards, mostly in the same direction of the preceding change (78%) and even by the same scale (45%). Only a tiny proportion of the quote changes involves a change in the quote spread or quote size.

2.5.2 Quote status

Table 2.12 suggests the market makers who straddle the yellow strip are usually more than those who post the best quotes. Table 2.14 provides further evidence of the behaviour of market makers. The quote status of market makers is classified into five categories: alone on both sides of the yellow strip, on both sides with another market makers, on one side alone, on one side with another market makers, and straddling the yellow strip. As will be shown shortly, market makers behave very differently when they are alone on the best quote.

The left-hand side of Table 2.14 reports the average time during a day for different quote status. For example, a market maker who posts quotes in a two-market-maker stock is on average on both sides of the yellow strip alone for two minutes per day, and with the other market maker for four hours and thirty-five minutes. Note that the sum of the duration of each market-maker group may be less than eight hours (the length of mandatory quote period) as only the time when all market makers post the quotes is counted. The less the market makers, the longer they are present on both sides of the yellow strip. Most of the market makers of the same stock always quote the same spread for any given stock. If everybody quotes the same spread, then one market maker is on both sides of the best quote if and only if everybody is on both sides.[6] Thus, "not alone on both sides" often implies everybody quotes the same bid and ask prices. When everybody quotes the same price, the touch spread is much wider than its average. This property is also found with the quotes on NASDAQ (Chan, Christie, and Schultz 1995). However, such occasions are not very common when there are a large number of market makers. "Alone on both side", on the other hand, implies the quote spread of the market maker is narrower than those of the others, and it rarely occurs. The time when market makers are on one side of the best quote declines with the number of market makers. The next two columns show the time when market makers are on one side of the yellow strip. Again, market makers are much less likely to stand alone on the strip. As the number of market makers increases, they spend more time straddling the yellow strip. When there are more than fifteen market makers, a typical market maker straddles the yellow strip for half of the day.

The right-hand side of the table shows the daily average frequency of which a market maker moves into a new quote status. A new status includes the opening status, the change from one of the five status to another, and although rarely occurred, the change between the best bid and the best ask

[6] Suppose M market makers quote bid and ask price $(a_1, b_1), (a_2, b_2), \cdots, (a_M, b_M)$ respectively. The best bid is $\max(b_1 \cdots b_M)$, and the best ask is $\min(a_1 \cdots a_M)$. Now suppose everybody quotes the same spread, that is, $a_i - b_i = a_j - b_j, \forall\, i, j \in \{1, 2 \cdots, M\}$. If $\exists\, k$, $a_k = \min(a_1 \cdots a_M)$ and $b_k = \max(b_1 \cdots b_M)$, then $\forall\, i$,

$$a_i - b_i = a_k - b_k = \min(a_1 \cdots a_M) - \max(b_1 \cdots b_M) \leq a_i - b_k.$$

Hence, $b_i \geq b_k = \max(b_1 \cdots b_M)$. Similarly,

$$a_i - b_i = a_k - b_k = \min(a_1 \cdots a_M) - \max(b_1 \cdots b_M) \leq a_k - b_i.$$

$a_i \leq a_k = \min(a_1 \cdots a_M)$. That is, for any given i, $a_i = a_k$, $b_i = b_k$.

within the "on one side" categories. In general, the number of frequency increases with the average time of staying in the status, but it is not always the case. The most notable exception is the frequency of straddling the best quote is always less than that of on one side (not alone), even when the length of time of straddling is longer. It implies market makers move out of the on-one-side status quicker than straddling. The last column of the Table 2.14 sums up the daily average frequency. The changes in quote status are more than the changes in quote posted shown in Table 2.12. Every change in quote status results from the change of quotes either by the market maker himself or by somebody else, so the quote status may change when the quotes are unchanged.

Although Table 2.14 shows that an average market maker straddles the yellow strip most of the time, the time when only one market maker stands on one side of the yellow strip is very long. For example, the table shows an average market maker in 19-market-maker group stays on one side of the yellow strip sixteen minutes a day. It means the time when there is only one market maker on one side of the yellow strip is $16 \times 19 = 304$ minutes. Suppose the time of being alone on best bid and best ask are equal, then on average there is only one market maker on either side for more than two and half hours a day. While most of the market makers are straddling the strip, some market makers appear to stay on the strip very often. This finding is consistent with Board et al. (1997) in which they find a few market makers act as price leaders and update the quote rather aggressively.

2.5.3 Willingness to post the best quote

The distinction of the source of change is important especially when the market maker is alone on the yellow strip. If a market updates either the lowest ask or the highest bid to stand alone on the yellow strip, then he is a price leader. On the other hand, if a market maker is left alone on the yellow strip, it may imply that he is a slow mover. The attempt to make the distinction between the two sources of quote change is shown in Table 2.15. Each category of quote status is further divided into two according to the source of the change: "S" represents the change is made by the market maker himself, and "O" represents the change is made by the others. Only the frequency of quote status is reported. First, the difference in the numbers of status between the S and O categories on both sides of the yellow strip increases proportionally to the increase in the number of market makers. The difference can be seen clearly by looking at the O/S ratio, defined as the number of O changes divided by that of S changes. The ratio rises from 1.00 when there are two market makers, to 9.34 when there are ten, and to 17.89 when there are nineteen. If there are M market makers, then the O/S ratio in this column is close to $M - 1$, which is a direct result from the fact that "one market maker quotes the best bid and ask price" often implies that "the rest of $M - 1$ market makers quotes the best bid and ask price, too."

A smaller difference between S change and O change is in the status in which the market maker is on one side of the yellow strip. In the "alone on one side" category, market makers are slightly more often moved into the status by others than by themselves. S changes are less frequent than O changes when the market makers are few, and are nearly as much as O changes when market makers are more. S changes are less frequent in "not alone on one side" category. The highest O/S ratio in the category is 3.1, where there are nineteen market makers. However, the difference in this category is the major contribution to the gap between the total sum of S and O changes: market makers do not voluntarily move to one side of the best quote very often. Instead, their status as on one side often results from the quote changes by the others. Finally, the move to straddle the yellow strip is also slightly more likely by O changes than S changes.

After the source of the change is examined, the next step is naturally to investigate what ends the current quote status. The end of the quote status may not only result from the quote change by the market maker himself or by somebody else, but it may also result from the withdrawal of the quote at the end of the day. As the proportion of withdrawal does not vary substantially across different quote status categories, the focus is to compare the numbers of S changes and O changes.[7] Table 2.16 contains the O/S ratio of *subsequent* changes. If the ratio is smaller than one, it implies the status is ended by the market maker himself more often than by somebody else. If the ratio is greater than one, the status is changed more often by somebody else. If the source of the current status is an S change, that is, the market maker moves into the status by his own choice, then he should be happy to stay in the status, and it is more likely that somebody else forces him out. On the other hand, if the current status is an O change, the market maker is moved into the status involuntarily, then he is more likely to end the status by himself. To sum up, an S change is expected to be followed by a big O/S ratio, and an O change to be by a small O/S ratio.

Table 2.16 reports the subsequent O/S ratios by the quotes status and by whether current quote is an O change or an S change. First, examine the two "on one side" categories. The ratios in the two S columns under "on one side" categories are all greater than one, and the ratios in the two O columns are smaller than one with only three exceptions. Moreover, the ratios of the O column on the "alone" category are mostly between 0.19 and 0.31, while those in "not alone" column are at least 0.50. It means that a market maker is much more likely to get out of the best quote voluntarily when he is left alone than when he is there with somebody else. The ratios are very similar for the two S columns.

What if the market maker is on both sides of the yellow strip? The little evidence on the two "alone" categories does not provide too many clues. However, the numbers in the two "not alone" columns are all greater than one with only one exception. Once the market maker is moved to be on both sides of the best quotes with other market makers, than regardless of the source of the change, the market maker is not likely to move to another status by himself. The same is true in the straddle category: once moved in, the market maker often stays in the status until somebody else forces him

[7] Nonetheless, a substantial proportion of the next quote change is withdrawal, for example, it is 59% when there are only two market makers. If there are more market makers, there are more changes in the quote status during the day, and the proportion of withdrawal falls to 5% when there are nineteen market makers.

out of it.

If a market maker ends the quote status by himself, to which direction does he move the quote? As has been shown in Table 2.13, 99.95% of the quote changes move the mid-quote only; the quote spread and the quote size remains unchanged. To examine the direction of the change, all the quote status preceding quote changes which only move the mid-quote are classified into seven categories: alone on both sides, not alone on both sides, straddle, alone on ask, not alone on ask, alone on bid, and not alone on bid. Take the means of the subsequent changes in the mid-quote by each current status in each stock to perform t-tests provided that there are at least five changes in the category. 656,737 quote changes in 1,537 stocks are used to perform 6,153 t-tests. Instead of presenting more than six thousand numbers, Table 2.17 groups the t-statistics by their sizes. The break-up points are 0, ± 2 and ± 3.

If the quote status does not affect the way in which a market maker changes his quote, then the means of the changes in mid-quotes are similar across the status categories.[8] The test result shows it is not true. If market makers change the quotes when they are straddling or on the both sides of the yellow strip, then it is unclear whether they will move the mid-quote upwards or downwards. Most of the t-statistics in those categories are between -2 and 2, neither of which is significantly greater than or smaller than zero.

If the market maker changes the quote when he is at the best ask but not the bid, then he is more likely to move the quote upwards instead of downwards. In fact, the means are significantly greater than zero in overwhelming majority of the stocks. The reverse is true when the market maker is on the best bid: he is more likely to move the quote downwards than upwards in most of the stocks. Moreover, the tendency is more pronounced when the market maker is alone on the yellow strip. 95% of the t-values of change in mid-quote are greater than two when market makers are alone on the best ask, while the proportion is 91% when market makers are with somebody

[8] The means of the change need not be zero. In a bull market, most of the quotes are adjusted upwards. In a bear market, most of the quotes are adjusted downwards.

else. When the market makers are on the best bid, the proportions of t-values less than two are 91% in the "alone" category and 79% in the "not alone" category. Both moving up from best ask and moving down from the best bid means moving *away* from the best quote, which is consistent with the other findings earlier in this section that market makers prefer straddling the yellow strip to standing on one side of the quote.[9]

To assemble the evidence together, a typical market maker exhibits certain patterns in posting the quotes. He spends a lot of time to stand on both sides of the yellow strip when there are few market makers, and to straddle the strip when there are a lot of market makers. He does not often move to one side yellow strip voluntarily. When he is forced to be on one side of the yellow strip, he often moves away. On the other hand, he often chooses to straddle the yellow strip, and he is willing to straddle even it is not by his own choice.

2.6 Causes of quote change

2.6.1 Regression analysis

Section 2.3 and 2.4 examine what the consequences of quote changes are, Section 2.5 documents the patterns of quote changes, and it is time to investigate what trigger the quote change. The technique used here is cross-sectional regressions. The dependent variable is the difference of mid-quote, and the independent variables are the inventory levels, the previous sum of trades, the previous trading volumes, and the preceding quote changes. The regression is run on the full sample, so all of the variables are transformed to avoid the possible distortion brought by the difference of the stocks. For example, liquid stocks have more trades and smaller quote change; illiquid stocks have less trades and bigger quote change. If the regression were run on the unad-

[9] Table 2.13 shows the popular absolute quote changes are to move the mid-quote by one or two pence, while the popular quote spreads are five, seven, three and ten pence. Consequently, a quote change is generally not enough to move from one side of the best quote to the other.

justed variables, a negative and false relationship between trades and quote change would be detected. The dependent variable of the regression model is $\Delta m_{t,i}$, the mid-quote of market maker i after the change at time t minus the mid-quote before the change. Note that the mid-quotes of market makers are not necessarily the mid-touch. The quote changes which does not move the mid-quote are not included in the analysis. More than 99.95% of the quote change is to move up or down both bid and ask prices by the same amount, so only a few observations are dropped. To standardise $\Delta m_{t,i}$, the variable is divided by the mean of the absolute value of quote change of the stock.

The independent variables can be classified into six groups:

1. The number of trades executed by market maker i between the current quote change and the preceding quote. Basket trades, put-throughs and crosses are now included as they may affect the decision of the change in quotes. Trades are classified according to the trade directions and the time of the trades, and then the number of trades are divided by the daily average number of the trades of the stock executed by the firm in the sample. Six variables are included in the regression: $x_{i,b,0}$ is the (standardised) number of buys executed by market maker i that take place less than one minute before the quote change, $x_{i,b,1}$ is the number of buys between one and two minutes before the quote change, and $x_{i,b,2}$ is the number of trades between two minutes before the quote change and the time when the preceding quote is posted. The preceding quote is the latest quote posted in the market before the quote change, and it may be posted by another market maker. The number of sells are defined in a similar way and denoted as $x_{i,s,0}$, $x_{i,s,1}$, and $x_{i,s,2}$.

2. The trading volumes executed by market maker i between the current quote change and the preceding quote. Volumes are classified according to the trade directions and the time of the trades, and then are divided by the average daily volumes of the stock executed by the firm in the sample. Six variables are included in the regression, and they are defined in a similar way as when the number of trades are defined: $v_{i,b,0}$ is the standardised buy volumes by market maker i taking place

less than one minute before the quote change, $v_{i,b,1}$ is the buy volumes between one and two minutes before the quote change, and $v_{i,b,2}$ is the volumes between two minutes before the quote change and the time when the previous quote is posted. The numbers of sell volumes are defined in a similar way and denoted as $v_{i,s,0}$, $v_{i,s,1}$, and $v_{i,s,2}$.

3. The number of trades without the involvement of the market maker who is changing the quote. The trades are classified according to the time of the trades and "trade directions". When the market maker does not execute the trades, he cannot observe the trade directions directly, and the directions are identified by matching the trade prices with the mid-touch. That is, the trade is a buy if its price is above the mid-touch, it is a sell if the price is below the mid-touch, and the trades whose prices equal the mid-touch are excluded from the analysis. The number of trades is then divided by the daily average number of trades of the stock without the involvement of the market maker. Six variables are included in the regression: $x^o_{i,b,0}$, $x^o_{i,b,1}$, $x^o_{i,b,2}$, $x^o_{i,s,0}$, $x^o_{i,s,1}$ and $x^o_{i,s,2}$ are the buys and sells without the involvement of market maker i. They are defined in a similar way as the counterparts $x_{i,*,*}$.

4. The trading volumes without the involvement of the market maker who is changing the quote. Volumes are classified according to the time of the trades and "trade directions", which are determined by the prices of the trades and the mid-touch when the trades take place. Then the volumes are standardised by divided by the daily average volumes of the stock without the involvement of the market maker. Six variables are included in the regression: $v^o_{i,b,0}$, $v^o_{i,b,1}$, $v^o_{i,b,2}$, $v^o_{i,s,0}$, $v^o_{i,s,1}$ and $v^o_{i,s,2}$. They are defined in a similar way as the counterpart $v_{i,*,*}$.

5. The preceding quote change of the stock, which may be made by the same or different market maker. Three variables are included according to the time when the preceding change occurs. If the two changes are less than one minutes apart, then $\Delta m_{t-1,0} = \Delta m_{t-1}$, $\Delta m_{t-1,1} = 0$, and $\Delta m_{t-1,2} = 0$. If the time difference of the two changes are between one

and two minutes, then $\Delta m_{t-1,0} = 0$, $\Delta m_{t-1,1} = \Delta m_{t-1}$, and $\Delta m_{t-1,2} = 0$. If the time difference of the two changes are more than two minutes, then $\Delta m_{t-1,0} = 0$, $\Delta m_{t-1,1} = 0$, and $\Delta m_{t-1,2} = \Delta m_{t-1}$. If there is no preceding quote change, for example, the preceding quote is an opening quote, then $\Delta m_{t-1,0} = \Delta m_{t-1,1} = \Delta m_{t-1,2} = 0$.

6. The inventory of the firm immediately before the quote change. Assume the inventory is zero in the beginning of the sample period. The inventory level is defined as the accumulated sell volumes minus buy volumes.[10] Two variables are included. I_t is the standardised inventory, the inventory subtracted from the average end-of-the-day inventory of the market maker and then divided by the standard deviation (see Hansch et al. 1998.) If there is no trade at all, then $I_t = 0$. The second variable is ΔI_t, the change of inventory between the current quote change and the preceding quote by market maker i.

   The unadjusted change in inventory is the sum of sell volumes minus the sum of buy volumes between the current quote change and the preceding quote, so including ΔI_t and the variable group two together in the regression may cause colinearity. However, while ΔI_t is the sum of volumes between two quotes of market maker i, the volumes in the second group of variable only include those between the current change and the preceding quote, which may not be made by the same market maker. The correlation coefficients between ΔI_t and each variable in group two are small, and the correlation coefficient between ΔI_t and $\sum(v_{i,s} - v_{i,b})$ is -0.31.

According to inventory hypothesis, $x_{i,b,*}$, $v_{i,b,*} > 0$ and $x_{i,s,*}$, $v_{i,s,*}$, I_t, $\Delta I_t < 0$. According to information hypothesis, if trades contain information, then $x_{i,b,*}$, $v_{i,b,*}$, $x^o_{i,b,*}$, $v^o_{i,b,*} > 0$ and $x_{i,s,*}$, $v_{i,s,*}$, $x^o_{i,s,*}$, $v^o_{i,s,*}$, I_t, $\Delta I_t < 0$.

[10] Note that buy and sell are defined from the customers' point of view.

2.6.2 Evidence

The result of the regression of the full sample is presented in Table 2.18. Two regression models are used. The first one employs all of the six groups of independent variables, and the second one omits Group 2 variables to avoid the possible colinearity. The coefficient values and t-statistics are shown. The statistics are adjusted for heteroscedasticity by White's (1980) approach. The results from the two models are very similar. All of the coefficients of $x_{i,b,*}$ are positive and of $x_{i,s,*}$ are negative, which means the quotes are likely to go up after executing buy trades and are likely to go down after sell trades. The absolute value of $x_{i,*,*}$ is decreasing as the time difference increase, which means the quote changes are more likely to associated with the nearest trades. However, the second group of variables, $v_{i,*,*}$, does not exhibit similar properties. $v_{i,b,*}$ is not always positive, and $v_{i,s,*}$ is not always negative. The t-statistics are not significant except for $v_{i,b,2}$. The estimates of Group 3 and Group 4 variables exhibit a similar pattern to those in group one. All of the coefficients of $x^o_{i,b,*}$ and $v^o_{i,b,*}$ are positive and all of $x^o_{i,s,*}$ and $v^o_{i,s,*}$are negative. The t-values of Group 3 variables are all significant at 0.0001 level, so are most of the t-values of Group 4 variables. Furthermore, recall that the sample consist of more sells than buys. Given the fact that the coefficients of buy trades and buy volumes are not much bigger than those of sell, it follows that the quotes were moved downwards in the sample period. This is exactly what happens with most of the stocks. There are more downward changes of the stocks with more than fifteen market makers, where most of the trades occur. The most important variables, however, are the preceding change in the mid-quote. The coefficients of $\Delta m_{t-1,0}$, $\Delta m_{t-1,1}$, and $\Delta m_{t-1,2}$ are all positive and the t-values are much bigger than zero. The coefficient value of $\Delta m_{t-1,0}$ is greater than that of $\Delta m_{t-1,1}$, which is greater than that of $\Delta m_{t-1,2}$. The decreasing estimates imply the more closely the current quote change to the preceding change is, the more likely the former resembles the latter. In fact, the preceding change explains more than 40% of the variation of the quote change. After introducing group five variables, the rest of the variables only improve the R^2 marginally. In contrast, the inventory variables do not have much effect on the quote change. The coefficient estimates are very close to zero, and the t-statistics are not significant at all. The second regression still does not show the inventory variables explain the quote change very well, either. The bottom of the table shows the R^2 and the number of observations (N) used in the models.

It is interesting to examine whether the results from the regression of the full sample hold for the subsets of the sample. Table 2.19 reports the regression results of the quote changes during the day when there are two, ten and nineteen market makers present. Group two variables are not included in the regressions. The results are very similar to those reported in Table 2.18. All of the coefficients associated with buy trades and buy volumes are positive with only two exceptions. All of the coefficients associated with sell trades and sell volumes are negative without any exceptions. The t-statistics are not as significant as those in Table 2.18, but most of them are still significantly different from zero. The preceding quote changes, $\Delta m_{t-1,*}$, are still the most influential variables of the models. The magnitudes of the coefficients of $\Delta m_{t-1,*}$ are not very different from the counterpart in Table 2.18. The inventory variables, however, show some difference. In the two-market-maker sample, the coefficients of inventory variables are significantly smaller than zero, which indicates the inventory play a role in determining the change of quotes. For the other two samples, most of the coefficients of I_t and ΔI_t are negative, but they are not significantly different from zero.[11]

A great difference of the three samples is that market makers weigh the trade information differently. The magnitudes of the coefficients of group one variables ($x_{i,*,*}$) of the two-market-maker sample are often bigger than

[11] No distinction has been made between customer trades and inter-dealer trades in this chapter. However, Section 3.6 will show that IDB trades and IMM trades affect the quote change differently from customer trades. In general, IDB trades do not move the quotes and IMM trades do considerably. The parameters estimated in the regression may depend on the proportion of IDB and IMM trades in the sample. Since the number of inter-dealer trades is small, the aggregation of customer trades and inter-dealer trades may not even affect the result of the nineteen-market-maker case, let alone the other stocks. See further discussions in Chapter 3.

those of the other two samples. For example, the coefficients of $x_{i,b,0}$ equals 0.8312 in the two-market-maker sample, which is greater than 0.4077 in the ten-market-maker sample and 0.6571 in the nineteen-market-maker sample. On the other hand, the coefficients of $x_{i,s,0}$ equals -0.9728 in the two-market-maker sample, which is smaller than -0.3618 in the ten-market-maker sample and -0.4136 in the nineteen-market-maker sample. In contrast, the magnitudes of the coefficients of Group 3 and Group 4 variables ($x^o_{i,*,*}$ and $v^o_{i,*,*}$) of the nineteen-market-maker sample are often bigger than those of the other two samples. One of the reasons is probably that the more market makers in the market, the less market shares a market maker has, and the more the market makers value the trade information of the others.

Table 2.20 presents the regression results on the subsamples of three market making firms. Again, all of the coefficients associated with buy trades and buy volumes are positive, and all of the coefficients associated with sell trades and sell volumes are negative with four exceptions. $\Delta m_{t-1,*}$ still explain most of the quote change and all of the coefficients are significantly different from zero. Inventory variables are significant in the Firm A sample, but not in Firm B sample. ΔI_t is also significant in the Firm C sample. The significance of the inventory variables corresponds to the significance of those variables in the two-market-maker sample, as both Firm A and Firm C quote a large number of small stocks. The most notable difference among the regression results of the three firms is that Firm B values the information of the trades not executed by itself more than the other two firms do. The coefficients of $x^o_{i,*,*}$ and $v^o_{i,*,*}$ of Firm B are often bigger than those of Firm A or Firm C, and the coefficients of $x_{i,*,*}$ of Firm B are often not significantly different from zero.

Table 2.21 reports the regressions of the quote change according to the quote status after the change. The four-way classification is used: quotes are either on both sides of the yellow strip (on-both), only on the best ask (on-ask), only on the best bid (on-bid) or neither on best ask nor best bid (straddle). The coefficients associated with buy trades and buy volumes are all positive with two exceptions, and the coefficients associated with sell trades and sell volumes are negative with three exceptions. $\Delta m_{t-1,*}$ are still positive and significantly greater than zero, but the coefficient values of on-ask and on-bid samples are very different from those of on-both and straddle samples. The coefficients of $\Delta m_{t-1,*}$ of on-both and straddle category are closer to those reported in the full sample, while those of $\Delta m_{t-1,0}$ and $\Delta m_{t-1,1}$ of the on-bid and on-ask regressions fall by near 0.2 to between 0.4 and 0.5, and those of $\Delta m_{t-1,2}$ fall to between 0.21 and 0.24. Therefore, when market makers change to quote only on one side of the yellow strip, the preceding change is not as important as when market makers change to quote both sides or to straddle the strip. Inventory variables are significant in the on-both sample, which may also reflect the results in Table 2.19 and Table 2.20 that it is more common for the market makers to quote both sides in the small stocks. As $\Delta m_{t-1,*}$ cannot explain the quote change in on-ask and on-bid sample very much, the R^2 has dropped to around 0.21. On the contrary, R^2 is between 0.55 and 0.59 in the on-both and straddle samples.

Table 2.18, 2.19, 2.20 and 2.21 further provide diagnostic statistics of colinearity and autocorrelations. The Condition Indices are tiny, which indicate there is no evidence of colinearity. The Durbin-Watson statistics reveal that the autocorrelations of the residuals of the regressions are small. Therefore, the regression models are well-specified.

Finally, a series of Wald tests are performed to examine whether the quote change is more likely to be affected by the trades by the market maker who changes the quote (group one variables $x_{i,*,*}$), or by the trades without the involvement of the market maker (Group 3 variables $x^o_{i,*,*}$). Because the variables are standardised by the total number of trades of the stock, their coefficients can be compared directly. Table 2.22 summarises the results of compared the absolute values of $x_{i,*,*}$ and $x^o_{i,*,*}$. An "S" indicates the absolute value of the coefficient of $x_{i,*,*}$ is bigger than that of $x^o_{i,*,*}$, and an "O" indicates the opposite is true. When all of the observations are used, the magnitude of the coefficient of $x_{i,*,*}$ tends to be smaller than its counterpart of $x^o_{i,*,*}$. When the observations are grouped according to the number of market makers of the stock, the $x_{i,*,*}$ moves the quote more for the two-market-maker group, and $x^o_{i,*,*}$ moves more for most of the ten- and nineteen-market-maker group. When the observations are grouped according to the market makers who change the quote, the results of Firm A are mixed, the scale of the coefficient of $x^o_{i,*,*}$ is bigger for Firm B with only one exception, and that of $x_{i,*,*}$ is bigger for Firm C. These results are consistent with the hypothesis that market makers pay attention to the trade information of the market as a whole. When there are only two market makers, the average market share is big, and they value the information of trades by themselves more than by the rivals. When the number of market makers increases, the market share of each one is getting smaller, and they value the trade information from the market more than the trade by themselves. The last panel shows the results of the regression models of which the observations are grouped by the quote status. The results are intriguing. $x^o_{i,*,*}$ are more important for "on-both" and "straddle" changes, but $x_{i,s,*}$ is more important for "on-ask" change and $x_{i,b,*}$ is more important "on-bid" changes.

2.6.3 Interpretations

It appears there are mainly two different types of quote change. The first type is to straddle the yellow strip or to quote both sides of the strip. The second type is to quote the best bid or the best ask only. Each consists of about half of the quote changes. The decision is essentially whether or not to quote only one side of the yellow strip. Section 2.3 shows that market makers are more likely to receive unbalanced order flows while on one side of the yellow strip. Although quoting one side attracts more orders (Table 2.3), market makers may prefer to straddle the yellow strip to receive relatively balanced orders. To quote both sides is another way to obtain more balanced order flows. Recall that a market maker quotes both sides often means that everybody else quotes the same bid and ask prices. It follows that before a market maker quotes both sides of the yellow strip, the rest of the market makers have already quoted the same bid and ask prices. The last market maker who joins the rest essentially changes the quote from being alone on best bid or best ask. Table 2.5 and Table 2.6 show that being on one side of the yellow strip alone attracts extremely unbalanced order flows, and that being on the yellow strip with everybody else receives balanced orders. Consequently, the decision of a market maker to quote both sides implies that he prefers to receive more balanced orders.

The regression results reveal the preceding quote change plays an important role in determining the current quote change, and it is much more important when market makers decide to straddle or to quote both sides of the yellow strip. When market makers observe they have been or they are about to be on one side of the yellow strip, they choose to straddle the strip or to quote the same prices as everybody else. The decision to avoid being on one side of the yellow strip may partly explain why inventory variables are often not significant in most of the regression models. A lot of the market makers do not choose to expose themselves to unbalanced order flows in the first place, so they do not put themselves into an awkward position with huge unbalanced inventory. Consequently, market makers do not have to change the quote to balance the inventory. Inventory management is difficult when only two firms make the market. The stocks are less liquid, which implies the inventory imbalance may last for a long time. Trade publication ensures both market makers know exactly each other's inventory, so the market maker with unbalanced inventory has no intention to conduct inter-dealer trading. Consequently, changing the quotes in response to the change in the inventory level is a reasonable option in the two-market-maker case. The duopoly literature may shed some lights on this result (Sutton 1991; Hay and Morris 1991; Darrough 1993; Kyle and Wang 1997).

The insignificance of the inventory variables seems to be contradicted to the result from Hansch et al. (1998). However, they do not consider the information effect, and it is not clear whether the inventory effect in their study would exist had other effects been controlled. Moreover, they show that the inventory effect is strongest when the inventory levels of market makers are extremely unbalanced. The extreme inventory imbalance is by definition very uncommon, which indicates the inventory effect does not matter very much most of the time. It is therefore not surprising the regression analysis in the previous section does not find strong inventory effects. Furthermore, although the almost-all positive coefficients of buy trades $x_{i,b,*}$ and the almost-all negative coefficients of sell trades $x_{i,s,*}$ are consistent with both inventory and information hypothesis, it can never be over-emphasised that the inventory effect and information effect of trades move the quote to the same direction. When an upward revision of quote coincides with a number of buy trades and the accompanied inventory reduction, it is always sensible to ask whether the change is motivated by inventory or information consideration. The regression results are in favour of information effects. Repeatedly more trade variables $x_{i,*.*}$ are significantly different from zero than volume variables $v_{i,*.*}$ are. If inventory were the determining factor to move the quotes, then each share traded would be of the same importance and volumes would be a better explanatory variable of quote changes. More importantly, market makers do not only care about the trades by themselves. The quote changes are related to the trades not executed by the market makers who are changing the quotes, and this information is valued more when there are a lot of firms making the market or when the market share of the firm is small. It means that market makers consider the trade information of the market as a whole and change the quotes accordingly. Therefore, it is possible that the significance of inventory variables in the regressions of some sub samples is partly attributed to the information effect of the trades. It is also possible that some of the inventory effects detected by Hansch, Naik, and Viswanathan (1998) are really information effects.

The most influential group of variables of the quote change is the preceding change. Quite a large number of quote changes take place soon after another market maker revises the quote. There are two ways to explain the impacts of preceding change along the line of information theory. First, the quote changes close to one another reflect the reactions of market makers to some public news in the market. Because the news is not transmitted to all of the market makers simultaneously, some of them are bound to react to the news quicker than the others. As there is consensus of their reactions, almost all of the market makers revise the quote to the same direction with similar scale. The second explanation is that some of the market makers do not actively collect information. They learn from the quotes of the price leaders in the market (Black 1996; Board et al. 1997). When new information arrives at the market, the price leaders revise the quotes, and the others follow suit. Without spending too many resources on collecting information, the price followers tend to avoid exposing themselves to the yellow strip, as they are very vulnerable to unbalanced order flows brought by informed investors. Therefore, the price followers are likely to straddle the strip, which explains the big R^2 of the regression of straddling the yellow strip in Table 2.21. On the other hand, the price leaders react to the news and lead the quote revision of the market. They are willing to quote the best bid or ask to avoid order flows of informed investors on one side, while attracting uninformed orders on the other side. Therefore, the preceding quote change has less influence on the current change to one side of the yellow strip, and the R-squares of the respective regressions are much smaller.

Do the quotes fully reflect market information? Although the answer to the question may be beyond the scope of this chapter, a few points are worth mentioning. The quick quote changes by the price followers indeed "fully reflect" the changes by price leaders, so the answer rests on whether the quotes by price leaders fully reflect market information. The explanatory power of the preceding quote change is greatly reduced when market makers decide to quote only one side of the yellow strip. The information concern seems to be a good candidate to explain the change to be on only one side. No matter the quotes fully reflect the information or not, the bottom line is, when market makers change the quotes, they take trade information into account.

2.7 Conclusions

The quote revisions made by market makers on the London Stock Exchange may be classified into two categories. The first one is to quote one side of the yellow strip in order to attract unbalanced order flows. The quotes are often made by price leaders in response to public information. The upward quote revisions are often associated with the arrival of buy orders, and the downward revisions are often with sell orders. Because volume variables do not perform better than trade-number variables in the regression models, and because trades made by another market makers affect the quote revisions as well, it seems market makers change the quotes according to information rather than inventory. Inventory affects the decision only when there are few firms making the market. The second category of quote change consists of the quote on both sides of the yellow strip or straddling the strip. The quotes are often made by price followers in response to the quote change of price leaders. Since straddling the strip attracts relatively balanced order flows, the price followers do not face inventory unbalance very often.

The benefit of straddling the yellow strip is to receive balanced order flows; the drawback is the order flows tend to be less than standing on one side of the strip. The benefit and drawback of standing on one side of the strip is the other way round: receiving more order flows, and the orders are unbalanced. Evidence from Section 2.5 and Board et al. (1997) both suggest the majority of the market makers are price followers while a small number of market makers post the quote rather aggressively. One of the directions of future research is to examine under what circumstances will a market maker becomes price leader or price follower. Moreover, the quoting behaviour of market makers is not fully studied in this chapter. For example, almost all of the market makers maintain the same quote spread for a stock, a property which is repeatedly used in this chapter without questioning the reason. It would be interesting to know why market makers do not narrow the quote spread to stand on both sides of the yellow strip. Theoretical work may contribute insights to the questions. Information consideration apparently directs the decision of market makers to change the quotes, while the existing models of heterogeneous market makers (Ho and Stoll 1983; Kandel and Marx 1997) address the issues of inventory or liquidity. A theoretical model with emphasis on information may help explaining the behaviour of heterogeneous market makers in multiple dealership markets.

Table 2.1: Summary Statistics of Trade data

		number of stocks	number of trades	sum of volumes in £millions	average size per trade in pounds
total		1,832	1,219,767	80,952	66,367
buy		1,822	505,905	40,259	79,578
sell		1,829	713,862	40,693	57,004
median stock			226	2	9,948
market	A	1,132	336,000	10,749	31,990
maker	B	21	271	53	193,899
	C	933	81,931	708	8645
trades excluded:					
basket trades			13,565	2,644	194,919
crosses			1,670	1,019	610,083
put-throughs			6,081	1,398	229,856

Table 2.2: Exploring the relationship between order flows and quote status by two-way classification

Δ^N of a particular trade category is defined as the percentage of the number of buys in the category minus the percentage of the number of sells. Similarly, Δ^V of a particular trade category is defined as the percentage of buy volumes in the category minus the percentage of sell volumes.

	Number of trades					Volumes in £ millions				
	buy	%	sell	%	Δ^N	buy	%	sell	%	Δ^V
on best bid	193,222	39.4	297,229	60.6	-21.2	13,610	44.1	17,232	55.9	-11.7
off best bid	312,683	42.9	416,633	57.1	-14.3	26,650	53.2	23,461	46.8	6.4
on best ask	226,536	47.2	253,312	52.8	-5.6	15,550	55.9	12,246	44.1	11.9
off best ask	279,369	37.8	460,550	62.2	-24.5	24,709	46.5	28,446	53.5	-7.0

Table 2.3: Orders and quote duration in three-way classification
Daily average duration is express in (hour):(minute). Zeros (0) in the percentage columns indicate the numbers in their corresponding columns, i.e., number of trade, volumes or duration, are smaller than 0.5 % of the category, so the percentages are rounded to zero.

	quote status	number of trades	%	volumes (£1000)	%	average duration	%
Panel 1. The full sample							
	both sides	111,460	9	3,343,094	4	1:21	17
	one side	747,379	61	51,952,371	64	4:41	60
	straddling	360,928	30	25,656,229	32	1:49	23
Panel 2. By different number of market makers							
2	both sides	37,955	54	488,613	54	4:35	56
	one side	32,397	46	414,315	46	3:38	44
	straddling	90	0	585	0	0:01	0
10	both sides	1,708	4	128,282	6	0:24	5
	one side	25,346	67	1,531,194	69	4:50	62
	straddling	10,994	29	546,702	25	2:31	33
19	both sides	659	0	39,536	0	0:02	1
	one side	83,482	55	8,006,531	58	3:18	45
	straddling	68,968	45	5,741,592	42	3:58	54
Panel 3. Selected market makers							
A	both sides	20,738	6	433,978	4	1:27	18
	one side	208,262	62	6,970,784	65	4:59	63
	straddling	107,000	32	3,343,834	31	1:26	18
B	both sides	2	1	581	1	0:06	1
	one side	178	66	33,674	64	2:56	39
	straddling	91	34	18,292	35	4:30	60
C	both sides	33,588	41	306,084	43	3:29	43
	one side	46,174	56	385,698	54	4:26	54
	straddling	2,169	3	16,506	2	0:14	3

Table 2.4: Trades in the four-way classification

		number of trades					volumes in £1,000				
		buy	%	sell	%	Δ^N	buy	%	sell	%	Δ^V
Panel 1. The full sample											
	on both	54,589	49	56,871	51	-2	1,645,789	49	1,697,305	51	-2
	on ask	171,947	47	196,441	53	-7	13,904,703	57	10,548,962	43	14
	on bid	138,633	37	240,358	63	-27	11,963,862	44	15,534,844	57	-13
	straddle	140,736	39	220,192	61	-22	12,744,679	50	12,911,550	50	-1
Panel 2. By different number of market makers											
2	on both	21,177	56	16,778	44	12	240,455	49	248,158	51	-2
	on ask	10,595	66	5,510	34	32	102,408	52	94,012	48	4
	on bid	7,536	46	8,756	54	-7	106,130	49	111,764	51	-3
	straddle	48	53	42	47	7	179	31	406	69	-38
10	on both	734	43	974	57	-14	61,330	48	66,952	52	-4
	on ask	6,024	48	6,610	52	-5	403,192	56	322,029	44	11
	on bid	4,689	37	8,023	63	-26	377,192	47	428,781	53	-6
	straddle	4,528	41	6,466	59	-18	251,851	46	294,851	54	-8
19	on both	225	34	434	66	-32	19,470	49	20,066	51	-2
	on ask	16,177	38	26,502	62	-24	2,395,437	62	1,491,880	38	23
	on bid	12,817	31	27,986	69	-37	1,570,713	38	2,548,501	62	-24
	straddle	25,533	37	43,435	63	-26	2,817,411	49	2,924,182	51	-2
Panel 3. Selected market makers											
A	on both	9,126	44	11,612	56	-12	208,105	48	225,873	52	-4
	on ask	42,414	40	64,845	61	-21	1,846,666	55	1,512,055	45	10
	on bid	36,845	37	64,158	64	-27	1,574,049	44	2,038,014	56	-13
	straddle	39,385	37	67,615	63	-26	1,778,964	53	1,564,870	47	6
B	on both	2	100	0		100	581	100	0		100
	on ask	38	75	13	26	49	5,859	69	2,642	31	38
	on bid	36	28	91	72	-43	6,897	27	18,276	73	-45
	straddle	38	42	53	58	-17	6,436	35	11,856	65	-30
C	on both	18,599	55	14,989	45	11	149,995	49	156,089	51	-2
	on ask	12,888	59	9,051	41	18	86,472	49	88,434	51	-1
	on bid	12,479	52	11,756	49	3	102,956	49	107,832	51	-2
	straddle	1,130	52	1,039	48	4	7,724	47	8,782	53	-6

Table 2.5: Trade imbalance (Δ^N)after adjustment

The definition of adjusted Δ^N is in Section 2.3.5. T-statistics are computed under the null hypothesis that Δ^N is zero of the full sample, of the number-of-market-maker groups, and of the market-maker groups. An asterisk (*) after the t-statistics indicates the value is at 0.05 significant level, and two asterisks (**) indicate it is at 0.01 level.

	on-both		on-ask		on-bid		straddle	
	mean	t-value	mean	t-value	mean	t-value	mean	t-value
Panel 1. The full sample								
	-0.08	-0.11	22.04	45.14**	-15.21	-34.13**	2.68	4.64**
Panel 2. By number of market makers								
2	-1.31	-1.36	26.39	15.77**	-21.14	-12.93**	-24.26	-0.89
3	-2.81	-2.29*	18.76	12.28**	-17.48	-12.56**	-2.43	-0.60
4	0.00	0.00	19.22	12.40**	-18.33	-12.52**	-1.70	-0.59
5	-1.26	-0.58	21.22	12.92**	-18.24	-11.41**	-3.42	-1.27
6	-0.34	-0.11	17.55	9.40**	-19.64	-11.64**	-3.41	-1.39
7	2.22	0.64	22.43	10.41**	-15.33	-7.73**	-0.93	-0.39
8	-3.74	-1.12	19.91	10.61**	-16.50	-9.77**	-1.69	-0.88
9	0.33	0.10	19.83	12.62**	-17.12	-12.80**	-2.74	-1.73
10	-4.35	-1.01	18.03	8.11**	-16.07	-8.49**	-2.32	-1.17
11	-2.74	-0.59	23.10	6.98**	-11.30	-3.57**	2.56	0.76
12	-4.81	-0.97	17.50	5.91**	-10.66	-4.07**	2.42	0.90
13	10.27	2.19*	30.14	10.93**	-0.92	-0.32	12.85	5.06**
14	-6.05	-1.06	24.47	7.13**	-19.27	-6.60**	5.24	1.79
15	1.42	0.29	15.67	5.87**	-13.06	-5.52**	4.96	2.08*
16	1.88	0.45	26.66	14.19**	-11.88	-7.19**	8.62	5.65**
17	9.79	2.20*	19.02	9.30**	-10.37	-5.63**	6.45	3.61**
18	8.58	1.91	31.58	13.53**	-6.05	-3.09**	11.54	6.07**
19	14.70	1.94	40.15	13.07**	-3.24	-1.23	17.68	7.76**
Panel 3. Selected market makers								
A	-1.83	-1.28	10.27	9.16**	-16.17	-15.10**	-5.32	-3.74**
B	159.90	10.29	83.47	4.58**	-20.29	-1.80*	18.51	1.13
C	1.09	1.00	10.60	6.83**	-5.91	-4.16**	1.12	0.28

Table 2.6: Volume imbalance (Δ^V) after adjustment

The adjusted Δ^V is defined in a similar way to Δ^N in Section 2.3.5. T-statistics are computed under the null hypothesis that Δ^V is zero of the full sample, of the number-of-market-maker groups, and of the market-maker groups. An asterisk (*) after the t-statistics indicates the value is at 0.05 significant level, and two asterisks (**) indicate it is at 0.01 level.

	on-both		on-ask		on-bid		straddle	
	mean	t-value	mean	t-value	mean	t-value	mean	t-value
Panel 1. The full sample								
	-6.84	-8.66 **	15.13	26.13 **	-18.95	-34.42 **	-4.76	-6.88 **
Panel 2. By number of market makers								
2	-3.21	-2.47 *	23.43	11.56 **	-16.46	-8.61 **	-24.11	-0.66
3	-6.58	-4.02 **	15.56	8.51 **	-15.20	-8.92 **	-0.12	-0.02
4	-3.57	-1.62	14.32	7.51 **	-19.63	-10.57 **	-9.46	-2.77 **
5	-9.05	-3.33 **	14.24	6.61 **	-19.63	-9.37 **	-13.21	-4.19 **
6	-6.38	-1.83	10.56	4.36 **	-19.92	-8.88 **	-12.21	-3.98 **
7	-1.99	-0.46	14.13	5.47 **	-13.95	-5.53 **	-6.24	-2.02 *
8	-10.97	-2.77 **	10.07	4.39 **	-16.80	-7.52 **	-4.27	-1.68
9	-2.82	-0.73	16.07	8.74 **	-16.67	-9.10 **	-4.84	-2.30 *
10	-20.51	-3.88 **	14.11	5.45 **	-19.80	-7.28 **	-10.00	-3.60 **
11	-12.43	-2.01 *	11.44	3.05 **	-19.52	-5.46 **	-2.51	-0.60
12	-18.53	-3.13 **	4.94	1.34	-22.04	-6.55 **	-6.99	-2.12 *
13	-5.29	-0.88	19.45	5.77 **	-14.75	-4.60 **	-0.35	-0.11
14	-16.00	-2.48 *	16.80	4.54 **	-21.99	-6.56 **	-1.89	-0.58
15	-13.62	-2.30 *	8.98	3.02 **	-22.85	-8.25 **	0.55	0.20
16	-5.67	-1.18	22.11	10.61 **	-19.20	-9.81 **	2.73	1.64
17	-2.61	-0.51	9.30	3.88 **	-23.29	-10.40 **	-3.51	-1.67
18	-16.83	-3.22 **	15.35	6.13 **	-23.32	-9.86 **	-3.11	-1.41
19	-6.68	-0.73	24.09	7.94 **	-28.78	-9.41 **	-4.49	-1.56
Panel 3. Selected market makers								
A	-12.19	-7.11 **	14.54	10.32 **	-20.79	-14.36 **	-11.93	-5.74 **
B	101.02	267.77 **	60.87	3.24 **	-42.22	-3.20 **	-13.83	-0.78
C	-2.72	-1.88	5.55	2.98 **	-5.22	-3.02 **	-0.03	-0.00

Table 2.7: Volume imbalance with different number of market makers on the best ask

When there are less than five market makers, either one market maker is alone on the best ask (category (a)), or no less than half of the market maker are on the best ask, so no trades fall into category (b), and the t-statistics in category (c) are the difference of Δ^V between category (a) and (c). A asterisk (*) after the t-statistics indicates the value is at 0.05 significant level, two asterisks (**) indicate it is at 0.01 level, and a negative sign (-) indicates the sign of the mean difference is different from what is expected.

	(a) alone on ask	(b) with less than half of the market makers on ask		(c) with more than half of the market makers on ask		(d) not on ask; only on bid or straddle	
	mean	mean	t-value (a)-(b)	mean	t-value (b)-(c)	mean	t-value (c)-(d)
Panel 1. The full sample							
	29.68	15.72	11.55 **	0.39	14.91 **	-12.69	18.52 **
Panel 2. By number of market makers							
2	23.56			-3.84	11.48 **	-16.51	5.47 **
3	18.96			3.36	5.75 **	-11.62	6.91 **
4	30.36			1.63	9.15 **	-15.90	7.57 **
5	27.90	12.73	3.28 **	0.97	3.17 **	-16.88	6.75 **
6	27.83	15.29	2.41 *	-1.80	4.05 **	-16.35	5.01 **
7	26.31	14.45	2.13 *	3.76	2.36 *	-10.15	4.06 **
8	28.02	10.65	3.54 **	0.01	2.69 **	-10.70	3.64 **
9	39.18	21.31	4.42 **	2.23	5.86 **	-10.71	4.99 **
10	29.08	15.84	2.16 *	-6.29	4.59 **	-14.90	2.40 *
11	12.40	18.31	-0.71 -	-5.58	3.88 **	-10.79	1.09
12	22.47	6.78	2.06 *	-10.05	2.83 **	-14.31	0.95
13	39.43	20.69	2.71 **	2.57	3.26 **	-7.19	2.25 *
14	27.25	25.27	0.26	-0.89	4.50 **	-11.64	2.37 *
15	35.04	10.85	4.29 **	-4.07	3.14 **	-10.39	1.67
16	45.12	22.26	5.93 **	6.95	4.31 **	-7.77	5.21 **
17	34.08	9.40	5.20 **	-0.53	2.58 **	-12.89	4.00 **
18	41.12	14.59	5.78 **	-2.22	4.18 **	-12.80	3.21 **
19	47.37	21.36	4.41 **	4.60	3.01 **	-16.31	4.35 **
Panel 3. Selected market makers							
A	25.11	7.42	5.72 **	0.29	2.64 **	-15.94	8.73 **
B	62.92	33.38	0.59	44.99	-0.28 -	-28.03	2.38 *
C	5.89	11.48	-0.78 -	0.06	1.66	-4.06	1.92

Table 2.8: Volume imbalance with different number of market makers on the best bid

When there are less than five market makers, either one market maker is alone on the best ask (category (a)), or no less than half of the market maker are on the best ask, so no trades fall into category (b), and the t-statistics in category (c) are the difference of Δ^V between category (a) and (c). A asterisk (*) after the t-statistics indicates the value is at 0.05 significant level, two asterisk (**) indicates it is at 0.01 level, and a positive sign (+) indicates the sign of the mean difference is different from what is expected.

	(a) alone on bid	(b) with less than half of the market makers on bid		(c) with more than half of the market makers on bid		(d) not on bid; only on ask or straddle	
	mean	mean	t-value (a)-(b)	mean	t-value (b)-(c)	mean	t-value (c)-(d)
Panel 1. The full sample							
	-28.44	-22.75	-4.80 **	-11.29	-11.61 **	6.31	-25.09 **
Panel 2. By number of market makers							
2	-16.26			-3.49	-5.54 **	23.12	-11.02 **
3	-19.85			-8.78	-4.21 **	11.81	-9.37 **
4	-23.94			-12.31	-3.45 **	5.64	-7.72 **
5	-28.59	-23.69	-1.06	-13.52	-2.79 **	2.43	-6.03 **
6	-24.18	-18.12	-1.17	-13.82	-1.06	-0.02	-4.77 **
7	-23.89	-25.94	0.37 +	-4.66	-5.09 **	4.02	-2.58 **
8	-28.03	-22.96	-1.03	-11.37	-3.07 **	2.98	-5.03 **
9	-35.84	-18.42	-4.23 **	-9.19	-2.82 **	5.61	-5.76 **
10	-26.80	-21.49	-0.89	-19.75	-0.38	2.10	-6.25 **
11	-26.68	-26.36	-0.04	-15.68	-1.81	4.23	-4.25 **
12	-33.06	-23.55	-1.34	-19.72	-0.68	-1.20	-4.30 **
13	-27.85	-19.38	-1.26	-9.98	-1.80	8.95	-4.44 **
14	-35.08	-27.40	-1.04	-15.63	-2.14 *	6.98	-5.12 **
15	-45.67	-23.54	-4.11 **	-13.66	-2.15 *	4.47	-4.88 **
16	-42.60	-19.88	-5.69 **	-11.20	-2.58 **	11.90	-8.44 **
17	-41.01	-23.50	-3.79 **	-10.82	-3.38 **	2.55	-4.32 **
18	-41.84	-28.99	-2.80 **	-16.99	-3.11 **	5.67	-7.05 **
19	-46.03	-24.69	-3.70 **	-13.39	-2.05 *	9.07	-4.87 **
3. Selected market makers							
A	-29.81	-23.85	-1.89	-11.41	-4.71 **	4.93	-8.79 **
B	-74.90	-41.44	-1.16	-3.53	-1.33	17.29	-0.77
C	-2.24	-11.31	1.01 +	-4.97	-0.72	4.29	-4.23 **

Table 2.9: Trades, volumes, and sizes

$[a, b)$ means the size of the trade is no smaller than a times MQS and smaller than b time MQS. (a, b) means the size is bigger than a and smaller than b.

	number of trades				volumes in £millions			
	buy	%	sell	%	buy	%	sell	%
MQS in								
$(0, 0.05)$	211,527	42	410,418	57	784	2	1,399	3
$[0.05, 0.1)$	42,724	8	52,677	7	392	1	545	1
$[0.1, 0.2)$	44,993	9	47,685	7	694	2	779	2
$[0.2, 0.5)$	64,166	13	62,563	9	2,069	5	2,078	5
$[0.5, 1)$	39,370	8	38,196	5	3,558	9	3,456	8
$[1, 2)$	42,053	8	41,089	6	6,834	17	6,641	16
$[2, 3)$	22,728	4	22,408	3	6,013	15	5,690	14
$[3, 6)$	20,129	4	20,487	3	7,764	19	7,609	19
$[6, 75)$	17,503	3	17,622	2	10,706	27	11,097	27
$[75, \infty)$	712	0	717	0	1,444	4	1,401	3

Table 2.10: Adjusted Δ^N and Δ^V by sizes of trades

An asterisk (*) after the t-statistics indicates the value is at 0.05 significant level, and two asterisks (**) indicate it is at 0.01 level. The definitions of $[a, b)$ and (a, b) are in Table 2.9.

	on-both		on-ask		on-bid		straddle	
	mean	t-value	mean	t-value	mean	t-value	mean	t-value
Panel 1. Δ^N								
$(0, 0.05)$	-1.98	-1.57	6.45	7.45 **	-13.62	-17.13 **	-3.45	-4.04 **
$[0.05, 0.1)$	0.43	0.28	9.28	9.22 **	-8.92	-9.19 **	1.41	1.32
$[0.1, 0.2)$	0.49	0.35	14.64	15.32 **	-9.47	-10.41 **	5.58	5.17 **
$[0.2, 0.5)$	3.22	2.75 **	23.74	28.29 **	-11.67	-14.60 **	8.51	8.61 **
$[0.5, 1)$	1.05	0.77	36.88	38.46 **	-21.28	-23.35 **	12.80	10.39 **
$[1, 2)$	-0.39	-0.29	45.72	49.47 **	-28.06	-31.95 **	13.99	11.21 **
$[2, 3)$	-0.37	-0.22	38.20	33.02 **	-19.37	-17.36 **	14.19	9.16 **
$[3, 6)$	-2.46	-1.51	24.06	19.45 **	-5.26	-4.40 **	14.73	9.00 **
$[6, 75)$	-0.13	-0.07	3.49	2.54 *	11.93	9.26 **	6.61	3.42 **
$[75, \infty)$	-0.50	-0.10	-11.48	-2.19 *	1.27	0.27	31.10	2.56 *
Panel 2. Δ^V								
$(0, 0.05)$	-14.09	-9.77 **	-6.83	-6.98 **	-26.95	-29.09 **	-19.26	-20.02 **
$[0.05, 0.1)$	-7.59	-4.46 **	-2.63	-2.40 *	-21.67	-20.78 **	-13.85	-12.46 **
$[0.1, 0.2)$	-5.75	-3.81 **	4.30	4.17 **	-20.94	-21.43 **	-8.92	-8.02 **
$[0.2, 0.5)$	-2.50	-1.98 *	14.01	15.86 **	-22.22	-26.43 **	-5.05	-5.06 **
$[0.5, 1)$	-2.78	-1.94	27.74	28.59 **	-30.71	-32.84 **	-0.74	-0.61
$[1, 2)$	-4.42	-3.19 **	36.12	39.86 **	-37.33	-42.50 **	0.37	0.30
$[2, 3)$	-2.44	-1.49	29.71	26.47 **	-28.05	-26.02 **	0.73	0.49
$[3, 6)$	-5.64	-3.53 **	15.23	12.76 **	-13.81	-12.22 **	0.90	0.58
$[6, 75)$	-4.04	-2.74 **	-5.67	-4.52 **	4.65	4.00 **	-4.69	-2.59 **
$[75, \infty)$	-1.52	-0.37	-14.37	-3.38 **	-2.59	-0.65	24.13	2.26 *

Table 2.11: Effective spreads and order flows

An asterisk (*) after the t-statistics indicates the value is at 0.05 significant level, and two asterisks (**) indicate it is at 0.01 level. The definitions of $[a,b)$ and (a,b) are in Table 2.9. The blanks of the t-values in the bottom row of Panel 2 indicates that there is only one group in the $[75,\infty)$ category, and the t-statistics cannot be calculated.

	correlation coefficients			number	volumes
	mean	t-value		of trades	£millions
Panel 1. Grouped by day					
total	0.04	10.25	**	722,549	12,029
MQS in					
$(0,0.05)$	0.16	35.24	**	575,871	2,130
$[0.05,0.1)$	0.02	2.69	**	47,688	596
$[0.1,0.2)$	-0.02	-1.87	*	31,520	582
$[0.2,0.5)$	-0.05	-4.38	**	34,861	1,107
$[0.5,1)$	-0.06	-3.38	**	10,524	858
$[1,2)$	-0.10	-4.74	**	9,896	1,366
$[2,3)$	-0.12	-4.19	**	4,860	1,241
$[3,6)$	-0.05	-1.68	*	4,133	1,950
$[6,75)$	-0.07	-1.97	*	3,186	2,149
$[75,\infty)$	0.33	0.50		10	51
Panel 2. Grouped by hour					
total	0.14	41.74	**	427,373	2,991
MQS in					
$(0,0.05)$	0.16	48.86	**	405,850	1,610
$[0.05,0.1)$	0.02	0.90		6,592	89
$[0.1,0.2)$	0.01	0.28		4,441	81
$[0.2,0.5)$	-0.05	-1.88	*	5,249	170
$[0.5,1)$	-0.05	-1.09		1,393	120
$[1,2)$	-0.04	-0.87		1,922	213
$[2,3)$	-0.11	-1.83	*	911	204
$[3,6)$	-0.20	-2.81	**	629	247
$[6,75)$	-0.04	-0.42		383	253
$[75,\infty)$	-1.00			3	2

Table 2.12: Summary statistics of quotes sample

The medians quote and yellow strip are the medians of all the spreads of all of the stocks in the market-maker category. The quote spread is defined as the proportion of the mid-quote in percentage points. The touch spread is the proportion of the mid-touch in percentage points. The quote size median is the median of the average of the ask and bid quote size in pound sterlings. The median of the number of quotes is the median of the number of quotes by a market maker per stock per day. The daily number of market makers are time-weighted average of those who are respectively on best ask only, on best bid only, on both best bid and ask, and straddling the yellow strip.

number of market makers	number of stocks	Quote spread median	Inside spread median	Quote size median	number of quotes median	mean daily number of market makers on ask	on bid	on both	straddle
2	611	5.41	4.08	2,400	1	0.43	0.43	1.12	0.01
3	432	4.08	2.94	4,045	1	0.88	0.90	1.11	0.10
4	254	3.35	2.11	8,450	1	1.34	1.36	0.96	0.34
5	161	2.55	1.53	13,150	1	1.65	1.70	0.92	0.72
6	114	2.47	1.32	15,125	1	2.03	2.10	0.62	1.25
7	73	2.43	1.31	21,000	1	2.16	2.30	0.59	1.95
8	74	2.21	1.16	25,125	1	2.40	2.70	0.54	2.35
9	77	1.44	0.79	56,125	2	2.79	2.89	0.39	2.91
10	32	1.83	1.10	67,625	2	2.97	2.93	0.53	3.55
11	20	1.90	0.97	91,750	3	3.02	2.99	1.04	3.94
12	23	1.85	0.82	100,875	3	3.05	2.96	0.60	5.36
13	23	1.57	0.64	132,750	3	3.07	3.33	0.78	5.80
14	21	1.22	0.69	157,500	4	3.55	3.72	0.58	6.11
15	36	1.43	0.59	186,250	5	3.51	3.89	0.41	7.17
16	46	1.29	0.53	204,250	6	3.43	3.78	0.34	8.42
17	47	0.98	0.54	206,250	6	3.66	3.84	0.37	9.10
18	32	0.79	0.41	382,750	8	3.95	4.17	0.36	9.49
19	14	0.93	0.40	378,000	8	3.76	4.05	0.14	11.01

Table 2.13: Top ten popular quote and trade practices

Ticks are the minimum differences of the prices among the quotes or trades records. The type of quote change is examined in three dimensions. First, compared with the previous quote, the spread may be the same, wider or narrower. Second, compared with the previous quote, the quote change may move the mid-quote upwards, downwards, or the mid-quote may be unchanged. Finally, compared with the preceding quote change, "same change" means the change in both bid and ask prices are both identical to the preceding change. "Same direction" means the direction of the change in mid-quote is the same as the preceding change, but the magnitude of the change of either bid or ask is not the same. "Opposite" means the directions of the two changes are not the same. "Opening" means the previous quote is an opening quote and hence no preceding change to compare with. "Size change" means the current change is a change in quoted shares. The zero (0) in the percentage columns implies the proportion is less than 0.5% and hence the percentage is rounded to zero.

quote spread pence	%	quote tick pence	%	trade tick pence	%	absolute quote change pence	%	quote change type: spread	mid-quote change	compared with preceding change	%
5.0	33	1.00	67	1.00	61	1.00	56	same	up	same change	23
7.0	20	10.00	12	0.50	10	2.00	28	same	down	same change	22
3.0	15	5.00	11	10.00	9	3.00	7	same	up	same direction	17
10.0	12	25.00	3	5.00	9	0.50	3	same	down	same direction	16
4.0	11	0.50	2	2.50	3	5.00	3	same	down	opposite	9
2.0	3	50.00	2	0.25	2	4.00	1	same	up	opposite	8
1.5	1	100.00	1	25.00	2	0.25	0	same	up	opening	3
20.0	1	2.50	1	0.01	2	10.00	0	same	down	opening	2
15.0	1	0.25	1	100.00	1	7.00	0	same	same	size change	0
1.0	1	0.01	0	50.00	1	6.00	0	wider	up	opposite	0
	98		98		100		99				100

Table 2.14: Daily quote status statistics

The daily average time of a market maker in different quote status is expressed in (hour):(minute). Zero minute (0:00) means the average time is less than thirty seconds and rounded to zero. Similarly, zero quote frequency (0.0) means the average frequency is less than 0.05 and rounded to zero. Blank means no observation has been found in the quote status.

number of market makers	daily average time					daily average frequency					
	on both sides		on one side		straddle	on both sides		on one side		straddle	sum
	alone	not alone	alone	not alone		alone	not alone	alone	not alone		
2	0:02	4:35	3:32	0:12	0:02	0.0	0.9	0.8	0.1	0.0	1.7
3	0:00	3:03	1:57	2:56	0:16	0.0	0.7	0.5	0.9	0.1	2.1
4	0:00	1:56	1:19	4:05	0:40	0.0	0.5	0.4	1.1	0.2	2.3
5	0:00	1:28	0:59	4:22	1:09	0.0	0.4	0.4	1.4	0.4	2.6
6	0:00	0:49	0:50	4:38	1:39	0.0	0.3	0.4	1.5	0.6	2.8
7	0:00	0:39	0:44	4:17	2:12	0.0	0.3	0.4	1.5	0.7	2.9
8	0:00	0:31	0:35	4:26	2:19	0.0	0.3	0.4	1.8	0.9	3.3
9		0:20	0:30	4:29	2:33		0.3	0.4	2.3	1.2	4.3
10		0:24	0:29	4:05	2:45		0.3	0.5	2.9	1.7	5.4
11		0:43	0:30	3:41	2:45		0.6	0.6	3.8	2.1	7.1
12		0:23	0:27	3:21	3:25		0.4	0.6	3.4	2.3	6.6
13		0:27	0:25	3:19	3:25		0.6	0.8	4.8	3.1	9.3
14		0:18	0:21	3:35	3:19		0.6	0.8	5.5	3.5	10.3
15	0:00	0:12	0:21	3:21	3:37	0.0	0.5	0.9	6.4	4.4	12.2
16	0:00	0:09	0:20	3:02	3:57	0.0	0.5	1.2	8.0	5.9	15.6
17		0:09	0:19	2:57	4:01		0.4	1.1	7.6	5.7	14.9
18		0:08	0:17	3:03	3:55		0.5	1.3	10.4	7.8	20.3
19		0:03	0:16	2:44	4:14		0.3	1.5	9.6	7.8	19.1

Table 2.15: Source of change in quote status

Both the frequency of S and O changes are of daily average per firm. O/S ratio is calculated by dividing O frequency by S frequency. Zeros (0) mean the frequency is less than 0.05. Blanks mean that no such changes exist.

number of	on both side						on one side						straddle			Sum	
makers	alone			not alone			alone			not alone							
makers	S	O	O/S	S	O	O/S	S	O	O/S	S	O	O/S	S	O	O/S	S	O
2	0	0	1.61	0.4	0.4	1.00	0	0	1.00	0.4	0.4	0.61	0	0	0.61	0.8	0.8
3	0	0	3.70	0.2	0.4	2.01	0.2	0.3	1.42	0.3	0.6	2.20	0	0.1	1.76	0.7	1.4
4	0	0	1.50	0.1	0.4	2.97	0.2	0.3	1.62	0.4	0.8	2.21	0.1	0.2	1.94	0.7	1.6
5	0	0	3.50	0.1	0.3	3.93	0.2	0.3	1.65	0.4	0.6	2.33	0.1	0.3	2.07	0.8	1.8
6	0			0.1	0.3	5.15	0.1	0.2	1.69	0.4	1.1	2.39	0.2	0.4	2.13	0.8	2.0
7	0	0	4.17	0	0.2	5.76	0.1	0.2	1.68	0.5	1.1	2.36	0.2	0.5	2.30	0.8	2.0
8	0			0	0.2	6.89	0.1	0.2	1.53	0.5	1.3	2.45	0.3	0.6	2.09	1.0	2.3
9				0	0.3	7.89	0.2	0.2	1.29	0.7	1.6	2.23	0.4	0.8	2.02	1.3	2.9
10				0	0.3	9.34	0.2	0.3	1.25	0.9	2.0	2.31	0.6	1.1	1.75	1.7	3.7
11				0.1	0.6	10.19	0.3	0.4	1.34	1.2	2.6	2.25	0.8	1.2	1.44	2.3	4.7
12				0	0.3	11.08	0.3	0.3	1.21	0.9	2.4	2.57	0.9	1.4	1.56	2.1	4.5
13				0	0.5	11.10	0.3	0.4	1.30	1.4	3.4	2.35	1.4	1.7	1.23	3.2	6.1
14				0	0.5	13.55	0.3	0.4	1.23	1.6	3.8	2.38	1.6	1.9	1.17	3.6	6.7
15	0			0	0.4	13.31	0.4	0.5	1.14	1.7	4.7	2.75	2.1	2.3	1.07	4.3	7.9
16	0			0	0.4	13.96	0.6	0.6	1.03	2.1	5.9	2.80	2.9	3.0	1.05	5.6	10.0
17				0	0.4	14.64	0.5	0.6	1.03	1.9	5.7	2.95	2.9	2.9	1.00	5.4	9.5
18				0	0.5	16.98	0.7	0.8	1.03	2.6	7.8	2.99	4.0	3.8	0.96	7.4	12.9
19				0	0.3	17.89	0.7	0.7	0.97	2.3	7.3	3.10	3.8	4.0	1.07	6.9	12.3

Table 2.16: O/S ratios of subsequent quote changes

Quote changes are classified according to the quote position and the source of the change. The quote status may be ended by a quote change of the market maker himself (S change) or somebody else (O change). The change by the withdrawal of the quote at the end of the day is not considered here. Blank in the table implies either current of subsequent changes are not available for the quote status.

number	current change is by the market maker himself (S) or others (O)									
of	on both side				on one side				straddle	
market	alone		not alone		alone		not alone			
makers	S	O	S	O	S	O	S	O	S	O
2	0.92	0.20	1.47	0.68	3.53	0.28	1.19	0.87	5.00	1.08
3	4.00	0.75	2.88	1.65	4.18	0.29	5.18	1.91	3.98	2.52
4	1.00		4.45	2.64	4.46	0.28	4.09	1.21	5.30	2.71
5		1.00	6.44	3.45	5.36	0.26	4.15	1.01	6.22	3.20
6			8.76	4.51	5.22	0.27	4.05	0.91	6.03	3.11
7	2.00	1.56	7.86	4.95	5.20	0.30	3.72	0.83	5.94	2.91
8		2.00	8.32	6.29	4.36	0.30	3.42	0.82	6.42	3.33
9			10.77	7.50	3.41	0.28	2.82	0.77	6.39	3.35
10			11.84	8.86	3.40	0.27	3.04	0.73	7.85	3.97
11			12.08	10.17	3.40	0.24	3.69	0.60	9.25	5.36
12			19.60	10.52	2.05	0.31	2.97	0.70	8.82	5.18
13			11.88	10.05	2.87	0.21	3.27	0.54	10.02	5.39
14			20.45	12.84	2.93	0.19	3.26	0.50	10.96	5.81
15		1.00	12.22	14.09	2.10	0.22	2.66	0.54	11.83	6.58
16		0.50	12.62	13.31	1.90	0.23	2.66	0.54	10.13	5.78
17			15.30	13.48	1.82	0.24	2.55	0.54	10.28	6.02
18			14.15	17.27	2.03	0.22	2.55	0.51	12.44	6.30
19			11.10	17.66	1.81	0.24	2.22	0.60	10.03	5.24

Table 2.17: Summary of t-statistics of subsequent quote changes
Quote changes are classified into seven categories according to whether the quotes are on the yellow strip immediately before the changes. The mean of the change of mid-quote is calculated for each category of each stock. If the number of changes in the category is five or more, then the t-test is performed. All the available t-statistics are grouped into six categories according to their values. A Zero (0) in the percentage columns means the proportion of the t-tests in the group is less than 0.5% and hence rounded to zero. A blank means the t-statistics is unavailable.

quote	number of stocks of which the t-statistics are in												
status	$(-\infty,-3)$		$[-3,-2)$		$[-2,0)$		$(0,2]$		$(2,3]$		$(3,\infty)$		Sum
		%		%		%		%		%		%	
on both													
alone					1	20	2	40	2	40			5
not alone	78	7	53	5	354	30	441	38	95	8	154	13	1175
straddle	25	4	27	4	180	27	243	36	78	12	123	18	676
on ask													
alone	1	0	1	0	6	0	47	4	50	4	1129	91	1234
not alone			3	0	8	1	82	9	64	7	806	84	963
on bid													
alone	958	84	83	7	84	7	18	2			2	0	1145
not alone	679	71	78	8	160	17	34	4	2	0	2	0	955

Table 2.18: Causes of quote changes – full sample

One asterisk (*) means the t-value is different from zero at 0.05 significance level, two asterisks (**) mean the t-value is at 0.01 significance level, and three asterisks (***) mean at 0.0001 significance level.

		all variables		group two omitted	
		coefficient	t-value	coefficient	t-value
	intercept	0.0044	3.70**	0.0045	3.77**
group 1:	$x_{i,b,0}$	0.7115	24.72***	0.7358	29.11***
self trades	$x_{i,b,1}$	0.4780	15.18***	0.4461	12.11***
	$x_{i,b,2}$	0.0707	6.99***	0.1134	12.74***
	$x_{i,s,0}$	-0.6365	-31.26***	-0.6453	-37.97***
	$x_{i,s,1}$	-0.3296	-14.52***	-0.3194	-15.32***
	$x_{i,s,2}$	-0.1812	-23.01***	-0.1741	-24.60***
group 2:	$v_{i,b,0}$	0.0169	1.89		
self volumes	$v_{i,b,1}$	-0.0113	-1.02		
	$v_{i,b,2}$	0.0443	8.85***		
	$v_{i,s,0}$	-0.0055	-0.73		
	$v_{i,s,1}$	0.0034	0.67		
	$v_{i,s,2}$	-0.0024	-0.52		
group 3:	$x^o_{i,b,0}$	0.9908	14.68***	0.9881	14.63***
other trades	$x^o_{i,b,1}$	0.7929	11.86***	0.7926	11.83***
	$x^o_{i,b,2}$	0.1304	10.80***	0.1196	9.90***
	$x^o_{i,s,0}$	-0.9834	-13.51***	-0.9816	-13.49***
	$x^o_{i,s,1}$	-0.9263	-16.04***	-0.9294	-16.08***
	$x^o_{i,s,2}$	-0.0203	-10.62***	-0.0201	-10.51***
group 4:	$v^o_{i,b,0}$	0.1214	3.23**	0.1220	3.24**
other volumes	$v^o_{i,b,1}$	0.1826	4.61***	0.1833	4.62***
	$v^o_{i,b,2}$	0.1692	23.10***	0.1726	23.41***
	$v^o_{i,s,0}$	-0.0874	-4.11***	-0.0882	-4.17***
	$v^o_{i,s,1}$	-0.0295	-2.14*	-0.0298	-2.18*
	$v^o_{i,s,2}$	-0.1030	-20.30***	-0.1020	-20.11***
group 5:	$\Delta m_{t-1,0}$	0.7497	282.72***	0.7498	282.73***
previous change	$\Delta m_{t-1,1}$	0.7125	162.22***	0.7127	162.26***
	$\Delta m_{t-1,2}$	0.4725	194.06***	0.4726	193.87***
group 6:	I_t	0.0001	0.80	0.0001	0.81
inventory	ΔI_t	-0.0006	-0.62	-0.0013	-1.03
R^2		0.4841		0.4840	
number of observations (N)		642,932		642,932	
Condition Index (CI)		3.26		2.86	
Durbin-Watson (DW)		2.27		2.27	

Table 2.19: Causes of quote changes – by different number of market makers

One asterisk (*) means the t-value is different from zero at 0.05 significance level, two asterisks (**) mean the t-value is at 0.01 significance level, and three asterisks (***) mean at 0.0001 significance level.

	two market makers		ten market makers		nineteen market makers	
	coefficient	t-value	coefficient	t-value	coefficient	t-value
intercept	0.0542	4.45***	0.0035	0.47	-0.0001	-0.03
$x_{i,b,0}$	0.8312	7.60***	0.4077	4.90***	0.6571	6.06***
$x_{i,b,1}$	0.4820	2.53*	0.2275	2.65**	0.0968	0.65
$x_{i,b,2}$	0.2462	3.81**	0.0569	1.57	0.0513	1.37
$x_{i,s,0}$	-0.9728	-18.75***	-0.3618	-4.12***	-0.4136	-4.30***
$x_{i,s,1}$	-0.9146	-13.53***	-0.0389	-0.48	-0.0626	-0.75
$x_{i,s,2}$	-0.2674	-6.72***	-0.0939	-3.04**	-0.1806	-5.79***
$x^o_{i,b,0}$	0.2515	2.60**	3.6515	7.94***	21.2119	14.51***
$x^o_{i,b,1}$	0.1380	1.65	2.3799	4.70***	18.3593	11.82***
$x^o_{i,b,2}$	-0.0156	-0.40	0.1676	2.61**	0.1235	0.97
$x^o_{i,s,0}$	-0.3865	-3.05**	-2.6705	-4.87***	-18.2349	-14.08***
$x^o_{i,s,1}$	-0.4567	-6.12***	-2.1967	-5.62***	-12.3292	-9.26***
$x^o_{i,s,2}$	-0.0160	-0.88	-0.0633	-4.27***	-0.0214	-2.60**
$v^o_{i,b,0}$	0.1053	2.16*	-0.1984	-2.55*	0.4072	1.42
$v^o_{i,b,1}$	0.1345	3.61**	0.2992	2.71**	0.4826	1.76
$v^o_{i,b,2}$	0.1487	7.42***	0.2157	5.84***	0.8810	13.31***
$v^o_{i,s,0}$	-0.1164	-1.25	-0.1537	-0.61	-0.2759	-3.95***
$v^o_{i,s,1}$	-0.0613	-5.34***	-0.3233	-2.04*	-0.1847	-0.59
$v^o_{i,s,2}$	-0.0708	-4.35***	-0.1045	-5.81***	-0.3896	-7.07***
$\Delta m_{t-1,0}$	0.7227	52.81***	0.7488	101.47***	0.7193	186.84***
$\Delta m_{t-1,1}$	0.5544	9.40***	0.7709	39.08***	0.6276	53.71***
$\Delta m_{t-1,2}$	0.4543	31.99***	0.5210	45.85***	0.3497	38.28***
I_t	-0.0350	-5.64***	0.0003	0.05	-0.0000	-0.23
ΔI_t	-0.0874	-3.50**	-0.0083	-0.41	-0.0040	-0.91
R^2	0.4358		0.4887		0.4789	
N	19,310		19,094		59,733	
CI	3.00		3.47		4.46	
DW	1.78		2.28		2.31	

Table 2.20: Causes of quote changes – by selected market makers
One asterisk (*) means the t-value is different from zero at 0.05 significance level, two asterisks (**) mean the t-value is at 0.01 significance level, and three asterisks (***) mean at 0.0001 significance level.

	Firm A		Firm B		Firm C	
	coefficient	t-value	coefficient	t-value	coefficient	t-value
intercept	0.0107	2.04*	-0.0257	-2.62**	0.0463	4.92***
$x_{i,b,0}$	0.6055	13.68***	0.0306	0.14	0.7508	7.78***
$x_{i,b,1}$	0.2929	4.60***	0.3717	0.91	0.5863	7.31***
$x_{i,b,2}$	0.0615	2.20*	0.1369	1.44	0.0950	1.88
$x_{i,s,0}$	-0.5038	-10.67***	0.6989	1.16	-0.9664	-17.29***
$x_{i,s,1}$	-0.0676	-1.26	0.5674	1.07	-0.6455	-6.43***
$x_{i,s,2}$	-0.1882	-7.70***	-0.3163	-2.09*	-0.2331	-6.64***
$x^o_{i,b,0}$	0.7887	3.97***	11.5657	3.18**	0.3435	2.26*
$x^o_{i,b,1}$	0.7575	5.49***	11.6720	3.10**	0.3361	2.73**
$x^o_{i,b,2}$	0.2108	5.55***	0.1570	0.58	0.0354	0.84
$x^o_{i,s,0}$	-1.0947	-6.40***	-8.8357	-3.09**	-0.6841	-3.76**
$x^o_{i,s,1}$	-0.9482	-5.19***	-10.7185	-2.92**	-0.4927	-5.24***
$x^o_{i,s,2}$	-0.0438	-3.80**	0.0035	0.27	-0.0409	-0.72
$v^o_{i,b,0}$	0.2425	2.51*	0.3073	0.37	0.1354	1.22
$v^o_{i,b,1}$	0.1013	1.69	0.4581	0.48	0.0922	1.14
$v^o_{i,b,2}$	0.1215	6.44***	0.8327	5.03***	0.1301	7.97***
$v^o_{i,s,0}$	-0.0190	-0.39	-0.3187	-5.49***	0.1129	1.22
$v^o_{i,s,1}$	-0.0175	-0.25	-0.7007	-0.74	-0.0649	-4.53***
$v^o_{i,s,2}$	-0.0864	-6.29***	-0.6060	-3.67**	-0.0758	-4.27***
$\Delta m_{t-1,0}$	0.8571	77.03***	0.7515	74.41***	0.6609	38.19***
$\Delta m_{t-1,1}$	0.7778	37.50***	0.7471	28.84***	0.6786	28.95***
$\Delta m_{t-1,2}$	0.4989	60.12***	0.5163	24.52***	0.4059	29.89***
I_t	-0.0177	-4.32***	0.0003	0.84	-0.0093	-1.75
ΔI_t	-0.1109	-5.74***	-0.0002	-0.13	-0.0568	-3.26**
R^2	0.4242		0.5960		0.4396	
Condition Index	3.15		3.67		3.02	
Durbin-Watson	2.13		2.17		1.97	
N	51,659		8,056		20,155	

Table 2.21: Causes of quote change – by quote status after the change
One asterisk (*) means the t-value is different from zero at 0.05 significance level, two asterisks (**) mean the t-value is at 0.01 significance level, and three asterisks (***) mean at 0.0001 significance level.

	on-both		on-ask		on-bid		straddle	
	coefficient	t-value	coefficient	t-value	coefficient	t-value	coefficient	t-value
intercept	0.0404	6.67***	-0.5875	-173.02***	0.6117	191.86***	-0.0100	-6.42***
$x_{i,b,0}$	0.6910	13.22***	0.6498	17.87***	0.5269	17.37***	0.6819	9.76***
$x_{i,b,1}$	0.5136	7.17***	0.2692	4.60***	0.4034	14.93***	0.4472	12.16***
$x_{i,b,2}$	0.1709	7.31***	0.0889	6.72***	0.1490	9.74***	0.0753	6.89***
$x_{i,s,0}$	-0.7755	-16.73***	-0.5408	-27.92***	-0.4572	-15.88***	-0.5742	-17.67***
$x_{i,s,1}$	-0.5928	-9.03***	-0.4765	-18.07***	-0.1747	-6.98***	-0.3077	-8.05***
$x_{i,s,2}$	-0.2960	-11.79***	-0.3008	-26.59***	-0.0707	-6.65***	-0.1394	-13.73***
$x^o_{i,b,0}$	0.1613	1.62	2.3598	9.06***	0.1258	1.66	1.8079	9.38***
$x^o_{i,b,1}$	0.2005	2.83**	2.1220	11.68***	0.0366	0.39	1.4463	7.33***
$x^o_{i,b,2}$	0.0374	1.14	-0.1491	-6.45***	0.1226	7.94***	0.0736	3.44**
$x^o_{i,s,0}$	-0.3269	-2.96**	-0.1988	-2.68**	-1.2452	-6.34***	-1.1543	-5.59***
$x^o_{i,s,1}$	-0.4037	-6.88***	-0.1246	-1.66	-1.5988	-10.06***	-1.3264	-8.70***
$x^o_{i,s,2}$	-0.0399	-3.51**	-0.0523	-14.73***	0.0392	9.50***	-0.0104	-5.10***
$v^o_{i,b,0}$	0.0082	0.10	0.1377	1.08	0.1310	3.03**	-0.0398	-0.55
$v^o_{i,b,1}$	0.0848	1.65	0.1309	1.67	0.2025	3.63**	0.1442	1.48
$v^o_{i,b,2}$	0.1101	4.82***	0.2011	12.22***	0.0820	11.15***	0.2125	13.79***
$v^o_{i,s,0}$	-0.0335	-0.43	-0.0608	-3.04**	0.0345	0.57	-0.0756	-1.36
$v^o_{i,s,1}$	0.0057	0.36	-0.0308	-2.06*	-0.0475	-0.81	-0.0095	-0.27
$v^o_{i,s,2}$	-0.0799	-6.45***	-0.0687	-10.91***	-0.0614	-7.32***	-0.0945	-11.85***
$\Delta m_{t-1,0}$	0.7700	122.13***	0.4958	82.35***	0.4598	100.06***	0.6864	189.56***
$\Delta m_{t-1,1}$	0.6362	11.89***	0.4064	53.35***	0.4155	52.57***	0.6969	143.90***
$\Delta m_{t-1,2}$	0.5302	48.97***	0.2155	54.70***	0.2374	62.34***	0.5300	151.06***
I_t	-0.0110	-2.62**	-0.0007	-1.35	-0.0002	-0.79	0.0001	1.01
ΔI_t	-0.0293	-2.17*	-0.0153	-3.45**	-0.0007	-0.58	0.0011	1.85
R^2	0.5559		0.2127		0.2067		0.5912	
N	21,952		160,084		176,336		284,560	
CI	3.00		2.91		3.02		2.99	
DW	1.90		2.03		1.99		2.05	

Table 2.22: Summary of Wald statistics for the trade variables

This table summarises the results of Wald tests of the models in Table 2.18, 2.19, 2.20 and 2.21. The null hypothesis of any Wald test is the *absolute value* of the coefficient of standardised trades executed by market maker i ($x_{i,*,*}$) equals that of standardised trades without the involvement of market maker i ($x^o_{i,*,*}$). "O" in the "which is bigger" column indicates the coefficient of $x^o_{i,*,*}$ is bigger, and "S" indicates the coefficient of $x_{i,*,*}$ is bigger. One asterisk (*) means the t-value is different from zero at 0.05 significance level, two asterisks (**) mean the t-value is at 0.01 significance level, and three asterisks (***) mean 0.0001 significance level.

compared the scales of	which is bigger	Wald statistics	which is bigger	Wald statistics	which is bigger	Wald statistics
Panel 1. Regressions of the whole sample (Table 2.18)						
	all variable		without group two			
$x_{i,b,0}$ vs $x^o_{i,b,0}$	O	13.81**	O	11.71**		
$x_{i,b,1}$ vs $x^o_{i,b,1}$	O	17.69***	O	19.72***		
$x_{i,b,2}$ vs $x^o_{i,b,2}$	O	11.37**	O	0.14		
$x_{i,s,0}$ vs $x^o_{i,s,0}$	O	20.51***	O	19.79***		
$x_{i,s,1}$ vs $x^o_{i,s,1}$	O	88.80***	O	94.75***		
$x_{i,s,2}$ vs $x^o_{i,s,2}$	S	368.29***	O	410.94***		
Panel 2. Regressions by number of market makers (Table 2.19)						
	2 MMs		10 MMs		19 MMs	
$x_{i,b,0}$ vs $x^o_{i,b,0}$	S	16.08**	O	47.07**	O	195.26***
$x_{i,b,1}$ vs $x^o_{i,b,1}$	S	2.83*	O	17.39***	O	135.87***
$x_{i,b,2}$ vs $x^o_{i,b,2}$	S	8.21**	O	1.88	O	0.27
$x_{i,s,0}$ vs $x^o_{i,s,0}$	S	18.35***	O	16.60***	O	186.81***
$x_{i,s,1}$ vs $x^o_{i,s,1}$	S	21.04***	O	27.87***	O	84.33***
$x_{i,s,2}$ vs $x^o_{i,s,2}$	S	34.79***	O	0.72	S	22.58***

(continued to the next page)

(continued from the previous page)

compared the scales of	which is bigger	Wald statistics	which is bigger	Wald statistics	which is bigger	Wald statistics
Panel 3. Regressions by selected market makers (Table 2.20)						
	Firm A		Firm B		Firm C	
$x_{i,b,0}$ vs $x^o_{i,b,0}$	O	0.79	O	10.01**	S	5.03*
$x_{i,b,1}$ vs $x^o_{i,b,1}$	O	9.13**	O	8.92**	S	2.86*
$x_{i,b,2}$ vs $x^o_{i,b,2}$	O	7.89**	O	0.00	S	0.52
$x_{i,s,0}$ vs $x^o_{i,s,0}$	S	10.72**	O	10.38**	S	2.24
$x_{i,s,1}$ vs $x^o_{i,s,1}$	O	21.44***	O	9.49**	S	1.21
$x_{i,s,2}$ vs $x^o_{i,s,2}$	S	24.12***	S	4.36*	S	6.93**
Panel 4. Regressions by quote status (Table 2.21)						
	on-both		on-ask		on-bid	
$x_{i,b,0}$ vs $x^o_{i,b,0}$	O	23.71***	O	42.18***	S	22.28***
$x_{i,b,1}$ vs $x^o_{i,b,1}$	O	9.77**	O	92.29***	S	13.70**
$x_{i,b,2}$ vs $x^o_{i,b,2}$	O	9.11**	S	65.48***	S	1.06
$x_{i,s,0}$ vs $x^o_{i,s,0}$	O	14.25***	S	19.16***	O	15.75**
$x_{i,s,1}$ vs $x^o_{i,s,1}$	O	4.59*	S	18.99***	O	77.44***
$x_{i,s,2}$ vs $x^o_{i,s,2}$	O	76.79***	S	404.06***	S	83.51***
	straddle					
$x_{i,b,0}$ vs $x^o_{i,b,0}$	O	27.57				
$x_{i,b,1}$ vs $x^o_{i,b,1}$	O	24.39**				
$x_{i,b,2}$ vs $x^o_{i,b,2}$	O	0.00**				
$x_{i,s,0}$ vs $x^o_{i,s,0}$	O	7.51**				
$x_{i,s,1}$ vs $x^o_{i,s,1}$	O	39.40***				
$x_{i,s,2}$ vs $x^o_{i,s,2}$	S	145.21***				

Chapter 3

COST OF TRADING WITH MARKET MAKERS ON THE LONDON STOCK EXCHANGE

3.1 Introduction

Although the growing use of computer trading systems has assisted the trading in dealership markets, a large number of trades, especially those with non-trivial sizes, are still carried out by negotiation among traders. If the traders have been trading in the market for a long time, they may know their trading parties very well. During negotiation, the trading parties may gather more information about the security than merely the price and the size of the trade. This chapter focuses on the comparison of the cost of trading with market makers on the LSE. The data constructed from the settlement records consists of not only the prices and quantities of the trades, but also the identities of the trading parties and trading conditions. Although it is not possible to identify the end customers behind brokers, the data provides sufficient clues to understand how the market makers differentiate their customers.

Dealership markets are often designed to encourage dealers to compete with one another to attract customers, yet in practice people often wonder how competitive those dealers are. For example, there are suggestions that NASDAQ dealers collude (Christie and Schultz 1994; Christie, Harris, and Schultz 1994). Dutta and Madhavan (1997) argue the entry costs to a dealership market include establishing reputations and forming long-term relationship with brokers, which can be very expensive. In London, the difference in trading costs is alleged to be the result of price discrimination (Securities and Investments Board 1995a), which implies market makers exercise monopolistic power against the investors. It is interesting to investigate how much evidence there is to support the allegation.

Moreover, an inter-dealer-broker (IDB) market runs parallel to SEAQ to provide liquidity for market makers. Board and Sutcliffe (1995) and Reiss and Werner (1996, 1997) have shown that market makers trade with one another in both markets frequently, but the cost of trading of the two markets differ substantially. As market makers trade the same stocks with each other in both markets, the London market provides a laboratory to observe the similarity and difference between the two markets. The comparison between the two markets may shed some light on the ongoing discussion of the trading mechanisms.

The rest of this chapter is organised as follows. Section 3.2 gives the institutional background about the cost of trading on the LSE. Section 3.3 presents some descriptive statistics for the data investigated in this chapter. Section 3.4 reviews the literature of the theoretical arguments and empirical findings in the literature about the cost difference and the structural of dealership and auction markets. Section 3.5 shows some of the cost difference can be explained by the theories reviewed in Section 3.4. Section 3.6 investigates the information contents of different classes of trades, and Section 3.7 concludes the findings.

3.2 Institutional background

Section 1.1 has given a description of the operation of the LSE. The market makers in the Exchange trade domestic equities mainly in two trading systems: the Stock Exchange Automated Quotation system (SEAQ) and the Inter Dealer Broker (IDB) market. All members of the Stock Exchange are able to trade with market makers directly in SEAQ; they may act as principals to trade on their own account, or they can act as agents on behalf of the clients. On the other hand, IDB market is one to which only market makers have exclusive access. A market maker may contact one of the four inter-dealer brokers (IDBs) to place limit orders, and the IDB will put the best buy and sell limit orders on their computer screens. If another market makers wish to take the orders, they may call the IDB and the trades are executed.

Although this chapter will investigate the cost structure in greater details, it is initially motivated by the disparity of the trading costs in SEAQ and IDB market, a widely accepted allegation made by market participants in the survey by Securities and Investments Board (1995a). The survey finds the existence of IDB market appears to generate some resentment of the broker-dealers. Neither IDBs nor market makers are allowed to reveal any information in the IDB market. The broker-dealers suspect the best limit orders in IDB screens contain useful information, they believe the market makers use IDB market for price discovery, so the market makers have information advantage against them. More importantly, the brokers observe that a large number of trades take place at the prices well inside the touch, which are much better than the prices they can get from market makers. They believe some of those trades are IDB trades, so the conclude the market makers have costs advantage by trading with each other in the IDB market. Securities and Investments Board (1995b), however, supports the view of market makers that the exclusive access of the market makers to the IDB market is essential for the market makers to lay off their risk positions rapidly and anonymously, otherwise "the efficiency distribution of risk and liquidity would be seriously impeded (p.47)."

The empirical evidence of the cost difference between trading in SEAQ and in IDB is documented in Board and Sutcliffe (1995). They classify trades into three categories: the trades among market makers in the IDB market (which they call IDB trades), the trades between market makers and customers in SEAQ (customer trades) and the trades between market makers in SEAQ (IMM trades). An interesting result emerges: of the 42 stocks they survey, IDB trades are the cheapest in every stock, and IMM trades are the most expensive in 35 stocks (p.77). The result is robust even after the trades are grouped by size or by buy and sell. Using a similar methodology, Reiss and Werner (1997) concentrate on the comparison between IDB and IMM trades, and again, IDB trades appear to be cheaper than IMM trades. Reiss and Werner (1996) further define apparent spreads and find IDB trades are cheapest while costs of IMM trades and customer trades are the same. To sum up, the brokers' allegation that IDB trades cost less is confirmed, but the findings of Board and Sutcliffe and Reiss and Werner bring about more questions. The practitioners are fully aware of the difference in costs, some of them even offer plausible explanations, but the answers to those questions may not be as straightforward as they seem to be.

3.3 Data description

The transaction records of thirty-nine stocks are selected from the CD-ROM *Transaction Data Service* provided by the Quality of Market Group of the London Stock Exchange. The sample includes thirty-nine stocks, nineteen of which are FTSE-100 stocks, and twenty of which are randomly selected from the rest of FTSE-All share stocks, see Appendix A.3.2 for details. The procedure of editing the data is in Appendix A.4. A trade record consists of the price and quantity of the trade, the time when the trade was executed, the identities of the trading parties, and the flag of the "dealer capacity" of the trading parties. According to the dealer capacity, trades are classified into five categories: brokers on behalf of their clients (class A), private clients of market makers (N), inter-dealer brokers (I), market makers (M) and non-market-maker principals (P). Further description of the data is in Appendix A.2.

This chapter is concerned with the trades which involves market makers and which take place in mandatory quote period, and trade categories are defined according to the dealer capacities. In other words, class "A" trades are the trades between market makers and brokers, "I" trades are the trades between market makers and inter-dealer brokers (IDB trades), "M" trades are trades between market makers in SEAQ (IMM trades), and so on. The trades that involve market makers but take place outside mandatory quote period are classified as T trades. The trades without the involvement of market makers, such as agency cross, are classified as X trades. Figure 3.1 illustrates the relationship between market participants and the trade classes, and Table 3.1 presents the summary statistics of the trades. Three quarters of the trades in the sample are between market makers and brokers on behalf of small investors (A trades), but they only contribute to 20% of the volume. The average trade size of A trades is tiny compared with the size of the other categories. N trades are trades between market makers and their private clients.[1] Those trades contribute to more than 30% of the volume, but they only account for less than 8% of the trades. The number of "M" trades (IMM trades) and the trading volumes are doubled to make them comparable with IDB trades.[2] The market makers use IDB market to trade with each other very often, but they also trade with one another in SEAQ. Although the volume and the number of I trades are greater than M trades, the size of M trades tends to be bigger. In fact, the average size of M trades is the biggest of all categories.

There are two different ways to measure the cost of trading with market makers, both of which have been used in Chapter 2. The first approach is to classify the trade according to the best bid and ask prices in the market, for example, the trade is executed at, inside or outside the touch. (Neal 1992; Wells 1992; Huang and Stoll 1996). A slightly more sophisticated measure is further developed by Board and Sutcliffe (1995) and Reiss and Werner (1997) to calculate the gains to the traders by the distance between the price and

[1] They are labelled as "internalised trades" in Hansch et al. (1999).

[2] An M trade involves two market makers, and an I trades involve a market maker and an inter-dealer broker. Reiss and Werner (1997) halve I trades to make it comparable M trades, implicitly assuming a "full" I trade involving an inter-dealer broker, one buy market maker and one sell market maker. However, more than 20% of I trades in the sample involve several buyers and/or sellers, and halving the volumes of IDB trades makes the interpretation difficult. Doubling M trades is another way to make I trades and M trades comparable. See Appendix A.4.5 for further issues about the IDB data.

the best quote. Let p be the price of the trade, m_t be the mid-touch, and a_t and b_t be respectively the best ask and bid quote when the trade occurs, then the gain from the trade is defined as

$$\text{gain} = \begin{cases} (a_t - p)/m_t & \text{for a buy} \\ (p - b_t)/m_t & \text{for a sell} \end{cases} . \tag{3.1}$$

Market makers make two-way firm quotes during mandatory quote period. If the customer buys the stock with a price below the best ask price, or the customer sells the stock above the best bid price, then the customer is said to gain from the trade.[3]

Measuring the gains or costs according to the touch prices is fine as long as the touch spread is constant. The problem arises when the touch spread varies, which can be illustrated in Figure 3.2. The example is taken from the trade records of Bank of Scotland on 25 April 1996. The best bid and ask prices are respectively 251 and 252 pence during most of the morning, and are 250 and 253 pence after 16:15. A trade at the price 252 pence would have been classified as at the touch if occurred in the morning, but would be classified as inside the touch if occurred in the late afternoon. This is exactly what happened: a market maker bought 25,000 shares directly from another market maker at 9:32 at 252 pence (an M trade), and he bought another 25,000 shares again at 16:27 from the third market maker via IDB market at 252.13 pence (an I trade). If the definition in (3.1) is used, the gain from the morning M trade is zero, and then the gain from the afternoon I trade is 0.35%. Certainly the market maker cannot claim to gain more by paying more for the IDB trade. If a certain class of trades occur when the touch is wider than the other classes, then the gains from this class of trades will be bigger than the other trades, other things equal. For example, Reiss and Werner (1997) observe IMM trades concentrate in the morning IDB trades concentrate in the afternoon. They show that IMM trades occur when the touch is narrow and IDB trades occur when the touch is wide. What happened with the stock of Bank of Scotland fits perfectly in their

[3] The use of the term "gain" follows Reiss and Werner (1997). Board and Sutcliffe (1995) call it "Reltouch".

analysis: the touch was narrow in the morning and an IMM trade took place at the touch, and when the touch was wide in the afternoon, an IDB trade took place inside the touch. The market maker "gained" more from IDB trades, but he in fact paid a higher price to buy the stocks in IDB market.

The second approach to measure the costs is to define the costs as the difference between the price and the mid-touch, or the difference divided by the mid-touch. Reproducing (1.1),

$$\text{cost} = s^f = \begin{cases} (p - m_t)/m_t & \text{for a buy} \\ (m_t - p)/m_t & \text{for a sell} \end{cases}, \text{ where } m_t = (a_t + b_t)/2. \quad (3.2)$$

The direction of the trade has to be identified in order to compute the costs. If the trade involves only one market maker, then the buy and sell can be simply identified as the direction of the trade of his counter party. The direction of IDB trade can be identified as the direction of the hitter (Board and Sutcliffe 1995; Reiss and Werner 1997). However, the direction of an M trade has to be decided by the mid-touch: the M trade is a buy if its price is above the mid-touch, and is a sell if its price is below the mid-touch. If the M trade occurs at the mid-touch, then the direction of the trade is irrelevant in computing the costs of the trade, as the costs would be zero regardless of the direction. It appears equation (3.2) measures the costs of the trade better than equation (3.1), so the former will be the measure the cost difference hereafter.

Table 3.2 summarises the average cost of trading under different trading categories. Broker trades (A trades) have the highest costs, and IDB trades (I trades) have the lowest. Private clients of the market makers pay less than the brokers, consistent with the finding in Hansch et al. (1999). M trades cost more than I trades, which is consistent with the finding in Board and Sutcliffe (1995) and Reiss and Werner (1997) despite a different measure is used. Board and Sutcliffe find the costs of customer trades are between those of I trades and M trades, but a different picture emerges in this sample after N trades are separated from A trades: the latter are more expensive than M trades whereas the former are cheaper. The results hold in both FTSE-100 and non-FTSE-100 stocks.

The third panel is a further breakdown of the costs of the trades according to their normal market sizes. For the trades with the sizes less than six times NMS, the results are similar to the full sample: A trades are more expensive than M trades, which are more expensive than N trades, which are more expensive than P trades, and so on. The difference in costs between A trades and N trades becomes small when the sizes of the trades are between 6 and 75 time NMS. The next panel shows the percentages of trades under different category of trades. 87.15% of the A trades and 35.24% of the N trades are of the size less than 0.1 times NMS, but very small number of inter-dealer trades fall into this category. The sizes of M trades tend to be larger than those of I trades. P trades tend to be bigger than N trades. However, the percentages of big N trades are bigger than the other trades.

Finally, the last panel shows the average half touch weighted by the number of trades. The touches of FTSE-100 stocks when the trades occur are about half of the non-FTSE-100 stocks. If all of the trades occurred at the touch, then the average cost would be simply the same as half of the average touch. As more of the trades occur inside the touch than outside, the average cost of all classes of the trades is smaller than half the average touch. The average touch when I trades occur is among the biggest, and the average touch when M trades occur is the smallest, which is consistent with the finding by Reiss and Werner (1997) that IDB trades occur when the touch is narrow, and IMM trades occur when the touch is wide.

3.4 Explanations for the cost difference

The empirical evidence from the previous section shows that costs of trades with market makers vary with the trading parties. On the other hand, there have been extensive discussions in the literature regarding the cost of trading.

3.4.1 Order processing, inventory and adverse selection costs

Traditionally, the spreads are believed to consist of three components: order processing costs, inventory bearing costs and adverse selection costs (O'Hara 1995; Huang and Stoll 1997). The cost difference among the different classes of trades may result from the difference in at least one of the three components. If some trades contain more information than the others, then the former should have bigger spread than the latter. Numerous researchers subscribe to this view (Copeland and Galai 1983; Glosten and Milgrom 1985; Easley and O'Hara 1987; Glosten 1989). Applying the theory to the LSE, if the order-processing costs and the inventory bearing costs are the same among all classes of trades, then the difference of costs would be attributed to the information contents of costs: broker trades contain more information than client trades, IMM trades contain more information than IDB trades, and so on.[4] Easley and O'Hara (1987) argue the adverse selection costs increase with the size of the trade, as the informed trader wishes to take the most advantage of the information. In contrast, Naik et al. (1999) argue the market maker is a quasi-insider in the market before the trade is published, so the spread could decline with the information conveying in the trade in order to attract insider to reveal more information during the negotiation. The price schedule of the dealer therefore contains the size and the information. In fact, the evidence from Table 3.2 shows the costs of the trades are not always increasing with size for the trades. It appears size alone cannot explain the difference in the costs of the trades — either the market makers regard different classes of trades contain different information, or the other components of the trades are different.

The inventory of the dealer affects the spreads in two ways. The inventory level before the trade affects the placement of the quote, and the size of the current trade affects the magnitude of the bid-ask spread. Moreover, the inventory level affects the spread indirectly through the marginal utility of the

[4] This is a completely different view of some market practitioners, who believe IDB trades contain information (Securities and Investments Board 1995a).

wealth of the dealer (O'Hara and Oldfield 1986). The argument holds under both the assumption of a monopolistic specialist (Amihud and Mendelson 1980; Ho and Stoll 1981) and of multiple market makers (Ho and Stoll 1983). The inventory effect may explain the cost difference if, for example, cheap trades and expensive trades occur at different times or different directions when market makers have unbalanced inventory. Figure 3.3 illustrates that there are two market makers with different inventory levels: market maker A is short of inventory, so he raises the bid and ask price to attract more sell orders. Market maker B, on the other hand, has too many shares, so he lowers the quotes to attract more buy orders. Suppose market maker A attracts a sell order from the private client (an N trade), and a buy order from a broker arrives at the same time (an A trade). As a result, it would appear that the N trade is cheaper than the A trade, measured by the distances between the prices and the mid-touch. The same applies to when market maker B attracts an N buy and an A sell.

The order-processing cost of the spread is discussed in Demsetz (1968), Tinic (1972) among others. The costs include not only the costs to process the order, but they also include the gains from the monopolistic power of market makers. The "pure" order processing costs in a dealership market may contain a fixed component (de Jong et al. 1995), so the costs per share are decreasing in the size of the trade. It may explain why the costs of the small SEAQ trades may be bigger than the big ones. The monopolistic power of the dealers will be analysed in more details later.

3.4.2 Cartel

Like NASDAQ dealers, the market makers one the LSE are not immune from the accusation of collusion. The suggestion of the collusion in NASDAQ comes from the observations that NASDAQ dealers do not make odd-eight quotes as often as the even-eighth quotes (Christie and Schultz 1994; Christie, Harris, and Schultz 1994). Though the LSE does not have rules of minimum tick, a few practitioners (Securities and Investments Board 1995a) and Snell and Tonks (1998) support the claim that the market makers exploit the dealer. They suggest that market makers trade with one another low prices and collude to charge the customers higher prices. Huang and Stoll (1996) and Dutta and Madhavan (1997) also suggest inter-dealer trading helps market makers offseting the inventory so easily that they have little incentive to reduce the spread to attract small investors. Furthermore, market makers are able to receive preferenced order flow, which does not provide their incentive to reduce the spread, either. The collusion hypothesis appears to explain the low costs of IDB trades: the market makers offer favourable prices to each other in IDB market, and then exploit their customers in SEAQ. However, IMM trades are as expensive as, if not more than, the customer trades. If market makers collude, why do they offer unfavourable prices to each other in SEAQ?

Dutta and Madhavan argue that under implicitly collusion, the spread increases with the volume up to a critical point, and decreases with volume afterwards. The volume in their model is essentially the demand for trading, that is, the bigger the demand, the bigger the spread. On the other hand, too large demand for trading provides the incentive of market makers to cut the price to induce more business, and the collusion collapses. The relationship between volume and spread may indicate whether market makers collude.

3.4.3 Market mechanism

The above discussion focuses on the difference between market makers and the other market participants. The arguments are based on the belief that market makers are a special class of market participants, and the difference in costs between IDB trade and the rest of the trades reflects the difference in the characteristics between market makers and the rest of the traders. However, the difference in transaction costs may at least partly result from the difference in the two trading systems. IDB market differs from SEAQ in many aspects. IDB is an order-driven system. The limit orders of market makers are displayed on the IDB screen, and trades take place via an inter-broker dealer. On the contrary, SEAQ is a quote-driven system. Except for the small trades, buyers and sellers communicate with each other through telephone calls, and hence the market makers know the identities of the traders. The comparison of two trading systems is considered by Madhavan (1992). The market makers are assumed to offer a price schedule to the investors. Equilibrium exists if the information asymmetry is not very severe. Madhavan finds continuous auction is more robust to information asymmetry than a dealership system, but the prices in the auction system are not efficient and more volatile.

Vogler (1993) considers the scenario in which a market maker acts as a quasi-insider in the market. The news of the trade with customers comes to the market with some delays, so the market maker can exploit the temporary information advantage by trading with the other market makers. As a result inter-dealer trading has information contents. If there is no private information, Vogler (1997) finds the spread of the inter-dealer trade is smaller than the customer trade, mainly because the spread is increasing with the risk (the size) of the trade and the size of inter-dealer trading is smaller than the customer trade. The difference in spread is to compensate the additional risk of the dealer who trades with the liquidity trader, since in the end the dealer cannot diversify the risk asset away completely. The risk-sharing scenario may explain the small costs of IDB trades, but why IMM trades are so expensive is difficult to be fit in the story. Pagano and Roëll (1992) argue that a dealer market is less transparent than a batch auction market, so the market makers are forced to set spreads wider to guard against the information advantage. On the other hand, they suggest that an electronic market does not provides a mean for communication, whereas the market makers may be able to recognise the insiders as well as liquidity traders, so the spread may be smaller.

Board and Sutcliffe (1995) give a brief explanation about the existence of IMM trade and the reason for the difference in costs (p.74). They suggest that the IDB market does not provide enough immediacy, so it occurs that market makers need to pay a premium to buy immediacy in SEAQ from time to time. Reiss and Werner (1997) found IDB and IMM trades are likely to be substitutes of each other. They plot the IDB and IMM trading and find

IMM trades peak in the morning and IDB trades peak in the afternoon. They argue that there are relatively few IDB posting in the morning, so the market makers have to resort to IMM trades. Second, they report IMM trades are more common at narrow touches than IDB trades. They said the market makers may be more likely to use IMM trades when the touch is narrow. If spreads are narrow enough when IMM trades occur, then there is very little room for the market maker to manoeuvre.

Finally, numerous researches have been conducted on the comparison of the execution costs in dealer markets and auction markets. Lee (1993), Affleck-Graves et al. (1994), and Huang and Stoll (1996) compare the costs of NYSE and NASDAQ, Neal (1992) compares the costs of AMEX and CBOE, and de Jong et al. (1995) compare the cost of trading French shares in London and Paris market. Most of the work show that the cost of trading in auction markets are equal to or lower than dealership markets.

3.5 Determinants of the spreads: a regression analysis

Section 3.3 shows cost of trading with market makers vary substantially with the customers, Section 3.4 reviews the literature which offers explanations for the difference in costs, and this section is to examine to what extent the cost difference can be explained by the existing theories. The dependent variable of the regression models is the effective spread defined in equation (3.2). There are two groups of independent variables. The first group consists of dummy variables of the class of trade. Let d_I, d_M, d_N and d_P be respectively dummy variables of the I, M, N and P trades: 1 indicates the trade belongs to the class, and 0 otherwise. The second group consists of the "literature" variables, which are suggested in the literature to explain cost of trading:

1. The variables that measure the competition of the market: MM, the number of market makers of the stock at the day when the trade takes place. FTSE, the dummy variable of FTSE-100 stocks: 1 indicates the

stock is in the group of FTSE-100, and 0 otherwise. The third variable is sp_t, the size of half touch divided by the mid-touch. Both FTSE and MM are expected to be negatively related to the costs of trade, and sp_t is positively related to the costs.

2. sd_t, the standard deviation of the prices of the trades from the beginning of the day till one minute before the trade. It is computed first by taking the average of the standard deviation of buy prices and that of the sell prices, then the average is standardised by the average standard deviation at the time of the day. If there are only buy or sell trades, then use the standard deviation which is available without average. If there is no or only one previous trade in both buy and sell sides, sd_t cannot be computed and the observation is removed from the analysis. If traders are risk-averse, then the spread is increasing in the risk of the underlying securities (Stoll 1978b; Ho and Stoll 1981; Ho and Stoll 1983; Easley and O'Hara 1987).

3. The variables associated with the inventory effect. Following Madhavan and Smidt (1991), the pricing function of the market maker may be written as follows:

$$p = \begin{cases} \mu_t - \alpha(I_t - I_d) + s(\ldots) & \text{for a buy} \\ \\ \mu_t - \alpha(I_t - I_d) - s(\ldots) & \text{for a sell} \end{cases}, \tag{3.3}$$

where μ_t is the "true" price of the underlying security, I_t is the inventory level at time t, I_d is the ideal inventory level, and s is half of the bid-ask spread, a function of many variables. Combine equation (3.2) and (3.3),

$$\text{cost} = \begin{cases} (p - m_t)/m_t &= \frac{1}{m_t}[\mu_t - m_t - \alpha(I_t - I_d) + s(\ldots)] & \text{for a buy} \\ \\ (m_t - p)/m_t &= \frac{1}{m_t}[-(\mu_t - m_t) + \alpha(I_t - I_d) + s(\ldots)] & \text{for a sell} \end{cases}. \tag{3.4}$$

Furthermore, let d_b be 1 when the trade is a buy, and -1 when the trade is a sell, then (3.4) can be written as

$$s = \frac{1}{m_t}[d_b(\mu_t - m_t - \alpha(I_t - I_d)) + s(\ldots)]$$

Let I_0 be the inventory level at the beginning of the sample period, v_t be the net number of shares bought by the market makers between the beginning of the period and time t, then $I_t = I_0 + v_t$. Assume the ideal inventory level is the average of daily closing inventory level

$$I_d = \sum_{s=1}^{125} (I_0 + v_{s,e})/125,$$

where $v_{s,e}$ is the accumulated trading volumes at the end of date s, and there are 125 trading days in the sample. Hence

$$I_t - I_d = v_t - \sum_{s=1}^{125} v_{s,e}/125,$$

and

$$s = \beta_0 d_b + \beta_1 d_b (v_t - \sum_{s=1}^{125} v_{s,e}/125) + s(\ldots).$$

That is, the inventory effect is captured by d_b, and the multiplication of d_b and v_t. The inventory theory predicts $\beta_1 > 0$, and β_0 can be either positive or negative.[5]

The regression will be run on the sample of 39 stocks with different market makers. Market makers may have different ideal inventory levels of different stocks, so the standardised inventory is used; see Hansch et al. (1998).

4. The non-linear effect of time of the trade is modelled by cubic splines. Cubic splines are a set of cubic polynomials which are differentiable and continuous in the domain of independent variable. The technical details are described in Appendix B.1. Three knots points of the splines are 8:30, 12:30 and 16:30 respectively. To avoid the colinearity, the value of time spline at 8:30 is restricted to zero, and the two remaining variables are T1 and T2 respectively. The variable is to capture the intraday effect observed in some of the markets; see Section 4.2.4 for details.

[5] Some of the M trades are executed at the mid-touch. Their trade direction cannot be identified (d_b is unknown) and are removed from the analysis.

5. The size of the trade measured by NMS multiples. To capture the potential non-linear effect, the regression spline is used. The knot points are 0, 1, 6, 75, and 360 times NMS.[6] The spline of zero times NMS is restricted to zero to avoid colinearity, and the remaining variables are Q1, Q2, Q3, and Q4 respectively.

6. The trading volumes of the stock before the trade take place. According to Dutta and Madhavan (1997), if market makers collude implicitly, cost of trading increases with market volumes first and then decreases. As the accumulated volumes increase during the day, the variable is divided by the average accumulated volumes at the time of the day to measure the excess demand for trade. The volume effect is again modelled by cubic splines with knots point -1.464, 0, 4.407 and 10.278.[7] The spline of the first knot is restricted to zero to avoid colinearity, and the remaining variables are V1, V2, and V3 respectively.

To compare the effect of class dummies with the effect of "literature" variables on execution costs, three regression models are tested. Model (1) uses all of the variables, Model (2) has only "literature variables", and Model (3) uses trade-class dummies as independent variables. If the second group of variables explains the cost difference well, then Model (2) will perform better than Model (3), and the class dummies in Model (1) will be statistically insignificant.

The results of the regressions are shown in Table 3.3. Parameter estimates, t-values of the variables, and the R-squares of the models are presented.[8] By comparison of their R-squares, Model (3) apparently performs much poorer than Model (2) and Model (1). Indeed, as all of the class dummies of broker trades are zero, Model (3) merely reflects the average cost difference among different classes of trades: I trades are 17 basis points

[6] 360 times NMS is the maximum size observed in the sample. One, six and seventy-five times NMS correspond to the critical points of trade publication rules; see Section 1.1.

[7] -1.464 and 10.278 are respectively the maximum and minimum volumes observed in the sample, and 4.407 is the average of them.

[8] T-values are adjusted for heteroscedasticity using White's (1980) approach.

cheaper than A trades, M trades are two basis points cheaper than A trades, and so on. The cost differences among the trades are not exactly same as those reported in Table 3.2 because some of the observations used in the previous table are excluded from the regression analysis.

A glance at the coefficients associated with the second group of variables in Model (2) and Model (1) reveals the behaviour of execution costs are partly consistent with most of the existing literature. For example, the coefficients of both FTSE and MM are expected to be negative, but the coefficient of FTSE is positive in Model (1) and the coefficient of MM is positive in Model (2). Moreover, only the coefficient of FTSE in Model (1) is significant. The half touch, sp_t, is the most significant variable of the model. It reflects the fact that 71% of the trades in the sample occur at the best bid or best ask. The costs are increasing with sd_t, which is consistent with the theory that risk-averse market makers demand more compensation when the market is volatile. The inventory theory does not impose any restriction on d_b, but $d_b * v_{d,t}$ is expected to be negatively correlated to the cost, which is contradicted with the result of both Model (1) and Model (2). The coefficients of time spline, size spline and volume spline should not be considered separately; they should be recovered from the knots and the coefficients. Figure 3.4 plots the three estimated splines. The cost of trade is gradually increasing throughout the day, which are contradicted to the U-shape theory (Brock and Kleidon 1992).[9] The size spline declines sharply in the beginning until 5 times NMS, then continues to decline slowly to 60 time NMS, and then increases. The size effect is consistent with the finding by Hansch et al. (1999), in which they find the execution cost (the effective spread) is positively related to the sizes of big trades and negatively related to the sizes of small trades. The volume spline increases between zero and eight times standardised volumes, and is flat elsewhere. It does not support the collusion hypothesis by Dutta and Madhavan (1997). However, both the time and volume effect are of little economic significance. The difference between the maximum and minimum of time spline is less than one, that is, the time of the trade does not change

[9] Chapter 4 will examine the existence of intraday effects in more details.

the cost by more than one basis point. The range of volume spline is slightly larger, but it is still less than 1.5 basis points. In contrast, the range of size spline is more than fifty basis points. Even if the majority of the trades are less than 60 time NMS, the range of the spline with the size less than 60 times NMS is still bigger than ten basis points. Thus, the costs of trades are more sensitive to the sizes than volumes or times. The significance the coefficients of inventory-related variables, time, size and volume splines cannot be considered separately. The second panel of Table 3.3 presents the results of Wald tests of jointly significance of the splines. All of the splines are statistically significant, but χ^2 of the size spline is much bigger than those of the other two. Furthermore, the differences among the coefficients of class dummies in both Model (1) and Model (3) are all significant.

The coefficients of the same variable in Model (2) and Model (1) are often very similar. The introduction of class dummies in Model (1) has changed the signs of coefficients of FTSE and MM, and the magnitudes of coefficients of size splines and volume splines are reduced.[10] The biggest difference between the coefficients in Model (3) and Model (1) (and Model (2)) is the intercept: the variables suggested in the literature capture most of the effects of the intercept in Model (3). The absolute values of the coefficients of d_I and d_N are smaller in Model (1) than in Model (3), which means the difference of cost of trading among these classes of trades are not as big as it appears to be, after a great deal of effects discussed in the existing literature have been taken into account. d_P does not change very much. However, the dummy variable of M trades, d_M, changes from significantly smaller than zero in Model (3) to significantly bigger than zero in Model (1). It may imply that M trades are in fact the most expensive class of trades.[11] Finally, the Condition Index (CI) of Model (1) and Model (2) are high, which is mainly due to the fact that MM and sp_t are highly correlated. However, this does not affect the ability

[10] The splines of Model (2) are very similar to those of Model (1) except the ranges of volume and size splines have increased to two and seventy basis points respectively.

[11] However, those M trades which occurred at the mid-touch are not included in the analysis as the directions of trades cannot be identified. The exclusion of those trades may over-estimate the costs of M trades.

of the model to explain the costs. The Durbin-Watson statistics (DW) of the two models do not reveal the residuals of the two models are autocorrelated. In contrast, there is no sign of colinearity in Model (3), while its residuals are highly autocorrelated.

Although Model (2) performs better than Model (3), it does not mean the class dummies are unimportant in determining the costs of trades. In the last panel of Table 3.3, F statistics are used to test if the coefficients of the class dummies are jointly significant in Model (1). The F-statistics of the class dummies is 41,889.11, which confirms the class dummies still play an important role in determining the costs in Model (1). The variables proposed in the literature do explain the cost of trading very well, but they still cannot explain why there is difference among the costs of the different classes of trades.

3.6 Trades and quote changes

The previous section shows that a large part of the execution costs can be explained by the theories in the existing literature. On the other hand, the significance of trade dummies in Model (1) implies that market makers appear to treat different class of trades differently. An explanation of the significance of the dummies is that different classes of trades have different informational contents, which are captured in the class dummies in the regression models. If market makers believe that certain trades contain information, they may revise their beliefs about the fundamental value of the stock. The changes in beliefs may lead them to adjust the quotes. The aim of this section is to investigate if particular classes of trades affect the change in quotes.

Quotes of each market maker are provided in the CD-ROM *Transaction Data Service.* During the mandatory quote period in the first half of 1996, twenty-two market makers make 315,364 quote changes of the thirty-nine stocks in the sample. Chapter 2 has shown that market makers often change the quote just because another market makers have done so. Therefore, the quote changes are removed from the sample if there is another quote change preceding the current change by 30 seconds in the same direction, and only 141,225 quote changes are included in the analysis.

Define the dependent variable m_t the last mid-quote between t minute zero second and t minute 59 second by the market maker, then the change in quote during minute t is defined as

$$\Delta m_t = (m_t - m_{t-1})/m_{t-1}. \tag{3.5}$$

Note that here m_t is re-defined as the mid-quote of a market maker, not the mid-touch as in the previous sections of the chapter.

Two regression models are used. The first model includes independent variables such as number of buy and sell trades and volumes of different classes of trades, and the model can be written as

$$\begin{aligned}\Delta m_t &= \alpha_0 + \sum_{c1}\sum_{n=t-3}^{t}(\alpha_{c1,bx,n}x_{c1,b,n} + \alpha_{c1,bv,n}v_{c1,b,n} + \alpha_{c1,sx,n}x_{c1,s,n} + \alpha_{c1,sv,n}v_{c1,s,n}) \\ &+ \sum_{c2}\sum_{n=t-3}^{t+3}(\alpha_{c2,bx,n}x_{c2,b,n} + \alpha_{c2,bv,n}v_{c2,b,n} + \alpha_{c2,sx,n}x_{c2,s,n} + \alpha_{c2,sv,n}v_{c2,s,n}) \\ &+ \sum_{c3}\sum_{n=t-3}^{t}(\alpha_{c3,bx,n}x_{c3,b,n} + \alpha_{c3,bv,n}v_{c3,b,n} + \alpha_{c3,sx,n}x_{c3,s,n} + \alpha_{c3,sv,n}v_{c3,s,n}). \end{aligned} \tag{3.6}$$

The regression will detect which classes of trades have more influence on the quote change.[12] The subscript $c1$ indicates the trade class A, N, I, M and P when the initiator of the trade is the customer. $C2$ indicates the trade class I(h) and M(h) when the initiator of an I trade or an M trade is the market maker who changes the quote. $C3$ indicates the trade class S(o) and I(o), of which the trade either takes place in SEAQ or in IDB market and of which the trade does not involve the market maker who changes the quote. x is the number of the trades at time n, and v is the trading volumes (in NMS multiples) at time n.[13] The subscript b indicates the x (v) is of buys and s indicates x (v) is of sells. If trades are initiated by customers (class

[12] Alternatively, one may regress the realised spread, defined as the difference between the current price and the mid-touch at the end of the day, on the different classes of trades; see Easley et al. (1997).

[13] Both trades and volumes are not standardised. The regression model depends on a certain degree of homogeneity among stocks and market makers.

$c1$), a quote change during minute t may result from the trading before t, so lag trades at time t, $t-1$, $t-2$ and $t-3$ are included in the model. If the inter-dealer trade is initiated by the market maker who changes the quote (class $c2$), then not only the trades preceding the change may affect the change, but the market maker may also change the quote before initiating trades. Hence trades at $t+1$, $t+2$ and $t+3$ are included as independent variables. If market makers do not involve in the trade (class $c3$), only the lag trades may affect the change of the quote. As market makers do not observe the direction of SEAQ trades which they do not involve, they may not observe the directions of S(o) trades, and the directions are identified by the mid-touch. On the other hand, market makers have access to IDB screen, so they can identify the direction of I(o) trades.

If a customer initiates a buy trade (class $c1$), then the market maker may move up the quote after the trade according to either inventory theory or adverse selection theory, so the coefficients $\alpha_{c1,bx,n}$ and $\alpha_{c1,bv,n}$ are expected to be positive. A sell trade will cause the market maker to move down the quote, so $\alpha_{c1,sx,n}$ and $\alpha_{c1,sv,n}$ are expected to be negative. If a buy trade is initiated by the market maker who changes the quote (class $c2$), then he is likely to move up the quote if there is information effect. Similarly, the information effect will cause him to move down the quote as well as initiating a sell trade. The inventory effect, however, moves the quote to the different direction. Once the market maker initiates a buy trade, his inventory level rises, and the market maker will move down the quote to reduce the inventory. If the market maker initiates a sell trade, the inventory level falls, and the market maker will move up the quote to increase the inventory. If the inventory effect is greater than the information effect, then $\alpha_{c2,bx,n}$ and $\alpha_{c2,bv,n}$ are expected to be negative and $\alpha_{c2,sx,n}$ and $\alpha_{c2,sv,n}$ are to be positive. If the information effect is greater than the inventory effect, the outcomes are expected to be the opposite. Finally, the trades which do not involve the market maker (class $c3$) have no effects on his inventory. If there is information effect of the trades by others, then $\alpha_{c3,bx,n}$ and $\alpha_{c3,bv,n}$ are expected to be positive and $\alpha_{c3,sx,n}$ and $\alpha_{c3,sv,n}$ to be negative. The relationship between trades and quote changes are illustrated in Figure 3.5.

The result of the regression is in Table 3.4. Both the coefficients and their t-statistics are reported.[14] Recall the coefficients of $\alpha_{c1,bx,n}$ and $\alpha_{c1,bv,n}$ are expected to be positive and $\alpha_{c1,sx,n}$ and $\alpha_{c1,sv,n}$ to be negative under inventory models and adverse selection models. Most of the coefficients associated with the numbers of A, N, P and M trades behave as expected. The coefficients of number of trades tend to be more significant than the coefficients of volume of trades except for class A trades. Quite a few coefficients of volume variables are not consistent with information or inventory models. Trades at time $t-2$ and $t-3$ are related to the change in quote at t more closely than the trades at time t or $t-1$. It implies that market makers do not respond to trades instantly; they need time to digest the impacts brought by the trades.[15] The sizes of the coefficients associated with buys and those of the corresponding sells apparently do not exhibit any patterns: the absolute value of $\alpha_{c1,bx,n}(\alpha_{c1,bv,n})$ may be bigger or smaller than that of $\alpha_{c1,sx,n}(\alpha_{c1,sv,n})$. The most significant coefficients are those related to the number of M trades. In contrast, the coefficients associated with I trades are surprising. The coefficients of volume variables, $\alpha_{I,bv,n}$ and $\alpha_{I,sv,n}$, are not significant and with different signs. However, most of the $\alpha_{I,bx,n}$ are significantly positive and all of the $\alpha_{I,sx,n}$ are significantly negative. It implies market makers move down (up) the quotes after another market makers buy (sell) from them in IDB market, which are inconsistent with the theories. As a result, most of the signs of the coefficients of the number of trades are different from those of the volumes of trades. It appears market makers treat the trades in IDB market very differently from the rest of the trades when considering changing quotes.

If a market maker hits the limit order in IDB market (trade class I (h)), then he is likely to move up the quote after a buy or to move down the quote after a sell. The result is consistent with the information theory, or at least the information effect is bigger than the inventory effect. The signs of the

[14] The statistics are adjusted for heteroscedasticity using White's (1980) approach.

[15] The results here is somewhat contradicted to those found in Section 2.6.2, in which the magnitudes of most of the trades coefficients at time t are greater than those at $t-1$.

coefficients of number of trades and those of their corresponding volumes are not always the same, but most of the coefficients of volumes are not significant. The results are similar when the market maker trades with another one in SEAQ (trade class M (h)). Most of the coefficients of volume are not significant. Three significant volume variables are of the sells, two of which support inventory theories. However, all signs of $\alpha_{M(h),bx,n}$ are positive and all signs of $\alpha_{M(h),sx,n}$ are significantly negative, which strongly indicate the information effect is greater than inventory effect. The coefficients of M (h) class are more significant than their counterparts of I (h) class in general. Moreover, while there is evidence that market makers initiate M trades before changing the quotes, the practice appears to be less common when market makers initiate I trades: the coefficients of lead trades are more significant in M trades than those in I trades. If the trade takes places in IDB market and does not involve the market maker himself (trade class I(o)), then the trade hardly affects the decision to change the quote. Most of the coefficients of this class of trades are very small and insignificant. In contrast, all the coefficients associated with the trade class S(o) are significant and consistent with the adverse selection theory. It may imply that market makers regard trades in the market contain information and adjust the quotes accordingly.

The second regression model groups the trades by their sizes:

$$\begin{aligned}
\Delta m_t &= \alpha_0 + \sum_{c1} \sum_{n=t-3}^{t} \left(\alpha_{c1,xb,n} x_{c1,b,n} + \alpha_{c1,xm,n} x_{c1,m,n} + \alpha_{c1,xs,n} x_{c1,s,n}\right) \\
&+ \sum_{c2} \sum_{n=t-3}^{t+3} \left(\alpha_{c2,xb,n} x_{c2,b,n} + \alpha_{c2,xm,n} x_{c2,m,n} + \alpha_{c2,xs,n} x_{c2,s,n}\right) \\
&+ \sum_{c3} \sum_{n=t-3}^{t} \left(\alpha_{c3,xb,n} x_{c3,b,n} + \alpha_{c3,xm,n} x_{c3,m,n} + \alpha_{c3,xs,n} x_{c3,s,n}\right) \\
&+ \alpha_{S(o),xb,t-59/62} x_{S(o),b,t-59/62}. \qquad (3.7)
\end{aligned}$$

The reason of examining this regression model is that different classes of trades have different sizes (See Table 3.1 and Table 3.2), and trade sizes are believed to contain information (Easley and O'Hara 1987). M trades in the model (3.6) appear to affect the quote changes more than the other classes of trades, but M trades are also bigger than another trades. Thus, it is worthwhile to separate size effects and class effects. Trading volumes are not used in (3.7) because the results from the previous model reveal that volume variables are often not statistically significant. $x_{*,*,n}$ in (3.7) is no longer the number of buy or sell trades. Instead, it is the number of buy trades minus the number of sell trades at time n.[16] $c1$, $c2$, and $c3$ are the same classes of trades defined as before, while subscript b, m and s respectively indicate the trades are greater than or equal to six time NMS (big trades), between one and six times NMS (medium trades), and less than one times NMS (small trades). Because the publication of big trades is delayed for one hour, the model includes a new variable $x_{S(o),b,t-59/62}$, which is the difference between buy and sell trades taking place in SEAQ between 59 and 62 minutes before the quote chage.[17] According to the information theory, all α in model (3.7) are expected to be positive. According to inventory theory, $\alpha_{c1,*,n}$ are expected to be positive and $\alpha_{c2,*,n}$ are expected to be negative.

The regression results are shown in Table 3.5. The coefficients of most of the variables are positive. The biggest exceptions are those associated with I trades initiated by counter parties (Class c1.) However, coefficients of I(h) trades are positive, and those of I(o) trades are not significant. Some of the M and P trade variables are occasionally negative, especially for the big trades. One of the explanations is that some of the trade directions in the two classes may be incorrect. For the M trades, the directions are identified by the mid-touch. If the mid-touch does not always reflect the true value of the stock, then the directions of M trades may not always be correct. For the P trades, market makers occasionally look for another member firms for liquidity, and the trade direction in such a situation should be the direction of market

[16] Although the sample used in this chapter is different from the previous one, the fact that sell trades are more than buy trades remains the same. However, the magnitudes of $\alpha_{*,bx,n}$ and $\alpha_{*.bv,n}$ are not bigger than those of $\alpha_{*,bx,n}$ and $\alpha_{*.bv,n}$, so the use of net trades is justified. Like the sample in Chapter 2, downward changes of quote are more than upward changes among FTSE-100 stocks.

[17] Those big trades involving market makers themselves are known when the trades are executed, and market makers observe all the trades in the IDB market, so only the big trades in SEAQ of c3 class may not be known by the market maker who changes the quote immediately after the execution of the trade.

makers, not of principals. If this conjecture is correct, then market makers reverse the quotes after P trades are consistent with inventory theories.

The effects of medium trades appear to be bigger than both small trades and big trades in most of the trade classes. The effects of big trades are the strongest of trade class A. Table 3.1 shows that big A trades are rare. It implies that market makers regard big A trades as unusual and change the quote accordingly. Similarly, the average size of M trades is the biggest, small M trades are unusual, and market makers respond to small M trades more than the other classes of small trades. Finally, $\alpha_{S(o),xb,t-59/62}$ is very small compared with the other $\alpha_{S(o),xb,n}$. It may imply the delay for trade publication has little impact on the quote change. Even though big trades are formally published one hour after the trades are executed, the market makers who do not executed the trades learn the information from somewhere else and change the quotes soon after the trades take place.

Finally, the Condition Index of either (3.6) or (3.7) does not show any sign of colinearity. On the other hand, the low Durbin-Watson statistics indicate the residuals may be autocorrelated. This is consistent with the result of Chapter 2, which shows that the preceding change two minutes ago still explain the current change.

3.7 Discussions and conclusions

This chapter compares the costs of executing different categories of trades. The small investors, who have to use brokers to trade with market makers, appear to pay more than the private clients of the market makers. Member firms of the Exchange who act as principals have favourable price, and they may act as liquidity providers of the market makers. The trades in IDB market are remarkably cheaper than the trades in SEAQ, while IMM trades are remarkably expensive. The regression analysis shows the theories in the existing literature successfully explain part of the effective spreads, but they cannot explain all of the cost difference among different classes of trades. The regression analysis of quote changes reveals that market makers change quotes after customer trades, which imply those trades have certain inventory or information implication. The hitters in IMM trades tend to adjust the quotes before the trades, and their counter parties tend to adjust the quotes after the trades. The hitters in IDB trades do not adjust the quotes as enthusiastic as they do in IMM trades, and the posters of IDB trades change their quotes to the direction opposite to what is predicted by the literature.

My interpretation of the evidence is as follows. The weak evidence of the effect of inventory on the costs of trades in Section 3.5 suggests inventory consideration may not play an important role in determining the costs of trades. According to Section 3.6, the initiator of inter-dealer trading appears to consider the information effect more than the inventory effect. Moreover, most of the volume variables do not perform as good as the variables of number of trades in model (3.6). Although some of the results in the regression models do not separate the inventory effect from the information effect completely, there is evidence supporting information effect, such as the strong relationship between S (o) trades and quote changes in Section 3.6, whereas inventory effect is either small or over-shadowed by information effect. Therefore, most of the cost difference is likely to be attributed to information costs, order processing costs and the monopolistic power of market makers. The support for the implicit collusion of market makers does not exist. The volume spline in Figure 3.4 does not exhibit an inverse "V" shape as predicted by Dutta and Madhavan (1997) for implicit collusion. Furthermore, according to Model (3) in Section 3.5, M trades are the most expensive trades all possible considerations discussed in the literature have been taken into account. Therefore, market makers do not give one another favourable prices in SEAQ.

Based on the evidence that inventory consideration does not affect the effective spreads in a substantial way, the close relationship between quote change and trades is largely attributed the information brought by the trades. Market makers regard the customer trades, including trades with brokers and with clients, as valuable information, and they move the quotes accordingly. Market makers are also very sensitive to M trades. Table 3.4 suggests the effects of M trades on the quote change are more pronounced than the other classes of trades. If market makers cannot wait for trading with customers and rush into M trades, it indeed signals something unusual to their counter parties. Moreover, the initiators of M trades adjust their quotes to the same direction of the trades even before the trades take place, which strongly indicates the information effect outweighs the inventory consideration.

The significance of M trades from regression models (3.6) and (3.7) is consistent with the information hypothesis that M trades have more information contents than the rest of the trades. However, the regressions still cannot explain why A trades are more expensive than N trades, nor why N trades are more expensive than P trades. How to solve this enigma? Here is a possible explanation. According to Table 3.1, the sizes of the trades are small compared with the trades with other market participants. Compared with other market participants, investors who use the brokers are the least frequent traders in the market. The aggregate volumes of broker trades may be substantial, but the volumes of a typical investor are small. As a result, a small investor may not be able to make the broker negotiate very hard to get a good deal from the market maker. Evidence from Chapter 2 even suggests that most of the small trades are preferenced. Brokers are loyal to their clients to the extent in which they can obtain the price as good as the best bid or ask quote in the market, but their willingness to help the clients beyond "best execution" is somewhat limited. In contrast, a private client of market makers negotiates the trades with them herself, she is a frequent participant in the market, so she can get better prices from the market makers.

Is it a practice of price discrimination? Price discrimination is the practice of which the supplier of the market charges the price differently according to the demand of the consumers. However, the cost difference between the brokers and the private clients of market makers mainly reflect the different bargaining power of the consumers, not their demands. On the other hand, the huge costs of M trades may suggest the practice of price discrimination. As Board and Sutcliffe (1995) and Reiss and Werner (1996, 1997) have argued, market makers use M trades only as the last resort. Market makers commit more capital to the market than any other market participants do. If a market maker needs to buy or sell stocks with other market makers for any reasons, it may imply that he cannot depend on brokers or private clients to satisfy his need to trade. Given the low costs of IDB trades, the initiator of M trades clearly reveals his urgent demand for shares. With IDB trading unavailable, the counter party of an M trade realises it is very difficult for the initiator to get a deal if it is necessary to execute the trade immediately. The counter party can therefore charge the initiator a high cost.

A few explanations may account for the low costs of P trades. P trades are trades of market makers with member firms of the Exchange, who are also frequent participants of the market, and their bargaining power may be as strong as private clients of market makers. Second, occasionally market makers rather than principals may initiate P trades. 11% of P trades occur below the mid-quotes, which is a proportion bigger than any classes of trades. It suggests that sometimes market makers may resort to non-market-maker members for liquidity. Third, since member firms can act as principal and agent, market makers are willing to reduce the costs of P trades in order to attract more A trades. It is interesting to examine the relationship between the costs of P trades and the level of order preferencing in future studies.

IDB market appears to succeed in providing market makers a platform to trade with one another at low costs. The success, however, partly rest on the inability of IDB market to provide liquidity timely. Reiss and Werner (1997) have mentioned the market makers observed the IDB best orders are not always available on the screens. Table 3.1 shows the total number of IDB trades is 54,318. Dividing it by 39 stocks and 125 trading days, the average number of users of IDB trades for each stock every day is about 11, including both posters and hitters. In other words, the four IDBs altogether broke less than six deals per stock per day, which does not seem to be a remarkable achievement. Market makers place limit orders to IDBs without knowing when the trades will be executed. Therefore, they place limit orders only when they can afford waiting for unwinding their inventories. The posting in

IDB market clearly signals the trades only have inventory implication; there is no information content of IDB trades. Market makers generally subscribe to the view. Table 3.4 show that market makers may adjust their quotes when somebody else trades in SEAQ. In contrast, the market makers hardly take any notice to the trades in IDB markets. The hitters of I trades do move their quotes, but the relationship between quote change and I trades is not as strong as M trades, which is consistent with the conjecture that information content of I trades are not as big as M trades. On the other hand, the IDB posters are aware the IDB hitter may possess information, and in response they do not post big orders to IDB market. Table 3.1 shows the average size of IDB trades is less than 60% of the IMM trades. The effect of adverse selection makes IDB market less liquid than SEAQ, and it makes the costs of I trades smaller.

The investigation so far has generated some policy implications:

1. *Is there any device to improve the costs of broker trades in SEAQ?* The cost difference between A trades and N trades is due to the strong bargaining power of institutional investors. Unless the individuals are able to improve their bargaining power, cost difference remains between the two classes of trades.

2. *Is there a need for a central limit order book?* A central limit order book open to all investors is likely to improve the costs of individual investors.[18] The price of the market will be improved by the will of individual investors to place limit orders, and the execution costs of small investors are likely to be improved. However, an order-driven system that is open to all of the participants to the markets may be more liquid than the current IDB market. It will induce more information trades to hit the limit orders than what is currently observed in IDB markets. The information component of the effective spreads in the limit order book must be bigger than the current IDB market.

[18] SETS, the electronic order book for FTSE-100 stocks has been operated since October 1997, but it does not accept small orders.

3. *How about the demands for big trades?* The relatively small average size of IDB trades indicates big trades are not likely to take place in an order-driven market. If the adverse selection effect is more severe in the central limit order book than in IDB market, one would expect the size of the order will be even less. The fourth panel of Table 3.2 indicates the market participants, especially institutional investors, have strong demands for trades with size of six times NMS or above. Big limit orders are not many in IDB market, and they may not appear very often in an order-driven system. Moreover, the Exchange allows the trades with six times NMS or above to be published one hour after the trades are executed. Before big trades are published, market makers are able to use inter-dealer trading or other means to offset the effect of the trades. If the big trades are required to enter the limit order book, those trades may not take place in the first place. The customers will therefore split the orders to enter the limit order book. Because of the adverse selection effect, the sizes of those orders must be smaller than what are observed in IDB market. As a result, customers may have to wait much longer to finish the trades. From the market's point of view, the splits in orders implies the price formation process may be no sooner than is currently observed despite the delay for publication. A central limit book is not a satisfactory mechanism to facilitate big trades; another device is needed to assist the order-driven system.

To sum up, the quote-driven trading system in London has largely satisfied the needs for the frequent market participants to trade with one another, probably at the expense of infrequent investors. The infrequent investors pay more for the trading because they have less bargaining power against the big players in the market. The low costs of IDB market may not be the product of implicit collusion but of the infrequent trading, which makes trades in IDB market have little information content. A change from a dealership market to an order-driven system is likely to benefit the infrequent investors but not big players, it may not reduce the cost of trading to what is observed in IDB market, and it may also delay the price formation process.

Table 3.1: Summary statistics of sample stocks

The data includes the trades of thirty-nine stocks during the first six months of 1996. Trades without the involvement of market makers are classified as "X" trades. Trades which involve market makers take place outside the mandatory quote period (MQP) are classified as "T" trades. The rest of trades are further classified into five categories: "A" trades are trades between market makers and brokers, "N" trades are trades between market makers and their private clients, "I" trades are trades between market makers and inter-dealer brokers, "M" trades are trades between market makers themselves, and "P" trades are trades between market makers and non-market-maker member firms. Volumes and the number of "M" trades are doubled to be comparable with "I" trades. Average trade sizes are computed by dividing the volumes by the numbers of trades.

trade categories: market maker trade with	trade class	number of trades	volumes in £1,000	average trade size in £
brokers	A	594,684	13,642,700	22,943
private clients	N	60,045	20,325,343	338,473
inter-dealer brokers	I	54,822	14,774,112	269,452
market makers	M	30,656	14,178,948	462,404
principals	P	7,643	1,657,000	216,755
outside MQP	T	25,934	2,028,409	78,202
other trades	X	6,969	1,375,887	197,405
total		780,753	67,982,399	
of which				
FTSE-100 stocks		680,526	59,164,894	86,933
non-FTSE-100 stocks		100,227	8,817,505	87,954

Table 3.2: Cost of trading

Cost of trading are computed according to equation (3.2). Average half touch is one half of the difference between best bid and ask prices divided by the mid-touch. The unit of costs and half touch is a hundredth percent. The costs of I trades are costs of posters. The costs of IDB hitters are 11.18 basis points for the FTSE-100 stocks, 23.55 for non-FTSE stocks, and 13.25 on average. The difference is the commission fees of IDBs.

	trade category				
	A	N	I	M	P
average costs	25.92	20.02	8.82	22.29	15.22
of which					
FTSE-100 stocks	23.51	17.46	6.83	18.41	11.21
non-FTSE-100 stocks	44.59	32.82	18.82	38.94	28.42
breakdown of costs by NMS					
0 - 0.1	25.78	20.81	0.00	22.78	18.30
0.1 - 1	27.75	20.09	8.02	25.16	15.31
1 - 6	24.31	19.72	9.36	22.06	13.38
6 - 75	18.98	17.79	11.13	21.32	12.09
75 and above	22.91	24.46	3.75	39.37	26.67
Percentages of trades by NMS					
0 - 0.1	87.15	35.24	0.00	0.76	19.63
0.1 - 1	10.22	32.77	47.72	8.85	47.51
1 - 6	2.08	20.64	46.81	82.25	28.81
6 - 75	0.55	11.16	5.46	8.09	3.98
75 and above	0.01	0.18	0.01	0.05	0.08
half average touch					
FTSE-100 stocks	26.76	27.22	27.30	23.27	25.63
non-FTSE-100 stocks	56.71	55.73	57.33	46.82	58.55

Table 3.3: Result of regression models of costs

The dependent variable of the regression models is the cost of trading defined in (3.2). Model (3) uses only trade-class dummies as independent variables. Model (2) uses variables suggested in the literature. Model (1) uses all the variables in (2) and (3). An asterisk (*) indicates the t-statistic is significant at 0.05 level, two asterisks (**) indicate the statistic is significant at 0.01 level, and three asterisks (***) indicate at 0.0001 level.

Independent	Model (1)		Model (2)		Model (3)	
variables	coefficient	t-value	coefficient	t-value	coefficient	t-value
Intercept	3.5944	16.54***	3.6261	16.73***	25.8635	1159.70***
d_I	-14.6484	-119.27***			-17.0841	-178.27***
d_M	5.0318	34.00***			-2.5356	-17.09***
d_N	-3.6123	-36.42***			-5.9487	-56.29***
d_P	-10.1137	-36.72***			-10.6923	-38.10***
FTSE	0.2294	2.20*	-0.1910	-1.81		
MM	-0.0007	-0.08	0.0065	0.65		
sp_t	0.7128	261.17***	0.7082	261.62***		
sd_t	0.2116	11.12***	0.2000	10.42***		
d_b	-0.7098	-39.28***	-0.8702	-46.74***		
$d_b v_{d,t}$	0.0001	2.90**	0.0001	3.12**		
T1	0.7359	11.61***	0.7212	11.11***		
T2	0.9261	15.98***	0.6620	11.12***		
Q1	-3.6993	-39.32***	-6.1189	-85.66***		
Q2	-9.7694	-54.90***	-14.4084	-96.38***		
Q3	-12.9615	-4.20***	-17.6725	-5.82***		
Q4	39.2747	0.66	56.2279	1.00		
V1	-0.0119	-0.18	-0.0472	-0.69		
V2	0.9013	9.85***	0.9538	10.20***		
V3	1.3088	2.58**	1.8081	3.49**		
R^2	0.4942		0.4674		0.0414	
CI	26.05		25.65		1.52	
DW	1.84		1.83		1.02	

Wald tests of	Model (1) χ^2 value	Model (2) χ^2 value	Model (3) χ^2 value	Degrees of Freedom
Significance of size spline	3321.07***	13821.05***		4
Significance of time spline	277.02***	168.14***		2
Significance of volume spline	125.34***	151.69***		3
$d_I < d_P$	247.24***		469.79***	1
$d_P < d_N$	510.19***		252.96***	1
$d_N < d_M$	171.61***		361.91***	1

Tests of jointly significance of variables	Sum of Square Errors	Degrees of Freedom
Model (1)	120461778	701969
Model (2)	126850996	701973
Model (3)	228313582	701984

F-values of classs dummies = 41889.11

F-values of all other variables = 9308.00

Table 3.4: Result of the analysis of quote changes – trades and volumes
This table shows the result of regression model of equation (3.6):

$$\begin{aligned}\Delta m_t &= \alpha_0 + \sum_{c1}\sum_{n=t-3}^{t}(\alpha_{c1,bx,n}x_{c1,b,n} + \alpha_{c1,bv,n}v_{c1,b,n} + \alpha_{c1,sx,n}x_{c1,s,n} + \alpha_{c1,sv,n}v_{c1,s,n}) \\ &+ \sum_{c2}\sum_{n=t-3}^{t+3}(\alpha_{c2,bx,n}x_{c2,b,n} + \alpha_{c2,bv,n}v_{c2,b,n} + \alpha_{c2,sx,n}x_{c2,s,n} + \alpha_{c2,sv,n}v_{c2,s,n}) \\ &+ \sum_{c3}\sum_{n=t-3}^{t}(\alpha_{c3,bx,n}x_{c3,b,n} + \alpha_{c3,bv,n}v_{c3,b,n} + \alpha_{c3,sx,n}x_{c3,s,n} + \alpha_{c3,sv,n}v_{c3,s,n}).\end{aligned}$$

Δm_t is the quote change defined in (3.5). The independent variables are number of trades (x) and volumes (v) around the change in quotes. Subscript $c1$ are of the trade classes initiated by customers, $c2$ are of the classes initiated by the market maker, and c3 indicates the trades without the involvement of the market maker who changes the quote. Subscript b (s) indicates buys (sells). The coefficients and their t-statistics are presented. An asterisk (*) indicates the t-statistic is significant at 0.05 level, two asterisks (**) indicate the statistic is significant at 0.01 level, and three asterisks (***) indicate at 0.0001 level.

		buy trades		buy volumes		sell trades		sell volumes	
		$\alpha_{c,bx,n}$	t-value	$\alpha_{c,bv,n}$	t-value	$\alpha_{c,sx,n}$	t-value	$\alpha_{c,sv,n}$	t-value
class $c1$									
A	t	0.0085	0.97	0.0210	4.51***	0.0065	0.99	-0.0311	-3.68**
	t-1	0.0063	0.71	0.0147	3.30**	0.0098	1.04	-0.0137	-3.55**
	t-2	0.0281	3.17**	0.0262	3.59**	-0.0382	-6.16***	-0.0521	-4.95***
	t-3	0.0345	3.42**	0.0480	3.90***	-0.0175	-2.76**	-0.0235	-3.16**
N	t	0.0533	1.83	-0.0041	-0.85	-0.1087	-4.62***	-0.0076	-1.89
	t-1	0.0792	3.81**	0.0052	1.70	-0.0696	-3.14**	-0.0084	-1.40
	t-2	0.1369	5.67***	0.0023	0.47	-0.1746	-8.38***	-0.0046	-1.11
	t-3	0.1407	5.86***	0.0067	1.52	-0.1282	-6.09***	-0.0008	-0.35
P	t	0.0194	0.35	-0.0235	-8.57***	0.0668	0.63	0.0271	1.30
	t-1	0.0744	1.35	0.0075	0.24	-0.2036	-4.52***	0.0128	1.15
	t-2	0.1995	5.08***	0.0504	3.10**	-0.2854	-6.10***	0.0573	1.50
	t-3	0.2115	4.01***	-0.0005	-0.03	-0.1724	-2.83**	0.0234	0.32
I	t	0.0991	0.84	-0.1771	-1.60	0.2639	5.13***	-0.0402	-1.36
	t-1	-0.4948	-4.62***	0.1437	1.39	0.2167	6.13***	0.0070	0.28
	t-2	-0.2046	-6.63***	-0.0791	-2.54*	0.3035	13.54***	0.0127	0.43
	t-3	-0.3233	-9.04***	-0.0185	-2.03*	0.2860	14.33***	-0.0201	-1.51
M	t	0.1197	1.93	-0.0632	-3.30**	-0.0943	-1.55	0.0264	2.11*
	t-1	0.2630	5.85***	-0.0496	-5.87***	-0.1898	-4.62***	0.0365	2.40*
	t-2	0.3823	22.56***	0.0038	0.74	-0.3566	-20.22***	0.0298	3.91***
	t-3	0.4109	27.51***	0.0011	0.34	-0.3989	-24.50***	0.0217	2.98**

(continued to the next page)

(continued from the previous page)

		buy trades		buy volumes		sell trades		sell volumes	
		$\alpha_{c,bx,n}$	t-value	$\alpha_{c,bv,n}$	t-value	$\alpha_{c,sx,n}$	t-value	$\alpha_{c,sv,n}$	t-value
	class $c2$								
I	t+3	0.1697	3.42**	-0.0007	-0.05	0.0367	0.74	-0.0318	-1.93
(h)	t+2	0.0878	2.28*	0.0044	0.33	-0.1933	-4.76***	0.0297	1.75
	t+1	0.1893	3.40**	0.0075	0.39	-0.0497	-1.64	-0.0078	-3.45**
	t	0.1562	4.26***	-0.0112	-0.92	-0.2358	-6.04***	0.0211	1.39
	t-1	0.1243	3.18**	0.0080	0.49	-0.0654	-1.46	0.0091	0.59
	t-2	0.1939	10.20***	0.0015	0.28	-0.2047	-7.64***	0.0273	2.02*
	t-3	0.1965	11.54***	0.0081	3.79**	-0.1913	-9.23***	0.0063	1.02
M	t+3	0.1019	2.10*	0.0194	1.09	-0.1080	-3.26**	-0.0249	-2.49*
(h)	t+2	0.2531	6.65***	-0.0186	-0.91	-0.0989	-2.77**	-0.0429	-2.58**
	t+1	0.1005	2.88**	0.0155	0.66	-0.1735	-3.56**	-0.0122	-1.13
	t	0.2527	7.27***	-0.0410	-2.22*	-0.1191	-1.63	-0.1377	-1.63
	t-1	0.2992	6.12***	0.0175	0.99	-0.3456	-6.09***	0.0734	1.60
	t-2	0.2662	10.54***	-0.0294	-1.88	-0.3076	-10.65***	-0.0142	-0.90
	t-3	0.2576	10.51***	-0.0088	-0.58	-0.2768	-15.06***	-0.0253	-2.93**
	class $c3$								
I	t	0.0180	1.76	-0.0022	-0.32	0.0220	2.03*	-0.0030	-0.41
(o)	t-1	0.0177	1.83	0.0054	0.63	0.0202	1.94	0.0046	0.52
	t-2	-0.0162	-1.87	0.0221	2.23*	-0.0125	-1.29	0.0217	2.05*
	t-3	0.0030	0.48	0.0108	2.33*	0.0090	1.43	0.0092	1.97*
S	t	0.0206	11.68***	0.0094	9.53***	-0.0086	-5.04***	-0.0136	-11.67***
(o)	t-1	0.0199	11.84***	0.0082	10.09***	-0.0176	-11.25***	-0.0118	-13.62***
	t-2	0.0294	16.84***	0.0128	19.40***	-0.0202	-14.23***	-0.0158	-19.69***
	t-3	0.0509	23.63***	0.0131	14.46***	-0.0369	-23.51***	-0.0099	-5.60***
R^2			0.1298						
Condition Index			6.02						
Durbin Watson			0.44						

Table 3.5: Result of the analysis of quote changes – by the size of trades

This table shows the result of regression model of equation (3.7):

$$\begin{aligned}
\Delta m_t &= \alpha_0 + \sum_{c1}\sum_{n=t-3}^{t}(\alpha_{c1,xb,n}x_{c1,b,n} + \alpha_{c1,xm,n}x_{c1,m,n} + \alpha_{c1,xs,n}x_{c1,s,n}) \\
&+ \sum_{c2}\sum_{n=t-3}^{t+3}(\alpha_{c2,xb,n}x_{c2,b,n} + \alpha_{c2,xm,n}x_{c2,m,n} + \alpha_{c2,xs,n}x_{c2,s,n}) \\
&+ \sum_{c3}\sum_{n=t-3}^{t}(\alpha_{c3,xb,n}x_{c3,b,n} + \alpha_{c3,xm,n}x_{c3,m,n} + \alpha_{c3,xs,n}x_{c3,s,n}) \\
&+ \alpha_{S(o),bx,t-59/62}x_{S(o),b,t-59/62}.
\end{aligned}$$

Δm_t is the quote change defined in (3.5). The independent variables are number of big (b), medium (m) and small (s) trades around the change in quotes. Subscript $c1$ are of the trade classes initiated by customers, $c2$ are of the classes initiated by the market maker, and $c3$ indicates the trades which do not involve the market maker who changes the quote. An asterisk (*) indicates the t-statistic is significant at 0.05 level, two asterisks (**) indicate the statistic is significant at 0.01 level, and three asterisks (***) indicate at 0.0001 level.

		big trades		medium trades		small trades	
		$\alpha_{c,xb,n}$	t-value	$\alpha_{c,xm,n}$	t-value	$\alpha_{c,xs,n}$	t-value
class $c1$							
A	t	0.2431	5.75***	0.0402	0.97	0.0049	0.94
	t-1	0.1478	2.64**	0.0870	2.69**	0.0008	0.10
	t-2	0.3035	6.34***	0.3258	13.16***	0.0259	5.29***
	t-3	0.4008	4.22***	0.2826	8.17***	0.0194	3.73**
N	t	0.1077	2.33*	0.1215	3.56**	0.0431	2.00*
	t-1	0.1966	4.67***	0.1701	6.38***	0.0031	0.19
	t-2	0.1524	4.15***	0.2725	12.37***	0.0506	2.92**
	t-3	0.1959	4.48***	0.1670	7.40***	0.0831	4.26***
P	t	-0.2074	-7.96***	-0.0142	-0.22	-0.0797	-1.14
	t-1	-0.1653	-2.83**	0.1741	2.98**	0.0675	1.91
	t-2	0.1883	2.45*	0.2254	4.03***	0.2045	6.15***
	t-3	0.2046	1.14	0.2304	4.69***	0.1885	4.32***
I	t	-1.5451	-1.28	-0.1474	-5.32***	-0.1212	-2.25*
	t-1	0.7809	0.89	-0.2832	-5.85***	-0.2450	-8.59***
	t-2	-0.7952	-2.60**	-0.3072	-7.31***	-0.2730	-15.88***
	t-3	-0.3246	-2.45*	-0.2702	-9.17***	-0.2990	-13.23***
M	t	-0.0916	-1.61	-0.0476	-1.53	0.1895	2.40*
	t-1	-0.1622	-2.01*	0.0735	3.12**	0.1827	3.42**
	t-2	0.2950	4.21***	0.2636	25.92***	0.3415	11.80***
	t-3	0.3567	5.56***	0.2780	29.30***	0.4297	8.49***

(continued to the next page)

(continued from the previous page)

		big trades		medium trades		small trades	
		$\alpha_{c,,n}$	t-value	$\alpha_{c,,n}$	t-value	$\alpha_{c,,n}$	t-value
	class $c2$						
I	t+3	0.3389	3.70**	0.0996	3.72**	-0.0644	-1.08
(h)	t+2	0.1298	1.86	0.1072	5.94***	0.1706	3.08**
	t+1	0.2661	2.29*	0.1253	5.63***	0.0102	0.17
	t	0.0316	0.32	0.1559	8.51***	0.2587	5.56***
	t-1	0.1444	1.29	0.0757	3.61**	0.1851	3.89***
	t-2	0.1101	1.47	0.1782	15.86***	0.2166	5.42***
	t-3	0.2424	3.87**	0.1942	17.26***	0.1733	3.47**
M	t+3	0.1909	2.54*	0.1961	6.82***	0.0234	0.65
(h)	t+2	0.1789	1.25	0.2065	6.96***	0.1580	5.38***
	t+1	0.4085	4.99***	0.1326	3.11**	0.1110	2.80**
	t	0.9221	1.55	0.2325	6.82***	0.1813	7.12***
	t-1	0.0286	0.10	0.3162	4.94***	0.2777	8.06***
	t-2	0.2488	2.25*	0.2693	8.41***	0.2827	16.93***
	t-3	0.5324	4.85***	0.2122	10.99***	0.2897	17.96***
	class $c3$						
I	t	-0.0020	-0.03	-0.0013	-0.20	-0.0026	-0.34
(o)	t-1	0.0275	0.27	-0.0062	-0.85	0.0046	0.63
	t-2	0.0135	0.11	-0.0043	-0.65	0.0016	0.33
	t-3	0.0229	0.46	-0.0052	-1.13	-0.0023	-0.58
S	t	0.0957	13.94***	0.1175	34.95***	0.0084	6.62***
(o)	t-1	0.0962	13.87***	0.1143	36.34***	0.0122	9.87***
	t-2	0.1188	20.22***	0.1411	59.82***	0.0099	9.57***
	t-3	0.1106	18.22***	0.1620	74.55***	0.0199	14.74***
	t-59/62	0.0130	3.15**				
R^2			0.1639				
Condition Index			1.86				
Durbin Watson			0.45				

Figure 3.1: Participants in the London Stock Exchange

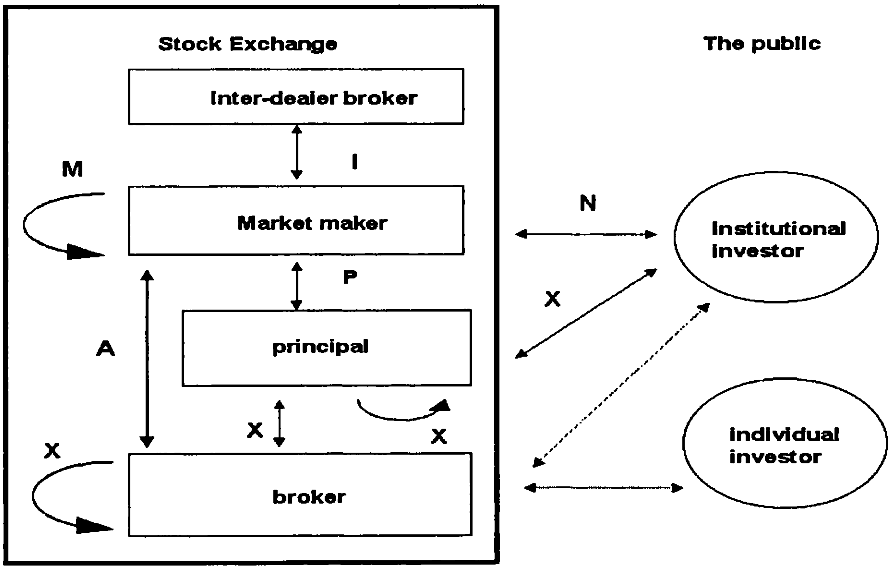


Figure 3.2: Measuring cost of trading

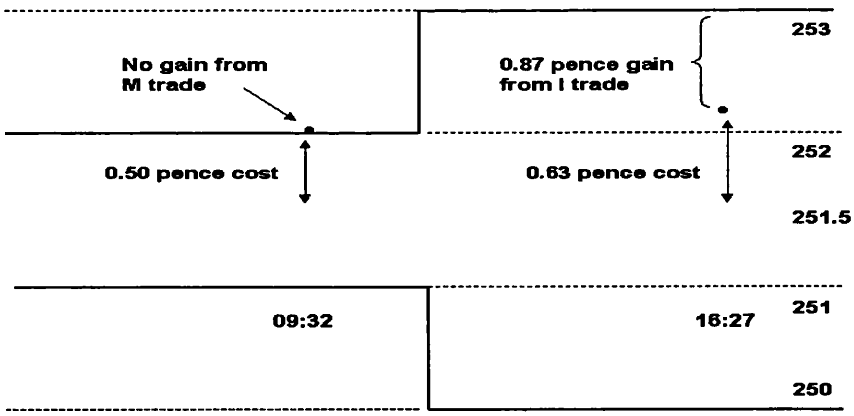


Figure 3.3: How inventory may affect trading costs

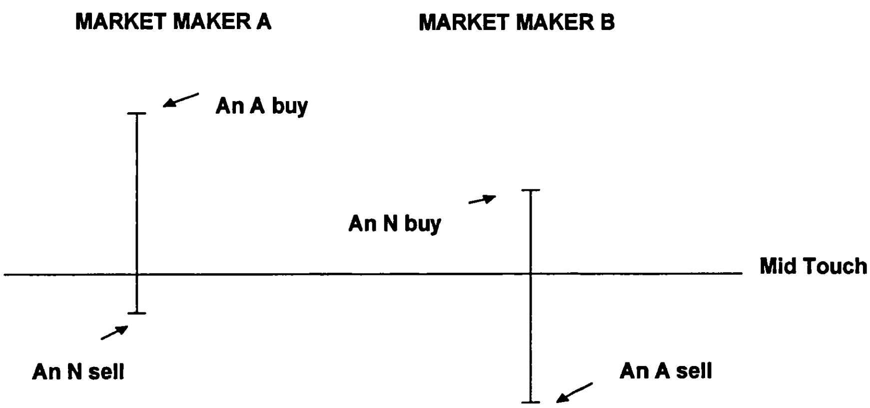


Figure 3.4: Effects of time, size and volume on the half effective spread

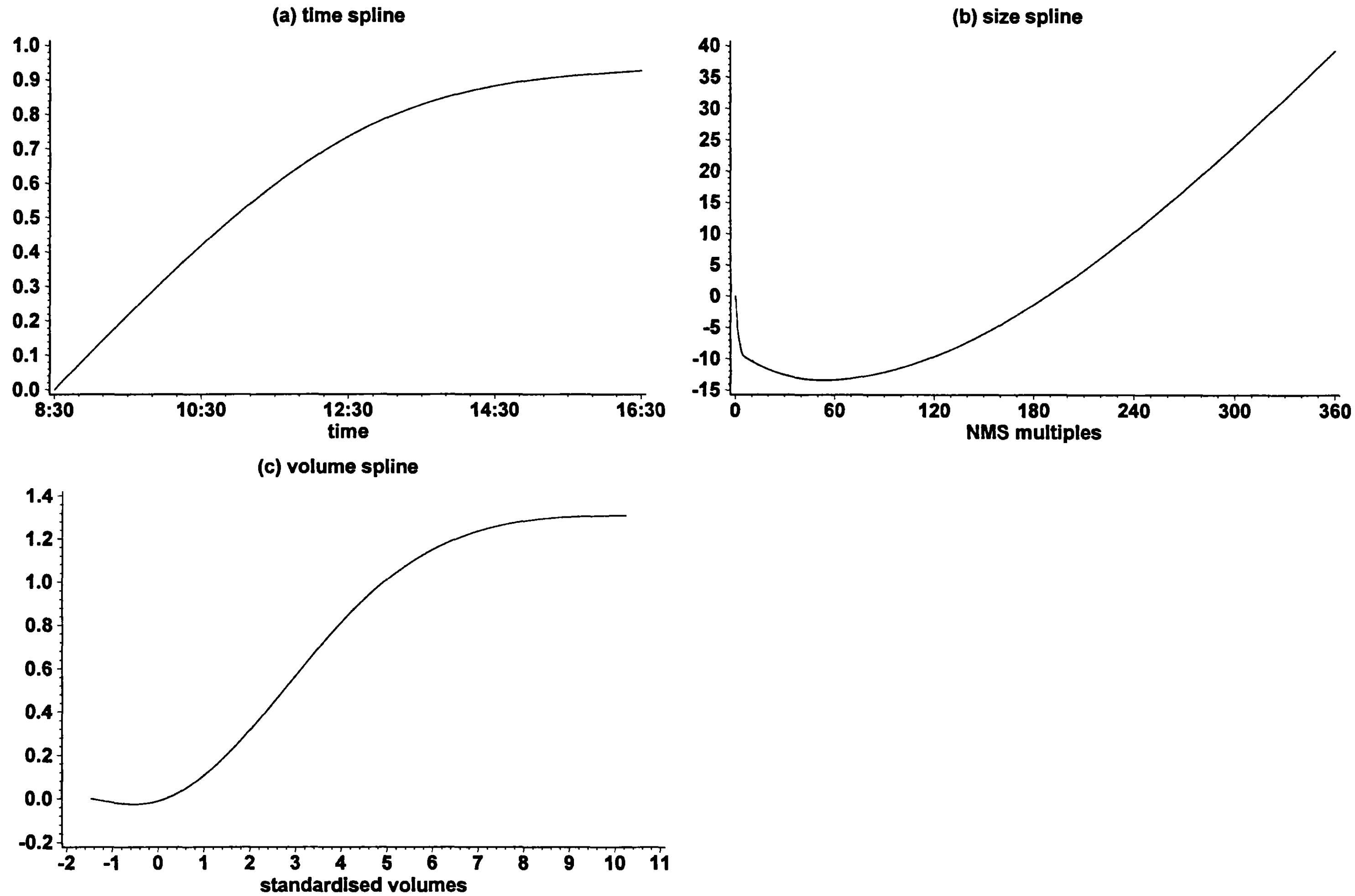
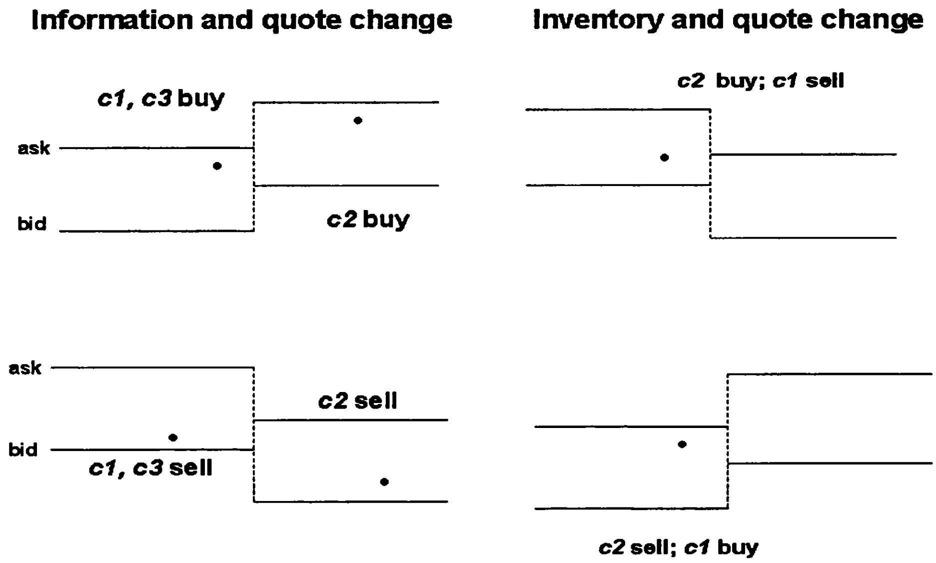


Figure 3.5: Quote change in response to inventory and information



Chapter 4

MEASURING BID-ASK SPREADS IN MULTIPLE DEALERSHIP MARKETS

4.1 Introduction

The price of securities in financial markets consists of two main components: the fundamental price and the spread. Both components cannot be observed. The standard approach of estimating the spread is based on the autocovariance structure of the differenced price data or on specific regression techniques. It follows that the prices must be ordered sequentially over time: at each time period only one price is observed.

There are two notions of data sequentiality: one refers to the market and the other refers to the available data set. It can be argued that securities traded at markets with a single market maker, such as the specialist on the New York Stock Exchange (NYSE), and securities traded at markets in which liquidity is provided by the public limit order book, such as Paris Bourse and Tokyo Stock Exchange, have a price for each time period. Data collected from such markets usually maintain the property of sequentiality.

For multiple dealership markets, such as Chicago Mercantile Exchange and NASDAQ, several dealers negotiate and complete multiple trades simultaneously. Therefore, different prices of the same security float within the market at the same time. The property of price sequentiality may still hold for data sets collected from such markets because of the method of the data manipulation. For example, the average of traded prices within a period may be considered. Data sequentiality remains in this approach but it may lead to a serious loss of information. Different prices are associated with different quantities that have a considerable effect on the spread. Moreover, such a practical solution is theoretically not satisfactory and it is desirable to employ a model-based approach.

The transaction data used in this thesis is collected from the London Stock Exchange (LSE). Most of the securities traded on the LSE have multiple market makers, and several trades of the same security occur simultaneously on a regular basis. It is not straightforward to identify the fundamental price in such markets due to the market itself and due to the lack of sequentiality of the data. First, the London market is not transparent. Trades are negotiated between several dealers on the phone, and delays in the publication of new transactions take place regularly. Participants in the market do not necessarily observe most of the trading in the market. Therefore, even if the completions of some trades may be a few seconds apart, those trades are essentially executed simultaneously. Second, data collected from the LSE is intrinsically non-sequential. There is no unique price within the time interval of one minute. The standard approach to estimate the spread using differenced price is clearly not applicable to the data. This has been the motivation to develop a new method to model bid-ask spreads for non-sequential trade markets.

The rest of this chapter is organised as follows. Section 4.2 gives an overview of the contributions in the literature on the empirical features of bid-ask spreads. The statistical model for bid-ask spreads is presented in Section 4.3. The section also includes discussions of the technical features of the model and some possible extensions. Section 4.4 presents an application of the model by using the transaction data of three stocks traded on the LSE: Glaxo Wellcome, British Telecommunications and Shell Transport and Trading. It is shown that the proposed model is capable of identifying the fundamental price and the spread in a straightforward manner. Section 4.5 compares the proposed model with alternative approaches, especially the use of mid-touch as the fundamental price. Section 4.6 concludes this chapter.

4.2 Bid-ask spreads

4.2.1 Estimation using differenced prices

The fundamental price and the spread of a security are unobserved components of the trade price. By making appropriate assumptions about the dynamic behaviour of the fundamental price and the spread, they can be estimated from the observed price. There are many ways to obtain the spread, but most of them require some sequential ordering of the data. Huang and Stoll (1997) give a detailed overview of two main approaches. The first approach is initiated by Roll (1984) and followed by, among others, Stoll (1989) and George, Kaul, and Nimalendran (1991). They estimate the spread by the autocovariance structure of the price and the quote differences. The second approach taken by Glosten and Harris (1988) and Madhavan and Smidt (1991) is based on regression techniques using trade indicators. This approach requires the use of differenced prices, quotes and trade-direction indicators, as dependent or independent variables.

Most data in the studies of bid-ask spread are ordered sequentially: daily data is ordered by day while intra-daily data is usually ordered by unique time stamps. Examples of studies based on daily data are those by Roll (1984) and by Affleck-Graves et al. (1994). Studies of foreign exchange markets often use quote records (Bollerslev and Domowitz 1993; Bollerslev and Melvin 1994; Bollerslev, Domowitz, and Wang 1997), which are ordered sequentially, so the standard approach may be adopted. Note that trade records are rarely available for these markets. Studies of futures or options markets analyse trades which are recorded sequentially by the clerks in the exchange; see Chung (1991) and Locke and Venkatesh (1997). It should be stressed here that the observed data sequence may not necessarily be identical to the sequence of trading. For example, Hasbrouck et al. (1993) report that members of the NYSE are responsible to enter at least 90% of the trades to the Consolidated Tape System within 90 seconds of execution, so it is likely that the reporting sequence is not consistent with the trade sequence.

4.2.2 Non-sequential data

In a multiple dealership market, one may expect dealers to negotiate the trades simultaneously. Multiple trades of a liquid security may be executed nearly at the same time as a result. The price series is not sequential and estimation based on differenced data is not possible. One notable and important example is the trade records retrieved from the LSE's *Transaction Data Service.* Figure 4.1 shows a simplified data format of transaction records of British Telecom of 1 April 1996. There were six transactions at 11:09 and ten at 11:10 with different prices and quantities bought and sold by different market makers. If the approach by, for example, Huang and Stoll (1997) is used, one has to decide which trades at 11:09 and at 11:10 should be used: big trade or small trade? Buy or sell? By which market maker? At what prices? Any decision to manipulate the sixteen trades does not only lose but also distort the information.

To avoid losing serious amount of information, early investigations of the LSE trades rarely use the approaches of differencing data. Instead, the fundamental price is replaced by the mid-touch, the average of best bid and best ask, and the spread is defined as twice the difference between the transaction price and the mid-touch. This approach is carried out by, among others, Board and Sutcliffe (1995) who study the impacts of the rules of trade publication, and Hansch et al. (1999) who study the costs of preferenced order flows. Chapter 2 and Chapter 3 also use mid-touch in several occasions. Thus the spread is not measured from changes in prices. Drawbacks of using the mid-touch are that the time series properties of the fundamental prices have not been considered in full, and more importantly, whether the mid-touch is indeed a good approximation of the fundamental price remains in doubt; see Reiss and Werner (1996) and Hansch et al. (1999) for the discussion about the use of the mid-touch. Moreover, the mid-touch may not be available outside the mandatory quote period, which is one of the main reasons why these outside trades are excluded in all the studies of the LSE market.

This chapter presents a model preserving the time series properties while data sequentiality is not required. The time series properties are modelled explicitly in order to obtain estimates of the underlying fundamental prices and spreads. Also, the model takes explicitly account of a number of effects, which may cause the variation of the spread such as size of trade and time of trade. The structure of the model is general and it is easy to include more explanatory factors into the model. The remainder of this section discusses the empirical features affecting the spread which have been emphasised by other contributions in the literature.

4.2.3 Trade size effect

The size of the trade is closely related to its order processing, inventory and information costs. Easley and O'Hara (1987) have argued that the spread increases with the size of the trade because large orders may indicate new information. On the other hand, the market maker deviates from the optimal portfolio by taking the order from the customer, so the spread is increasing in the size of the trade to compensate the loss; see Amihud and Mendelson (1980) and Ho and Stoll (1981, 1983). Furthermore, other authors have emphasised that spread size decreases with trade size because fixed costs are associated with each trade (Stoll 1978a; de Jong, Nijman, and Roëll 1995; Reiss and Werner 1996). Therefore, the size effect is expected to have a so-called U-shape: the spread is big for both small and big trades, and it is small for medium trades.

4.2.4 Intraday effect

The dominant feature of the bid-ask spread in many data sets is the intraday variation. Several contributions in the literature have studied the importance of intraday variation on the NYSE (Wood, McInish, and Ord 1985; McInish and Wood 1992; Brock and Kleidon 1992). It is argued that the spread is bigger in the early morning because traders have strong demands to re-establish their optimal portfolios at the start of the trading period. Also, market makers may fear that investors have private information at the start of the trading period. When the market is about to close, traders prepare for the non-trading period by adjusting their portfolios in an appropriate way.

Outside the NYSE, empirical evidence of the intraday pattern of spreads is mixed. Wang et al. (1994) use data obtained from the Chicago Mercantile Exchange and they find that the spreads of the S&P 500 index futures have an intraday U-shape. Lee et al. (1993) find that effective spreads of stocks from NYSE and American Stock Exchange (AMEX) exhibit U-shape patterns. On the other hand, Werner and Kleidon (1996) study cross-listed stocks in the US and the UK and they conclude that the spread in the UK declines during the day. Chan, Chung, and Johnson (1995) observe that the spreads of the Chicago Board Option Exchange are smaller at the end of the trading period. Chan, Christie, and Schultz (1995) find the inside spreads of NASDAQ stocks remain relatively constant in the morning and decline in the afternoon, and they suggest that differences of intraday patterns are due to institutional differences. Finally, van Ravenswaaij (1997) finds that only the first hour of trading on Paris Bourse has a considerable impact on the size of the spread.

4.2.5 Other effects

The level of competition of market makers is negatively related to the spreads; see Demsetz (1968). Competition level may be measured by the number of market makers of a security. The spread is positively related to the risk of returns to compensate market makers when market participants are risk-averse; see Stoll (1978b) for theoretical arguments and Bollerslev and Melvin (1994) for empirical evidence. Stoll (1978b) further argues that factors such as the wealth of market makers and holding periods of securities also affect the spread. Easley and O'Hara (1992) argue that high volume of trades signals an information event so that market makers have to increase the spread to protect themselves from adverse selection. On the other hand, Dutta and Madhavan (1997) argue the positive relationship between volume and spread may result from the monopolistic power of market makers. Finally, the structure of an individual market may determine the spreads. Multiple dealership markets are not transparent and, therefore, market makers increase the spread to protect themselves from information advantage; see Pagano and Roëll (1992). In contrast, Naik et al. (1999) suggest the market makers in the dealership markets may narrow the spread to solicit informed trades.

4.2.6 Volatility

The variances for the disturbances may not be time invariant. It has been emphasised in the literature that for many financial time series, the disturbances are heteroscedastic and that it can be modelled by autoregressive conditional heteroscedastic (ARCH) structures or by stochastic volatility (SV) specifications, see Anderson and Bollerslev (1997) for discussions. However, this chapter will concentrate on the intraday volatility; the volatility is not treated as a stochastic process but as a deterministic effect on variances of different time periods within the day.

Many contributions in the literature have found a U-shape for the intraday volatility of security returns (Wood, McInish, and Ord 1985; Park 1993; Chan, Christie, and Schultz 1995; Chan, Chung, and Johnson 1995; Werner and Kleidon 1996). French and Roll (1986) suggest that trading itself creates volatility and Slezak (1994) argues that the information asymmetry during market closure contributes the uncertainty to the opening. Furthermore, the works of Admati and Pfleiderer (1988) and Foster and Viswanathan (1990) indicate that volatility and volume are correlated. Empirical evidence of the positive relationship between volume and volatility is provided by Jain and Joh (1988). All these indications hint that the heavy trading at the beginning and at the closing of the trading period may be the source of the U-shape volatility. Finally, Jones et al. (1994) point out that the volatility-volume relationship is essentially a volatility-transaction relationship: the number of transactions is correlated more closely to the volatility than the volume is.

4.3 Statistical model for bid-ask spreads

The statistical model for prices of multiple dealership markets with multiple market makers presented below can be used generally for observations which are not necessarily sequentially ordered over time.

4.3.1 The main structure of the model

Suppose that N_t transactions have occurred at time t, where N_t is a non-negative integer. The trade prices of a security at time t are stacked into the vector p_t and the associated trade quantities are stacked into the vector v_t in the same order as p_t. The elements of vectors p_t and v_t are denoted by $p_{t,i}$ and $v_{t,i}$, respectively, for $i = 1, \dots, N_t$. When observations are not available for some time point $t = \tau$, the observation τ is treated as missing and $N_t = 0$. The estimation method employed later handles missing observations in a straightforward manner.

The underlying fundamental price of an equity is denoted by the scalar μ_t and it applies to all trade prices within time period t. The model is thus given by

$$\begin{array}{rclrclr} p_{t,i} & = & \mu_t + s_{t,i} + \varepsilon_{t,i}, & \varepsilon_{t,i} & \sim & N\left(0, \sigma_\varepsilon^2\right), & i = 1, \dots, N_t \\ \mu_t & = & \mu_{t-1} + q_t + \eta_t, & \eta_t & \sim & N\left(0, \sigma_\eta^2\right), & t = 1, \dots, n. \end{array} \tag{4.1}$$

The first equation states the trade prices consist of fundamental price μ_t, the half-spread $s_{t,i}$, and the pricing error $\epsilon_{t,i}$. The second equation states the fundamental price μ_t is the fundamental price in the previous period μ_{t-1} adjusted for the information from the order flows, q_t, and for the information from other sources, η_t. The specification for $s_{t,i}$ and q_t are discussed below. The normal distributed disturbances $\varepsilon_{t,i}$ are mutually independent and uncorrelated with the normal distributed disturbances η_t. The structure of the model is similar to the ones used by Glosten and Harris (1988) and Huang and Stoll (1997) but model (4.1) allows the number of trades to vary with time t. Also, the particular specifications for the spread and the adverse selection effect are different; see below. Finally, unlike Huang and Stoll (1997), the inventory effect is not considered for two reasons. First, market makers on the LSE heavily depend on the inter-dealer market to balance the inventory, see Hansch et al. (1998) and Reiss and Werner (1998). Second, the investigations in Chapter 2 and Chapter 3 have not found that inventory consideration determines the placement of the quotes nor the size of effective spread.

The assumption of normality for $\varepsilon_{t,i}$ may not be very realistic partly because of the nature of the observations which have been subjected to rounding functions. Problems related to rounded data have been given some attention in the literature; see Harris (1991, 1994), Hausman et al. (1992) and Chrodia and Subrahmanyam (1995). Furthermore, statistical techniques have been developed to deal with the rounding problem and they can be applied to the model. Hasbrouck (1997) considers a discrete bid-ask price model and estimates the model using a non-linear filtering technique. Manrique and Shephard (1997) extend this analysis by presenting a Bayesian treatment using Markov chain Monte Carlo methods. However, in view of the aim of this chapter and the large sample size, the rounding problem is not the crucial issue in the analysis.

4.3.2 The spread

The specification of the spread effect depends on a number of fixed (non-stochastic) but unknown parameters. The half spread at time t of the i-th trade is denoted by $s_{t,i}$ and its specification is given by

$$s_{t,i} = d_{t,i}\left(z_t'\gamma + w_v'\delta\right), \qquad v = v_{t,i}, \tag{4.2}$$

where $d_{t,i}$ is set equal to unity when the i-th trade at time t is a buyer-initiated trade and it is set equal to minus unity when it is a seller-initiated trade:

$$d_{t,i} = \begin{cases} 1, & \text{trade } t,i \text{ is buyer-initiated} \\ -1, & \text{trade } t,i \text{ is seller-initiated} \end{cases}.$$

When the data set contains information about the identities of the traders, as is the case in the sample, the direction of the trade can be easily identified. When such information is not available, the direction of the trade can also be identified by μ_{t-1} of (4.1):

$$d_{t,i} = \begin{cases} 1, & \text{if } p_{t,i} > \mu_{t-1} \\ -1, & \text{if } p_{t,i} < \mu_{t-1} \end{cases}.$$

The size of the i-th trade at time t is denoted by $v_{t,i}$.

The model does not attempt to decompose the spread into three components as Huang and Stoll (1997) do. Instead, it will uncover how the spread varies with the trade size and the time of the day. The part of (4.2) in the brackets represents a regression equation with the parameter vectors γ and δ. The explanatory variables z_t and w_v are constructed vectors which are based on the time-of-day t and the size $v_{t,i}$, respectively. This representation allows the introduction of piece-wise regression effects which have different parameters for different intervals within the range of independent variables; see Johnston (1984, Chapter10, Section 2). This specification can be further generalised to regression cubic spline functions which join the discrete jumps of the parameter coefficients to a twice differentiable smooth function; see Poirier (1973, 1976). Figure 4.2 graphically displays a regression line, a piece-wise regression line and a cubic spline function for an artificial set of random points. The cubic spline regression gives the best fit compared to the piece-wise regression with the same number of parameters. Some technical details of regression cubic splines are discussed in Appendix B.1.

Following the previous theoretical and empirical work on the bid-ask spread, the model specification for the spread captures the following two features:

1. The intraday effect $z_t'\gamma$ is modelled by a regression cubic spline, in which the x-scale is time. A limited number of knots are equally distributed between 0:00 and 23:59 hours. The spline is restricted to sum to zero to avoid confounding with the total effect.

2. The size effect $w_v'\delta$ is also modelled by a regression spline but here the x-scale is size. The knots are placed between the possible minimum and maximum sizes of trades. When no institutional boundaries for trade sizes exist, the minimum and maximum can be determined from the data.

An alternative approach for modelling non-linear effects is to transform the underlying variables into a vector of dummy variables and to estimate the corresponding coefficients either by generalised least squares techniques (Lehman and Modest 1994; Werner and Kleidon 1996) or by general method of moments (Sheikh and Ronn 1994; Chan, Chung, and Johnson 1995). This approach is unsatisfactory because of the discontinuity of parameters for different intervals. It is also cumbersome when a lot of dummy variables are required as is usually the case.

4.3.3 Adverse selection effect

The adverse selection effect is the effect of signed volumes on the fundamental price μ_t, and it is modelled by

$$q_t = r_t'\beta = \sum_{j=1}^{S} \beta_j r_{j,t} \qquad \text{where} \qquad r_{j,t} = \sum_{i=1}^{N_{t-j}} d_{t-j,i}\, v_{t-j,i}, \qquad j = 1, \ldots, S, \tag{4.3}$$

with $\beta = (\beta_1, \ldots, \beta_S)'$ as a fixed unknown vector of coefficients. The vector $r_t = (r_{1,t}, \ldots, r_{S,t})'$ contains the sum of the trade volumes multiplied by trade dummies $d_{t,i}$'s. Note that r_t is the sum of the signed volumes of the market, not of any single market maker.

After trades are executed, market makers revise their beliefs about the fundamental price according to the volumes of trades. If there are more unexpected buy volumes, then the security is probably undervalued and the price must be adjusted upwards. Similarly, if there are more unexpected sell volumes, then the price will be adjusted downwards. However, the adjustment q_t is not only a function of the volumes in the preceding period, but it also depends on the volumes even earlier for two reasons. First, market makers revise the beliefs of fundamental price by unexpected order flows (Hasbrouck 1991). Order flows are likely to be serially correlated because of, for example, price stickiness or order fragmentation. As the expected order flow a function of lagged order flows, so is the unexpected flow, which may be written as

$$r_{j,t} - E(r_{j,t} \mid r_{j-1,t}, r_{j-2,t}, \ldots) = r_{j,t} - \phi_1 r_{j-1,t} - \phi_2 r_{j-2,t} \ldots .$$

Second, the security market is rarely transparent. For example, the publication of trades may delay, so the unexpected trades a few minutes back may still affect the current price. As a result, the adverse selection effect can be written as a function of past trades, that is,

$$q_t = \sum_{j=1}^{S} \theta_j [r_{j,t} - \sum_{k=1}^{j} \phi_k r_{j-k,t}] \equiv \sum_{j=1}^{t} \beta_j \; r_{j,t}. \tag{4.4}$$

4.3.4 State space representation

Model (4.1) can be placed into state space form; see, among others, Harvey (1993) for a general discussion. The state space form consists of a transition and a measurement equation; they are respectively given by

$$\begin{aligned} \alpha_t &= T_t\alpha_{t-1} + W_t\lambda_w + R_t\eta_t, \quad \eta_t \sim N(0, Q_t), \qquad t = 1, \ldots, n, \quad (4.5) \\ p_{t,i} &= Z_{t,i}\alpha_t + X_{t,i}\lambda_x + \varepsilon_{t,i}, \quad \varepsilon_{t,i} \sim N\left(0, \sigma_{t,i}^2\right), \quad i = 1, \ldots, N_t, (4.6) \end{aligned}$$

where α_t is the $m \times 1$ state vector. The observation $p_{t,i}$ for time t and subject i is modelled as a linear function of the state vector α_t, the explanatory variable vector $X_{t,i}$ and the disturbance $\varepsilon_{t,i}$. The state vector follows a vector autoregressive process with transition matrix T_t, explanatory matrix W_t and selection matrix R_t for the disturbance vector η_t. The disturbances are mutually independent and uncorrelated with each other. The parameter vectors λ_x and λ_w associated with explanatory variables $X_{t,i}$ and W_t, respectively, allow the inclusion of fixed effects in the model. The matrices T_t, R_t and Q_t and the vectors $Z_{t,i}$ and $X_{t,i}$ are referred to as system matrices and vectors which are assumed deterministic and known. However, a small number of elements within the system matrices and vectors may be unknown. Denote ψ the vector of these elements. The parameter vector ψ can be estimated by maximum likelihood methods. See de Jong (1991) and Koopman and Durbin (1998) for more detailed discussions of the state space form.

It is straightforward to put model (4.1) into state space form. The state vector is the scalar μ_t so that $R_t = 1$ and $Q_t = \sigma_\eta^2$ in (4.5). The adverse selection effect q_t of (4.3) is modelled via the regression effect $W_t\lambda_w$ so that $W_t = r_t'$ and $\lambda_w = \beta$. The observation equation has $Z_{t,i}$ equals unity and the regression effect $X_{t,i}\lambda_x$ is used to model the spread (4.2) so that $X_{t,i} = (z_t', w_v')$ and $\lambda_x = (\gamma', \delta')'$. The variance $\sigma^2_{t,i}$ is the constant σ^2_ε. The initial state requires a diffuse prior condition, that is

$$\alpha_1 \sim N\left\{0, \kappa I\right\}, \tag{4.7}$$

where the diffuse prior κ represents a large scalar value, for example 10^5. The large variance is required because the fundamental price μ_t is modelled as a nonstationary time series process; see Koopman (1997).

The standard Kalman filter recursions evaluate the mean of the state vector α_{t+1} conditional on the vectors observations $p_1, \ldots, p_t$, that is $a_{t+1} = E\left(\alpha_{t+1} | p_1, \ldots, p_t\right)$, where

$$p_t = \begin{pmatrix} p_{t,1} \\ \vdots \\ p_{t,N_t} \end{pmatrix},$$

together with the variance matrix $P_{t+1} = var\left(\alpha_{t+1} | p_1, \ldots, p_t\right)$; see Anderson and Moore (1979). Koopman and Durbin (1998) argue that considerable computing savings can be achieved by treating the vector series p_1, ..., p_n as the univariate series $p_{1,1}$, ..., p_{1,N_1}, $p_{2,1}$, ..., p_{n,N_n} and applying the Kalman filter to the univariate series. The exact treatment of diffuse priors within the initial state vector variance matrix for the Kalman filter is also simplified considerably using the univariate approach.

The Kalman filter evaluates one-step and multi-step predictions of the state vector and it evaluates one-step-ahead prediction errors including their variances. These predictions are interpreted as minimum mean squares linear estimators. The smoothed estimator of the state vector, that is $\hat{\alpha}_t = E\left(\alpha_t | p_1, \ldots, p_n\right)$, and its variance matrix can be computed using a smoothing algorithm which is associated with the Kalman filter. The likelihood function can be constructed using the prediction errors via the prediction error decomposition; see Harvey (1993). The Kalman filter and the associated smoothing algorithm are discussed further in Appendix B.2.

4.4 Empirical results for the London Stock Exchange

4.4.1 The LSE data set

This chapter considers the trade data of three securities: Glaxo Wellcome, British Telecom and Shell Transport and Trading in 125 trading days between January and June 1996. The data used in this chapter does not include any IDB or IMM trades; every observation in the sample comes from a trade by a market maker and another non-market-maker participant. Therefore agency crosses are also excluded. Most of the trades are executed between the mandatory period, but 3.08% of all trades take place outside this time. Most of the prices in the sample are rounded: 91.22% of the trade prices are measured in pence so the vast majority of observations is discrete.

Standard procedures of data editing are applied; see Appendix A.4 for details. Some descriptive statistics are given in Table 4.1. The settlement records are time-stamped in minutes. There are occasions at which many trades are executed, for example, 89 transactions of British Telecom are executed on 17 May 1996 at 8:46, after shape trades and potential shape trades have been grouped (Appendix A.4). Including all of the data in the sample will increase the computational burden substantially, so only the ten biggest trades are included in the final sample if there are more than ten trades within a minute. As a result the data set is reduced by 2.42%.

4.4.2 Details of the model

Each trade record contains time of the trade t, the security price $p_{t,i}$ and the trade size $v_{t,i}$. As a result, N_t and q_t of model (4.1) are known. The unit of $p_{t,i}$ is in pence, the unit of $v_{t,i}$ is in 100 shares and the time t is in minutes. The trade records also contain the capacity flags of the buyer and the seller. Since market makers are marked as "M", the direction of the trade is the action of the counter party of the market maker: $d_{t,i} = 1$ if the sellcap is "M", and $d_{t,i} = -1$ if buycap is "M".

The spread is modelled as two regression spline functions: one for the time-of-day effect and one for the size effect. The spline functions require a set of knots. Although a larger number of knots gives a better fit, the optimal number is based on the right balance between fit and parsimony. The knots are therefore determined by Akaike information criterion (AIC)

$$\text{AIC} = 2\left(k - \log L\right),$$

where L is the value of the likelihood function of the estimated model and k is the number of parameters which is defined as the sum of the sizes of the vectors α_t, λ_x, λ_w and ψ. The model with the lowest AIC value is selected as the appropriate model. The knots are

$$\begin{array}{cccc} & \text{Glaxo} & \text{BT} & \text{Shell} \\ \text{time spline} & \begin{bmatrix} 0 \\ 510 \\ 750 \\ 990 \end{bmatrix} & \begin{bmatrix} 0 \\ 510 \\ 750 \\ 990 \end{bmatrix} & \begin{bmatrix} 0 \\ 510 \\ 750 \\ 990 \end{bmatrix} \\ \text{size spline} & \begin{bmatrix} 0 \\ 16000 \\ 32000 \end{bmatrix} & \begin{bmatrix} 0 \\ 22500 \\ 45000 \end{bmatrix} & \begin{bmatrix} 0 \\ 9000 \\ 18000 \end{bmatrix} \end{array}.$$

Three knots for size spline are based on the maximum trade sizes observed in the data. Four knots for time spline are based on the official trading hour: 8:30 and 16:30 are respectively 510 and 990 minutes after the mid-night. In the same way as for the number of knots for splines, the maximum time lag S can be determined by the AIC decision rule, and S is set to be 3. The inclusion of more lags than 3 increases the computational burden and it does not increase the fit of the model.

4.4.3 Disturbances

The disturbances in the model (4.1) are normally distributed. The variance of η_t can be expressed as a ratio of the variance of ε_t which will be referrd to as the signal-to-noise ratio, that is

$$\omega = \sigma_\eta^2/\sigma_\varepsilon^2.$$

This unknown parameter is fixed for all time points except for certain time periods, and the variances are multiplied by certain constants for these special periods. Alternatively, different values for σ_ε^2 and σ_η^2 can be applied to different time periods, but this approach is not desirable as it increases the number of parameters to be estimated.

The signal-to-noise ratio ω is estimated via numerical optimisation of the likelihood function. After estimating ω, the Kalman filter is used to compute one-step ahead prediction residuals together with their variances. The standardised error $u_{t,i}$ is defined as

$$u_{t,i} = v_{t,i}/F_{t,i}^{\frac{1}{2}}, \tag{4.8}$$

where prediction error $v_{t,i}$ and its variance $F_{t,i}$ are computed as in equation (B.3) in the Appendix B.2. The absolute values of $u_{t,i}$ are grouped by the time of the trade in ten-minute intervals for which the means are calculated. Figure 4.3 presents the average of absolute standardised errors in the sample. The errors are much larger at the opening and at the closure of the trading period as it is suggested in Section 4.2.6. Therefore, adjust the variances of the model to take account of these empirical phenomena. Define the variance multipliers a_t and b_t so that the variances of the disturbances in model (4.1) become time varying and they are given by

$$\sigma_{\varepsilon,t,i}^2 = a_t\sigma_\varepsilon^2, \qquad \sigma_{\eta,t}^2 = \omega b_t \sigma_\varepsilon^2, \qquad t = 1, \ldots, n, \qquad i = 1, \ldots, N_t. \tag{4.9}$$

The variance multipliers for the model are as follows:

1. *The beginning and the end of trade period.* Section 4.2.6 points out that the volatility may be relatively large at the beginning and the end of the trading period, so the variance σ_ε^2 is doubled ($a_t = 2$) at the opening of the trading period between 8:31 a.m. and 8:50 a.m. and at the closure of the trading period between 4:00 p.m. and 4:28 p.m.

2. *Outside official trade period.* Most exchange markets impose official trading sessions but sometimes it is allowed to trade outside these periods. For example, special arrangements exist for overseas clients who

wish to trade outside the official period. The volatility is even larger than the beginning or the end of the official trading period, so the variance σ_ε^2 is multiplied by eight ($a_t = 8$) when trades take place before 8:30 am, and it is multiplied by six ($a_t = 6$) when trades take place after 4:29 pm.

3. *Period of fundamental price correction.* Price adjustments and large fluctuations in prices are expected when new information becomes available to traders. This requires relatively larger values for η_t and ε_t and therefore the corresponding variances should be increased in these periods. Indeed, some standardised one-step-ahead residuals are relatively large. When the absolute values of more than two consecutive residuals are greater than a certain number, say eight, a_t and b_t are given values larger than one. Although the adjustments are somewhat arbitrary, most of these adjustments coincide with news reports on, for example, earnings or mergers published by the *Financial Times*; see Table 4.2. This implies that the fundamental price must be given the flexibility to adjust for changes during these limited periods. Hence the variance adjustments are well justified.

4. Otherwise, $a_t = 1$ and $b_t = 1$.

Satisfactory results are obtained with these settings. It should be noted that final results were not very sensitive to a different set of adjustments. By taking into account the variance adjustments, the model is re-estimated.

4.4.4 Model specification: time, size and volume effects

This section investigates whether the size and intraday effects explain the spread and whether the volume effect explains the fundamental price. The likelihood ratio (LR) test and the AIC are used. AIC has been defined in the previous section, and the LR test statistic is given by

$$\mathrm{LR} = -2 * \log(L_0/L) = 2 * (\log L - \log L_0),$$

where L is the likelihood of the estimated model (4.1) and L_0 is the likelihood of the estimated restricted model. Three restrictions are considered: (i) no existence of size effect on the spread, that is $w_v'\delta = 0$, (ii) no existence of intraday effect on the spread, that is $z_t'\gamma = 0$, and (iii) no existence of volume effect on the fundamental price, that is $q_t = 0$. Under the restricted hypothesis and assuming normality for the disturbances, the likelihood ratio test statistic is chi-square distributed with degrees of freedom equals the extra number of variables for the unrestricted model. It is already mentioned that the assumption of normality is weak for the data set and therefore the LR statistic should be regarded as an indicative diagnostic rather than as a formal test.

The distribution of the initial state is diffuse implying that the initial state variance is arbitrarily large. The Kalman filter may behave unstable for the first few observations. The likelihood function is evaluated via the Kalman filter and it is decided to exclude the first percent of observations from the summation operators of the log likelihood function; see Appendix B.3. The LR statistics for the three restrictions are reported in Table 4.3. The restrictions of no size effect and no volume effect are clearly rejected. The restriction of no time-of-day effect for British Telecom (BT) securities cannot be rejected. Moreover, although the hypothesis of no intraday effect for Glaxo Wellcome (Glaxo) and Shell Transport and Trading (Shell) may be rejected in the conventional confidence level, the LR statistics of no intraday effect are much smaller than those under the other two hypotheses.

Table 4.3 also reports the AIC value for the estimated model (4.1) and the three AIC values for the estimated models with the three subsequent restrictions of no size, no time and no volume effects. The theory of AIC's suggests to select the model with the lowest AIC value, that is, the estimated model (4.1) for Glaxo and Shell, and the estimated model without the intraday effect for BT. However, the AIC values for the estimated model without any restriction and the estimated model without the time-of-day effect are not very different from each other in the case of Glaxo and Shell. The exclusion of other effects causes larger AIC values.

A graphical display of the time-of-day spline $z_j'\gamma$ for $j = 1, \ldots, 1440$, that is, 24 hours of 60 minutes, is presented in Figure 4.4. In the left column of graphs, the splines are displayed for $j = 1, \ldots, 1440$ and, in the right column, the splines are plotted against $j = 510, \ldots, 990$ which represents the trading period. From the left column of plots it may be argued that the Glaxo time spline has the familiar U-shape, the spread of BT is large during the trading hours, and the Shell spline is of a U-shape with a small hump around noon. On the other hand, the shapes of time splines during the trading hours are very different for the three securities as can be observed from the right column of pictures in Figure 4.4. The spread of Glaxo is large around noon and small at ten o'clock and three o'clock, the spread of BT is almost flat everywhere with two small humps, and the time spline of Shell has an inverse U-shape during the official trading period. The variation of the spline is very small in all cases. The difference between the maximal and minimal value of the splines is at most 1.8 pence and only 0.1 pence during trading hours. Like some of the work on dealership market mentioned in Section 4.2.4, the evidence does not support a U-shaped intraday pattern. Unlike those work which suggest a declining spread, no significant intraday pattern can be applied to all three stocks investigated in this chapter. Indeed, Section 4.5 will argue that intraday pattern only exists in the *touch* spread.

In contrast, there is a strong size effect for the spread. Figure 4.5 plots the estimated size spline functions $w_v'\hat{\delta}$ for possible values of the trade size. The left column shows the size splines of the three stocks studied in this chapter, and the other two columns show the splines of another six stocks for comparison. The middle columns is of three middle-range FTSE-100 stocks in terms of market capitalisation, and the right column is of three bottom-range FTSE-100 stocks. The size splines are measured by the basis points of the average trade price, and they are restricted to be zero when the trade size approaches zero. The trade sizes are measured relative to the stocks' NMS. Although there are some stocks with big trades, the plots only show the size up to 40 times NMS to accommodate the fact that the maximum trade size of Shell is only 36 times NMS. The size splines for the nine securities exhibit broadly the same pattern: the spread is decreasing and then increasing with the size. The phenomenon of an increase of the spread when the size gets large is consistent with the inventory and information models of spreads in the literature. The initial fall of the spread may be attributed to the fixed component of the transaction costs. The splines reach their minimum at around 20 times NMS. However, the slopes of the splines are very different. It is consistent with the argument made in Chapter 3 that size alone may not account for all of the variations of the spreads. It is also likely that the NMS is not a good measure to standardise the trade size. The use of the modal quote size may be considered. On the other hand, it should be emphasised that there are very few big trades, and the upward-sloping part of the size spline is less reliable. Moreover, the estimation is not immune from the criticism of Franks and Schaefer (1995) that it may under-estimate the spread of big trades should they are protected.

4.4.5 Model misspecification: estimated disturbances

The average of the standardised one-step ahead predictions residuals within each minute are presented in Figure 4.6 and some summary statistics are shown in Table 4.4. The average residual is defined as

$$\bar{u}_t = \frac{1}{N_t} \sum_{s=1}^{N_t} u_{t,i}, \tag{4.10}$$

where $u_{t,i}$ is defined in equation (4.8). It can be concluded that the residuals are not normally distributed and exhibit weak autocorrelations. It is mentioned in Section 4.3.1 that the assumption of normality is not very realistic and therefore it is not surprising that the skewness and kurtosis in the table point to a departure from normality. However, it does not mean the model is mis-specified. Kalman filter performs general least square, and the estimation is still the minimum mean square linear estimator should the normality do not hold (Harvey 1989). Even if the residuals are not normally distributed, the log-likelihood ratios are still asymptotically chi-square distributed. Figure 4.7 presents the correlograms of the residuals for the three securities.

The function is defined as

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)} = \frac{\sum_{t=\tau+1}^{T} \bar{u}_t \bar{u}_{t-\tau}/(T-\tau)}{\sum_{t=1}^{T} \bar{u}_t^2/T}.$$

The first coefficient of the correlogram is modest in all three cases and the coefficients are negligible after ten lags. The existence of serial correlation can partly be explained because the residuals under consideration are an average of a set of residuals within one minute. This type of pooling may introduce some serial correlation in large data sets.

The residuals do not exhibit any ARCH-type effects. The squared estimated disturbances η_t possess no serial correlation. Moreover, an OLS regression of these squared residuals against the number of transactions N_t gives very weak fit (measured by R^2) in all stocks. Therefore, the imposed multiplication factors for the variances σ_ε^2 and σ_η^2 have successfully dealt with the possible volatility for the three securities.

4.4.6 Estimated parameter coefficients

Table 4.5 presents the estimated parameters of model (4.1) without the time spline. Firstly, the signal-to-noise ratio ω is reported which is estimated by maximum likelihood. This requires numerical optimisation of the likelihood function which is computed by the Kalman filter; see Appendix B.3. The estimated ratio ω is roughly the same for Glaxo and Shell and it takes the value of around 0.1. The ratio for BT is much smaller which indicates that the fundamental price evolves more smoothly than the price of the other equities. It may be concluded that new information other than from the order flows has less influence on the fundamental price for BT than for Glaxo and Shell. The Figures 4.8, 4.9 and 4.10 present the observed trade price $p_{t,i}$ and the smoothed fundamental price $\hat{\alpha}_t$ for the three equities and Figure 4.11 shows an example of the evolution of the fundamental price on a specific day for Shell.

Secondly, the estimated coefficients for the size effects are reported together with the corresponding t-statistics. The estimate for parameter vector δ is required to generate a graphical display of the size spline. Again, the size effects appear to be statistically significant since the t-statistics associated with the knot coefficients have values larger than 3.5. The estimated half spread can be obtained by using δ vector to recover $b_t' y^\dagger$ in (B.1). For example, when the size of the trade tends to zero, $b_t' y^\dagger$ collaspes to δ_1, so the half spreads of the very small trades of Glaxo, BT and Shell are 1.28 pence, 0.68 pence and 0.97 pence respectively.

Finally, the estimated coefficients of the adverse selection effect βs are reported. They are highly significant and they all take positive values. Although the structure of θ and ϕ of (4.4) is not imposed, the estimates for βs can still be interpreted as an indication of the depth of the market (Kyle 1985). For example, the sum of $\hat{\beta}$ of Glaxo Wellcome is about 0.000096, which means, roughly speaking, that the fundamental price of the security will increase one penny if the volume of buys in the market exceeds that of sells by $1/((\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3)/100) = 1,041,667$ shares.

4.5 Compared with traditional approaches

Section 4.2 argues that the use of autocovariance structure or data differencing to estimate the bid-ask spread requires the data to be ordered sequentially. The trade data in *Transaction Data Service* is not sequential, and neither of the approaches is theoretically appealing or empirically feasible without distorting information. An alternative method is to use the mid-touch as the fundamental price, and to define the effective spread as twice the difference of the trade price and the mid-touch, for example, (1.1) in Chapter 2. Since the mid-touch is unavailable or unreliable outside mandatory quote periods, the rest of the section will report the results based on the trades executed between 8:30 and 16:30. Denote m_t the mid point of the best bid and ask quote at time t, then half of the effective spread of the trade i at time t is

$$s_{t,i}^f = d_{t,i}(p_{t,i} - m_t).$$

The assumption that the spread is a non-linear function of the trade size is maintained, so the focus is to compare model (4.1) with the least-square regression models

$$
\begin{aligned}
s_{t,i}^f &= w_v'\delta \qquad\qquad + \varepsilon_{t,i}, \text{ and} && (4.11)\\
s_{t,i}^f &= w_v'\delta + z_t'\gamma + \varepsilon_{t,i}, && (4.12)
\end{aligned}
$$

where $w_v'\delta$ and $z_t'\gamma$ are defined in (4.2) and $\varepsilon_{t,i}$ in (4.1). The size splines of the two models are exactly the same as the counterpart in (4.2); that is, the constant terms are *not* restricted to zero, and the knot points are the same as described in Section 4.4.2. On the other hand, the time spline in (4.12) is no longer a periodic spline, and its only restriction is that its value is set to be zero when the time is 8:30 (510). Unlike those in the Kalman filter model, the knot points of the time spline in (4.12) are 8:30 (510), 10:30 (630), 12:30 (750), 14:30 (870) and 16:30 (900). The residuals of (4.11) and (4.12) are respectively defined as

$$
\begin{aligned}
\hat{\varepsilon}_{t,i} &= s_{t,i}^f \ -w_v'\hat{\delta} \qquad\qquad , \text{ and}\\
\hat{\varepsilon}_{t,i} &= s_{t,i}^f \ -w_v'\hat{\delta} \ -z_t'\hat{\gamma},
\end{aligned}
$$

where $\hat{\delta}$ and $\hat{\gamma}$ are the estimates of δ and γ respectively. The residual of model (4.1) is defined as

$$\hat{\varepsilon}_{t,i} = p_{t,i} - \hat{\mu}_t - \hat{s}_{t,i},$$

where $\hat{\mu}_t$ and $\hat{s}_{t,i}$ are respectively the estimated fundamental price and spread from Kalman filter smoother without the time spline in s.

Table 4.6 compares the residual obtained from (4.1) and from (4.11) of the three stocks. The sum of residuals of least-square regression is zero, and the sum of residuals from Kalman filter is close to zero. The variance of $\hat{\varepsilon}_{t,i}$ is always smaller in (4.1) than in (4.11). $\hat{\varepsilon}_{t,i}$ in (4.1) is more skewed for BT and for Shell but not for Glaxo Wellcome. The kurtosis of $\hat{\varepsilon}_{t,i}$ in (4.11) is bigger for Glaxo and Shell but slightly smaller for BT. Both models may produce serious mispricing: the error of (4.11) can be as big as -11.77 for Glaxo, and the error of (4.1) can be as big as 10.52 for Shell. However, (4.11) produces bigger absolute $\hat{\varepsilon}_{t,i}$ on average in all three stocks. To sum up, the residuals of (4.1) appear to behave better than those of (4.11).

Figure 4.12 plots the absolute values of $\hat{\varepsilon}_{t,i}$ averaged by the size of the trades. The solid bars represent the means of $|\ \hat{\varepsilon}_{t,i}\ |$ from model (4.1), and the dash bars represent those from (4.11). It appears the absolute residuals of (4.11) are bigger in most of the trade sizes for all three stocks. The mispricing is a more serious problem with big trade sizes, but it should be noted that there are very few big trades. For example, only 332 trades exceed ten times NMS, which is 75000 for BT and 50000 for the other stocks. Furthermore, only 27 trades are bigger than the half of maximum trade sizes observed in the sample. Figure 4.13 plots the absolute values of $\hat{\varepsilon}_{t,i}$ averaged across the time of the day. Both the residuals of (4.11) and (4.12) are plotted as well as those of (4.1), but adding a time spline in the least-square regression produces almost exactly the same residuals for Glaxo and BT, and even for Shell the difference is minor. It indicates that adding the intraday effect does not change the estimation of effective spreads dramatically. The mean absolute residuals from the three models are close to one another during the opening and closing hours, and those of (4.1) tend to be smaller between 9:00 and 16:00. The gap between Kalman filter residuals and mid-touch residuals is the biggest for BT and the smallest for Shell, which correspond to 0.06, 0.08 and 0.03 pence differences of $|\ \hat{\varepsilon}_{t,i}\ |$ of Glaxo, BT and Shell respectively in the bottom row of Table 4.6. In other words, (4.1) performs best for BT, and it does only marginally better than (4.11) and (4.12) for Shell.

Figure 4.14 provides further information for the estimation of Shell. The results from Glaxo and BT are not different from those from Shell and hence are omitted. The top-left panel plots the correlogram of the average residuals, $\hat{\varepsilon}_{t,i}/N_t$. The autocorrelation function of the average residuals of (4.11) is much bigger than that of (4.1) in the first few lags. The former gradually converges to the latter, but it is still slightly bigger after 200 lags. The top-right panel plots the size spline of model (4.1), represented by small triangles, the size splines of model (4.11), represented by small dots, and the average half effective spreads by the solid line. The two size splines are almost indistinguishable for the small trades, and the spline from (4.1) is bigger than (4.11) for the big trades. The bottom-left panel plots the intraday effects. Small triangles represent the estimated time spline of the Kalman filter model, and small dots represent the spline of (4.12). In addition, the five-minute average of the half effective spread (the dash line) and the half touch spread (the solid line) are plotted. All of the values are standardised to have zero means. There is minor difference between the two splines during the day, and they are almost identical after 14:00. Both splines pass through the average effective spread. The splines range between -0.04 and 0.04 pence, and the average half spread between -0.12 and 0.08 pence. However, the half touch spread exhibits the greatest variation. It has the peak at the opening with 0.32 pence, it declines sharply within half an hour, it does not vary very much during most of the day, and it is the lowest at the closing. The big spread in the opening is consistent with the finding by Werner and Kleidon (1996), and the decline in the closing is consistent with the ones reported in Chan, Chung, and Johnson (1995) and Chan, Christie, and Schultz (1995). All of their results are based on the quote data.[1] However, the big variation of quote spread does not contribute to effective spread nor Kalman filter spread. It appears that the costs charged by market makers are relatively invariant to the swings in the quote spread.

4.6 Discussions and conclusions

The available level of detail in databases of intra-daily transaction data for multiple dealership markets brings mixed blessings to the study of market microstructure. On the one hand, transaction prices and volumes become available for small time intervals, which provide the opportunity for financial analysts to have a better understanding about the market behaviour. On the other hand, the data set is not necessarily ordered sequentially and therefore standard techniques for estimation of the spread may not be applicable.

This chapter has presented a simple model to analyse the data of three

[1] Chan, Chung, and Johnson (1995) use Berkeley Options Data Base, which is a quotes and trade data set. Because the trading volumes are very low in the early morning (see their Figure 1 in page 337), the spreads obtained in the morning are presumable from quote data.

heavily traded stocks on the London Stock Exchange. The model includes components which allow for the time series properties of the data and the existence of non-linear effects. The problem of non-sequentiality is solved by putting the model into state space form and to estimate the model by maximising likelihood using the Kalman filter. The updating recursions of the Kalman filter do not require the dimension of the observational vector to be constant. The Kalman filter and associated algorithms can deal with missing observations in a straightforward manner. The underlying fundamental price of the security is extracted from the data, and at the same time the effect of volumes on the price and the explanatory factors of the spread are estimated. Strong evidence is found to support the fact that spread is a non-linear function of the trade size. The evidence of intra-daily effects is less strong. Apart from the weak significance of test statistics, the estimated time splines for the securities have three different shapes and the variation of the three splines are small. Intraday effects are only pronounced in touch spreads. Hence, it is plausible that the time-of-day effect is not a determinant factor of the spread on the London market. In conclusion, the analyses of the three securities have been successful. Model (4.1) can generate a wide range of statistics, which provide detailed information about the available transaction data. It also outperforms the traditional approach of using the mid-touch as the fundamental price. Empirical studies based on the model may contribute to a better understanding of the fundamentals of the bid-ask spread.

The basic model (4.1) can accommodate almost all features of the fundamental price and the spread. Variations of the model may be used to address other questions of the market microstructure. It may be interesting to modify the model specification and to improve the estimation techniques for further research. A three-way decomposition of bid-ask spread into order-processing, inventory and information costs will be in the agenda. Moreover, the strategy of imposing different disturbance variances for specific time periods is not satisfactory and it can be altered. The introduction of an intra-daily variance function can be considered and its specification may rely on dummy variables or smooth functions such as the cubic spline. Also, it is argued that the normality assumption is not realistic because the data is subjected to rounding functions. Furthermore, the erratic behaviour of financial time series may require error distributions with heavier tails. Thus a more in-depth analysis of the data requires more advanced estimation techniques. However, such improvements come with a price. The model specification in this chapter consists of one parameter which needs to be estimated by numerical optimisation. More sophisticated models will inevitably lead to an increase of computational costs.

Table 4.1: Descriptive statistics of sample stocks
The constructed sample consists of the first ten largest trades with the same time stamp. The directions of the trades are identified by the capacity flags. The unit of volumes is 1,000 shares, the unit of the size of trade is one share, and the price unit is one penny.

Company name		Glaxo	BT	Shell
Industry		Pharmaceutical	Utilities	Oil and Gas
Original sample				
number of observations		61,111	124,694	48,660
total trading volumes		474,829	912,699	297,926
max number of trades in 1 minute		38	89	56
Constructed sample				
number of trades	buy	31,734	29,198	13,301
	sell	28,476	91,797	34,273
	total	60,210	120,995	47,584
price	mean	852.00	360.09	873.84
	max	970.00	386.50	820.00
	min	765.00	326.00	949.00
	std	50.99	14.34	33.41
size of trade	mean	7,790	7,395	6,203
	max	3,200,000	4,484,000	1,798,800
	min	1	1	1
total trading volumes	buy	241,965	447,720	132,179
	sell	227,113	460,036	162,986
	total	469,078	907,756	295,165
number of market makers		19	19	19

Table 4.2: The timing and magnitude of changes in variances
This table shows the time when a_t and b_t, defined in equation (4.9), are adjusted. The magnitude of the adjustments is also shown. Events are news items of the companies reported in the *Financial Times* on the same day or the next day of the adjustments.

stock	date	time	a_t	b_t	events
Glaxo	31/01	08:03		100	Optimism that a cocktail of its AZT and 3TC drugs could be a weapon against HIV.
	06/03	09:13		40	ABN Amro HG downgrades the share.
		10:48-10:55	6.67	800	Glaxo may give up US patent fight. Disappointment on the full-year figures.
	26/03	10:05-10:20	2	20	Glaxo is about to link up with Pfizer.
	16/05	14:10-14:50	4	1000	A better-than-expected report on sales
		15:46-15:56		2	growth.
BT	06/03	11:45		50	The government plans to liberalise the overseas telephone call market
	08/03	13:35-14:25		8	
	28/03	15:59-16:06	2.5	200	Dealers on bid alert as speculation builds up
	29/03	07:46-08:07	50	600	BT and Cable &Wireless re-open merger talks
	02/04	11:50-12:00		5	France Telecom and Deutsch Telecom indicate they are not interested in bidding for Mercury
	18/04	08:40-09:00		5	BT and C&W may announce deal
	02/05	17:46-	20	1000	Shock news that BT had terminated merger talks with Cable and Wireless came after market hours.
	03/05	-8:02	20		BT marked down sharply from the outset on news of the groups' doomed merger.
	17/05	09:10		50	BT and BSkyB plan joint venture.
	10/06	09:12-09:45		6	Markets welcome ambitious venture with MCI.
Shell	15/02	10:16		80	23% rise in dividends
		13:15-14:50		5	Bad result for the 4th quarter 1995 Downgraded by Goldman Sachs and ABN Amro in New York opening
	08/03	13:55-14:35	4	150	The stock could not resist the pressure from Wall Street
	09/05	10:00-10:08	6.67	200	Recorded quarterly result of net income. Recommended by Goldman Sachs

Table 4.3: LR statistics for size, time and volume effects

The proposed model is equation (4.1) with size and time spline as in (4.2) and with adverse selection effect q_t defined in (4.3). The model without size effect is the same as the proposed model except that $w_v'\delta$ of (4.2) is dropped and replaced by a constant term for the spread. The model without time effect is the proposed model without $z_t'\gamma$ of (4.2). The model without volume effect is the proposed model with no q_t in (4.1). The log-likelihood is defined in Appendix B.3, the likelihood ratio is defined as LR$= -2*\log(L_0/L) = 2*(\log L - \log L_0)$, and df is the difference of the number of parameters between the proposed model and the model under the null hypothesis. The critical value for $\chi^2_{(3,0.99)}$ is 11.341 and $\chi^2_{(4,0.99)}$ is 13.277. AIC is defined as AIC $= 2\,(k - \log L)$, where k is the number of parameters, including the signal-to-noise ratio ω.

	Glaxo			BT			Shell			k	df
	$\log L$	LR	**AIC**	$\log L$	LR	**AIC**	$\log L$	LR	**AIC**		
Proposed model	-9886.44		19796.9	19966.08		-39908.2	4054.70		-8085.4	12	
No size effect	-9988.57	204.26	19995.1	19789.59	352.98	-39561.2	3908.42	292.56	-7798.8	9	3
No time effect	-9901.78	30.68	19819.6	19963.62	4.92	-39911.2	4036.69	36.02	-8057.4	8	4
No volume effect	-9958.42	143.96	19934.8	19848.35	235.46	-39678.7	3956.49	196.42	-7895.0	9	3

Table 4.4: Summary statistics for standardised residuals
This Table presents the summary statistics of the average standardised residuals $u_{t,i}$ which are defined in equation (4.10).

	Glaxo	BT	Shell
mean	0	0	0
variance	1	1	1
skewness	0.0975	0.1768	0.0870
kurtosis	6.0671	7.7270	6.0028
$\rho(1)$	0.1897	0.2659	0.1640

Table 4.5: Parameter estimates
This table presents the parameters of the final model, that is, 4.1 without time spline. The signal-to-noise ratio ω is $\sigma_\eta^2/\sigma_\varepsilon^2$, where σ_ε^2 and σ_η^2 are respectively the variance of $\varepsilon_{t,i}$ and η_t, as defined in (4.1). The estimates are reported of the elements δ_1, δ_2 and δ_3 of the size spline vector δ which is defined in (4.2). The estimates for β_1, β_2 and β_3, the elements of vector β of (4.3), are reported. The usual t-values are also reported.

	Glaxo		BT		Shell	
ω	0.115312		0.064824		0.116857	
σ_η^2	0.783025		0.219060		0.340653	
σ_ε^2	0.090291		0.014200		0.039808	
		t-value		*t-value*		*t-value*
δ_1	1.2759	343.23	0.6938	549.65	0.9664	307.02
δ_2	0.7906	3.86	0.4692	10.70	0.6631	8.57
δ_3	5.3886	4.42	4.6397	17.17	5.3573	12.02
β_1	6.7937e-05	8.98	9.1173e-06	6.98	7.6701e-05	8.82
β_2	-1.9276e-06	-0.21	4.6373e-06	2.90	8.4424e-06	0.79
β_3	3.0355e-05	4.04	1.0937e-05	8.43	2.0801e-05	2.43

Table 4.6: Mean residuals during mandatory quote periods
This table compares the summary statistics of the residuals $\hat{\varepsilon}_{t,i}$ estimated from model (4.1) and (4.11).

	Glaxo		BT		Shell	
$\hat{\varepsilon}_{t,i}$ of model	(4.1)	(4.11)	(4.1)	(4.11)	(4.1)	(4.11)
mean	-0.0187	0	-0.0006	0	-0.0006	0
variance	0.4288	0.5309	0.0815	0.1133	0.1871	0.2600
skewness	0.0868	-0.1979	0.5454	-0.1990	0.0205	0.0310
kurtosis	3.9779	5.8131	7.4636	7.1237	6.1197	9.3444
maximum	9.4369	7.7224	5.5140	5.2860	5.8374	10.5216
minimum	-7.3041	-11.7730	-3.7062	-5.6811	-6.7451	-6.9822
mean of $\mid \hat{\varepsilon}_{t,i} \mid$	0.4916	0.5573	0.2065	0.2872	0.3111	0.3465

Figure 4.1: An illustration of the *Transaction Data File*

DATE	TIME	BUY FIRM	SELL FIRM	BUY CAP	SELL CAP	PRICE	QUANTITY
01/04/96	11:09	ABC	DEF	M	A	379.00	55
01/04/96	11:09	ABC	FGH	M	A	379.00	440
01/04/96	11:09	IJK	LMN	M	A	379.00	400
01/04/96	11:09	IJK	OPQ	M	A	379.00	3045
01/04/96	11:09	RST	UVW	M	A	380.00	25000
01/04/96	11:09	OPQ	IJK	A	M	381.00	783
01/04/96	11:10	IJK	OPQ	M	A	379.00	880
01/04/96	11:10	IJK	IJK	M	N	379.00	1000
01/04/96	11:10	IJK	LMN	M	A	379.00	1270
01/04/96	11:10	XYZ	CAE	M	A	379.00	550
01/04/96	11:10	XYZ	OPQ	M	A	379.00	1500
01/04/96	11:10	RST	EWP	M	A	379.50	130000
01/04/96	11:10	FEN	RST	A	M	380.00	3000
01/04/96	11:10	OPQ	IJK	A	M	380.50	2267
01/04/96	11:10	LOU	ABC	A	M	381.00	700
01/04/96	11:10	OPQ	IJK	A	M	381.00	796

Figure 4.2: Example of regression, piece-wise regression and cubic spline

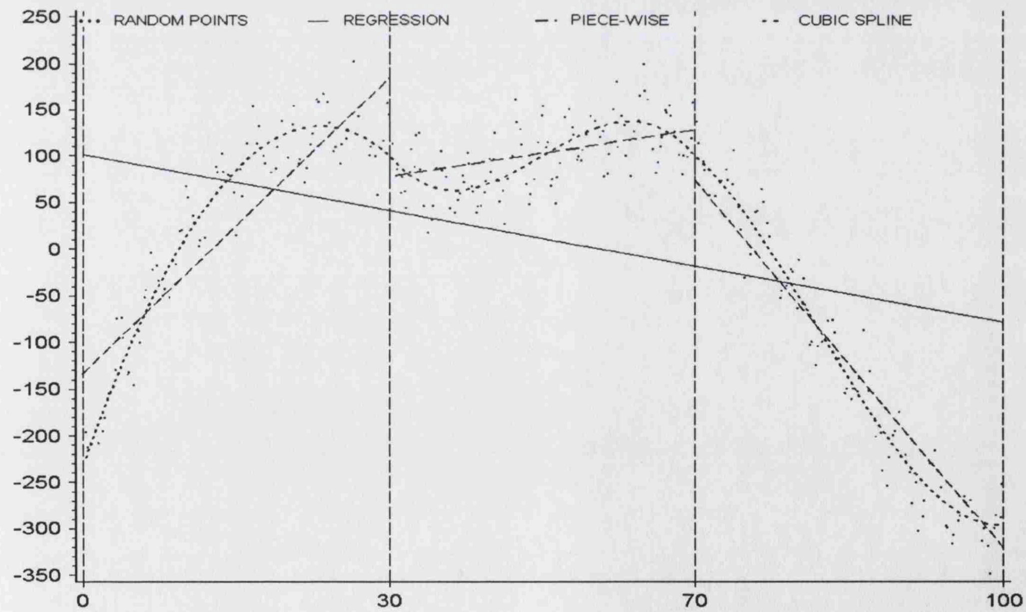


Figure 4.3: Means of absolute standardised errors of ten-minute intervals

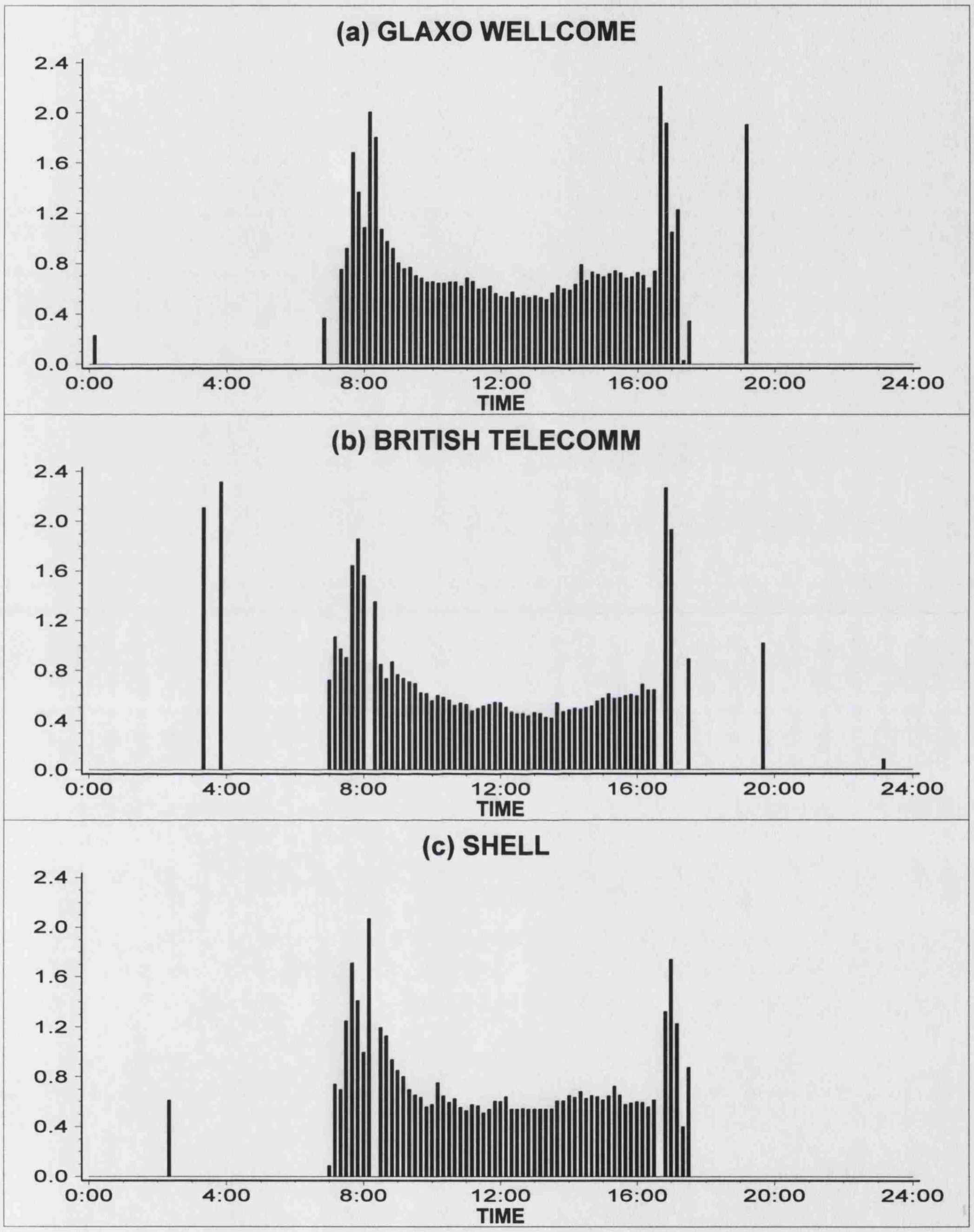


Figure 4.4: Estimated time splines of the spreads

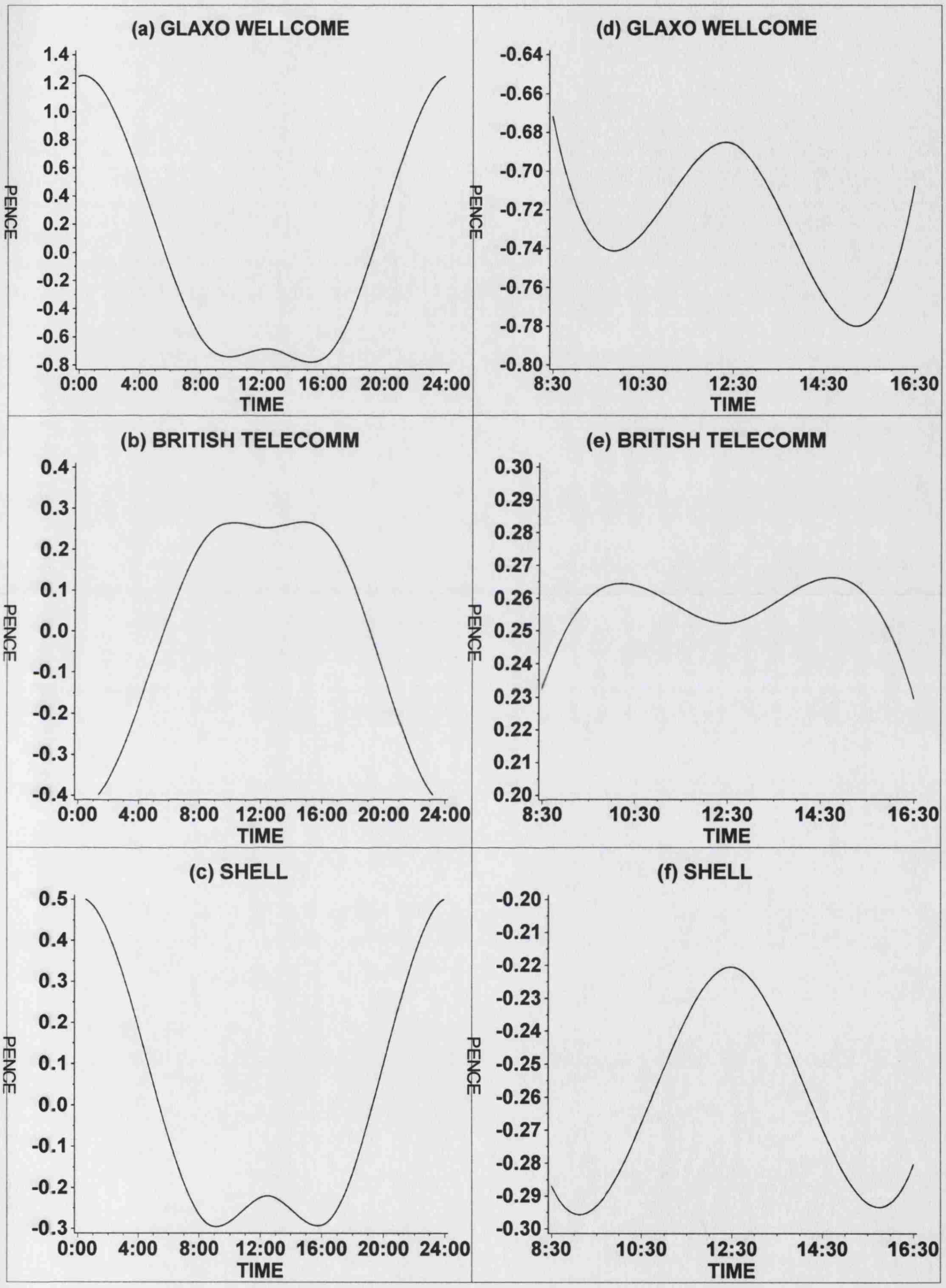
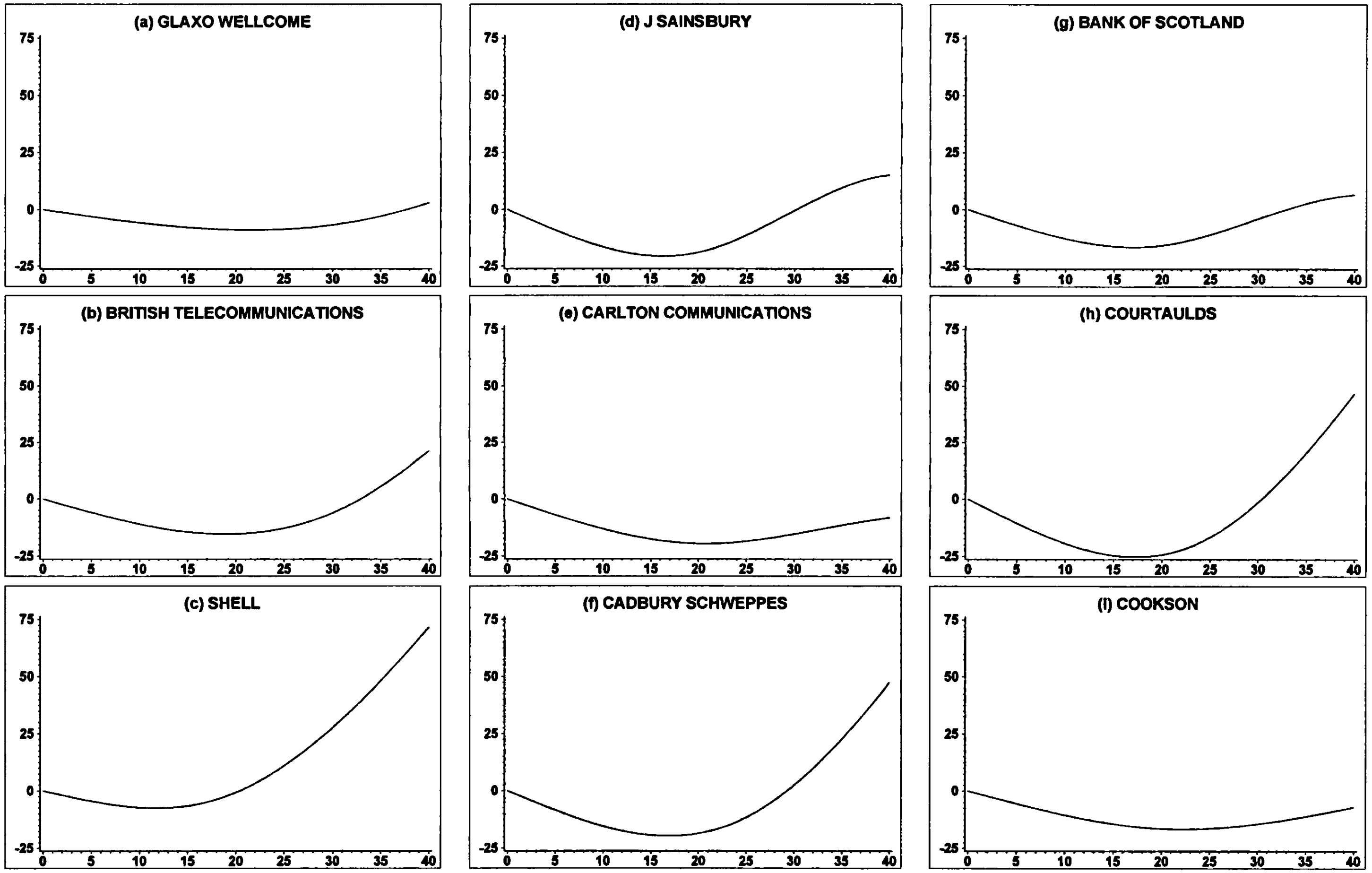


Figure 4.5: Estimated size splines of the spreads



TRADE SIZE UNIT: NMS SPLINE UNIT: BASIS POINT

Figure 4.6: Average of prediction residuals for each minute

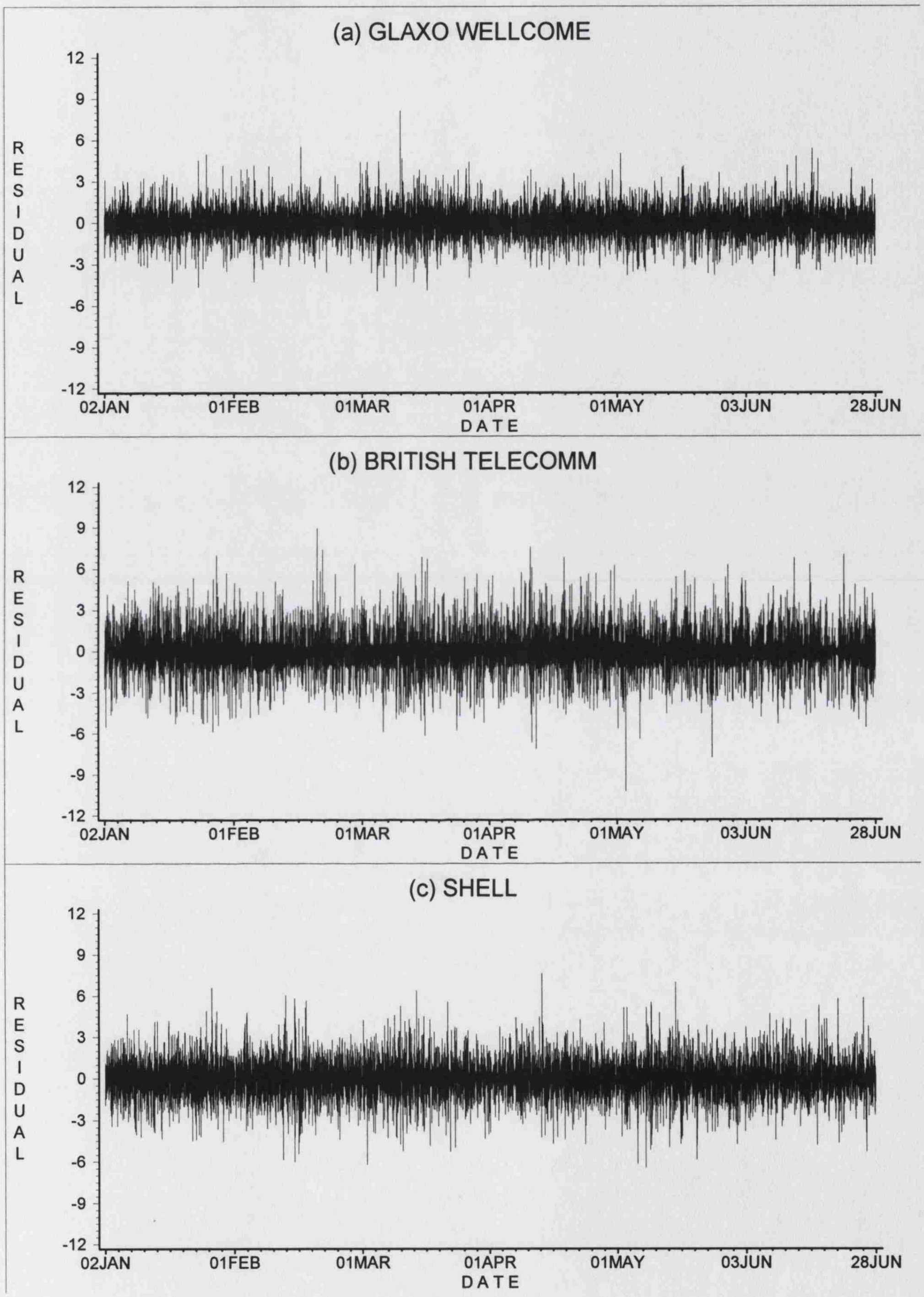


Figure 4.7: Correlogram for prediction residuals

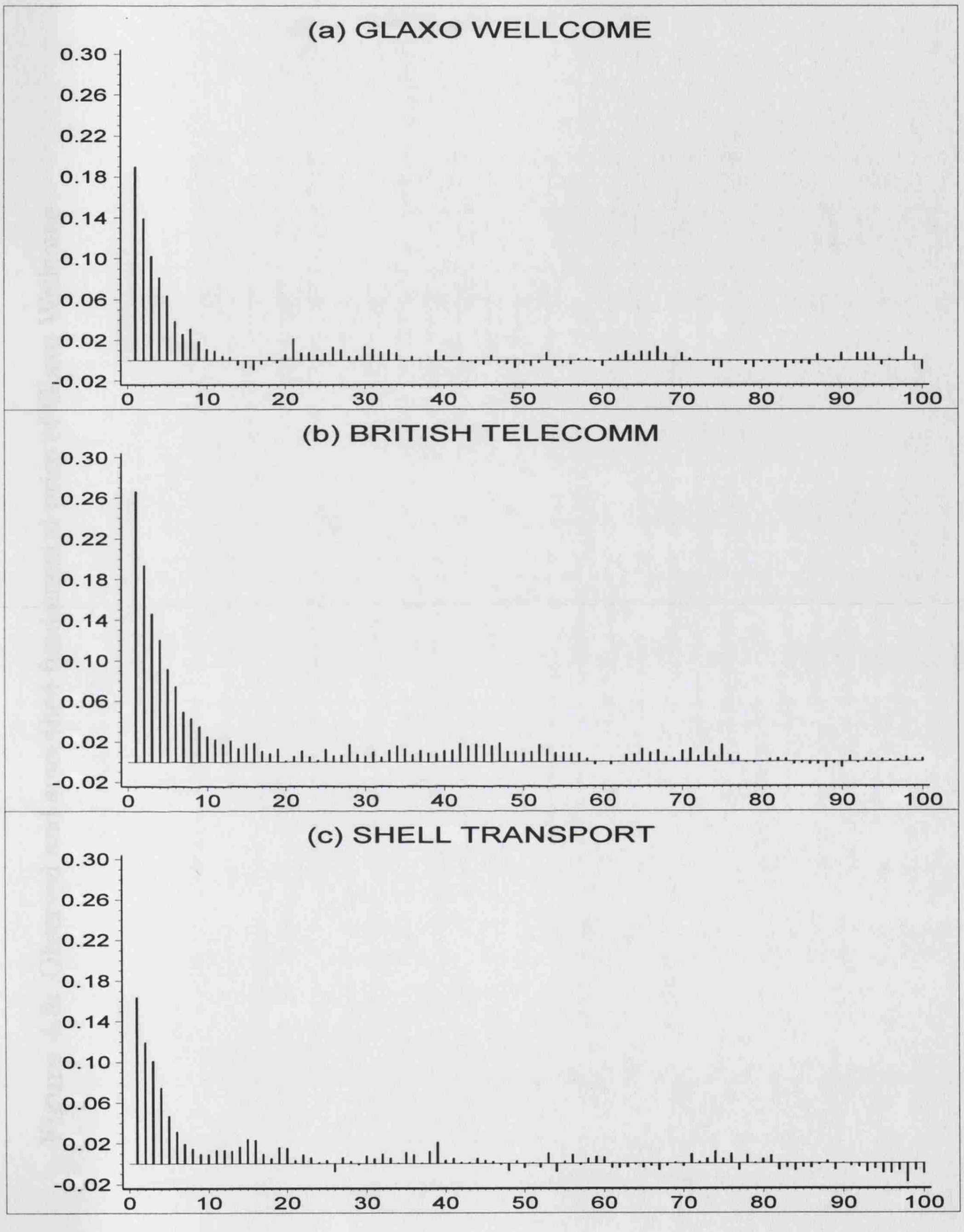


Figure 4.8: Observed and smoothed fundamental price of Glaxo Wellcome

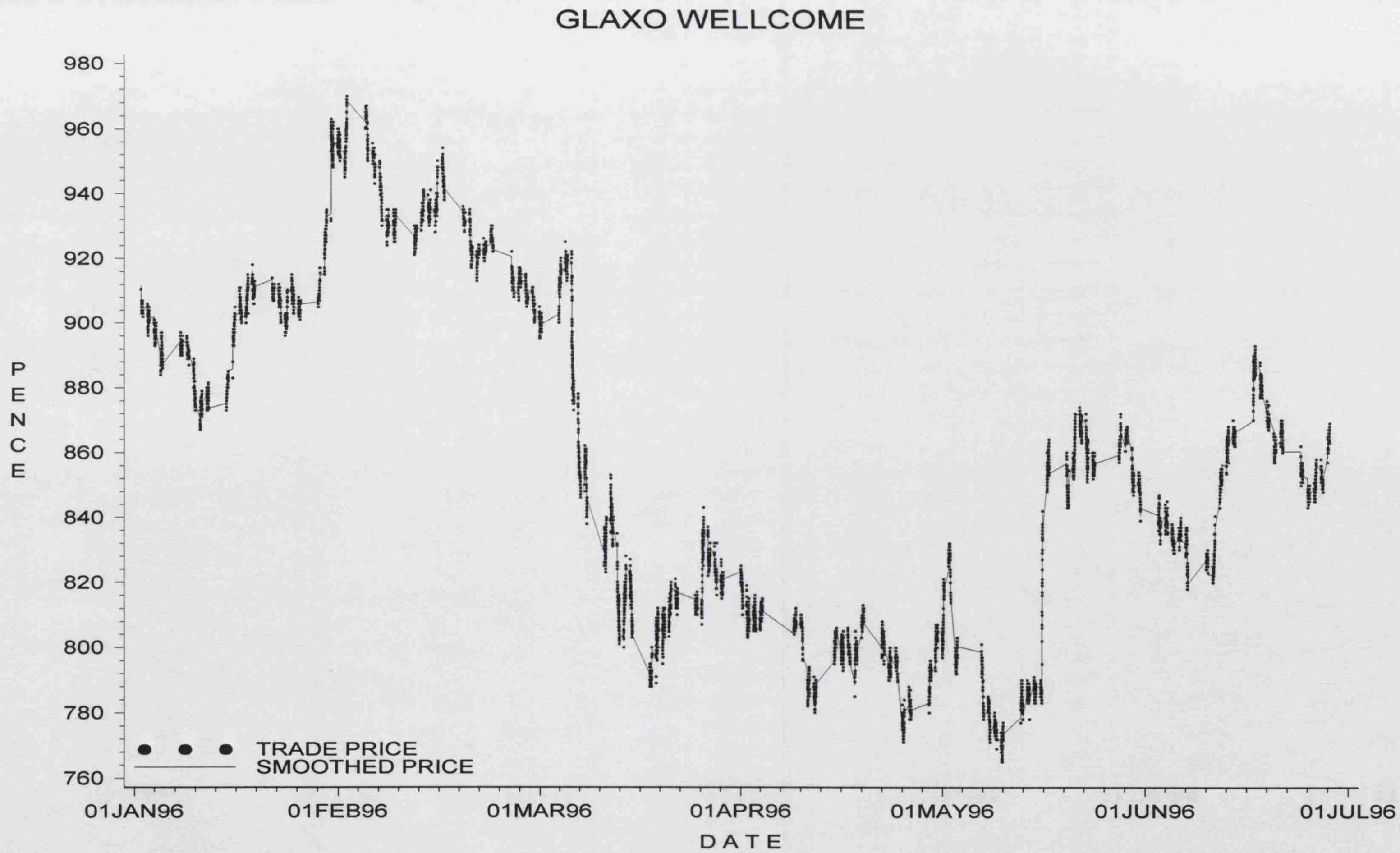


Figure 4.9: Observed and smoothed fundamental price of BT

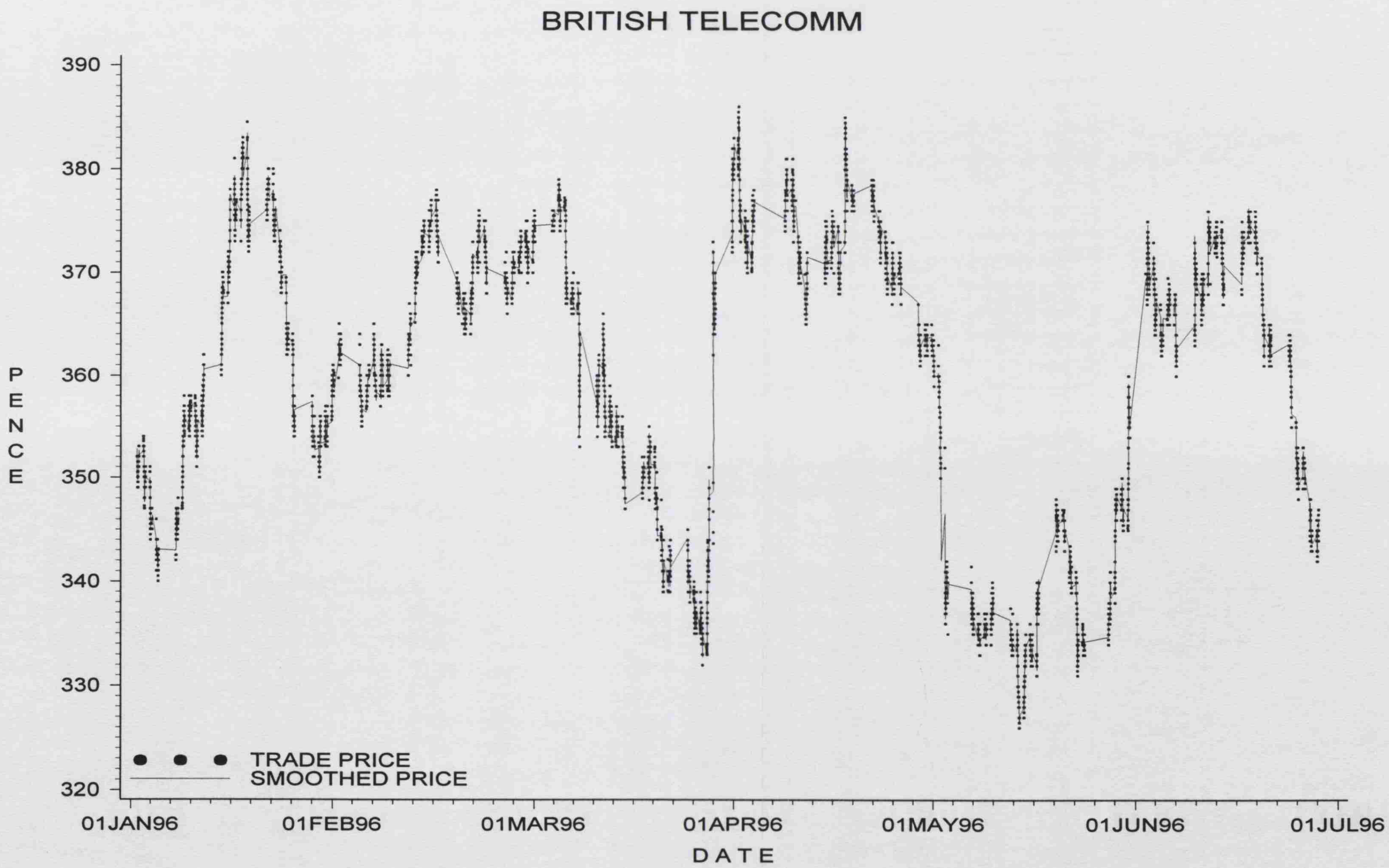


Figure 4.10: Observed and smoothed fundamental price of Shell

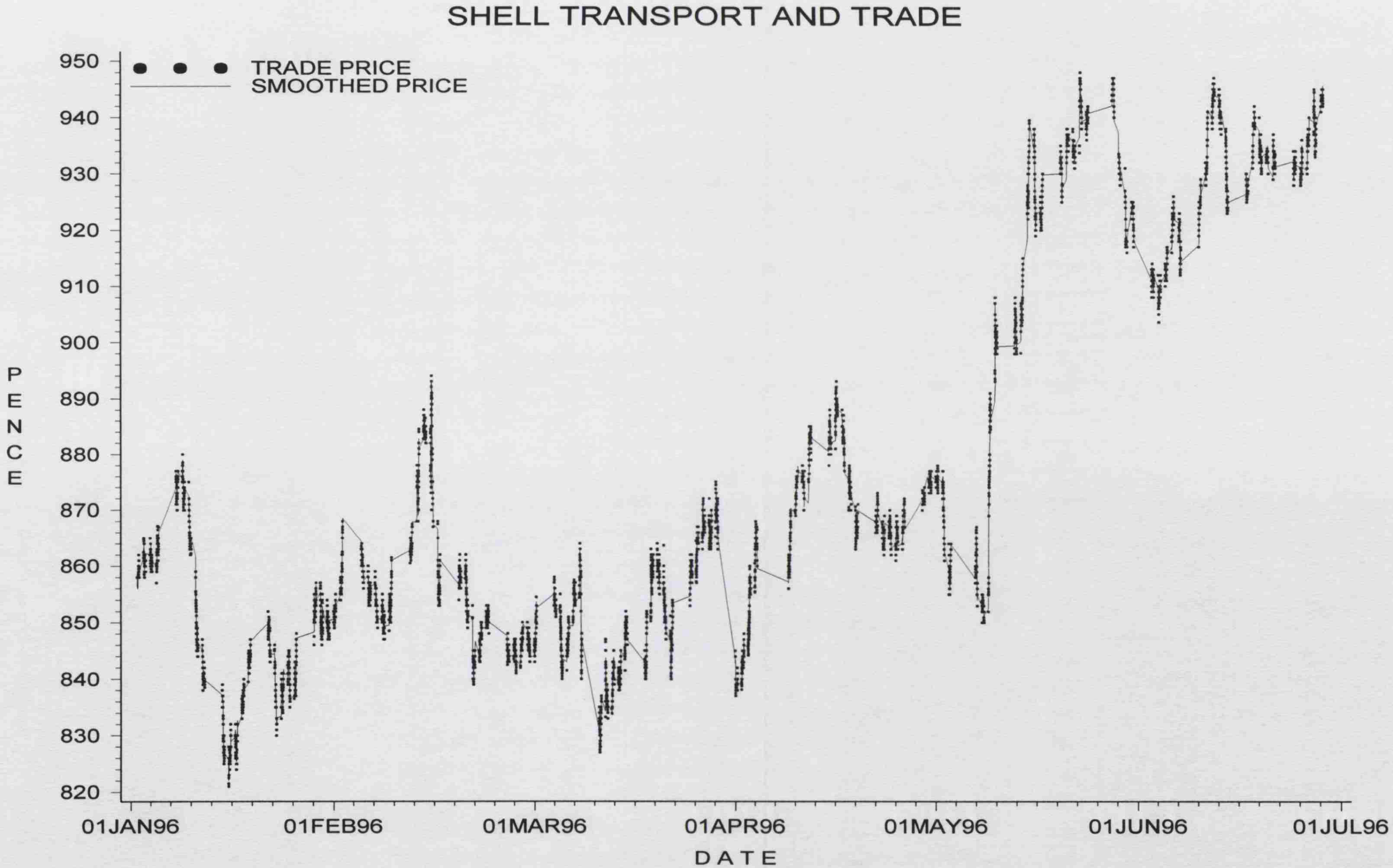


Figure 4.11: Fundamental price for one specific day

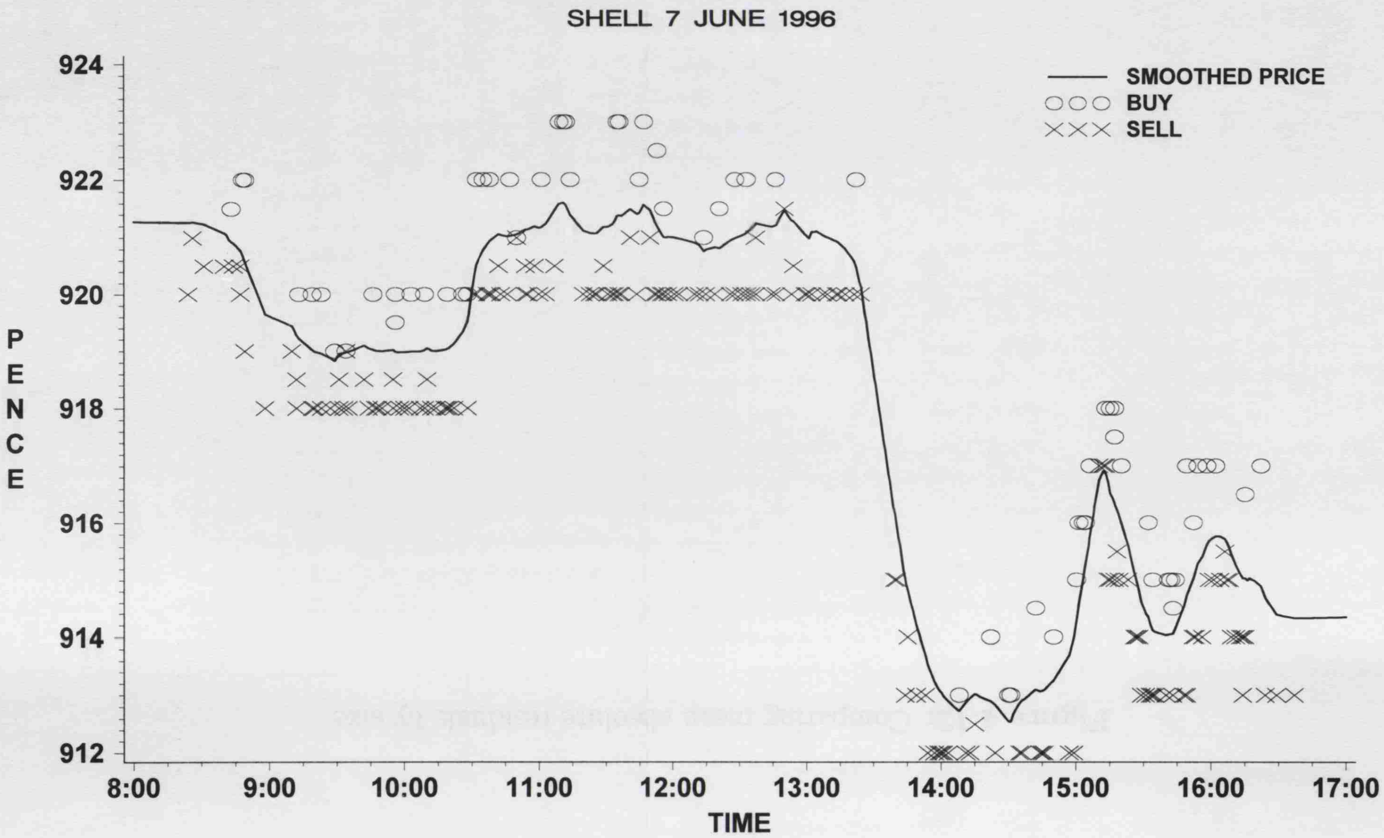


Figure 4.12: Comparing mean absolute residuals by size

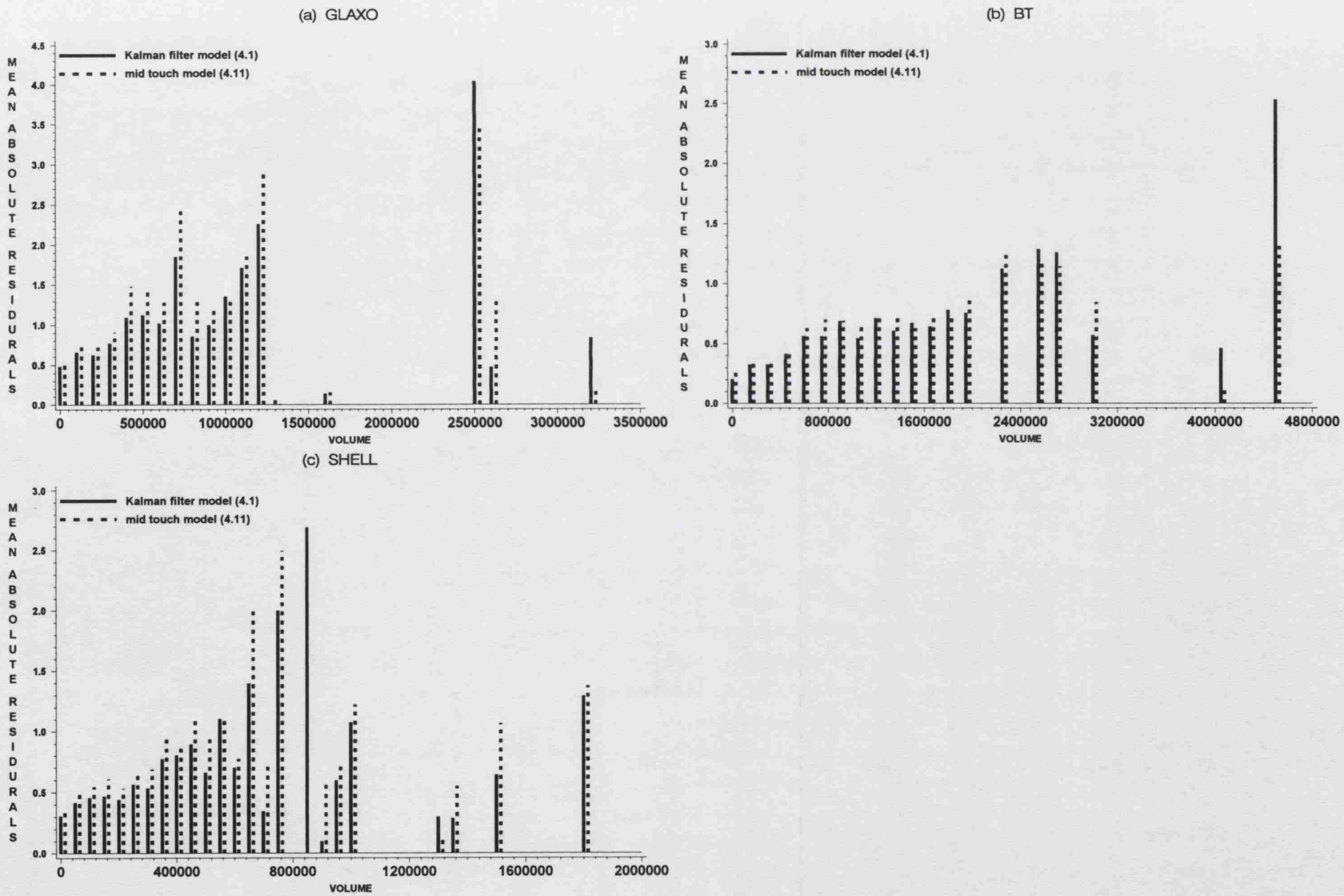


Figure 4.13: Comparing mean absolute residuals by time

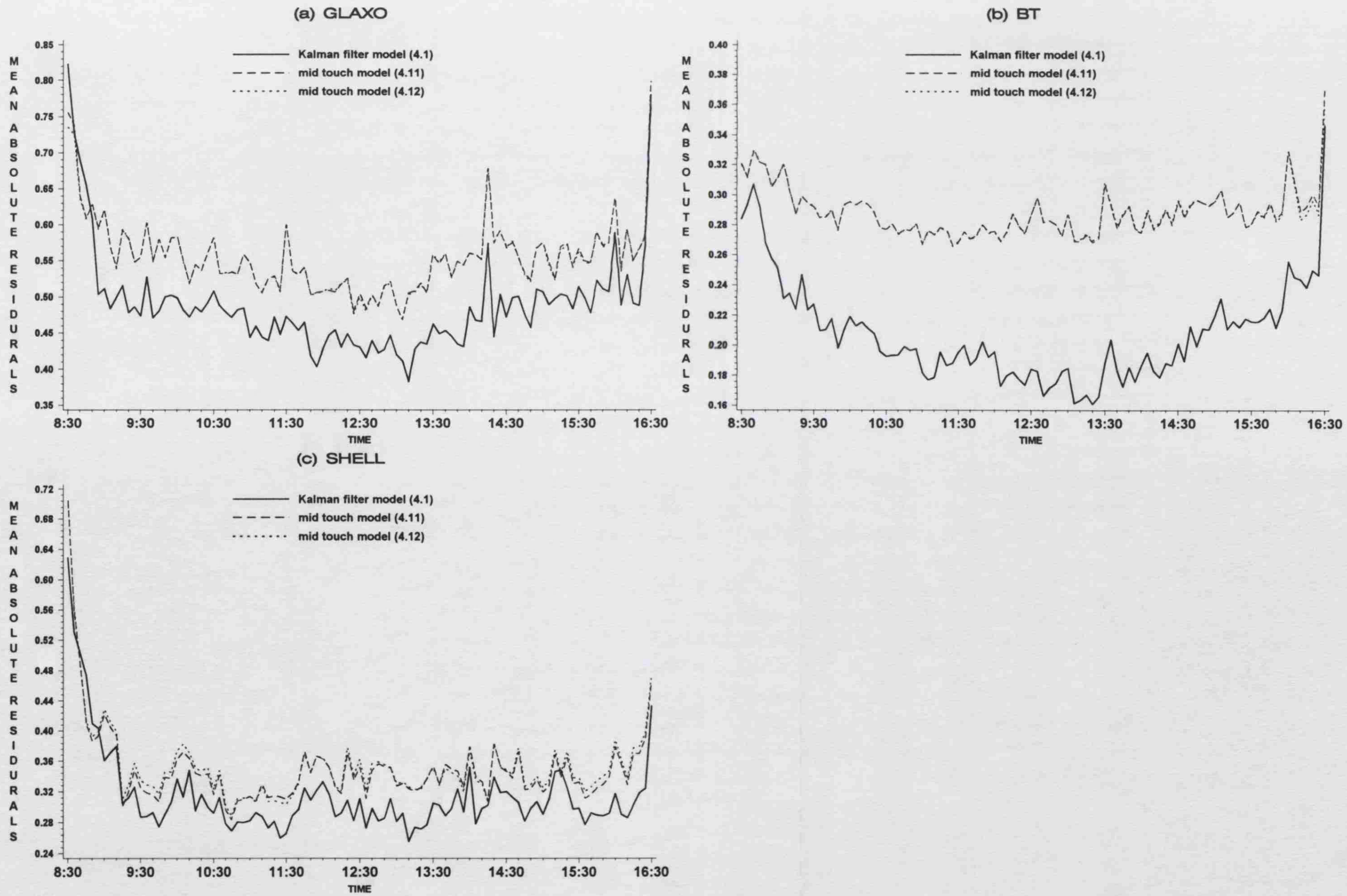
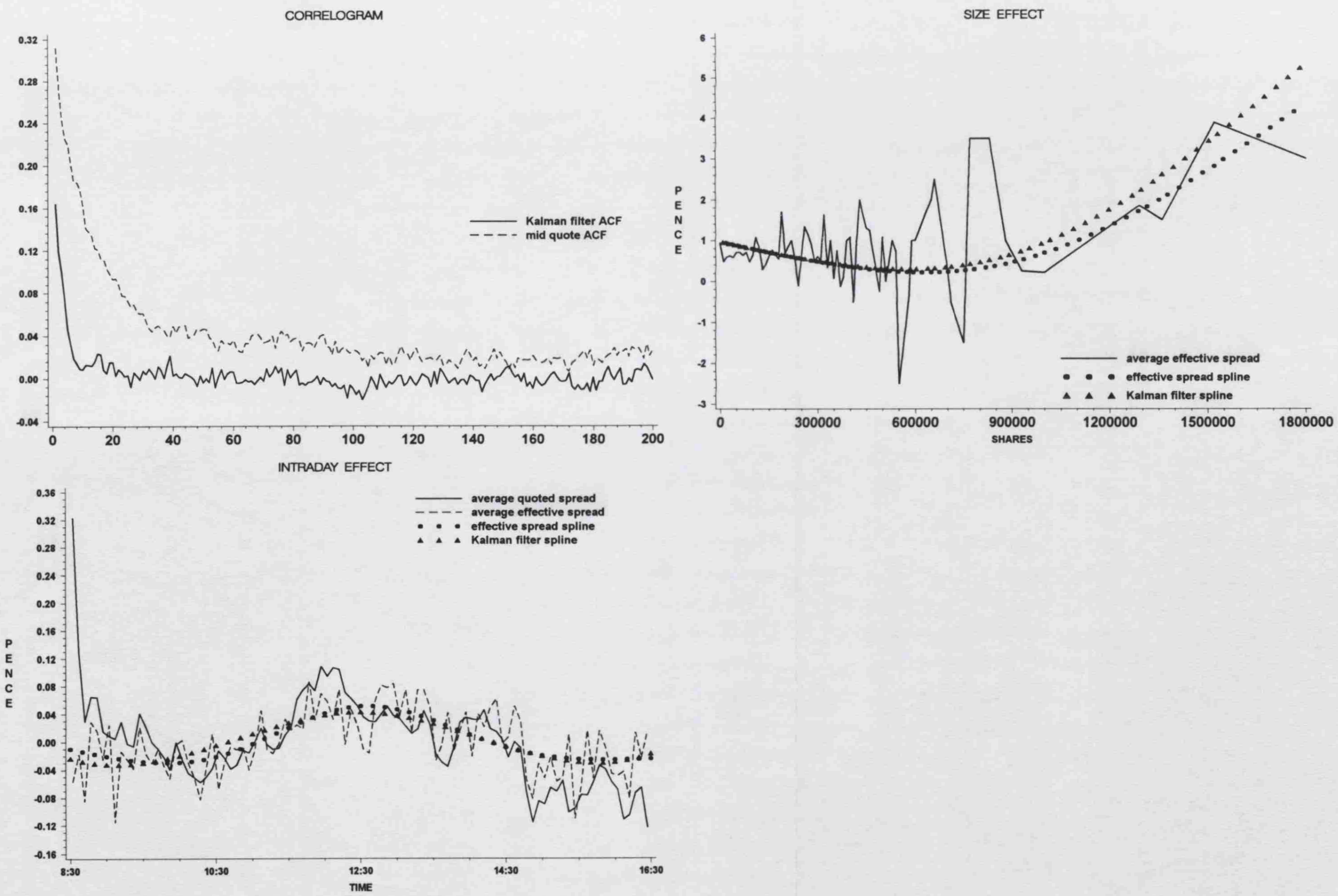


Figure 4.14: Residuals and splines of Shell Transport and Trading



Chapter 5

CONCLUSIONS

5.1 Main results

This thesis has investigated the trading and quoting behaviour of market makers extensively. The major contributions of this thesis are

1. This thesis clearly demonstrates the relationship between order flows and quote status. By taking account of the straddling quotes, the market order imbalance, and the number of market makers on the yellow strip, quote status clearly determines the amount and the balance of orders.

2. Posting the best quote attracts orders bigger than the quote size shown in the SEAQ screen, which implies the presence on the yellow strip has signalling effects. Furthermore, not only the quote spreads but also the effective spreads matter: the order flows are negatively related to the effective spreads for the medium and large trades.

3. There are two different types of quote revisions. The first type of the revision is to quote one side of the yellow strip aggressively in order to attract unbalanced order flows. The quotes are often made by price leaders in response to public information. The upward quote revisions are often associated with the arrival of buy trades, and the downward revisions are often with sell trades. The second type of quote revision is to deliberately straddle the yellow strip or to quote both sides of the strip. The revision is often made by a price follower soon after the

quote change of price leaders. The benefit of straddling the yellow strip is to avoid receiving unbalanced order flows; the drawback is the order flows received tend to be less than standing on the strip.

4. The variables suggested in the literature successfully explain part of the effective spreads, but they cannot explain most of the costs difference.

5. Market makers change quotes after customer trades, which imply those trades have certain inventory or information implication. IMM trades also tend to trigger the quote change which implies those trades contain information. On the contrary, the posters of IDB trades change their quotes to the opposite direction of the one predicted by the literature.

6. A model is proposed and implemented to estimate the spread by the non-sequential trade data. The model fully takes accounts of the time-series properties and the volume effects of the fundamental price, and the spread is estimated at the same time by considering the size effect and the intraday effect. The estimation uses Kalman filter to handle the with variables of non-constant dimensions. The model outperforms the alternative approach of using the mid-touch as fundamental price in several criteria. Nevertheless, the intraday effect does not exist in either the Kalman filter spread or the effective spread.

5.2 Other findings

By attempting to explain the quote revisions of the market maker, to resolve the puzzle of cost structure, and to achieve a better estimation of the fundamental price and bid-ask spread, this thesis has come across several issues which may help understanding the trading environment of the LSE.

5.2.1 The competitiveness of the LSE

The SEAQ screen is designed to encourage the competition among market makers. The market makers who post best bid and ask prices are highlighted on the yellow strip, right in the middle of the screen. Market participants can never avoid knowing that who are posting the best quotes. Furthermore, the best-execution rule enforces the brokers to find the best deals for their clients. The consequence is that almost all of the trades smaller than the quote sizes are executed at or inside the touch.

The fact that trades are at least as good as best quotes does not follow that customers get good deals. If the spreads are wide, then the deals are bad (Peterson and Fialkowski 1994). Huang and Stoll (1996) argue that preferenced ordering reduces the incentive for the market makers to narrow the spread. Chapter 2 reports the presence on the yellow strip has little effect on attracting small orders, so presumably order preferencing is most common among small trades. Chapter 3 reports that the majority of the small trades come from brokers on behalf of individual investors, and the costs of those trades are high. Moreover, market makers do not compete in spreads to attract order flows. Chapter 2 finds that most of the market makers are comfortable with straddling the yellow strip. They do not change the quote once somebody moves them off the strip, and once they are moved to the yellow strip they often tend to get off.

On the other hand, there is little evidence to suggest market makers collude. A small number of market makers appear to improve the yellow strip often and to stay on the strip for a long time. Second, results from the cross-sectional regression of the effective spread in Chapter 3 does not support the implicit collusion of market makers suggested by Dutta and Madhavan (1997). Most importantly, IMM trades turn out to be at least as expensive as agency trades, which could never happen had they collude. If there is anything that can be labelled as collusion, it is the order preferencing by broker-dealers. When they trade on behalf of investors outside the Exchange, the trading costs are high. When they trade on their own accounts, the costs are low. It is possible that broker-dealers route the customer orders to certain market makers in exchange for better deals for themselves. It is worth pursuing along this line in future research.

5.2.2 Variables which determine the spread

A few variables are used for the investigations in this thesis in various occasions. Chapter 2 and Chapter 3 use cross-sectional regression models to determine the sources of quote changes, and the number of trades appear to perform better than the trading volumes. Chapter 4 uses aggregate volumes to estimate the adverse selection effects. Although the results obtained from model (4.1) are satisfactory, it might be interesting to see whether the model can be further improved by replacing aggregate volumes with sum of trades in equation (4.3).

Strong evidence is found to support the fact that spread is a non-linear function of the trade size. The estimated size splines in Chapter 3 and Chapter 4 are remarkably similar: the spread is decreasing with the trade size when the size is small, and is increasing when the size is large. The decreasing part of the size spline can be attributed to the fixed component of the bid-ask spread, and the change in bargaining power between market makers and their counter parties. The increasing part of the spline is consistent with the inventory and information hypothesis in market microstructure theory.

The evidence of intra-daily effects is less strong. Apart from the weak significance of test statistics, the estimated time splines for the three securities in Chapter 4 and the one for effective spread in Chapter 3 have totally different shapes. Furthermore, the variations of those splines are very small compared with, for example, the size splines. The result is different from the intraday effects of *quote* spread found in Werner and Kleidon (1996) and Chan et al. (1995) that the spread is bigger in the morning than in the afternoon in dealership markets. In fact, Chapter 4 concludes that the wide touch spread in the early morning does not have a huge impact on either the effective spread calculated by the mid-touch or the Kalman-filter spread estimated by model (4.1).

5.2.3 The components of the spread

This thesis does not attempt to decompose the bid-ask spread as, for example, Huang and Stoll (1997) or Madhavan et al. (1997) do. However, some of the results of this thesis may shed some lights on the three components of the spread. Although some of the results in the regression models of Chapter 2 and Chapter 3 do not separate the inventory effect from the information effect completely, there is evidence supporting information effect, while there is little evidence suggesting strong inventory consideration plays an important role in determining the behaviour of market makers. The regressions of quote changes in Section 2.6 and Section 3.6 reveal that what changes the quote is not the change in inventory position, nor the trading volumes; instead, it is the number of trades of the market as a whole, especially the medium trades, that move the quote. Inventory affects the decision of quote change only when there are few firms making the market. Moreover, the inventory-related variables in the regression of effective spreads in Section 3.5 are not consistent with the inventory hypothesis by Amihud and Mendelson (1980) and Ho and Stoll (1981, 1983). As Kandel and Marx (1997) point out, inventory consideration should not dominate the decision of the firms who serve for providing liquidity. It appears either the successful inventory management reduces the possibility of market makers to face inventory overload or shortage, or inventory management is never an issue of the firm.

Chapter 2 concludes that information consideration plays an important role in determining the quote revisions of price leaders. Other results, including the adverse selection effects shown in Chapter 4, suggest that market makers regard the trading activities in the whole market as valuable information. If market makers change the quote according to the trades in the market, and if the change in fundamental price of the security is affected by the aggregate volumes of trades, then it implies that information determines the *positioning* of the spread. Although it is not equivalent to the statement that information determines the *size* of the spread, theoretical models such as the ones in Ho and Stoll (1981) and in Glosten and Milgrom (1985) suggest that what determines the position determines the size, too. Hence, the information component of the bid-ask spread is likely to be big. The upward sloping parts of the size splines in Chapter 3 and Chapter 4 also testify the importance of information costs.

Nevertheless, the large proportion of order-processing component can never be ignored. The downward-sloping part of the size splines is consistent with the hypothesis that there are substantial fixed components of order-processing costs (Stoll 1978a; Wells 1992), and this thesis will further argue that the variation of order-processing costs largely comes from the difference in the bargaining powers of the counter parties of market makers. The market makers who initiate IMM trades are most vulnerable, since they reveal that they cannot resort to liquidity elsewhere. The investors who use brokers are unable to request the latter to make efforts to negotiate for them harder, and their orders are often preferenced. The private clients of market makers and the broker-dealers of the Exchange have bigger bargaining power against market makers, and they obtain the most favourable deals.

5.3 Final comments

The information transmission on the London Stock Exchange is satisfactory. Transaction prices reflect the information of past volumes, market makers revise their quotes in response to trade information, and inventory consideration does not dominate quote revisions. On the other hand, the Exchange has been a trading environment in favour of big investors as they hold bigger bargaining powers. While willing to offer better prices for big market players, market makers have little incentive to narrow the spread. Orders from small investors are often preferenced, and the costs are high. A central limit order book opened to all of the market participants may benefit small investors. Before it happens, individual investors may be better off by investing in pension funds or unit trusts, instead of trading with a broker.

The fact that Kalman-filter spread is a better estimator than effective spread indicates that trade information is potentially useful than the quotes. The good news is that trade prices may reflect a great deal of market information. The bad news is that quotes may not be very reliable. Nevertheless, the quotes or the effective spreads are by no means useless. Quotes reflect the fundamental price with noises, and it is unlikely that quotes are very far away from the fundamental price. If trade information is scarce, for example, for the illiquid stocks, then one has to resort to quotes for information. If it is a liquid stock under investigation, then trade information may be much more useful.

Appendix A

TECHNICAL NOTES

A.1 The Definition of the Normal Market Size

The following is an excerpt from Quality of Market Quarterly, Summer Edition 1995, page 82.

Normal Market Size is based on a percentage of the stock's average daily customer turnover in the preceding year. It is intended to represent, on average, 2.5% of the daily trading volume. To calculate the NMS, define

$$\begin{array}{c}\text{average daily value}\\ \text{of customer turnover}\end{array} = \begin{array}{c}\text{value of customer turnover}\\ \text{in previous 12 months}\end{array} /250$$

$$\begin{array}{c}\text{value of normal}\\ \text{institutional bargain}\end{array} = \begin{array}{c}\text{average daily value}\\ \text{of customer turnover}\end{array} * 2.5\%$$

$$\begin{array}{c}\text{normal institutional bargain}\\ \text{in number of shares}\end{array} = \frac{\text{value of normal institutional bargain}}{\text{closing mid price on last day of quarter}}$$

Then the normal institutional bargain is rounded to one of the 12 NMS bands according to the table below.

NMS band	shares between
500	0–667
1,000	668–1333
2,000	1,334–2,400
3,000	2.401–3,750
5,000	3,751–6,667
10,000	6,668–12,000
15,000	12,001–18,000
25,000	18,901–33,000
50,000	33,001–60,000
75,000	60,001–93,000
100,000	93,001–160,000
200,000	$\geq$ 160,001

Note that the NMS data is not available in *Transaction Data Service.* The NMS used in Chapter 3 are obtained from the archive of the Exchange.

A.2 Files in *Transaction Data Service*

The CD-ROM *Transaction Data Service* is complied by Quality of Market Group of the LSE. It contains several files, and those which are relevant to this thesis are *the Transaction Data File*, *the SEAQ Price Quote File*, *the SEAQ Best Quote Price File*, *Securities Masterfile* and *Firms File.*

A.2.1 *Transaction Data File*

This file was constructed from the complete settlement records of transactions reported by the member firms to the Exchange. The relevant fields used for this thesis are:

1. **Action.** It indicates whether the transaction is reported by the buyer or the seller.
2. Stock Exchangd Daily Official List code; **SEDOL** for short. It is a unique code for each security.

3. The **date** when the transaction occurs.
4. The **firm** who reports the transaction. It is a three-digit code.
5. The three-digit code of the **counter party** of the transaction.
6. **Bargain conditions**. It indicates various conditions related to the transaction, for example, whether the underlying stock is ex-dividended. The field is used to identify and to delete option-related trades.
7. **Positive/negative flag**. It indicates whether the **volume** field should be positive or negative. For example, the flag of the contra transactions should be indicated as negative.
8. The **volume** of the transaction, the number of shares traded in the transaction. This is the "volumes" used in Chapter 4.
9. The **value** of the transaction. The unit of the value is a penny. This field is used to calculate "volumes" in Chapter 2 and Chapter 3.
10. The **time** when the transaction occurs.
11. The **source** of the record. The record may be retrieved from either Tailsman or Sequal.
12. **Buyer deal capacity**. The capacity of the buyer of the trade. There are five different capacities: A, I, M, N, and P. See Chapter 3 for details.
13. **Seller deal capacity**. The capacity of the seller.
14. **Deal type**. It could be a standard transaction, a SAEF transaction, a contra transaction, and so on. Stock loans are identified and deleted by this field.
15. **entry date**. The date in which the transaction was reported to the Exchange.
16. The **transaction number** given by the Exchange. The number is unique for each transaction in an **entry date** for a certain stock, so

one can identify any single transaction by using the fields **SEDOL**, **entry date** and **transaction number** together.

17. The **price** at which the deal is transacted. The price is rounded to pence, so it is replaced with 100 times **value** and divided by **volume**.

18. **deal currency**. The field is used to discard the transactions not dealt in Pound Sterling.

A.2.2 *SEAQ Price Quote File*

The quote file contains the bid and ask prices and quantities of the shares quoted by the market makers. Those fields used in the investigations include:

1. The **SEDOL** of the stock.

2. The **date** of the quote.

3. **Ontime**, the time at which the quote became valid. The times in the file are rounded to seconds.

4. **Offtime**, the time at which the quote ended.

5. The three-digit code of the **market maker** who posts the code.

6. The **ask price** in a hundredth pence.

7. The **ask size**.

8. The **bid price** in a hundredth pence.

9. The **bid size**.

10. The **indicator** which indicates whether the quote is an opening quote, a closing quote, or neither of them.

A.2.3 *SEAQ Best Quote Price File*

The file is constructed by Quality of Market Group to replicate the yellow strips of the stocks. Each records contains the following fields:

1. The **SEDOL** of the stock.
2. The **date** of the quote.
3. The **time** at which the quote became valid. Unlike price quotes file, the times are rounded to minutes.
4. The **ask price.**
5. The **bid price.**
6. The **indicator** which indicates whether the quote is an opening quote, a closing quote, or neither of them.

A.2.4 *Securities Masterfile*

The file contains the relevant information about the securities traded on the Exchange. Each record includes the **SEDOL** of the security, the name of the **company** which issues the security, the **type** of the security, for example, an ordinary share, a warrant, etc., and the status of the security in **FTSE** index, for example, whether the security is included in FTSE-100 index, FTSE-250 index, and so on.

A.2.5 *Firm File*

The file contains the relevant information about the member firms of the exchange. Each record includes the **name** of the firm, the four-letter **abbreviation** of the firm, the three-digit **code** of the firm, and the time when the code becomes valid. It is not uncommon for a firm with several **codes**, and five out of twenty-eight market makers in the data set have different codes in *the Transaction Data File* and in *the SEAQ Price Quote File.* Therefore, the information in *Firm File* is extremely important for identifying the link between order flows and quote status of the same market maker (Chapter 2).

A.3 Sample sections

A.3.1 Trade data used in Chapter Four

The choice of using Glaxo Wellcome, BT and Shell's trades in Chapter 4 is arbitrary. All of them are very liquid stocks, where several transactions occur at the same time is very common. BT is chosen because it is the most heavily traded stock during the sample period. The other two stocks are chosen because they belong to different industries.

A.3.2 Stocks used in Chapter Three

Nineteen of the 39 stocks used in Chapter 3 come from the constituents of FTSE-100 index. The names of the constituents of the indices are obtained from the Datastream International. It consists of ten stocks with the biggest trading volumes (measured in the values of transaction) during 1995. Another nine stocks of FTSE-100 constituents are with the smallest trading volumes during 1995. The data of trading volumes comes from London Stock Exchange (1996), in which there is a list of 100 stocks with the biggest trading volumes during 1996. FTSE-100 stocks are not necessary to be in the first 100 stocks with the biggest trading volumes, so all of the FTSE-100 stocks which are not in the list of London Stock Exchange (1996) are assumed to have smaller trading volumes than the ones in the list. There are eighteen FTSE-100 stocks which are not in the list, ten of them are randomly chosen, and one of them (Schroders) is dropped as the number of inter-dealer trades is too small. The list of the nineteen stocks is in Table A.1.

There are two stages to select the rest of the stocks. First, order the names of the stocks of FTSE-250 in random. Take the first thirty stocks and retain those which have trades everyday, and twenty-seven stocks remain in the sample. Repeat the procedure for the rest of FTSE-All Shares stocks by ramdomly seleting forty stocks and retain twenty-eight of them. In the second stage, the stocks with the sum of IDB and IMM trades less than 150 are excluded from the analysis, and only twenty stocks are left in the sample. The list of the stock is in Table A.2.

A.3.3 Quotes and trade data in Chapter Two

Chapter 2 uses the data from February and March 1996. The Quality of Market Group provides all the data available between January and June 1996, but the data are not complete. For example, quote data before 15 January are missing. To conduct the investigations in Chapter 2, it is essential to identify the quotes posted by certain market makers and the trades executed by them, so the January data are not used. The reason of not using the data after April 1996 is similar: the quote data of a medium-size market maker are available only between January and March 1996.

Does the missing quote of the firm affect the investigations in this thesis? Table A.3 shows the daily duration of the firm being alone on the yellow strip of the stocks used in Chapter 3 between February and March 1996. The firm is not often alone on the yellow strip, so the constructed yellow strip used in Chapter 3 and Chapter 4 is not very different from the true yellow strip.

A.4 Editing transaction data

The settlement records cannot be readily applied to data analysis; data editing is essential and is described below. Much of the editing procedure is not different from those documented in Hansch and Neuberger (1993), Board and Sutcliffe (1995), and Reiss and Werner (1996, 1997).

A.4.1 Shape trades

Both buyer and seller of a trade report to the settlement system, but not all of the trades have only one buy record and only one sell record. For example, sometimes the firms may wish to report several records for a trade for the convenience of bookkeeping. The trades which are recorded by more than a pair are called "shape trades". Shape trades can be easily identified because they always have the same **trade number** within the same **data-entry day**. There are approximately 17% of records of shape trades in the sample used for Chapter 3.[1] Most of the shape trades contain four records, and one trade even has 900 records. The shape trades are supposed to contain even number of records. If the firm wishes to report the trade with n records, its counter party will also report it with n records, but it will set the volume to zero for the $n-1$ "dummy" records. For example, if firm A buys 12,000 shares from firm B and wishes to report the trades with three records of 4,000 shares each, then the firm B will report with three records as well, one record of 12,000 shares and the other two of 0 shares.

Apart from the declared shape trades, 2.4% of the customer bargains have similar features as the shape trades but they are not declared as shapes. The similar features include that those trades occurred at the same time, with the same trading parties, with the same buyer/seller capacities, and with the same price. They are regarded as potential shape trades and are merged in the same way as merging the declared shape trades.

A.4.2 Contra trades

If a fault of a trade report is discovered, then the firm and its counter party should report a "contra trade" to the Exchange to cancel the trade. The records of contra trades are supposed to be exactly the same as the previous records except (a) that the mark of contra trades will appear in the **deal type** field, (b) that the buyers and the sellers are reversed, and (c) that there will be a negative sign in the **positive/negative flag** field. Both the contra trades and the faulted trades should be removed from the data. It can be done by tracing the normal trade records identical to the contra trades except for the three features above. If all of the contra trades were reported as described, all of the contra trades and the faulted trades would be removed from the data set. This is not the case. Contra trades constitute 1.7% of the data set, only three quarters of them can be removed according to the algorithm.

Several reasons result in the failure to identify incorrect records. For

[1] Unless otherwise mentioned, all of the numbers reported in this section are produced from editing the data set of Chapter 3.

example, there may be more than one fault in a faulted trades, some of which are failed to be identified by the firms. The contra trades may have their errors, too. The normal trade and the contra trade should have the same **SEDOL**, **deal date**, **deal time**, **buyer capacity**, **seller capacity**, **price** and **volume**. Because the restrictions of the fields do not detect one quarter of the contra trades, later the restriction of capacities, time, price and volume are relaxed one by one, and another fifteen percent of the contra trades and their corrsponding normal trades are identified. However, it is important to note that the relaxation of the rule runs the risk of identify the wrong trades. There is no objective way to decide whether the record should be eliminated, and any deviation from the standard rules involves arbitrary judgement.

A.4.3 Paired trades

After the trade is agreed between the buyer and the seller, both parties report the trades to the Exchange for the settlement. After shape trades are grouped, each trade consist of exactly one buy and one sell record. Since each trade has a unique **trade number** in the **data-entry day**, so a buy and its corresponding sell record can be easily paired.

A.4.4 Sell-and-buy-back trades and put-throughs

For bookkeeping reasons, member firms may agree to report a pair of trades with the same prices and quantities where the buyer and the seller are reverse. Such trades are called "sell and buy back". All of these trades are deleted.

A put-through is a cross trade under the name of market makers. The essence of the trade is that trader A buys some shares from trader B, but on the record there are two trades: trader B sells the shares to market maker C, and C sells the same amount of shares to A at the same time. There are less than 2,000 pairs of put-throughs in the sample of Chapter 2. Put-throughs are not used in Chapter 4, and they are used in Chapter 2 and Chapter 3 only for the regressions of quote changes.

A.4.5 IDB trades

If a firm approaches an inter-dealer broker (IDB) to hit the limit order displayed on the IDB screen, the IDB will inform the firm who places the limit order, and all of the parties report the trades to the Exchange. For example, if market maker X (the hitter) hits a sell limit order placed by market maker Y (the poster), then X will report a buy record to the Exchange, Y will report a sell record, and the IDB will report a buy record (as if it buys the shares from Y) and a sell record (as if it sells the shares to X.)

There is no need to match the IDB trades or to identify the hitters of the trades if one is only interested in the inventory or the order flows of the market makers, such as in Chapter 2. On the other hand, if one wishes to compute the costs of IDB trades, such as in Chapter 3, it is essential to identify the posters and the hitters of the trades. The posters are analogous to the market makers in SEAQ in the sense that they both prepare to trade with firm prices and quantities. The hitters are analogous to the customers in SEAQ as they are both regarded to initiate the trade. The costs of the trades are often defined as the costs of the posters, and the only way to identify the poster in IDB trades from the settlement records is to match the trades. As the hitter pays the commission to the IDB and the price before the commission is often rounded, the trading party who reports the odd price can be identified as the hitter.

Unfortunately, in the IDB market it is possible for a trade to have n buyers and m sellers, and to identify the hitter and poster the trade records have to be merged. Furthermore, there is no other way to match IDB trades except to compare the **time**, the **volume** and the **value** of one record with the rest of the records at the same day, in order to find the best-matched records. The difficulty results from the inconsistency of the report of the time, and the difference in commission. The hitter pays the commission which is on average 0.045% of the **value** (of the poster's record), but it is rarely exactly so. The commission can be as low as zero or as high as 0.1%. The inconsistency in the reported time is also severe. For the trades with two or three trading parties, the reported times are often the same or only have one minute difference. For the trades with many trading parties, the difference in **time** can be as big as half an hour or even bigger. As a result, identifying the trading parties of the trades can be arbitrary. The algorithm of matching the IDB trades used in the investigations is as follows:

1. To match as many small trades as possible. For example, suppose there are two records of buys, one of 10,000 shares and the other of 20,000 shares, and there are two records of sells of exactly the same size. Then they are identified as two trades with two trading parties, not one trade with four parties.

2. The records whose **times** are identical are matched first, and then the restriction of the time difference is relaxed gradually. This approach is in favour of the records whose **times** were reported closely. It is possible, however, for the wrong trades to be matched simply because one trade is falsely reported to occur at the **time** close to the other. The exception of the rule is allowed when the numbers of trading parties are exactly the trades to be matched. For example, if there are exactly two buying and one selling records to be matched at the day, and if the sum of the volumes of the buy records equals the sum of the volumes of sell records, then the three records are merged regardless of the time.

3. The trades will be matched only if the buyers pay for the shares to the IDB at the same price, and all the sellers receive the money for the IDB at another price. There is no exception for the rule.

4. The regulation of the Exchange states IDBs should avoid letting the market maker hit the limit order placed by itself. Without knowing if the rule is strictly enforced, the buy and sell record of the same market maker are not matched if less than five market makers are involved in a trade.

99% of the IDB trades can be identified by this algorithm.

A.4.6 Portfolio trades

Occasionally the customers may find it desirable to buy or sell several stocks as a portfolio. This is called "basket" or "portfolio" trading (Wells 1995). Because the trades are executed as a basket, the individual prices may be arbitrary. TDS does not record whether the trades belong to a basket, so only "potential" basket trades can be identified. The trades are identified as in a potential basket if the trading parties, trading capacities, and the reported times are the same, and a basket should consist of at least ten different stocks. Moreover, neither IDB trades nor IMM trades are identified as potential basket trades. Unlike Hansch et al. (1999), the prices of potential basket trades are not restricted to be equal to the mid-touches, but the percentage effective spreads of the trades, defined as (1.1), have to be the same. About 3% of the trades are identified as potential basket trades and are removed from the analysis.

A.4.7 Capacities

The trading parties report to the Exchange whether they act as principals or agents in the trades, and Quality of Market Group expand the capacity into five categories. The capacities reported by the member firms may not be correct, and it is extremely difficult if not impossible to correct the wrongly reported capacities. For the data used in this thesis, only the capacities of IDB and market makers are examined. There are only four inter dealer brokers in the markets, and their sole function is to act as IDBs. None of the other firms act as IDBs, so any "I" capacity reported by any other member firm is incorrect. Regarding the records of market makers, the quote data and IDB records are used to determine which firms may have acted as market makers. The suspicious records are sent to Quality of Market Group to confirm the validity of the capacities of trading parties.

A.4.8 Time of trades

Member firms report the **time** that they believe when the trade is executed. However, the buyers and sellers do not report the same **time** in about a third of trades. The differences are not always very large. Excluding IDB and contra records, only 1.5% of the records differ in time by more than 20 minutes, and the mean of the time differences, if there is any, is around 564 seconds.[2]. For the IDB trades, the times reported are the same by the market maker and the IDB in only 46.4% of the records, are one minute apart in 39.3% of records, and the means of the difference is 256 seconds. To determine the time of the trade, the approach is similar to Board and Sutcliffe (1995), which consists of the following rules:

1. The mandatory quote period of the market maker is between 8:30 and 16:30, and the trades which occur well beyond the time are rare. Therefore, if one party report the **time** which is between 8:00 and 17:00 and the other is not, the **time** of the former is used.

2. If the rule above is not applicable, and if one of the trading party is a market maker and the other party is an inter-dealer broker, then the **time** reported by the IDB is used. The rationale behind this rule is the **times** reported by IDBs are apparently more reliable than those by market makers. For example, the sample of Chapter 3 consists of 339 pairs of IDB trades of which the **times** reported by the IDBs and the market makers are at least one hour apart. Of those records, the **times** reported by IDBs are always between 8:00 and 17:00, whereas one of the **time** by market makers is after 17:00 and another 35 are before 8:00. It may imply the **times** reported by the IDBs are more reliable. Furthermore, if the **time** by IDBs are used, 28 trades are outside the touch, whereas 126 trades are outside the touch if the **times** by market makers are used.[3]

[2] The mean of the time difference is large because there are quite a few trades with apparently wrong **time**. For example, the buyer reports the trade is executed at 2pm, while the seller reports that it is executed at 2am.

[3] If the touch is not available at the time of reporting, the nearest touch is used.

3. If the rules above are not applicable, and if one of the trading party is a market maker and the other party is not, then the **time** reported by the market maker is used. The rationale behind this rule is the **time** reported by market makers are apparently more reliable. For example, the sample of Chapter 3 consists of 5219 pairs of SEAQ trades of which the **times** reported by the market makers and their counter parties are at least one hour apart. Of those records, the number of records which the **times** are before 08:00 and after 17:00 are, respectively, 90 and 33 by the market makers, and 224 and 78 by non-market makers. Furthermore, if the **time** by market makers is used, 817 trades are outside the touch, whereas 1620 trades are outside the touch if the times by the other parties are used.

4. If both or none of the trading parties are market makers, then sellers' **time** is used. The rule is the legal position by the Exchange whenever there is a dispute between the trading parties.

Moreover, the IDBs may reported different **times** for the same IDB trades, that is, the **time** of the trade with the hitter may not be the same as the **time** of the trade with the poster. The rules of determining the time of the merged IDB trades are:

1. An IDB trade may include m buyers and n sellers. If more than a half of the $(m+n)$ records are of the same **time**, then the **time** is chosen. The rule also applies when exactly half of the records are of the same **time** and the other half are of different **times**.

2. If exactly a half of the records are of one **time** and the other half are of another time, then the **time** which is not far beyond the trading hours is chosen. If both or neither are beyond the trading hours, the earlier **time** is chosen.

3. If neither of the rules above is applicable, then the **time** in the middle is chosen. For example, if there are five trading parties with five different **time**, then the third earliest **time**. If there are four trading parities with four different **time**, then the second earliest **time** is chosen.

Note that trading parties report to the Exchange the time when they believe the trade is executed, not the time when the trade is stamped, so there is no need to match the trade with the quote a few minutes earlier or later as Hasbrouck (1988) does. The experiment by Hansch et al. (1999) to match the trade and quote with different time shows their regression results are not improved by the change in time. Finally, de Jong, Nijman, and Roëll (1995) and Reiss and Werner (1997) believe the trades executed outside the touch have wrong reported time and change the time. They justify the approach by the best execution rule of the Exchange. However, this thesis holds the view that whether the best execution is strictly enforced remains to be investigated, that the change in reported time distorts the data, and that the time should never altered according to the touch.

A.5 Editing quote data

Not surprisingly, there are irregularities with quote data, too. Table A.4 is an example of how quote data are supposed to appear.

1. The **ontime** of the opening quote is the earliest of the **ontime** of the day, which is usually between 8:00 and 8:30. Late opening occurs occasionally.

2. The **ontime** of the closing quote is the latest of the **ontime** of the day, which is usually between 16:30 and 17:00. The **offtime** of the closing quote is 23:59:59, and the prices and quantities of the closing quote are zero. However, in 0.04% of the records, the market makers closed the quotes at the normal time, apparently reopened later, say at 18:00, and the quotes remained opened until the end of the day. Moreover, the market makers may not close the quotes at the end of the day, which happened in 0.03% of the records, and in such a case the last quotes have non-zero prices and quantities until mid-night. Late quotes are never used and do no affect the investigation of this thesis.

3. Between the opening quote and the closing quote, the times of the quotes are one after another; the **offtime** of one quote is just one second ahead of the **ontime** of its next quote. This regularity is never violated in the original data.

4. The quote should have the ask price bigger than the bid price. The quotes which have bids bigger than asks are removed, and the next quotes are assumed to start at the times when the deleted quotes began to be valid. The continuity of the quotes maintains, but the assignment of the time is arbitrary.

A.6 No change for best quote data

There is no revision of the best quote data. When the quote data is available, the data is used to construct the best quote. Otherwise the best quote data from the TDS is used. It appears that the best quote from the *Quote File* is more reliable than that from the *Best Quote File.* The main reason is that the former is recorded in seconds, and the latter is in minutes. When the best quote changes within one minute, the best quote data shows several observations of the same time stamp. Therefore, if the best quote data is absolutely necessary, for example, to match the trade before 15 January 1996, then one of the several "best quotes" has to be randomly selected to match the price.

Table A.1: FTSE-100 samples of Chapter Three

		Data availability		
		trade	quote	best-quote
1	British Gas	02/01–28/06	15/01–28/06	02/01–28/06
2	British Petroleum	02/01–28/06	15/01–28/06	02/01–28/06
3	British Telecommunications	02/01–28/06	15/01–28/06	02/01–28/06
4	BTR	02/01–28/06	15/01–28/06	02/01–28/06
5	Cable & Wireless	02/01–28/06	15/01–28/06	02/01–28/06
6	Glaxo Wellcome	02/01–28/06	15/01–28/06	02/01–28/06
7	Hanson	02/01–28/06	15/01–28/06	02/01–28/06
8	HSBC Holdings	02/01–28/06	15/01–28/06	02/01–28/06
9	Shell Transport & Trading	02/01–28/06	15/01–28/06	02/01–28/06
10	SmithKline Beecham	15/04–28/06	15/04–28/06	n.a.
11	Bank of Scotland	02/01–28/06	15/01–28/06	02/01–28/06
12	Burton	02/01–28/06	15/01–28/06	02/01–28/06
13	Cookson	02/01–28/06	15/01–28/06	02/01–28/06
14	Courtaulds	02/01–28/06	15/01–28/06	02/01–28/06
15	Guardian Royal Exchange	02/01–28/06	15/01–28/06	02/01–28/06
16	Next	02/01–28/06	15/01–28/06	02/01–28/06
17	Rentokil	02/01–28/06	n.a.	02/01–28/06
18	TI Group	02/01–28/06	n.a.	02/01–28/06
19	United Utilities	02/01–28/06	15/01–28/06	02/01–28/06

Table A.2: FTSE-All Share samples of Chapter Three

		Data availability		
		trade	quote	best-quote
20	BICC	02/01–28/06	15/01–28/06	02/01–28/06
21	Booker	02/01–28/06	15/01–28/06	02/01–28/06
22	British Land Co	02/01–28/06	15/01–28/06	02/01–28/06
23	Bunzl	02/01–28/06	15/01–28/06	02/01–28/06
24	English China Clays	02/01–28/06	15/01–28/06	02/01–28/06
25	Foreign & Col			
	Invest Trust	02/01–28/06	15/01–28/06	02/01–28/06
26	G.T. Japan Invest Trust	02/01–28/06	15/01–28/06	02/01–28/06
27	Govett Strategic Invest Trust	02/01–28/06	15/01–28/06	02/01–28/06
28	Hepworth	02/01–28/06	15/01–28/06	02/01–28/06
29	Lonrho	02/01–28/06	15/01–28/06	02/01–28/06
30	Lucas Industries	02/01–28/06	15/01–28/06	02/01–28/06
31	MEPC	02/01–28/06	15/01–28/06	02/01–28/06
32	MFI Furniture	02/01–28/06	15/01–28/06	02/01–28/06
33	Morgan Crucible	02/01–28/06	15/01–28/06	02/01–28/06
34	Scottish American Invest	02/01–28/06	15/01–28/06	02/01–28/06
35	TR Smaller Companies			
	Invest Trust	02/01–28/06	15/01–28/06	02/01–28/06
36	Unichem	02/01–28/06	15/01–28/06	02/01–28/06
37	Vickers	02/01–28/06	15/01–28/06	02/01–28/06
38	WPP Group	02/01–28/06	15/01–28/06	02/01–28/06
39	Yorkshire Water	02/01–28/06	05/01–28/06	02/01–28/06

Table A.3: Price leadership of the firm with missing quotes

The time is in (hour:minute).

FTSE-100 samples	time	other samples	time
1 British Gas	0:17	20 BICC	0:05
2 British Petroleum	0:10	21 Booker	0:01
3 British Telecommunications	0:03	22 British Land Co	0:05
4 BTR	0:18	23 Bunzl	0:08
5 Cable & Wireless	0:03	24 English China Clays	0:03
6 Glaxo Wellcome	0:02	25 Foreign & Col Invest Trust	0:37
7 Hanson	0:11	26 G.T. Japan Invest Trust	0:12
8 HSBC Holdings	1:33	27 Govett Strategic Invest Trust	0:10
9 Shell Transport & Trading	0:07	28 Hepworth	0:44
11 Bank of Scotland	0.22	29 Lonrho	0:17
12 Burton	0:06	30 Lucas Industries	1:02
13 Cookson	0:08	31 MEPC	0:26
14 Courtaulds	0:00	32 MFI Furniture	0:08
15 Guardian Royal Exchange	0:04	33 Morgan Crucible	0:15
16 Next	0:14	34 Scottish American Invest	0:02
19 United Utilities	0:02	35 TR Smaller Companies Invest Trust	0:05
		37 Vickers	0:51
		38 WPP Group	0:00
		39 Yorkshire Water	0:06

Table A.4: An example of "Standard" quote data

SEDOL	FIRM	DATE	ON TIME	OFF TIME	BIDP	BIDQ	ASKP	ASKQ
0067889	000	01FEB96	8:24:30	8:24:50	330	100000	335	100000
0067889	000	01FEB96	8:24:51	8:40:10	328	100000	333	100000
0067889	000	01FEB96	8:40:11	8:44:11	327	100000	332	100000
0067889	000	01FEB96	8:44:12	8:53:20	325	100000	330	100000
0067889	000	01FEB96	8:53:21	9:01:52	323	100000	328	100000
0067889	000	01FEB96	9:01:53	12:50:55	325	100000	330	100000
0067889	000	01FEB96	12:50:56	16:32:46	327	100000	332	100000
0067889	000	01FEB96	16:32:47	23:59:59	0	0	0	0

Appendix B

STATISTICAL NOTES

The following are complied from the appendices of Koopman and Lai (1998).

B.1 Regression spline functions

The regression spline function is defined as a smooth function through the data points y_t which are a response to the scalar series x_t, for which $x_t < x_{t+1}$ and $t = 1, \ldots, n$. The spline model is

$$y_t = \theta(x_t) + \varepsilon_t, \qquad E(\varepsilon_t) = 0, \qquad var(\varepsilon_t) = \sigma^2,$$

where $\theta(\cdot)$ is a smooth function which is based on $k+1$ knot points $(x_0^\dagger, y_0^\dagger), \ldots, (x_k^\dagger, y_k^\dagger)$. The smoothness of $\theta(\cdot)$ is created by setting its second derivative with respect to x as a linear function of $k+1$ coefficients, that is

$$\theta_i''(x) = [(x_i^\dagger - x)/d_i]a_{i-1} + [(x - x_{i-1}^\dagger)/d_i]a_i$$

with $d_i = x_i^\dagger - x_{i-1}^\dagger$ and $\theta_i(x) = \theta(x)$ for $x_{i-1}^\dagger < x < x_i^\dagger$ and $i = 1, \ldots, k$. The $k+1$ coefficients a_i are assumed fixed and they can be identified by solving a linear set of equations. These regression spline equations are obtained as follows:

1. use $\theta_i''(x)$ and standard integration rules to get expressions for $\theta_i(x)$;
2. enforce the spline function $\theta_i(x)$ at $x = x_i^\dagger$ to be equal to the known value of $y_i^\dagger$;
3. restrict the first derivative to be continuous by enforcing $\theta_i'(x_i^\dagger) = \theta_{i+1}'(x_i^\dagger)$ for $i = 1, \ldots, k-1$.

Step 2 leads to a linear expression for $\theta_i(x)$ in terms of $y_i^\dagger$ and a_i, for $i = 0, \ldots, k$. Step 3 leads to $k-1$ linear equations for the $k+1$ coefficients $a_0, \ldots, a_k$ in terms of $y_0^\dagger, \ldots, y_k^\dagger$. The "natural" restrictions $a_0 = a_k = 0$ allow solving this linear system with respect to the remaining coefficients a_i for $i = 1, \ldots, k-1$. The spline function can now be fully expressed in terms of $y_0^\dagger, \ldots, y_k^\dagger$ by

$$\theta(x_t) = \theta_i(x_t) = b_{0,t}y_0^\dagger + \ldots + b_{k,t}y_k^\dagger, \qquad x_{i-1}^\dagger < x_t < x_i^\dagger, \qquad , t = 1, \ldots, n,$$

where the weights $b_{0,t}, \ldots, b_{k,t}$ depend on the knot positions $x_0^\dagger, \ldots, x_k^\dagger$ and the value for (or the position of) x_t. For a given set of values $y_0^\dagger, \ldots, y_k^\dagger$, the spline function can be computed for any $x_0^\dagger < x < x_k^\dagger$. The regression spline can be expressed as

$$\theta(x_t) = b_t' y^\dagger,$$

where $b_t = (b_{0,t}, \ldots, b_{k,t})'$ and $y^\dagger = (y_0^\dagger, \ldots, y_k^\dagger)'$. Consequently, the spline model can be expressed as the standard regression model

$$y_t = b_t' y^\dagger + \varepsilon_t, \tag{B.1}$$

where parameter vector $y^\dagger$ can be estimated by least squares techniques. In the case of model (4.1), the parameter vectors for the two different splines are estimated by generalised least squares. More details are given by Poirier (1973, 1976). The generalisation of time-varying regression splines within the state space framework are developed by Harvey and Koopman (1993).

B.2 Kalman filter smoother

Consider the state space model (4.5) and (4.6). The Kalman filter evaluates the minimum mean squared linear estimator of the state vector, conditional on 'past' observations, together with its variance matrix. Follow the treatment of Koopman and Durbin (1998) and exclude the regression vectors λ_x and λ_w from the state space model. Define

$$\begin{aligned} a_{t,1} &= E\left(\alpha_t Y_{t-1}\right), & P_{t,1} &= var\left(\alpha_t Y_{t-1}\right), \\ a_{t,i} &= E\left(\alpha_t Y_{t-1}, y_{t,1}, \ldots, y_{t,i-1}\right), & P_{t,i} &= var\left(\alpha_t Y_{t-1}, y_{t,1}, \ldots, y_{t,i-1}\right), \end{aligned}$$

for $i = 2, \ldots, N_t$, where

$$Y_t = \{y_{1,1}, \ldots, y_{1,N_1}, y_{2,1}, \ldots, y_{t,N_t}\}.$$

The filtering equations are given by

$$a_{t,i+1} = a_{t,i} + K_{t,i}F_{t,i}^{-1}v_{t,i}, \qquad P_{t,i+1} = P_{t,i} - K_{t,i}F_{t,i}^{-1}K_{t,i}', \tag{B.2}$$

where

$$v_{t,i} = y_{t,i} - Z_{t,i}a_{t,i}, \qquad F_{t,i} = Z_{t,i}P_{t,i}Z_{t,i}' + \sigma_{t,i}^2, \qquad K_{t,i} = P_{t,i}Z_{t,i}', \tag{B.3}$$

for $i = 1, \ldots, p_t$ and $t = 1, \ldots, n$. This formulation has $v_{t,i}$ and $F_{t,i}$ as scalars and $K_{t,i}$ as a column vector. The transition from time t to time $t+1$ is achieved by the relations

$$a_{t+1,1} = T_t a_{t,p_t+1}, \qquad P_{t+1,1} = T_t P_{t,p_t+1} T_t' + R_t Q_t R_t'. \tag{B.4}$$

These forward recursions are initialised by $a_{1,1} = a$ and $P_{1,1} = P$ as given by (4.7).

Minimum mean squared linear estimators using all observations Y_n are evaluated by a smoothing algorithm which require output of the Kalman filter. The basic smoothing recursions operate backwards and the equations are given by

$$\begin{aligned} r_{t,i-1} &= Z_{t,i}'F_{t,i}^{-1}v_{t,i} + L_{t,i}'r_{t,i}, & N_{t,i-1} &= Z_{t,i}'F_{t,i}^{-1}Z_{t,i} + L_{t,i}'N_{t,i}L_{t,i}, \\ r_{t-1,N_t} &= T_{t-1}'r_{t,0}, & N_{t-1,N_t} &= T_{t-1}'N_{t,0}T_{t-1}, \end{aligned} \tag{B.5}$$

where $L_{t,i} = I - K_{t,i}Z_{t,i}F_{t,i}^{-1}$, for $i = N_t, \ldots, 1$ and $t = n, \ldots, 1$. The initialisations are $r_{n,N_n} = 0$ and $N_{n,N_n} = 0$. The equations for r_{t-1,N_t} and N_{t-1,N_t} do not apply for $t = 1$.

The output of recursions (B.5) can be used to construct the smoothed estimator of the disturbances, that is, for example, $\widehat{\varepsilon}_t = E(\varepsilon_t Y_n)$, together with their corresponding variances. The smoothed disturbances are computed by

$$\begin{aligned} \widehat{\varepsilon}_{t,i} &= \sigma_{t,i}^2 F_{t,i}^{-1}\left(v_{t,i} - K_{t,i}'r_{t,i}\right), & var\left(\widehat{\varepsilon}_{t,i}\right) &= \sigma_{t,i}^4 F_{t,i}^{-2}\left(F_{t,i} + K_{t,i}'N_{t,i}K_{t,i}\right), \\ \widehat{\eta}_t &= Q_t R_t' r_{t,0}, & var\left(\widehat{\eta}_t\right) &= Q_t R_t' N_{t,0} R_t Q_t, \end{aligned} \tag{B.6}$$

for $t = n, \ldots, 1$. The proofs and more general results for smoothed disturbances are given by Koopman (1993).

The smoothed state vector $\hat{\alpha}_t = E(\alpha_t Y_n)$ and variance matrix $V_t = var(\alpha_t Y_n)$ also use (B.5) and they can be evaluated by

$$\hat{\alpha}_t = a_t + P_t r_{t-1}, \qquad V_t = P_t - P_t N_{t-1} P_t, \tag{B.7}$$

for $t = n, \ldots, 1$. A substantial amount of additional storage space is required for a_t and P_t. Proofs of (B.5) and (B.7) are given by de Jong (1988) and Kohn and Ansley (1988). A more efficient algorithm for calculating the smoothed estimator of the state vector only is given by

$$\widehat{\alpha}_{t+1} = T_t \widehat{\alpha}_t + R_t \widehat{\eta}_t, \qquad t = 1, \ldots, n,$$

with $\widehat{\alpha}_1 = a + P r_0$ and $\widehat{\eta}_t$ is given by (B.6); see Koopman (1993) for a discussion.

The Kalman filter smoother also provides a general procedure to handle missing observations in time series. When no observations are available for a certain time period τ, or a sequence of time periods, the dimension $p_\tau = 0$ and the updating equation (B.4) is applied. The smoothing recursions adjust naturally to this situation. Compared to other treatments of missing observations in statistics, this approach is very simple.

B.3 Maximum likelihood estimation

Consider the state space model (4.5) and (4.6) with system matrices and vectors depending on the parameter vector ψ. For a given vector ψ, the output of the Kalman filter is used to construct the likelihood function. Harvey (1993) shows how the likelihood function of the state space model with normal distributed disturbances can be calculated via the prediction error decomposition. The log-likelihood function is given by

$$\log L(\psi) = \text{constant} - \frac{1}{2} \sum_{t=1}^{n} \sum_{i=1}^{p_t} \log F_{t,i} + v_{t,i}^2 / F_{t,i},$$

where $v_{t,i}$ and $F_{t,i}$ are obtained from the (B.3) which depend on parameter vector ψ. In the context of state space models, maximum likelihood estimation refers to numerically optimising the log-likelihood function with respect to ψ.

BIBLIOGRAPHY

Admati, A. and P. Pfleiderer (1988). A theory of intraday trading patterns: Volume and price variability. *Review of Financial Studies 1*(1), 3–40.

Affleck-Graves, J., S. P. Hegde, and R. E. Miller (1994). Trading mechanisms and the components of the bid-ask spread. *Journal of Finance 49*(4), 1471–1488.

Amihud, Y. and H. Mendelson (1980). Dealership market: Market making with inventory. *Journal of Financial Economics 8*(1), 31–53.

Anderson, B. and J. Moore (1979). *Optimal Filtering* (Second ed.). Englewood Cliffs, New Jersey: Prentice Hall.

Anderson, T. G. and T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: Uncover the long-run in high frequency returns. *Journal of Finance 52*(3), 975–1005.

Black, B. S. (1996). Comments. In A. Lo (Ed.), *The industrial organisation and regulation of the securities industries*, Chapter 5, pp. 171–174. London: University of Chicago Press.

Board, J. L. G., A. Fremault Vila, and C. Sutcliffe (1996). Market maker heterogeneity and order preferencing. Manuscript, London School of Economics.

Board, J. L. G., A. Fremault Vila, and C. M. S. Sutcliffe (1997). Market maker performance: the search for fair weather market makers. Discussion Paper 276, Financial Markets Group, London School of Economics.

Board, J. L. G. and C. Sutcliffe (1995). The effects of trade transparency in the London Stock Exchange. Project Report Jointly Commissioned by the London International Financial Futures snd Options Exchange and the London Stock Exchange.

Board, J. L. G. and S. Wells (1998). Sets: Still exactly the same? Manuscript, London School of Economics.

Bollerslev, T. and I. Domowitz (1993). Trading patterns and prices in the interbank foreign-exchange market. *Journal of Finance 48*(4), 1421–1443.

Bollerslev, T., I. Domowitz, and J. Wang (1997). Order flow and the bid-ask spread: An empirical probability model of screen-based trading. *Journal Economic Dynamics and Control 21*(8-9), 1471–1491.

Bollerslev, T. and M. Melvin (1994). Bid-ask spreads and volatility in the foreign exchange market—an empirical analysis. *Journal of International Economics 36*(3-4), 355–372.

Brock, W. A. and A. W. Kleidon (1992). Periodic market closure and trading volume: a model of intraday bids and asks. *Journal of Economic Dynamics and Control 16*(3), 451–489.

Chan, K., W. G. Christie, and P. Schultz (1995). Market structure and the intraday pattern of bid-ask spreads for NASDAQ securities. *Journal of Business 68*(1), 35–60.

Chan, K., Y. P. Chung, and H. Johnson (1995). The intraday behavior of bid-ask spreads for NYSE stocks and CBOE options. *Journal of Financial and Quantitative Analysis 30*(3), 329–346.

Christie, W. G., J. H. Harris, and P. Schultz (1994). Why do NASDAQ market makers stop avoiding odd-eighth quotes. *Journal of Finance 49*(5), 1841–1860.

Christie, W. G. and P. Schultz (1994). Why do NASDAQ market makers avoid odd-eighth quotes. *Journal of Finance 49*(5), 1813–1840.

Chrodia, T. and A. Subrahmanyam (1995). Market making, the tick size, and payment-for-order flow: Theory and evidence. *Journal of Business 68*(4), 543–576.

Chung, Y. P. (1991). A transactions data test of stock index futures market — efficiency and index arbitrage profitability. *Journal of Finance 46*(5), 1791–1809.

Copeland, T. E. and D. Galai (1983). Information effects and the bid-ask spread. *Journal of Finance 38*(5), 1457–1649.

Darrough, M. N. (1993). Disclosure policy and competition - Cournot vs Bertrand. *American Economic Review 68*(3), 534–561.

de Jong, F., T. Nijman, and A. Roëll (1995). A comparison of the cost of trading french shares on the Paris Bourse and on SEAQ International. *European Economic Review 39*(7), 1277–1301.

de Jong, P. (1988). A cross validation filter for time series models. *Biometrika 75*(3), 594–600.

de Jong, P. (1991). The diffuse Kalman filter. *Annals of Statistics 19*(2), 1073–1083.

Demsetz, H. (1968). The cost of transacting. *Quarterly Journal of Economics 82*(1), 33–53.

Dennert, J. (1993). Price competition between market makers. *Review of Economic Studies 60(3)*(204), 735–751.

Diamond, D. W. and R. E. Verrecchia (1991). Disclosure, liquidity, and the cost of capital. *Journal of Finance 46*(4), 1325–1359.

Dutta, P. K. and A. Madhavan (1997). Competition and collusion in dealer markets. *Journal of Finance 52*(1), 245–276.

Easley, D., N. M. Kiefer, and M. O'Hara (1997). One day in the life of a very common stock. *Review of Financial Studies 10*(3), 805–835.

Easley, D. and M. O'Hara (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics 19*(1), 69–90.

Easley, D. and M. O'Hara (1992). Time and the process of security price adjustment. *Journal of Finance 47*(2), 577–605.

Foster, F. D. and S. Viswanathan (1990). A theory of the interday variations in volume, variance, and trading costs in securities markets. *Review of Financial Studies 3*(4), 593–624.

Franks, J. and S. Schaefer (1995). Equity market transparancy on the London Stock Exchange. *Journal of Applied Corporate Finance 8*(1), 70–77.

French, K. R. and R. Roll (1986). Stock-return variances: the arrival of information and the reaction of trades. *Journal of Financial Economics 17*(1), 5–26.

George, T. J., G. Kaul, and M. Nimalendran (1991). Estimation of the bid-ask spread and its components: a new approach. *Review of Financial Studies 4*(4), 623–656.

Glosten, L. R. (1989). Insider trading, liquidity, and the role of the monopolist specialist. *Journal of Business 62*(2), 211–236.

Glosten, L. R. and L. E. Harris (1988). Estimating the components of the bid-ask spread. *Journal of Financial Economics 21*(1), 123–142.

Glosten, L. R. and P. Milgrom (1985). Bid, ask and transactions prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71–100.

Goodhart, C. A. and L. Figliuoli (1991). Every minute counts in financial markets. *Journal of International Money and Finance 10*(1), 23–52.

Grossman, S. J. and M. H. Miller (1988). Liquidity and market-structure. *Journal of Finance 43*(3), 617–637.

Hansch, O., N. Naik, and S. Viswanathan (1998). Does inventory matter in dealership markets: Evidence from the London Stock Exchange. *Journal of Finance 53*(4).

Hansch, O., N. Naik, and S. Viswanathan (1999). Preferencing, internalization, best execution, and dealer profits. *Journal of Finance forthcoming.*

Hansch, O. and A. Neuberger (1993). Block trading on the London Stock Exchange. Working Paper 182, Institute of Finance and Accounting, London Business School.

Harris, L. E. (1991). Stock price clustering and discreteness. *Review of Financial Studies 4*(3), 389–415.

Harris, L. E. (1994). Minimum price variations, discrete bid-ask spreads, and quotation sizes. *Review of Financial Studies 7*(1), 149–178.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Harvey, A. C. (1993). *Time Series Models* (Second ed.). London: Harvester Wheatsheaf.

Harvey, A. C. and S. J. Koopman (1993). Forecating hourly electricity demand using time-varying splines. *Journal of the American Statistical Association 88*(424), 1228–1236.

Hasbrouck, J. (1988). Trades, quotes, inventories and information. *Journal of Financial Economics 22*(2), 229–252.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance 46*(1), 179–207.

Hasbrouck, J. (1997). The dynamics of discrete bid and ask quotes. Working paper, Stern School of Business, New York University.

Hasbrouck, J., G. Sofianos, and D. Sosebee (1993). New York Stock Exchange systems and trading procedures. Working Paper 93-01, New York Stock Exchange.

Hausman, J., A. Lo, and A. C. MacKinlay (1992). An ordered probit analysis of transaction stock prices. *Journal of Financial Economics 31*(3), 319–379.

Hay, D. A. and D. J. Morris (1991). *Industrial Economics and Organization.* New York: Oxford University Press.

Ho, T. and H. R. Stoll (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics 9*(1), 47–73.

Ho, T. and H. R. Stoll (1983). The dynamics of dealer markets under competition. *Journal of Finance 47*(4), 1053–1074.

Huang, R. D. and H. R. Stoll (1996). Dealer versus auction markets: a pair comparison of execution costs on NASDAQ and the NYSE. *Journal of Financial Economics 41*(3), 313–357.

Huang, R. D. and H. R. Stoll (1997). The components of bid-ask spread: a general approach. *Review of Financial Studies 10*(4), 995–1034.

Jain, P. C. and G.-H. Joh (1988). The dependence between hourly prices and trading volume. *Journal of Financial and Quantitative Analysis 23*(3), 269–283.

Johnston, J. (1984). *Econometric Methods* (Third ed.). Singapore: McGraw-Hill.

Jones, C. M., G. Kaul, and M. L. Lipson (1994). Transactions, volume and volatility. *Review of Financial Studies 7*(4), 631–651.

Kandel, E. and L. M. Marx (1997). Nasdaq market structure and spread patterns. *Journal of Financial Economics 45*(1), 61–89.

Kohn, R. and C. Ansley (1988). A fast algorithm for signal extraction, influence and cross-validation in state space models. *Biometrika 76*(1), 65–79.

Koopman, S. J. (1993). Disturbance smoother for state space models. *Biometrika 80*(1), 117–126.

Koopman, S. J. (1997). Exact initial Kalman filtering and smoothing for nonstationary time series models. *Journal of the American Statistical Association 92*(440), 1630–1638.

Koopman, S. J. and J. Durbin (1998). Fast filtering and smoothing for multivariate state space models. Working paper, Tilburg University.

Koopman, S. J. and H. N. Lai (1998). Measuring bid-ask spread in competitive dealership markets. Discussion Paper 9832, Center for Economic Research, Tilburg University.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335.

Kyle, A. S. and F. A. Wang (1997). Speculation duopoly with agreement to disagree: Can overconfidence survive the market test? *Journal of Finance 52*(6), 2073–2090.

Lee, C. M. C. (1993). Market integration and price execution for NYSE-listed securities. *Journal of Finance 48*(3), 1009–1038.

Lee, C. M. C., B. Mucklow, and M. J. Ready (1993). Spreads depths and the impact of earnings information: An intraday analysis. *Review of Financial Studies 6*(2), 345–374.

Lehman, B. N. and D. M. Modest (1994). Trading liquidity on the Tokyo Stock Exchange: a bird's-eye view. *Journal of Finance 49*(3), 951–984.

Locke, P. R. and P. C. Venkatesh (1997). Futures market transaction costs. *Journal of Futures Markets 17*(2), 229–245.

London Stock Exchange (1996). Stock exchange fact book 1996. London.

London Stock Exchange (1997). Fact file 1997. London.

Madhavan, A. (1992). Trading mechanisms in security markets. *Journal of Finance 47*(2), 607–642.

Madhavan, A., M. Richardson, and M. Roomans (1997). Why do security prices change? a transaction-level analysis of NYSE stocks. *Review of Financial Studies 10*(4), 1035–1064.

Madhavan, A. and S. Smidt (1991). A Bayesian model of intraday specialist trading. *Journal of Financial Economics 30*(1), 99–134.

Manrique, A. and N. Shephard (1997). Likelihood analysis of a discrete bid/ask price model for a common stock. Manuscript, Oxford University.

Martinson, J. (1998). Exchange's electronic order book fails to win friends: Since its introduction, only 30 per cent of FTSE 100 shares traded have gone through SETS. *The Financial Times 19 August*, 9.

McInish, T. H. and R. A. Wood (1992). An analysis of intraday patterns in bid/ask spreads for NYSE stocks. *Journal of Finance 47*(2), 753–764.

Naik, N., A. Neuberger, and S. Viswanathan (1999). Trade disclosure regulation in markets with negotiated trades. *Review of Financial Studies Forthcoming.*

Neal, R. (1992). A comparison of transaction costs between competitive market maker and specialist market structures. *Journal of Business 65*(3), 317–334.

O'Hara, M. (1995). *Market Microstructure Theory.* Cambridge, Massachusetts: Blackwell.

O'Hara, M. and G. Oldfield (1986). The microeconomics of market making. *Journal of Financial and Quantitative Analysis 21*(1), 361–376.

Pagano, M. and A. Roëll (1992). Auction and dealership markets — what is the difference. *European Economic Review 36*(2-3), 613–623.

Park, H. Y. (1993). Trading mechanisms and price volatility — spot versus futures. *Review of Economics and Statistics 75*(1), 175–179.

Peterson, M. A. and D. Fialkowski (1994). Posted versus effective spreads: Good prices or bad quotes? *Journal of Financial Economics 35*(3), 269–292.

Poirier, D. J. (1973). Piecewise regression using cubic splines. *Journal of American Statistical Associations 68*(343), 515–524.

Poirier, D. J. (1976). *The Econometrics of Structural Change with Special emphasis on Spline Functions.* Amsterdam: North-Holland.

Reiss, P. C. and I. M. Werner (1996). Transaction costs in dealer markets: Evidence from the London Stock Exchange. In A. Lo (Ed.), *The Industrial Organisation and Regulation of the Securities Industries*, Chapter 5, pp. 125–169. London: University of Chicago Press.

Reiss, P. C. and I. M. Werner (1997). Interdealer trading: Evidence from London. Manuscript, Stanford University.

Reiss, P. C. and I. M. Werner (1998). Does risk sharing motivate interdealer trading? *Journal of Finance 53*(5), 1657–1703.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance 39*(4), 1127–1139.

Saporta, V. (1997). Which inter-dealer market prevails? an analysis of inter-dealer trading in opaque markets. Manuscript, Bank of England.

Securities and Investments Board (1995b). *Regulation of the United Kingdom Equity Markets.* London: SIB.

Securities and Investments Board (1995a). *Regulation of the United Kingdom Equity Markets: Market Views: a Digest of Responses to SIB'S Discussion Paper.* London: SIB.

Sheikh, A. M. and E. I. Ronn (1994). A characterization of the daily and intraday behavior of returns on options. *Journal of Finance 49*(2), 557–579.

Slezak, S. L. (1994). A theory of the dynamics of security returns around market closures. *Journal of Finance 49*(4), 1163–1211.

Snell, A. and I. Tonks (1995). Determinants of price quote revisions on the London Stock Exchange. *Economic Journal 105*(428), 77–94.

Snell, A. and I. Tonks (1998). Testing for asymmetric information and inventory effects in market maker behaviour on the London Stock Exchange. *Journal of Empirical Finance 5*(1), 1–26.

Stoll, H. R. (1978a). The pricing of security dealer services: An empirical study of NASDAQ stocks. *Journal of Finance 33*(4), 1153–1172.

Stoll, H. R. (1978b). The supply of dealer services in securities markets. *Journal of Finance 33*(4), 1133–1151.

Stoll, H. R. (1989). Inferring the components of the bid-ask spread: Theory and empirical tests. *Journal of Finance 44*(1), 115–134.

Sutton, J. (1991). *Sunk Costs and Market Structure.* Cambridge, Massachusetts: The MIT Press.

Tinic, S. M. (1972). The economics of liquidity services. *Quarterly Journal of Economics 86*(1), 79–93.

van Ravenswaaij, M. (1997). Intraday patterns and the bid-ask spread on the Paris Bourse. Working paper, Tilburg University.

Vogler, K.-H. (1993). Inter-dealer-trading. Discussion Paper 174, Financial Markets Group, London School of Economics.

Vogler, K.-H. (1997). Risk allocation and inter-dealer trading. *European Economic Review 41*(8), 1615–1634.

Wang, G. H. K., R. J. Michalski, J. V. Jordon, and E. J. Moriority (1994). An intraday analysis of bid-ask spreads and price volatility in the S&P 500 index futures. *Journal of Futures Markets 14*(7), 837–859.

Wells, S. (1992). Price improvement and best execution. *Stock Exchange Quarterly with Quality of Market Review*, 25–31.

Wells, S. (1995). Portfolio trades. *Stock Exchange Quarterly with Quality of Market Review*, 12–14.

Werner, I. M. and A. W. Kleidon (1996). U.K. and U.S. trading of British cross-listed stocks: An intraday analysis of market integration. *Review of Financial Studies 9*(2), 619–664.

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica 48*(4), 817–838.

Wood, R. A., T. H. McInish, and J. K. Ord (1985). An investigation of transactions data for NYSE. *Journal of Finance 40*(3), 723–739.