Order Flow, Information and Trading Behaviour in Foreign Exchange and Equity Markets

Jinhui Luo

London School of Economics and Political Science

Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor of Philosophy

University of London

February 2004

Acknowledgements

I am greatly indebted to Charles Goodhart, Jón Danielsson and Richard Payne for their inspirations, suggestions and comments throughout past years. I also want to thank Paolo Vitale who has led me into the areas of Market Microstructure.

During the course of my PhD study I have benefited enormously from the comments made by various academics and market participants I have met in seminars and conferences. In particular, I would like to thank Richard Lyons, Michael Melvin, Timothy Simin, Andrew Patton and Stephen Sapp. They have read one or more chapters of this work and contributed valuable comments and criticism.

This work was mainly completed while I was a research assistant in the Financial Markets Group (FMG). David Webb (director of FMG) and other colleagues and students in the FMG have never hesitated in lending their hands when they are needed and I feel deeply indebted to their support and help.

Most of the work contained in this dissertation has been presented in various seminars and conferences. These occasions have produced substantial feedback that leads to subsequent improvements of the work. I would like to express my gratitude to the participants in seminars at the London School of Economics, Oxford University, the Royal Holloway (University of London), the 8th Annual International Finance Conference (Atlanta, USA), the 2002 Annual Conference of Northern Finance Association (Alberta, Canada) and the 2002 Venice Summer Institute.

I greatly appreciated the funding support from the K.C. Wong Education Foundation and ESRC.

Finally, I would like to thank my families for their constant support throughout the years. Without their understanding it would have been much harder for me to start and finish the work presented here.

Abstract

It has been shown in the literature that under asymmetric information, trading process itself is a part of pricing mechanism and order flow is the vehicle of information transmission and has a profound impact on prices. This thesis is composed of three major closely connected parts on order flow economics: (1) exchange rate determination and inter-market order flow effect, (2) market conditions and order flow impact and (3) limit order execution and microstructure factors.

The first part of this thesis empirically investigates the price impact of order flow in four major currency markets and the results show that order flow has strong impact on exchange rates in all four markets and over various sampling frequencies. In a new result, inter-market effect is discovered where exchange rate movements in one market can be explained by the order flow in other relevant markets. In terms of forecasting ability, the order flow model out-performs random walk model that has so far beaten all macro-based exchange rate models.

The second part addresses the dependence structure between flow and price change in the FX markets and finds that flow-return relationship is not linear as assumed in the previous literature. Order flow tends to be more informative and has larger impact on prices when market spreads are large, volume is low or volatility is high. These results cannot be fully explained by existing microstructure models.

The last part of thesis studies how limit order execution probability is affected by microstructure factors. Using the tick data from the London Stock Exchange, it is demonstrated that price aggressiveness, spread and potential market pressure have significant impacts on the limit order execution.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Information, Order Flow and Asset Pricing

A vast number of financial assets are bravely priced by markets everyday. A principle task of finance theory is to understand how the assets are priced. In the standard asset pricing paradigm, to value an asset, one only needs to account for the *delay* and the *risk* of payments. John Cochrane puts this idea as "...price equals expected discounted payoff. The rest is elaboration, special cases and a closet full of tricks that make the equation useful for one or another application" (See Cochrane (2001), Page xiii).

Under textbook perfect market conditions where all information about the expected payoff is public and the mapping mechanism from information to price is common knowledge and effortless,[1] we can apply Cochrane's pricing formula to any asset easily. Under these assumptions, price formation, the mapping process from information to equilibrium price, is automatic and instant; all transactions take place at a publicly agreed equilibrium price to realize risk sharing and the

[1]These conditions are normally implicitly assumed for most macro exchange rate models and early asset pricing models.

trading mechanisms and trading activities are irrelevant to asset pricing.[2]

As it has been widely accepted now that these assumptions, however, are not necessarily true in reality. We have seen a large amount of researches, both theoretical and empirical, demonstrating that some information about the asset future payoff is not public and is dispersed among millions of investors as private information (see, for example, Bagehot (1971), Grossman and Stiglitz (1980), Kyle (1985), Glosten and Milgrom (1985), French and Roll (1986), Hasbrouck (1991a, 1991b), Payne (2003), to name a few.). It is also widely acknowledged that different investors might have different interpretation of the same piece of information and the mapping mechanism is not homogeneous. Isard (1995, pp. 182) describes this embarrassment in a context of foreign exchange determination:

> ...economists today still have very limited information about the relationship between equilibrium exchange rates and macroeconomic fundamentals. Accordingly, it is hardly conceivable that rational market participants...could use that information to form precise expectations about the future market-clearing levels of exchange rates.

When information structure is asymmetric, market is believed to play a role of information aggregation. How is the dispersed information aggregated or how the private information is incorporated into the prices? Expectations model (for example, Grossman (1976) and Grossman and Stiglitz (1980)) usually assume a fiction of representative Walrasian auctioneer who collects orders and finds the market clearing prices. The process starts with investors submitting their optimal demand conditional on their own private information, the auctioneer collects the orders and announces a potential price, investor revise their optimal demands conditional on the price and their information. The process of *submitting demand* $\Rightarrow$

[2]Such an agnostic view can be found in the rational expectation literature, see Radner (1982) and Blume and Easley (1990) for more discussion.

announcing potential price $\Rightarrow$ *revising demand* continues until there are no more demand revisions, the equilibrium price is found and information is aggregated. Clearly there is no transaction taking place outside the equilibrium price and the final trading activities play no role of price discovery but only serve as means of market clearing.[3]

Certainly there exist markets that bear at least an approximate resemblance to the Walrasian framework[4]. But there are many other financial markets where price-setting cannot be convincingly argued to fit in this picture. For example, in the foreign exchange markets and in the main trading periods of most stock exchanges, trading takes place continuously and there are specific market participants playing roles far more active in price-setting than the Walrasian auctioneer does. The market maker or specialist in many exchanges do not have the luxury of going through the repetitive process of *submitting demand* $\Rightarrow$ *announcing potential price* $\Rightarrow$ *revising demand* before setting the actual transaction prices. Also importantly, the insiders, such as those in Kyle (1985), might not even have the incentives to submit a fully informative demand and demand revision during the Walrasian auction process[5]. How are the "equilibrium prices" formed?[6] The easiest way to demonstrate the information incorporation mechanism is by using Glosten and Milgrom (1985) framework. Assume there are informed and liquidity traders in the market. The market maker sets the ask price a_t to the expected value of an asset after seeing a trader wishing to buy (i.e. $E_t[\widetilde{V}|Buy]$. a_t depends on the conditional probability that $\widetilde{V}$ is either lower or higher than his prior belief V_0 given that a trader wishes to buy. The bid price b_t is defined similarly given

[3]It seems that, in this Walrasian framework, the process of *submitting demand* $\Rightarrow$ *announcing potential price* $\Rightarrow$ *revising demand* is contributing to the price discovery.

[4]For example, the open and close auction call at the London Stock Exchange.

[5]This valuable incentive argument is suggested by Richard Lyons.

[6]"Equilibrium price" is usually used as a static concept. Applying it in a dynamic sense, we mean the regret-free price conditional on the information available at the time of setting price.

that a trader wishes to sell. If the liquidity traders are assumed equally likely to buy or sell whatever the information, good news will result in an excess of buy orders and bad news will result in an excess of sell orders. This implies that the conditional probability incorporates the new information that the market maker learned from observing the order flow and is hence a posterior belief about the asset value. The posterior will become a new prior in the next round of trading. As the trading takes place continuously, the updating process continues and the information "flows" into the prices.

In the above Bayesian learning process, trading activity is clearly relevant to the price discovery per se. Trading, the signed transaction quantities in particular, not only serves as a mechanism of market clearing, but also serves as a vehicle of information transmission. Dispersed private information is effectively incorporated into prices through trading order flow.

Order flow in the market microstructure literature is defined as the net of buy initiated orders and sell initiated orders. It is a measure of market"opinion". It should be noted that order flow is fundamentally different from trading volume which has no direction and has no clear relation with the price setting. It should be also pointed out that order flow is different from the classic demand-supply pressure. Classic demand-supply is associated with public information and price changes due to demand-supply imbalance do not need transactions. Clearly here price changes due to order flow as in the Bayesian learning process are associated with trading activities. It is through the trading activities that non-public information is incorporated into the prices.[7] In this sense, order flow is informative and can be used to explain price changes.

Turning to the empirical side, there is an increasing amount of evidence show-

[7]In the real world, there might be other channels through which private information can be incorporated into prices, e.g. limit order, as I will argue in the last section of this chapter, could be alternative channel. However, as shown in the vast microstructure empirical researches and thorough out this dissertation, order flow perhaps is one of the most important channels.

ing the strong impact of order flow on asset returns in both equity and foreign exchange markets. Hasbrouck (1991a, 1991b), Hasbrouck and Seppi (2001), French and Roll (1986) and Chordia, Roll and Subrahmanyam (2002) study equity markets and Evans and Lyons (2002a), Caia et al. (2001), Rime (2000), Danielsson, Payne and Luo (2002) and Payne (2003) study foreign exchange markets. The evidence from all these research works amounts to a simple point: order flow is informative and has significant impact on returns.

If price is driven by order flow, then what drives order flow? Like many other important issues in finance, economists still only have very limited understanding to this fundamental question. This question can be exploited in the following two sub-questions:

1. Can order flow convey public information or can public information drive order flow?

2. What information is behind the order flow that drives price?

The first question seems to be out of place at first glance. If information is public, there is no need to " convey" it and it should have been incorporated into the price once it is announced. In market microstructure theory, order flow is also always modelled to convey private information. It should be noted, however, that this argument is built up on an assumption that investors are homogeneous in the sense of knowledge background, information processing ability and the magnitude of rationality. If investors have different interpretations of and draw different insights from the same piece of information, then the price adjustment (to the public information) might not be sufficient and instant. In this situation, order flow might convey public information as well. Evans and Lyons (2003) and Love and Payne (2002) provide evidence that order flow is useful in transmitting (public) macro news to exchange rates.

The second question is very important but also more complex. Even though it is believed that order flow contains information valuable for forecasting price (see, for example, Goodhart (1988)), it is less clear what this information might be about. In a broad sense of asset pricing, there are two types of information that is relevant to asset price: information about future payoff and information about the discount factor or pricing kernel [8]. In theory, order flow is overwhelmingly assumed to convey the first type of information (for example, Kyle (1985), Glosten and Milgrom (1985) and Easley and O'Hara (1992)) but the empirical evidence is limited.[9] Evans and Lyons (2002a), Rime (2000) and Danielsson and Payne and Luo (2002) find strong contemporaneous correlations between flows and returns in foreign exchange markets, the horizons over which the returns and flows are measured are at relatively short (from five minutes to one week). Though authors working with high frequency data believe that results based on certain intervals, such as daily or weekly, are consistent with the story that order flow conveys fundamental information on future payoffs because the price impacts can be seen as 'permanent', those who believe fundamental information on future payoff should last much longer (above one month, for example) think that the existing results are not hard evidence and the evidence from much lower frequency data is still desired.

On the other hand, there is increasing amount of evidence showing that order flow conveys information about the discount factors.[10] Caia et al. (2001) find that order flow reveals private information about the portfolio shifts that have significant impact on the dramatic changes of USD/JPY rate in 1998. In their survey

[8]Campbell and Shiller (1988) demonstrate that a large part of variation in aggregate equity values appears to be due to changes in discount rates rather than changes in expected payoff.

[9]Evans and Lyons (2003) and Love and Payne (2002) find order flow is useful in conveying macro news which can be interpreted as news regarding future payoff.

[10]This argument would require imperfect substitutability among different assets in the market. For more discussion, see Lyons (2001).

paper, Gehrig and Menkhoff (2002) provide evidence that order flow conveys information on *semi-fundamentals*, such as short-term trading objective, liquidity consideration and portfolio shifts, which has short-run price impacts. Overall, so far we see more evidence that order flow conveys information relevant to short and medium term price movements than the evidence that order flow conveys information about future payoff, a view labelled by Froot and Ramadorai (2002) as *weak flow-centric view*.

If information conveyed by order flow is less important in the long run than in short or medium run, is it still useful to study order flow for a better understanding of asset pricing? The answer is yes for two reasons. Firstly, the concept of"fundamentals" of an asset is far more elusive than it appears to be. For example, in most theoretical models, the fundamentals refer to the final payoff of an asset and is assumed to be static. However, it is almost impossible for an investor to tell what is the final payoff of a stock or an exchange rate. Even for the same piece of public information (on future payoff), different investors can draw different conclusions and will take different actions based on their own understanding. Market order flow serves as a vote counter and aggregates information on economic factors, some of which even economists do not understand quite well. Given the dismal empirical result on the relationship between"fundamentals" and the actual asset prices (such as exchange rate disconnect puzzle, excess volatility puzzle), the order flow approach, humble but more practical, can lead us closer to understanding price behaviour. Secondly, the short to medium term price formation process, like the long-term relationship between asset prices and the ultimate economic variables, is also of profound importance for market participants.[11].

[11] Some fundamental factors have temporary feature, e.g. the temporary shift of fiscal policy. Other factors can be mistakenly thought to be transitory. For example, until very recently, liquidity has been thought to be a transitory effect and has been excluded in most classic asset pricing models. Academics have now started realizing that such 'short term' factors could be fundamental and pervasive risk factors priced by markets (see Pastor and Stambaugh (2003)).

Investors having short investment horizons obviously care about such short to medium term price behaviour. Policy makers, like central banks and regulators, are also turning to the micro market behaviour to better understand many classic financial issues like intervention and financial crisis.[12] Order flow, though itself is only a proxy for the economic variables that drive asset prices, it can nevertheless serve as a media which allows us to study price evolution on one hand and to telescope the ultimate and elusive economic variables on the another hand.

This thesis contains three closely connected articles on order flow economics: (1) exchange rate determination and inter-market order flow effects, (2) market conditions and order flow impact and (3) limit order execution and microstructure factors. These research works extend the existing literature on order flow and enhance our understanding of price formation process in foreign exchange and equity markets. The following sections in this chapter provide a brief outline of the motivations and main results obtained in subsequent chapters.

1.2 Order Flow Effects in FX Markets

Chapter two is based on co-joint work with Jon Danielsson and Richard Payne, "Exchange Rate Determination and Inter-Market Order Flow Effects". It further examines the price impact of order flow across different exchange rates and over a range of sampling frequencies, investigates information flow across related markets and tests the forecasting power of order flow models.

[12]e.g. Payne and Vitale (2000), Danielsson and Saltoglu (2003), and Dominguez (2003a, b) adopt micro approach to intervention and crisis.

1.2.1 Issues, Methodology and Data

Empirical models of exchange rate determination, especially at intermediate estimation horizons, have frustrated economists at least since the Meese and Rogoff (1983a, b) result that macro-based exchange rate models have little in-sample explanatory power and underperform a random walk in out-of-sample forecasting. In the empirical finance literature, there is, however, a long tradition of studying the higher frequency relationship between the prices and trading volume (see, for example, Clark (1973), Tauchen and Pitts (1983)). Such analysis cannot help resolve Meese-Rogoff puzzle, not least because volume is directionless and bears no information on the direction of exchange rate movements. Recently, researchers have investigated the price impact of order flow, a variable that is fundamentally different from volume. For example, Hasbrouck (1991a) and Madhavan and Smidt (1991) study equity markets. Cohen and Shin (2002) and Danielsson and Saltoglu (2002) study fixed income markets. Lyons (1995), Payne (2003), and Evans and Lyons (2002a) study foreign exchange rate markets. In these empirical microstructure researches order flow has been shown to be a key determinant in high frequency asset price changes.

From the perspective of benchmark rational expectations models of asset pricing or exchange rate determination, the importance of order flow is puzzling. Such models predict that prices should respond to new information without any consistent effect on order flow. Intuitively, when new information arrives every agent immediately revises his/her estimate of the value and thus there are no reasons/opportunities for one-sided trading. Thus one must look beyond these models to find a rationale for order flow's effects on prices.

In general terms, order flow can be a determinant of prices in environments where agents do not share the same information or where they do not agree on the model by which asset prices are determined. In this context, order flow can convey information about micro factors (e.g. shifts in hedging demands) and macro factors (e.g. public announcements). While the micro information is specific to individual agents, macro information can be also interpreted differently across agents. Consider an economy where agents have asymmetric information and/or disagree about the asset pricing model. In that case, disagreements between FX market participants regarding the manner in which exchange rates are determined or the non-public information relevant for exchange rate determination will be revealed in order flow and thus flow can be used to explain (and possibly forecast) exchange rates in the short to medium term. Recent empirical work supports this intuition, e.g. Evans and Lyons (2002a) who find strong dependence of daily exchange rate changes on daily order flow, even after accounting for macroeconomic fundamentals.

Chapter two is aimed to investigate the importance of order flow for exchange rate determination and extends the extant research in three areas:

1. Extend the earlier results of Evans and Lyons (2002a) who only consider the daily sampling frequency and two currency pairs. Using a unique data set which contains four currency pairs (EUR/USD, EUR/GBP, GBP/USD and USD/JPY) and spans nine months on average, this research investigates the relationship between order flows and exchange rate changes across frequencies, ranging from five minutes to one week and for four rates. The relatively long time-series dimension of our data and the fact that we have four exchange rates to consider, we can examine how the explanatory power of order flow changes with sampling frequency and whether the effects of flows are consistent across currency pairs.[13]

2. The flow-return relationship is investigated in an inter-market framework.

[13] Also note that order flow used in this study is measured from direct inter-dealer market and it may be different from that measured from brokered transactions.

Specifically, we investigate whether order flows in other relevant markets (e.g. EUR/USD) have explanatory power for exchange rate changes (e.g. GBP/USD) beyond its own order flow. This research is motivated by the following economic reasoning: a currency trader who has private information regarding the future value of a currency, say the dollar, could profit from this information in many possible currency markets e.g. GBP/USD, USD/JPY, EUR/USD. Given the informational links between all dollar related markets, order flow in one market has important implication for the asset in another market[14].

3. This work goes beyond prior work on the *contemporaneous* relationship between returns and flows to evaluate the forecasting power of order flows for exchange rates. In theory (e.g. Kyle (1985)), order flow generated by informed trading has the property of autocorrelation because informed traders behave strategically and split their demand in a sequence of trades. Here we seek to understand whether, using order flows and in some cases past returns also, we can generate forecasts that improve upon naive statistical alternatives. Thus we test whether empirical models based on order flow can pass the tests that standard macroeconomic models failed. Of course, the analysis conducted here is at much higher sampling frequencies than those considered by Meese and Rogoff (1983a), but the possibility that one can generate useful, non-trivial forecasts of rates at the daily level indicates that one might also be able to do so at lower frequencies.

[14]More recently, this idea has been used as a basic assumption to address international currency competition by Lyons and Moore (2003).

1.2.2 Main Results and Contribution

We study the explanatory and forecasting power of order flow for exchange rates changes for our major exchange rates at sampling frequencies ranging from 5 minutes to 1 week. The key results include;

- The contemporaneous relationship between flows and changes in rates is significant for all four rates but model explanatory power varies across sampling frequencies. At intra-day frequencies the flow-return relationship is strong for all four rates. At the daily and weekly level, there is still strong explanatory power of order flow for exchange rates changes for EUR/USD and USD/JPY but not for EUR/GBP and GBP/USD.

- However, when one examines the inter-market effects of order flows, one sees that price changes for EUR/GBP and GBP/USD are strongly affected by EUR/USD order flow and, taking these effects into account, at daily and weekly sampling frequencies overall flows have strong explanatory power for the Sterling rates.

- An analysis of the forecasting power of order flows, using the technique of Meese and Rogoff (1983a, b), demonstrates that flow analysis outperforms a naive benchmark across essentially all sampling frequencies for all rates.

- A simple true out-of-sample forecasting experiment using order flows does not provide terribly valuable exchange rate forecasts. There is, however, evidence that order flow can be forecast out of sample.

These results serve to emphasize the role played by order flow in FX, and possibly other, markets. We provide clear evidence that order flows can be used to explain and forecast rates at very high frequencies as well as observations intervals relevant to international macroeconomics. The results of the inter-market order flow effects are new and provide direct empirical evidence for the international currency competition model by Lyons and Moore (2003) who assume informed currency traders can exploit their information advantage by in-direct trading. Finally, the results from forecasting experiments indicate order flow has potential to forecast exchange rates and provide a step stone for a more sophisticated forecasting model development.

1.3 Market Conditions and Price Impact of Order Flow

Order flow drives price because it carries information. Order flow will have larger impact on prices when it is more informative, and vice versa. Chapter three, "Market Conditions, Order Flow and Exchange Rates Determination", investigates how informativeness of order flow changes under different market conditions in foreign exchange markets.

1.3.1 Issues, Methodology and Data

In comparison to macroeconomic models of nominal exchange rates, the market microstructure approach has been quite successful in explaining exchange rate changes over short to medium time horizons. Though there is reported evidence that fundamentals have explanatory and forecasting power to exchange rates over longer periods (for example, Mark (1995)), Meese-Rogoff puzzle has never been convincingly over-turned over short horizons. On the order hand, researchers who apply the microstructure approach have been more lucky. Evens and Lyons (2002a) claim that net order flow has substantial explanatory power for exchange rate changes on a daily basis and can explain 60 percent and 40 percent of return variation for DEM/USD and USD/JPY respectively. Rime (2000) shows that order flow can explain 30 percent of return variation for NOK/DEM on a weekly basis. Chapter two will provide further evidence that order flow has explanatory power for exchange rate changes for EUR/USD, GBP/USD, USD/JPY, and EUR/GBP at higher frequency levels and within an inter-market framework

The microstructure approach to exchange rate analysis is based upon a simple idea that order flow carries information and thus has explanatory power for price movements. The existing research work in this line has largely assumed that the informativeness of order flow is constant even if the market experiences dramatically different situations. Is this assumption valid? This question hasn't been addressed in the extant empirical literature.

Turning to theory, we don't have a better answer either. There is little consensus among the existing microstructure theories on at what market conditions order flow could be more informative. In the models of Admati and Pfleiderer (1988) and Subrahmanyam (1991), both the informed and the discretionary liquidity traders can strategically choosing their trading time and such strategic behavior will impact on the informativeness of order flow. Admati and Pfleiderer (1988) predict that informativeness of order flow will be positively correlated with bid-ask spreads and negatively correlated with volatility and trading volume.

In the information uncertainty model of Easley and O'Hara (1992), however, the occurrence of trades is associated with information events. When a trade occurs, the market maker will update his/her belief towards a higher probability of some information events and adjust the quotes accordingly. The model predicts that order flow is more informative when spreads are high, volatility is high and trading volume is high.

Given the order flow's significant impact on price and its increasing usage by practitioners in their decision making, characterizing the information transmission mechanism of order flow is important from the point of views of both academics and market participants. Chapter three attempts to address these issues by studying a simple empirical model:

$$\Delta P_t = \alpha + \beta * Q_t + \varepsilon_t \tag{1.1}$$

where ΔP_t is price change and Q_t is aggregated order flow within time interval $[t-1, t]$ respectively.

Generally speaking, the information effect of order flow will be captured by the regression coefficient β. In words, the more informative the order flow the larger the β, and vice versa. If order flow is not equally informative, β is expected to be a function of some structural parameter Z_t measuring market conditions, i.e, $\beta = \beta(Z_t)$.

The main objective of chapter three is to empirically characterize the function $\beta(Z_t)$, where three market condition measures are examined: bid-ask spread, volatility and trading volume. The data used in this study is transaction level data collected by Reuters D2000-2 in the inter-dealer spot FX markets. It covers four major currency pairs (EUR/USD, EUR/GBP, GBP/USD and USD/JPY) and spans nine months on average between 1999 to 2000.[15]

1.3.2 Main Results and Contribution

The major findings include:

- The informativeness of order flow is not constant under different market conditions as was assumed in the existing empirical microstructure literature. This result highlights the non-linearity in the relationship between

[15]the data set used in chapter three is the same as that used in chapter two except that chapter two only use transaction information while chapter three use information on both transactions and firm quotes.

order flow and price movement.

- Order flow tends to be more informative when bid-ask spreads are high, volatility is high or trading volume is low. This relationship is significant and persistent across different sampling frequencies and exchange rates.

- The relationships between the price impact of order flow and the market conditions are significantly captured by our Interaction model and Logistic Smooth Transition Regression (LSTR) model. In particular, the empirical results from the LSTR model also suggest that the price impact of order flow shifts within a relative small range of market conditions that the market is very likely to experience.

It should be noted that the measure of order flow informativeness in this research relies on the aggregation of order flow. Working on a transaction base, the regression model (1.1) might be distorted by the transitory liquidity effect. Through aggregation, the liquidity trades can be largely cancelled out over time. In this research we have run a set of robustness tests, including testing seasonality impact and experimenting regression (1.1) in an inter-market framework in which liquidity effect does not exist at all.

This research work challenges a widely adopted assumption in previous empirical market microstructure literature that order flow is equally informative under different market conditions. The evidence provided here indicates that the equal informativeness assumption is ungrounded. The non-linearity identified in this study implies that the flow models can be improved by taking into account the market conditions.

This work is among the first that test the theoretical predictions on the relationship between the order flow informativeness and market conditions. Clearly the results are neither fully consistent with the predictions of Admati and Pfleiderer's model nor with those of Easley and O'Hara's information uncertainty model. So far, to my knowledge, we haven't seen any theoretical models that explicitly take market conditions as an input element in pricing mechanism. Although to develop a theory that better explain the empirical results presented in chapter three is not within the research agenda of this thesis, I believe future research on this issue will be worthwhile.

1.4 Limit Order Execution Probability

Limit order is one of the major types of order commonly seen in electronic trading systems. With a limit order, one needs to specify the desired transaction price and the quantity/size to buy or sell. Once a limit order is submitted, it will sit in a queue usually called limit order book. For limit order trading, most exchanges maintain a strict price and time priority.[16] For a limit buy order, when market transaction price is at or below the pre-specified price of the limit order, the limit order will be partially or completely executed depending on the size of the incoming sell order. For a limit sell order, execution condition is similar. A salient feature of limit order is that limit orders are usually executed at better prices than the best offers (for buy limit order) or bids (for sell limit order) available at the time of limit order submission, but there is no guarantee of execution. Non-execution implies implicit cost. Chapter four investigates various factors that might impact on the execution probability of limit orders.

[16]Orders with more competitive prices will be first filled. For orders that have same prices, those are submitted earlier will be filled first.

1.4.1 Issues, methodology and Data

As more and more stock exchanges are adopting the automatic order-driven systems, limit order has been widely used as a major trading means. Comparing with market order, the benefits of limit order includes price improvement (buy at bid price or sell at ask price) and no price uncertainty. However, the benefit does not come without cost: execution is not secured. Since non-execution cost could be very high, investors trading with limit order need to evaluate the execution probability of their submission strategies and take this into account in forming the investment decisions.[17]

Chapter four analyzes how such probability can be affected by various factors and attempts to provide a method to evaluate the execution probabilities of different order submission strategies. The issue is clearly important to the market participants, especially to the trading desk of institutional investors for whom the trading cost is among the top concerns.[18] It is also important from market structure point of view: limit order placement has important impacts on market liquidity and spread dynamics. In the classic market microstructure models, informed traders are overwhelmingly assumed to use market orders to exploit their information advantage. However, there is experimental evidence showing that informed investors do not restrict themselves to market orders. In an experiment study, Bloomfield, O'Hara and Saar (2002) find informed traders can use more limit orders than liquidity traders do in certain market conditions. Actually the use of the limit order by informed investors who want to buy can be rationalized

[17]For example, Donald Keim and Ananth Madhavan wrote in the a article that "the implicit costs...associated with missed trading opportunities...are significant relative to explicit costs and realized portfolio returns". Source: Financial Times, 23 July 2002.

[18]Recently the fund management companies are under increasing pressure from regulatory organizations and investors to better manage their trading cost (see *BGI INSIGHT*, 2003, quarterly magazine from the Barclays Global Investors).

in following simple model (similarly for those who want to sell):

$$\begin{cases} \pi_m &= E[v|s] - (p^{ask} + \lambda Q^2) \\ \pi_l &= \Pi(\kappa)(E[v|s] - p^l) + (1 - \Pi(\kappa))0 \end{cases} \tag{1.2}$$

where the π_m and π_l are expected profit of trading quantity of Q by using market and limit order respectively. $E[v|s]$ is the expected value of the asset conditional on the private signal s. p^{ask} is current market ask price and p^l is the price of limit order. λ is the price impact coefficient in the sense of Kyle (1985). Π is the probability of execution of limit order. Π is a function of κ, a vector of factors which can impact on the limit order execution probability.

In the above setup, observation of private signal wouldn't necessarily lead to the use of market orders. By using a market order, informed investors can secure a transaction but have to pay the transaction costs, including bid-ask spread and price impact cost, which is measured by λQ^2. By using a limit order, though the transaction price could be lower ($p^l < p^{ask} + \lambda Q^2$), the execution is not secured ($\Pi < 1$). Under certain market conditions (for example, when execution probability and spread are large), informed trader might be better off by using limit order rather than market order. An implication of this simple model is that limit order could also convey information, a point that I believe is of profound importance in market microstructure but is largely ignored in existing literature.

Clearly, the execution probability, Π, is a key factor affecting investors' decision on the choice between different types of orders. A better understanding of how such probability is determined — the objective of this research — will not only have important implications for practitioners but also can enhance our knowledge of price discovery process in limit order trading systems.

In recent years, limit order trading has drawn increasing attention in market microstructure literature and it has been demonstrated that various factors can impact on limit order execution (Cohen, Maier, Schwartz and Whitcomb (1981), Angel (1994), Glosten (1994), Kumar and Seppi (1992), Seppi (1997), Handa and Schwartz (1996), Parlour (1998) and Foucault (1999), Chung, Ness and Ness (1999), Biais, Hillion and Spatt (1995) and Ahn, Bae and Chan (2001) ,etc). The focus of these studies, however, is not on the limit order execution itself but is mainly on the choice between market order and limit order and how such choice impacts on market equilibrium.

Works on limit order execution are relatively few and most concentrate on time-to-execution. Foucault, Kadan and Kandel (2001) develop an equilibrium model for time-to-execution of limit orders. In another work, Lo, Mckinlay and Zhang (2002) compare three different econometric models for the time-to-execution of limit orders and find the time-to-execution is sensitive to limit price and other explanatory variables. Current research focuses on the execution probability and can be seen as complementary study of this line of literature.

Research in chapter four is also related to microstructure literature on profitability study of order submission strategies. Handa and Schwartz (1996) and Harris and Hasbrouck (1996) compare the ex post profitability of different order submission strategies in NYSE. This chapter uses a sample of FT30 stocks from limit book system of London Stock Exchange (LSE) to analyze how limit order execution probability is affected by various factors and evaluate such probability ex ante. Obviously the analysis conducted in this chapter is a building block to extend the profitability analysis in real time.

Chapter four adopts a two-stage methodology to address the issue. In the first stage, potential factors, mainly derived from theoretical models or relevant empirical researches, are examined one by one against actual transaction data to test whether the factors can impact on the limit order execution probability as assumed in theory. In the second stage, the probability is modelled as a function of the all factors examined in the first stage. In this way the factors can be tested simultaneously. Moreover, the estimates can be feedback into to model to evaluate the execution probability of different limit order submission strategies in real time.

1.4.2 Main Results and Contributions

Chapter four empirically investigates the impacts of various microstructure factors on the limit order execution probability and models such probability of limit order submission strategies. The main results are,

- For a given stock, *price aggressiveness*, measured as the extent to which a limit order betters the best existing quote on the same side, has a significant positive impact on the limit order execution probability. In General, *spread* is negatively correlated with unconditional limit order execution probability. When price aggressiveness is controlled, spread impact is particularly strong. *Potential Market Pressure* (PMP), a difference between the buying pressure and selling pressure built up on the limit order book, has significant impact on the execution probabilities. For buy (sell) orders, the larger PMP, the smaller (larger) execution probability.

- Across different stocks, market liquidity, measured as trading volume of a stock, has strong impact on the limit order execution probabilities. Trading less liquid stock, a trader have to post a more aggressive limit price to get the same likelihood of execution as trading liquid stocks.

- Contrary to the intuition, order size and the time of the day do not have strong impacts on limit order execution as expected. Order size has a U-shape of execution probability with the middle-size orders have the smallest execution chance. Large orders are usually submitted more strategically by cutting spread aggressively when it is large.

- Volatility has positive impact on limit order execution but much weaker compared with price aggressiveness, spread or PMP.

- The limit order execution probability is modelled as a function of a vector of factors (price aggressiveness, spread, PMP, order size, volatility and time of day) in a probit model. The estimation results indicate that the proposed model is capable of capturing the features regarding the relationship between limit order execution probability and its major determinants and can be used to forecast the such probabilities of different limit order submission strategies.

In previous researches, a set of factors have been assumed to have important impacts on the order submission[19]. Chapter four attempts to test these predictions thoroughly by using a unique data set from the London Stock Exchange. The empirical results on the impact of spread and aggressiveness provide direct supporting evidence for the model predictions of Foucault et al. (2001). Volatility, which is the key driving force of market liquidity dynamics in theoretical models is found to have only slight impact on limit order execution while potential market pressure, which has been escaped the economists' scope, is found to have profound impact on the limit order execution in our sample. This research work tries to establish some "stylized facts" for further theoretical research on the liquidity demand-supply in the limit order trading system.

In the existing market microstructure literature, both theoretical and empirical, market order is believed to be the vehicle of information transmission. As argued in previous subsection, in equilibrium, the choice between limit order and market order depends on the expected execution probability which in turn depends on

[19]For example, volatility has been a key driving force of market liquidity dynamics in theoretical models by Foucault (1999) and Handa and Schwartz (1996). Order size (e.g. Harris and Houbrouck (1996)) and time of day have also been assumed to have negative impact on the limit order execution.

the quality of signal, spread, aggressiveness of limit order and current market pressure. If the execution probability is large enough, limit order can be used by informed investor and therefore can be informative. The strong impact of PMP on limit order execution probability indicates that PMP might have predictive power for the future price movements and can be explained as suggestive evidence on the hypothesis that limit order can also convey information. The formal modelling of information transmission by limit order is beyond the research agenda of this chapter but the result found in this research is encouraging for the future research in this direction.

Chapter 2

Exchange Rate Determination and Inter-Market Order Flow Effects

2.1 Introduction

Empirical models of exchange rate determination, especially at intermediate estimation horizons, have frustrated economists at least since the Meese and Rogoff (1983a, b) result that macro-based exchange rate models underperform a random walk in forecasting ability. In the empirical finance literature, there is, however, a long tradition of studying the higher frequency relationship between the price of financial assets and trading volume.[1] Such analysis cannot help resolve the Meese-Rogoff puzzle, not least because volume is directionless, i.e., a change in volume cannot predict the direction of FX changes. Recently, researchers have investigated the impact of *signed volume*, i.e., the decomposition of volume into transactions initiated by sellers and buyers, separately. The difference between seller and buyer initiated volume is termed *order flow*, a variable that is fundamentally different from volume. Order flow has been shown in empirical mar-

[1] See e.g. Clark (1973); Epps and Epps (1976); Tauchen and Pitts (1983); Karpoff (1987).

ket microstructure research to be a key determinant in high frequency asset price changes.[2] Several authors, e.g. Lyons (1995), Payne (2003), and Evans and Lyons (2002a) study the relationship between order flow and foreign exchange rates. The objective of this chapter is to investigate the relationship between order flow and FX rates, extending extant research in three aspects: providing augmented evidence regarding the dependence of prices on flows, examining inter-market order flow effects, applying Meese-Rogoff methodology to test the forecasting power of order flow model.

From the perspective of benchmark rational expectations models of asset price or exchange rate determination, the importance of order flow is puzzling. Such models predict that prices should respond to new information without any consistent effect on order flow. Intuitively, when new information arrives every agent immediately revises his/her estimate of value and thus there are no reasons/opportunities for one-sided trading. Thus one must look beyond these models to find a rationale for order flow's effects on prices.

In general terms, order flow can be a determinant of prices in environments where agents do not share the same information or where they do not agree on the model by which asset prices are determined. In this context, order flow can conveys information about micro factors (e.g. shifts in hedging demands) and macro factors (e.g. public announcements). While the micro information is specific to individual agents, macro information can also be interpreted differently across agents. Consider an economy where agents have asymmetric information and/or disagree about the asset pricing model. In that case, the agents' trading strategies and, in particular, aggressiveness, might reveal underlying information regarding future payoffs or change of risk premia and hence affect asset price

[2]For example, by Hasbrouck (1991a) and Madhavan and Smidt (1991) who study equity markets, and Cohen and Shin (2002) and Danielsson and Saltoglu (2003) who study fixed income markets.

changes. Recent empirical work supports this intuition. For example, Evans and Lyons (2002a) find strong dependence of daily exchange rate changes on daily order flow, even after accounting for macroeconomic fundamentals.[3]

The objective of this chapter is to investigate the importance of order flow for exchange rate determination. In this we extend the earlier results of Evans and Lyons (200a) who only consider the daily sampling frequency and one currency pair a time. We investigates the relationship between order flow and exchange rates across frequencies, ranging from five minutes to one week. Furthermore we study four currency pairs (EUR/USD, EUR/GBP, GBP/USD and USD/JPY) and explicitly model the impact of order flow across markets, e.g. investigating the impact of EUR/USD order flow on EUR/GBP exchange rate. Finally, we apply Meese-Rogoff methodology to test whether the order flow models beat a random walk in forecasting.

Our data derives from transaction-level information obtained from the Reuters D2000-2 electronic brokers and covers approximately 10 months for EUR/USD and GBP/USD and eight months for EUR/GBP and USD/JPY. The sample starts in 1999 and ends in 2000. Our analysis consists of three set of empirical exercises.

First, we evaluate how order flow is contemporaneously related to changes in exchange rates across sampling frequencies. Taking advantage of the relatively long time-series dimension of our data and the fact that we have four exchange rates to consider, we can examine how the explanatory power of order flow changes with sampling frequency and whether the effects of flows are consistent across currency pairs.

Second, we look at the dependence of exchange rate changes on order flows from other markets by investigating whether order flows in one currency pair have explanatory power for another currency pair.

[3]Similarly, Chordia et al. (2002) show that daily changes in US equity market levels are strongly related to market wide order flow measures.

Finally, in order to investigate order flow from a macroeconomic perspective, we evaluate the forecasting power of order flows for exchange rates. Here we seek to understand whether, using order flows and perhaps past returns, we can generate forecasts that improve upon naive statistical alternatives. Thus we test whether empirical models based on order flows can pass the Meese-Rogoff test.

The results from these three research questions provide new insights into the market microstructure analysis of high frequency exchange rates as well as the macroeconomics analysis of medium term exchange rate determination.

First, we demonstrate that, within a single market, contemporaneous order flow significantly explains exchange rate changes, across sampling frequencies. We however observe considerable differences in the explanatory power of the various regressions. For EUR/USD rate, R^2 hovers around 40% across frequencies, while for USD/JPY the R^2 increases with aggregation, from 6% at five minutes to 67% at one week. In contrast, R^2 for both GBP rates decreases with aggregation from 26% at five minutes to 1% at one week. Taken in isolation, the results from GBP regressions are somewhat puzzling.

We subsequently extend the model by including order flow from other markets. For the EUR/USD and USD/JPY, the inclusion of order flows from other markets makes little difference. However, for GBP rates, especially at lower frequencies, order flow from other markets has strong and significant impact, especially for the EUR/GBP rate where the EUR/USD order flow is found to be the primary exchange rate determinant. To understand why, here we offer two explanations, one from theoretical perspective and one from practical perspective. In theory, a currency trade can be achieved by direct trading or in-direct trading. Suppose a investor with Sterling Pound has private information about Euro (e.g. Euro will appreciate in the future). He can buy Euro directly in EUR/GBP market and in this case we will see the standard single market result, i.e., a positive impact of EUR/GBP order flow on EUR/GBP rate. However this investor can also trade GBP/USD and EUR/USD to exploit his information advantage if the indirect markets are liquid and price impacts are small[4]. Obviously the order flow in GBP/USD and EUR/USD markets conveys information about EUR/GBP. From the practical perspective, market players in the one market will keep their eyes open to the price changes and order flows of other markets. Dealers in EUR/GBP market will adjust the EUR/GBP quote when observing positive order flow in EUR/USD market or negative order flow in GBP/USD market, both indicating a appreciation of Euro against Sterling. In this case, we will observe price change in EUR/GBP rate without companying EUR/GBP order flow.

These results suggest that while basic own order flow model may be appropriate for the largest currencies, it is less so for smaller currencies with many traded exchange rates such as GBP.[5] While Rime (2000) finds that order flow are significant in explaining EUR/SEK, the EUR contract is the only traded currency for SEK. In contrast, there are multiple traded currency pairs for GBP. As a result, information regarding SEK will go through EUR/SEk order flow, while the information regarding GBP can flow through any traded currencies. Furthermore, this provides significant evidence of strong information links between FX markets, with small markets dominated by larger ones. This effects persist across our frequencies and strengthen with aggregations, suggesting that these information links may persist beyond our sampling frequencies.

The final key result is on the forecasting of exchange rates. First, we use Meese and Rogoff (1983a, b) framework, and find that the order flow models almost always yield a better forecast (in RMSE terms) than does a random walk model.

[4]In that case, we will see a negative order flow in GBP/USD market and a positive order flow in EUR/USD market.

[5]Indeed, according to the Bank of International Settlements (2002) in April 2001 the EUR/USD represented 30% of all spot FX trading, the USD/JPY 21%, GBP/USD 7% and EUR/GBP 3%. The first three exchange rates are the three largest currency pairs while EUR/GBP is only the eighth.

This result is consistent across sampling frequencies and currencies. Therefore, the order flow model passes the Meese-Rogoff test that macroeconomic models have failed so often. We note however that Meese-Rogoff test is not a genuine out-of-sample forecasting test. We run such a test with simple specification and find that order flow does not perform particularly well in forecasting exchange rates. We find however that order flow itself can be forecasted. This suggests that a more sophisticated specification for a pure forecast model for exchange rates may provide significant forecasting power.

In sum, our results suggest that order flow analysis can be very useful in understanding exchange rate determination. From a low frequency, macroeconomic perspective, order flows can contribute strongly to our ability to explain exchange rate changes while they allow one to improve exchange rate forecasts most dramatically at a microstructure level. While further work using longer data samples would be useful to verify and clarify our results, the analysis here clearly points to the information content of order flow.

The rest of this chapter is structured as follows. Section 2.2 outlines FX market, our data sources and our processing of the data. Section 2.3 presents our analysis of the explanatory power of order flow for exchange rates in single market and section 2.4 presents our inter-market analysis results. The following section presents the forecasting results. Some discussion of our findings is given in Section 2.6 and Section 2.7 concludes.

2.2 The FX Markets and the Data

2.2.1 The foreign exchange markets

The spot foreign exchange market is best described as a de-centralized multi-dealer market. In this market, market makers are large commercial banks located in major money centres, including London, New York, Tokyo, Zurich and Hongkong. These banks operate as dealers, trading with each other as well as with non-bank customers. Unlike equity market, FX market is a 24 hours market and has no opening and closing procedure. But, since the market activities (in the sense of quoting and trading) are very sparse on weekends and some holidays, it is practically viewed as closed during these periods.

Following Lyons (2001), we divided the spot FX market into three segments by their information structure characteristics: customer-dealer, brokered inter-dealer and direct inter-dealer. The customer-dealer segment is usually thought as the major source of information in the FX markets. However, due to the lack of transparency, the transactions between dealers and their non-bank customers largely remain private information to the dealer themselves. The brokered inter-dealer market is thought to be the most transparent part of the FX spot market and most of the transactions are conducted through EBS or Reuters D2000-2, the two major electronic dealing systems in this market.[6]

2.2.2 The Data

The data set used in the research comes from the brokered segment of the inter-dealer FX market and are drawn from the Reuters D2000-2 system. Thus the data contains no information on customer-dealer FX trades or on direct (i.e. non-intermediated) trades between dealers.[7] A subscriber to D2000-2 sees the following items on the trading screen, for up to 6 exchange rates:

- Best limit buy and sell prices

[6]EBS, Reuters D2000-2 are brokered inter-dealer systems and Reuters D2000-1 is a bilateral direct inter-dealer system. EBS claimed to handle 37% of the brokered trade in London and it is believed that Reuters has the same share. See Payne (2003).

[7]For a full description of the segments of the spot FX market and the data available from each segment see the excellent descriptions contained in Lyons (2001).

- The quantities available for trade at the best prices
- An indicator of the characteristics of the last trade rder flow can be forecast out of sample.

The raw data set is composed of tick level information (including transactions and firm quotes), covering four major floating rates: EUR/USD, EUR/GBP, GBP/ USD and USD/JPY. Each transaction record contains a time stamp for the trade, a variable indicating whether the trade was a market buy or sell and the transaction price. For the quotes, each record contains a time stamp, bid and ask prices.[8] Thus we do not need to make use of potentially inaccurate, ad hoc algorithms to assign trade direction. The samples for EUR/USD and GBP/USD cover a period of ten months from 28 September 1999 to 24 July 2000. Samples for EUR/GBP and USD/JPY cover a period of eight months from 1 December 1999 to 24 July 2000. One limitation of the data supplied is a lack of information about the size of each trade. Thus we cannot analyze whether the Dollar value of order flow matters over and above order flow measured simply in terms of numbers of trades.[9] Nevertheless this high frequency data set has two valuable characteristics: long sample periods and multiple exchange rates. The long sample period ensures reasonable statistical power for various econometric tests and the broad currency scope provides a platform to check the robustness of model estimation cross-sectionally on major floating exchange rates.

[8]Chapter two will only use transaction information and Chapter three will use information on both transaction and quote

[9]However, we do not expect the using the number of trades to proximate volume will alter the validity this research given the high correlation between the number of trade and volume reported in the literature, e.g Danielsson and Payne (2000) report a correlation of 0.94.

2.2.3 Filtering and time aggregation

For the later analysis, we time aggregate the transaction-level data to various degrees. Prior to time aggregation, however, we remove sparse trading periods from the data. Such sparse trading periods include the overnight periods, weekends, some world-wide public holidays and certain other dates where the feed from D2000-2 is very low.[10]

The analysis focuses on 8 different time aggregation levels: 5 minutes, 15 minutes, 30 minutes, 1 hour, 4 hours, 6 hours, 1 day and 1 week.[11] Note that the definition of one day in this research corresponds to a trading day defined as the interval between 6 and 18 DST. Thus one day covers 12 rather than 24 hours. Similarly, one week covers 5 trading days. The time aggregation is done as follows. First, The sample is scanned along calendar time. At every observation point the last transaction price is recorded along with the excess of the number of market buys over market sells since the last observations point. The logarithmic price changes is constructed from the price data.

After filtering and aggregation, 32 databases (8 sampling frequencies $\times$ 4 exchange rates) are created. These aggregated databases are the foundation for the model estimations and their properties are summarized in Table 2.1. As mentioned above, the long sample periods are valuable in that after filtration and aggregation we still have a decent number of observations at our lower sampling frequencies. For example, at the daily level, we have 201 observations for EUR/USD and GBP/USD and 160 observations for EUR/GBP and USD/JPY. The sample periods covers a time during which there was a depreciation of the EUR against the USD

[10]Overnight is defined as a period from 18:00 to 6:00 DST next day. It should be noted that this definition is only proper for the traders in London and New York, but not for the traders in Asian markets. It corresponds to the portion of the day when trade on D2000-2 is least intensive, even for USD/JPY.

[11]We have experimented with denser time aggregation levels and the results do not alter the pattern we reported here.

and GBP, a depreciation of GBP against USD and a depreciation of JPY against USD. These market trends are reflected in the columns of each panel in Table 2.1 that display mean returns. Comparing panel (b) with the other three panels, it is clear that the number of trades in USD/JPY is far less than for the other three markets. GBP/USD is the most heavily traded pair with EUR/USD and GBP/USD just behind. These numbers reflect two things. First, Reuters D2000-2 has poor coverage of JPY markets and, compared to its competing system (EBS), has a minority share in EUR/USD trade. In contrast, D2000-2 dominates trade in GBP rates.

2.3 Own Order Flow and Foreign exchange rate Determination

The study of the high frequency relationship between price changes and order flow has a long tradition in the market microstructure literature. In contrast, it is only recently that such relationships have been studied at lower sampling frequency pairs, such as daily or weekly.

We begin simply by tracking how the explanatory power of order flow for price changes varies across sampling frequencies and across currencies. To this end we run a set of regressions of the following form;

$$\Delta P(k)_{i,t} = \alpha(k)_i + \beta(k)_i Q(k)_{i,t} + \varepsilon_t \tag{2.1}$$

where $\Delta P(k)_{i,t}$ is the change in transaction price for currency pair i at sampling frequency k and $Q(k)_{i,t}$ is the order flow in the interval ending at t for currency pair i at sampling frequency k. While several authors have provided results from regressions such as that above for specific currencies at specific sampling frequencies, our analysis looks at four exchange rates from a sampling frequency of 5 minutes all the way through weekly.

Table 2.2 contains the estimation results for model (2.1) for our four exchange rates and over the entire spectrum of time aggregation levels.[12]

At the highest frequencies (less than one hour) there is significant effects from order flow for all currencies, with the strongest effects for EUR/USD and R^2 ranges from 33% to 45%. These results confirm what microstructure economists have long known — order flow carries information for high-frequency asset price determination. People who are sceptical to the information explanation may think that the regressions at very high frequencies are simple capturing the liquidity effects.

As a result, from the point of view of exchange rate determination, we move toward lower frequencies. Consider first results from the daily frequency frequency, initially for EUR/USD and USD/JPY in order to provide comparability with Evans and Lyons (2002a).[13] Their daily USD/DEM and USD/JPY regression R^2 were just over 0.60 and 0.40 respectively which are broadly consistent with our results. The results reported in this research directly corroborate those of Evans and Lyons (2002a).

However, our results on the GBP related exchange rates are much less supportive of their results. By looking at the low frequency regressions in the final two panels of Table 2.2, one can see that the explanatory power of order flow for GBP/EUR and GPB/USD is very poor. At sampling frequencies exceed-

[12]Since the normality of our return data is rejected by the Jarque-Bera test (not reported), we also experimented with a LAD estimator for these regressions, but the results were not qualitatively affected.

[13]Note that our definition of the aggregation time interval is slightly different from that in ?. Whilst their 'daily' aggregation interval is defined as a period from 4:00 pm to 4:00 pm next day our definition is a period from 6:00 am to 6:00 pm excluding overnight period. We also experimented with a interval definition that includes overnight period in this comparison study and find results that do not differ qualitatively from those we reported here.

ing one hour, in no single case does the regression R^2 exceed 0.10, although in five of the eight cases the order flow variable is statistically significant. Thus, at least for GBP, the assertion that order flow matters for exchange rate determination when one moves towards sampling frequencies that matter to international macroeconomists appears less secure than our EUR/USD and USD/JPY results suggested.

A graphical representation of these results using a somewhat more dense set of sampling frequencies is given in Figure 2.1. The figure clearly demonstrates the importance of order flow regardless of sampling frequency for EUR/USD and USD/JPY but also points to the declining effect of order flow in the Sterling markets.

2.4 Inter-market Order Flow Analysis

Most existing order flow research focuses on one asset at a time.[14] However, since exchange rates are relative prices of currencies, and three of our exchange rates form a triangle relationship, it is of interest to investigate how order flow in one currency pair might be used to explain the exchange rate of another currency pair. This is denoted as *inter-market order flow* analysis.

The reason for considering inter-market effect is the peculiar nature of currency markets. In theory, an informed currency trader can have multi-route trading strategies to exploit his information. In a broad sense, consider, e.g., a trader has superior information about the future value of USD, perhaps that USD can be expected to appreciate vis-a-vis other currencies. The trader can exploit this information by trading in USD/JPY, EUR/USD, GBP/USD, and so on. Holding others constant, an order flow in one market (e.g. GBP/USD) can convey valuable

[14] Evans and Lyons (2002b) is an exception.

information to other relevant markets (USD/JPY). Given the fact that three of the currency pairs used in this study forms a perfect triangle system, this multi-route trading possibility is particularly important because any trading within the triangle system can be realized by direct or in-direct trading. Suppose a investor has superior information about the relative value of EUR to GBP (e.g. Euro will appreciate in the future). He can buy EUR directly in EUR/GBP market and in this case we will see the standard single market result, i.e., a positive impact of EUR/GBP order flow on EUR/GBP rate. However this investor can also trade GBP/USD and EUR/USD to exploit his information advantage if the in-direct markets are liquid and price impacts are small. This reasoning is actually in line with the argument of vehicle currency literature that transaction cost is the major determinant on the choice of trading currencies (e.g. Matsuyama et al. (1993) and Rey (2001)). Under this hypothesis, order flow in GBP/USD and EUR/USD markets conveys information about EUR/GBP. Since liquidity suppliers in one market can observe the order flow in other market, they can revise their valuation of a currency without actual trading.

It is in this analysis that the multi-currency nature of our data set becomes extremely useful. The inter-market effects is investigated by extending (2.1) to include the order flow from all currency pairs, while still remaining within the linear specification that relates price changes in market i to contemporaneous order flows in all four markets. The specification is given below;

$$\Delta P(k)_{i,t} = \alpha(k)_i + \sum_j \beta(k)_{i,j} Q(k)_{j,t} + \varepsilon_{i,t}\,, \quad i,j = ED, SD, ES, DY \qquad (2.2)$$

where, as before, k indexes sampling frequency, i is the rate to be explained and the summation over j gives an explanatory term that is linear in all four order flow variables. Table 2.3 presents the main results from estimating (2.2) for all rates and all sampling frequencies, while the change in R^2 is shown in Figure 2.2.

Consider first the results for USD/JPY as it is the only JPY rate and because the other three rates form a triangulating relationship. We see that for USD/JPY, aside from the strong own flow effects uncovered in Section 2.3, there are few other significant flow variables. As one might expect, no GBP/EUR flow variable is significant. A couple of EUR/USD and GBP/USD flows are significant and, as expected given the definition of the rates, they enter with negative signs.[15] In all cases the improvement in the R^2 of the regressions as compared to the univariate specifications in Section 2.3 is small.

The results for other three exchange rates are striking. For EUR/USD, panel (a) of Table 2.3, the order flow coefficients of EUR/GBP and GBP/USD are, as expected, consistently positive and significant at the 1 percent level at relatively high frequencies. The significance of GBP/USD flow persists to the daily level. Also, USD/JPY flow is significant, with the expected negative coefficient, at very high sampling frequencies. Overall, these effects lead to improvements in explanatory power (change in R^2 or ΔR^2) up to 6% and for all specifications below the daily level this improvement is significant.

For GBP rates, the results are very interesting. Flows in the other GBP rate (EUR/GBP flow in the GBP/USD price change regressions and vice versa) are strongly significant at higher frequencies while USD/JPY flows have virtually no effects. However, the dominant new right-hand side variable in these regressions is EUR/USD flow. In each and every case for these two exchange rates, EUR/USD flows are strongly significant with a positive coefficient. These extended specifications show markedly improved explanatory power over the univariate models in Section 2.3, of between 5% and 35% with the largest improvements being at the lowest sampling frequencies.[16] In all cases, the extra right-hand side variables

[15] A negative flow in these rates means Dollar sales, thus driving the Yen price of a Dollar down.

[16] Note, even if there is substantial improvement of R^2 for both EUR/GBP and GBP/USD when

can be shown to significantly improve the explanatory power of the regression. It is for the EUR/GBP that the effects of EUR/USD flow is strongest, providing virtually all explanatory power at the lower frequencies.

Bring the the results of Table 2.2 and Table 2.3 together, a seemingly puzzle emerges: Rueters D2000-2 has the best coverage of Sterling based exchange rates, the explanatory power of Euro-Dollar order flow, which is not mainly covered by D2000-2 system, is much stronger than Sterling order flow. However, if price change is more due to information flow than duo to liquidity effects, flow coverage of a particular system might not be a critical determinant of flow explanatory power if the information carried by order flow is visible to market participants regardless where the flow is traded. Since Euro-Dollar is most liquid asset in FX market, information that drives other assets movements (e.g. Euro-Sterling) can very likely be represented in Euro-Dollar flows.[17] This conjecture is consistent with the results in Panel (c) and (d) of Table 2.3, where the ΔR^2 of (c) is much larger than ΔR^2 of (d), because if Euro-Sterling is less liquid than Sterling-Dollar information about Euro-Sterling is more likely to be explored through Euro-Dollar than in the case of Sterling-Dollar. This puzzle can be further investigated by confining the inter-market analysis to the two Sterling based exchange rates in our sample because D2000-2 has best coverage of them. I exclude Dollar-Yen and Euro-Dollar from the regression 2.2 and results are reported in Table 2.4. When sampling frequency is high, for both EUR/GBP and GBP/USD, adding the order flow of second rate does not change R^2 very much. When sampling frequencies are lower than 4 hours, ΔR^2 is larger when GBP/USD flow is added in EUR/GBP equation than when EUR/GBP flow is added in GBP/USD equation. This result is also consistent with our information hypothesis because GBP/USD is more liquid

other order flow is included in the regressions, R^2 for GBP/USD suffers decrease from 30% to 15% percent. This is probably due to the sample specific randomness.

[17]This hypothesis has be confirmed by my recent talks with the primary traders in Barclays Global Investors (BGI), one of the largest currency managers in the world.

than EUR/GBP and information can be more easily exploited through GBP/USD than EUR/GBP.

These results provide clear evidence of flow information being transmitted across linked exchange rate markets, especially for the less liquid markets. The EUR/USD exchange rate is the largest and most liquid in the world, and its order flow is shown to dominate across all three triangular currency pairs. This is especially apparent for the least liquid of these three currency pairs, EUR/GBP.

In sum, the presence of significant inter-market order flow effects indicates that information spills over linked markets and reinforces the notion that order flow is a significant determinant of exchange rates. Furthermore, the fact that the order flow from the largest currencies dominates the determination of the smaller currencies, suggests that new information first flows into the most liquid markets, i.e. where new information can be best exploited. This result has important implication for the studies on exchange rate determination, especially for less liquid exchange rates.

2.5 Forecasting Ability Analysis

The order flow models estimated above (2.1) and (2.2) used contemporaneous order flow to explain exchange rate changes. However, as argued by Frankel and Rose (1995,pp. 1702) *"Fitting exchange rates to contemporary observable variables, in-sample, is one thing. Forecasting out of sample is quite another"*. The forecast ability of exchange models is examined by Meese and Rogoff (1983a,b) who study the out-of-sample forecasting ability of various structural and time series models from 1 to 12 months and concluded that none of these models performed any better than a random walk model at short horizons (one month). This research provides a first investigation of the out-of-sample forecasting performance of the order flow models for exchange rates, across different sampling frequencies and using a variety of forecasting specifications. We first use the methodology proposed by Meese and Rogoff (1983a,b), and then extend this to genuine out-of-sample forecasts testing.

2.5.1 Meese-Rogoff Out-of-Sample Forecasts

The Meese and Rogoff (1983a,b) test requires using data up until time t to estimate the parameters of the relationship between price changes and order flow, and then using the estimated relationship to forecast the price change at $t+1$ based on observed order flow at $t+1$. The root mean squared error (RMSE) from order flow model is then compared to the RMSE of a random walk (RW) model with a drift. Thus the Meese-Rogoff test is not a genuine out-of-sample forecasting experiment as observed future order flow is used in the forecasting.

The sampling frequency analyzed ranges from 5 minutes to 1 week and for each sampling frequency we look at forecasting horizons from one to 12 observations.[18] The forecasting equation that is equivalent to the regression model (2.1) is thus given by;

$$\Delta\hat{P}(k)_{i,t+h} = \alpha(k)_{i,t} + \beta(k)_{i,t}Q(k)_{i,t+h} + \varepsilon_{t+h} \tag{2.3}$$

where $\Delta\hat{P}(k)_{i,t+h}$ is the return of a specific time interval (defined by the sampling frequency) *h-step-ahead* of time t and forecasted at time t. $Q(k)_{i,t+h}$ is the order flow of the time interval over which the return is forecasted. we have added time subscripts to the regression intercept and slope coefficients to emphasize that they are estimated using information until t only.

[18]Take a frequency of four hours and a horizon of 6 as an example. The forecast horizon, in terms of hours, is 24 hours (4×6). Since 12 hours represents one trading day (as the overnight period has been excluded) 24 hours represents a two-day forecast

The benchmark forecasting model is a drifting random walk where log price follows a random walk with drift. The *h-step-ahead* price change is forecast to be the average exchange rate change from the beginning of the sample till time t.

$$\Delta P(k)_{i,t+h} = \mu(k)_{i,t} + \eta_{t+h} \tag{2.4}$$

where $\mu_{i,k,t}$ is the estimated drift based on information up to time t only (for exchange rate i and sampling frequency k). The models are estimated and forecast recursively. The initial forecast estimation period is the first four months of the data in all cases. The criterion used to judge forecast accuracy is the RMSE, comparing that of the order flow model with that of the random walk.

The main results are reported in Table 2.5. The columns headed 'OF' and 'RW' are the RMSEs generated by forecast models (2.3) and (2.4) respectively. The t-stats comparing forecast accuracy are those given in Diebold (2001, pp. 293). The most striking feature of Table 2.5 is that the RMSEs generated by the order flow models are virtually all lower than those generated by RW model. Furthermore, for all exchange rates, this forecast improvement is significant at higher sampling frequencies. When sampling frequency goes towards the lower end, the forecast improvements become insignificant for Sterling based rates.[19] Nevertheless, if we put all four exchange rates, sampling frequencies and forecasting horizons together, we find that overall, order flow based forecasting analysis outperforms that of the macro models considered by Meese and Rogoff (1983a,b). Here we show that even at the daily and weekly sampling frequencies, heavily traded exchange rates such as EUR/USD and USD/JPY can be forecasted using order flow. Furthermore, since these results are generated only by using own order

[19]The lack of significance for Sterling based rates might be due to the insufficient explanatory power of Sterling order flow model itself discussed in section 2.4, i.e. information about Sterling rates might not manifested itself through Sterling flow directly.

flow, the GBP results would probably be improved considerably when flows from other markets are included.

2.5.2 Genuine Forecasting

Since the Meese-Rogoff test is no a genuine forecast test, this subsection extends the forecast results above by considering true forecasts of price changes, based not on order flow observations arriving after the forecast date, but only using order flow information available at the forecast date. Thus, it is expected that these forecasting results would be somewhat less impressive than those in the prior subsection. The focus here is on the one-step ahead forecasting for all of the previously studied sampling frequencies and for all four exchange rates. The order flow based forecasts are drawn from the following specification;

$$\Delta P(k)_{i,t+1} = \alpha(k)_{i,t} + \beta(k)_{i,t} Q(k)_{i,t} + \varepsilon_{t+1} \tag{2.5}$$

To emphasize, when constructing the forecast of the price change at $t+1$ from the prior specification, the intercept and slope in the regression are estimated using information available at t only and the order flow observation entering into the forecast is dated t also. Thus, these are genuine forecasts. Again we compare the ability of specification (2.5) to forecast price changes with the forecast produced by the random walk model (2.4). Results are presented in Table 2.6 for the entire spectrum of sampling frequencies and exchange rates.

Inspection of this table indicates that if there is any statistical significance in our somewhat naive linear specification then it is concentrated at the highest frequency, i.e. 5 minutes. For all of the regressions considered here, the RMSE of the order flow forecasting model is virtually the same as, or at most only marginally below, that of the random walk forecast. Thus, the explanatory power of our genuine forecasting regressions is poor and there is little evidence that these simple linear specifications contain true forecasting power. Only at the highest frequencies is the relationship between order flow at t and the one-period price change to $t+1$ positive and significant.

2.5.3 Order Flow Forecasting

The final forecasting exercise focusses on the predictability of order flow itself. The aim is to test whether flows can be forecast with past information on flows themselves and price changes. If this was the case, then another route to forecasting exchange rate changes would possibly exist. One could combine the strong contemporaneous relationship between price changes and order flows uncovered in Section 2.3 and an order flow forecast to construct a price forecast. The forecasting model considered here is specified as;

$$Q(k)_{i,t+1} = \alpha(k)_{i,t} + \sum_{j=1}^{J} \beta(k)_{j,i,t} \Delta P(k)_{i,t-j+1} + \sum_{l=1}^{L} \gamma(k)_{l,i,t} Q(k)_{t-l+1} + \varepsilon_{t+1} \quad (2.6)$$

i.e. for a given sampling frequency (k) and exchange rate (i), the flow at $t+1$ is regressed on it's own first L lags and on J lags of the price change. In the estimations both J and L is set at 2 after some experimentation with alternative lag lengths. Results from estimations are given in Table 2.7.

The results indicate that the majority of the statistical significance in the forecasting regressions comes at very high frequencies. Even though there is evidence of high-frequency positive dependence in order flow, in all cases the RMSE from the random walk model and (2.6) are virtually identical.

For the GBP exchange rates there is also evidence of negative dependence of current flow on past returns. Thus, when prices have been rising in the recent past, order flows tend to become negative — a manifestation of contrarian or negative feedback trading. This causality is reversed for the USD/JPY. Thus, in this case there would seem to be evidence of aggressive momentum type trades.

While there is significant relationship between current flows and past return and flow information, our simple linear specifications cannot be used to forecast price changes. The results do suggest however that there is some potential for the creation of a sophisticated forecast model for prices and flows.

2.6 Discussion

The current study has thrown up a number of new and interesting results on the explanatory power, forecasting ability and inter-market effects of order flow analysis. However, one clear implication of this work is an affirmation of results from previous analysis — order flow has strong explanatory power for exchange rate changes. This explanatory power is very strong at very high frequencies but persists to much lower frequencies, daily and weekly, that can genuinely be considered of interest to international macroeconomists. This research provides strong evidence that currency flows carry information, confirming the evidence contained in Payne (2003), Evans and Lyons(2002a) and Rime (2000) amongst others. Of course we cannot uncover whether order flow carries information regarding fundamentals, long-run risk premia or a mixture of the two, but the key result is that flows are informative to those looking from a trading perspective and those with a medium term macroeconomic view.

While the results on the longer-run relationship between flows and exchange rate changes in this chapter are similar to those derived by Evans and Lyons (2002a), there is some interesting and important differences between both their results and the data upon which they are based and those presented in this research. First of all, data used in this research is drawn from the electronically brokered segment of the market while theirs is from the direct trading segment. In the former case this implies that there is pre-trade anonymity but trades are published to the market at large. In the latter case, quoting and trading is clearly non-anonymous but the occurrence and details of trades are both kept private to the counterparties. Based on this, brokered trades may have different information content, and in this research, strong evidence of information effects is identified in the brokered segment. Thus results provided in this research provide strong corroborating evidence for the results of Evans and Lyons (2002a), especially when considering the different data sources and sample periods.

However, results in this chapter contain a very important difference to those in Evans and Lyons(2002a). The univariate regressions (equation (2.1)) of price changes on order flow for GBP exchange rates perform very poorly at lower sampling frequencies, with explanatory power close to zero. This appears to fly in the face of the preceding discussion — perhaps the USD/JPY and EUR/USD results are anomalous and order flow has no long run effect on exchange rates for the majority of currency pairs. While this is clearly a possibility (despite empirical evidence to the contrary for Scandinavian exchange rates in Rime (2000)), we feel that such a conclusion would be unwarranted. Indeed, it is demonstrated in the inter-market regressions (equation (2.2)) that once one allows for aggressive buying and selling pressure in related markets, order flows have strong effects in all four exchange rates at every sampling frequency considered in this study. This is a key new result. Therefore order flow carries information that not only affects exchange rate changes in its own market but also in the other markets. Empirically we see information instantly spilling over from market A (via order flow) to prices in market B.

It is interesting to note that the dominant flow variable in this data set is

EUR/USD flow. EUR/USD flow has clear and persistent effects on both EUR/GBP and GBP/USD rates. This is intuitive since because the EUR/USD is the most liquid and heavily traded currency pair globally, one can expect that any information about either EUR or USD would hit this market first due to its low transaction costs and massive participation. Thus those quoting prices in related pairs will very likely keep an eye on EUR/USD developments, including order flow, in forming of their prices.

A final point to note regarding the inter-market flow analysis carried out in Section 2.4 is that in this analysis we see prices for a given rate move in the absence of trade in that rate, as they are affected by flows occurring in *other markets*. Thus, one cannot explain away the importance of order flow in an inter-market context by simply asserting that aggressive buying or selling pressure is just temporarily moving prices due to low market liquidity and that after such "digestion effects" have run their course prices would revert — here there is nothing to digest aside from information. This, in our view, only serves to reinforce evidence that order flows do carry information and also information that is relevant at macroeconomic sampling frequencies.

The final part of this study is the forecasting power of order flows for exchange rates. Three sets of results are established. First, the order flow model beats the same random walk benchmark that macroeconomic models of 70s and 80s lost out to. The second is a true one-step ahead out-of-sample experiment. This experiment shows that order flow forecasts hardly reduce RMSEs relative to random walks (If there is any, it is only at the highest sampling frequencies i.e. 5 minutes). Finally, order flow is found to be predictable.

2.7 Conclusion

This chapter studies the explanatory and forecasting power of order flow for exchange rates changes at sampling frequencies ranging from 5 minutes to one week. It is demonstrated that order flow analysis has strong power to both explain and forecast exchange rate changes at virtually all frequencies. The key results of this research are as follows;

1. The contemporaneous relationship between flows and changes in rates is very strong at intra-day frequencies for all four rates.

2. At the daily and weekly level, there is still strong explanatory power of order flow for exchange rates changes for EUR/USD and USD/JPY. This is not the case for EUR/GBP and GBP/USD.

3. However, when one examines the inter-market effects of order flows, one sees that price changes for EUR/GBP and GBP/USD are strongly affected by EUR/USD order flow and, taking these effects into account, overall flows have strong explanatory power for the GBP rates. The result that EUR/USD order flow significantly explains EUR/GBP and GBP/USD rates suggests that a significant portion of information on GBP is revealed in the trading activities in EUR/USD market.

4. An analysis of the forecasting power of order flows, using the technique of Meese and Rogoff (1983a,b), demonstrates the flow analysis outperforms a naive benchmark across essentially all sampling frequencies for all rates.

5. A true out-of-sample forecasting experiment, however, demonstrates that the simple order flow models do not provide terribly valuable exchange rate forecasts aside from at extremely high sampling frequencies.

6. Order flow can be forecasted out of sample and this implies the possibility of the creation of some more sophisticated forecasting models.

These results serve to emphasize the role played by order flow in foreign exchange, and possibly other markets. This chapter provides clear evidence that order flows can be used to explain and forecast rates at trading desk sampling frequencies as well as observations intervals relevant to international macroeconomics. The information content of order flow implies that simple symmetric information, rational expectations models of exchange rate determination are not consistent with the data. Further work on modelling exchange rates to take account of these effects as well as further empirical work to clarify the role of order flow in exchange rate determination can only help move exchange rate analysis out of the cul-de-sac in which it has resided for the last two decades or so.

Table 2.1: Summary of time aggregated databases

In each panel of Table 1, k is sampling frequency. Obs is the total number of (derived) observations in that database. $\bar{r}$ is the average return for that sampling frequency. Returns are defined as $100 \times (log(P_t) - log(P_{t-1}))$. Columns headed Trades, Quotes, Buys and σ give the average number of trades, average number of quotes, average number of buys and standard deviation of returns for that frequency.

	EUR/USD(a)						USD/JPY(b)					
k	Obs	Trades	Quotes	Buys	$\bar{r}$	σ	Obs	Trades	Quotes	Buys	$\bar{r}$	σ
5m	29107	16	51	8	-0.0006	0.06	23148	1	7	1	0.0004	0.08
15m	9701	49	153	25	-0.0017	0.10	7715	4	21	2	0.0008	0.10
30m	4850	98	306	49	-0.0038	0.13	3857	7	41	4	0.0022	0.13
1hr	2424	196	611	99	-0.0050	0.20	1928	15	83	8	0.0038	0.18
4hr	605	782	2444	395	-0.0196	0.40	481	58	330	30	-0.0052	0.37
6hr	404	1174	3669	593	-0.0313	0.47	321	88	496	45	0.0089	0.41
12hr	201	2347	7317	1185	-0.0676	0.62	160	175	988	90	0.0351	0.56
1wk	42	11305	35831	5702	-0.3373	1.53	33	1024	5961	526	0.2785	1.22

	EUR/GBP(a)						GBP/USD(b)					
k	Obs	Trades	Quotes	Buys	$\bar{r}$	σ	Obs	Trades	Quotes	Buys	$\bar{r}$	σ
5m	23148	14	34	8	-0.0002	0.05	29107	17	44	9	-0.0004	0.04
15m	7715	43	103	23	-0.0007	0.09	9701	52	131	27	-0.0012	0.07
30m	3857	87	206	45	-0.0015	0.13	4850	104	263	53	-0.0029	0.09
1hr	1928	174	411	90	-0.0025	0.18	2424	208	525	106	-0.0049	0.13
4hr	481	694	1646	362	-0.0106	0.37	605	832	2098	424	-0.0200	0.26
6hr	321	1041	2468	542	-0.0160	0.45	404	1249	3150	636	-0.0291	0.32
12hr	160	2085	4944	1086	-0.0384	0.61	201	2496	6280	1271	-0.0603	0.45
1wk	33	10383	25506	5423	-0.0482	1.36	42	13245	35328	6753	-0.2299	0.92

Table 2.2: Explaining Exchange Rates with Order Flow

$$\Delta P(k)_{i,t} = \alpha(k)_i + \beta(k)_i Q(k)_{i,t} + \varepsilon_t$$

where $\Delta(k)P_{i,t}$ is price change at sampling frequency k for exchange rate i at time t and $Q(k)_{i,t}$ is order flow for the same exchange rate at the same sampling frequency. All t-values are constructed using the Newey-West estimator of the coefficient variance-covariance matrix. The order flow is scaled by 10^{-2}.

	EUR/USD			USD/JPY			EUR/GBP			GBP/USD		
k.	β	t-value	R^2	β	t-value	R^2	β	t-value	R^2	β	t-value	R^2
5m	0.40	72.39	0.33	1.08	24.71	0.06	0.41	60.30	0.26	0.29	65.07	0.26
15m	0.38	53.25	0.43	1.17	26.53	0.15	0.38	32.01	0.26	0.26	36.58	0.24
30m	0.36	45.30	0.45	1.19	20.96	0.25	0.33	20.51	0.21	0.23	21.30	0.21
1hr	0.36	29.91	0.38	1.25	18.98	0.30	0.30	12.85	0.16	0.21	13.95	0.16
4hr	0.34	17.63	0.38	1.14	9.70	0.30	0.16	3.00	0.05	0.13	3.95	0.05
6hr	0.34	15.35	0.38	1.21	10.59	0.42	0.10	2.00	0.02	0.11	3.66	0.05
12hr	0.30	11.04	0.35	1.17	10.70	0.50	0.02	0.36	0.00	0.14	4.12	0.08
1wk	0.31	5.51	0.45	0.91	11.43	0.67	0.06	0.59	0.01	0.05	0.70	0.01

Table 2.3: Inter-market Information flow

$$\Delta P(k)_{i,t} = \alpha(k)_i + \sum_j \beta(k)_{i,j} Q(k)_{j,t} + \varepsilon_{i,t}, \quad i = ED, DY, ES, SD$$

where k indexes sampling frequency, i is the rate to be explained and the summation over j gives an explanatory term that is linear in all four order flow variables. Columns headed ΔR^2 give the changes in R^2 between the model with and without order flow from other markets. The last column in each panel is the value of the F-test of the null $H_0 : \beta_j = 0$ for $j \neq i$. The order flow is scaled by 10^{-2}. ^{a,b,c} indicate the 1%, 5% or 10% significance level by using the Newey-West coefficient variance-covariance estimator.

	EUR/USD (a)						USD/JPY (b)					
k	$\hat\beta_{ED}$	$\hat\beta_{DY}$	$\hat\beta_{ES}$	$\hat\beta_{SD}$	ΔR^2	p-value	$\hat\beta_{ED}$	$\hat\beta_{DY}$	$\hat\beta_{ES}$	$\hat\beta_{SD}$	ΔR^2	p-value
5m	0.32^a	-0.05^a	0.18^a	0.14^a	0.054	0.01	-0.02^b	1.07^a	-0.01	-0.01	0.001	0.01
15m	0.31^a	-0.05^b	0.16^a	0.13^a	0.056	0.01	-0.02^c	1.17^a	-0.00	-0.01	0.001	0.05
30m	0.31^a	-0.10^a	0.13^a	0.11^a	0.048	0.01	-0.01	1.19^a	0.00	-0.01	0.001	> 0.10
1hr	0.32^a	-0.01^b	0.11^a	0.12^a	0.039	0.01	-0.02	1.24^a	-0.00	-0.03^b	0.005	0.01
4hr	0.35^a	-0.05	0.03	0.10^a	0.016	0.01	-0.04^c	1.13^a	0.05	-0.03	0.010	0.10
6hr	0.34^a	-0.09	0.03	0.12^a	0.029	0.01	0.01	1.22^a	0.01	-0.03	0.002	> 0.10
12hr	0.33^a	-0.03	-0.05	0.08^c	0.015	> 0.10	0.02	1.17^a	-0.02	-0.02	0.003	> 0.10
1wk	0.39^a	-0.01	-0.00	0.15	0.044	> 0.10	0.04	0.89^a	0.06	-0.12^b	0.057	> 0.10

	EUR/GBP (c)						GBP/USD (d)					
k	$\hat\beta_{ED}$	$\hat\beta_{DY}$	$\hat\beta_{ES}$	$\hat\beta_{SD}$	ΔR^2	p-value	$\hat\beta_{ED}$	$\hat\beta_{DY}$	$\hat\beta_{ES}$	$\hat\beta_{SD}$	ΔR^2	p-value
5m	0.21^a	-0.02	0.30^a	-0.13^a	0.099	0.01	0.10^a	-0.03^b	-0.11^a	0.27^a	0.041	0.01
15m	0.23^a	-0.02	0.26^a	-0.13^a	0.130	0.01	0.09^a	-0.03	-0.10^a	0.25^a	0.045	0.01
30m	0.23^a	-0.05	0.22^a	-0.11^a	0.142	0.01	0.08^a	-0.03	-0.10^a	0.22^a	0.048	0.01
1hr	0.23^a	-0.03	0.19^a	-0.09^a	0.146	0.01	0.10^a	-0.06	-0.09^a	0.20^a	0.055	0.01
4hr	0.26^a	-0.05	0.07	-0.02	0.220	0.01	0.10^a	-0.02	-0.07^b	0.10^a	0.065	0.01
6hr	0.28^a	-0.04	0.01	0.01	0.235	0.01	0.09^a	-0.09	-0.01	0.09^b	0.053	0.01
12hr	0.28^a	-0.11	-0.03	-0.04	0.263	0.01	0.10^a	-0.00	-0.02	0.12^a	0.054	0.01
1wk	0.29^a	-0.07	0.01	0.08	0.335	0.01	0.12^b	0.05	-0.00	0.07	0.107	0.05

Table 2.4: Inter-market information flow experiments

$$\Delta P(k)_i = \alpha(k) + \beta(k)_i Q(k)_i + \varepsilon(k)_i, , \quad i = ES, SD$$

$$\Delta P(k)_i = \alpha(k) + \beta(k)_{i,1} Q(k)_{ES} + \beta(k)_{i,2} Q(k)_{SD} + \varepsilon(k)_i, , \quad i = ES, SD$$

where k indexes sampling frequency and i is the exchange rate to be explained. Time index t has been suppressed for simplicity. $\Delta P(k)^i$ is price change at sampling frequency k for exchange rate i and $Q(k)^i$ is order flow for the same exchange rate and same sampling frequency. The columns under ΔR^2_{ES} and ΔR^2_{SD} give the difference of R^2 of the above two regression equations

Freq	ΔR^2_{ES}	ΔR^2_{SD}
5m	0.78%	0.79%
10m	0.87%	0.75%
15m	0.72%	0.78%
20m	0.55%	0.79%
25m	0.65%	1.04%
30m	0.50%	0.97%
1hr	0.08%	0.77%
2hr	0.07%	0.77%
4hr	0.50%	0.51%
6hr	1.90%	0.06%
12hr	0.41%	0.00%
1wk	3.31%	0.00%

Table 2.5: Meese-Rogoff (1983) forecasting experiments: Root Mean Squared Errors (RMSE)

h is forecast horizon in observations, *k* is the sampling interval. Forecasting horizon in real time is defined as $h \times k$. The columns under *OF* and *RW* give the RMSEs of the *h-step-ahead* return forecast for the order flow and random walk models (2.3) and (2.4). The *t*-statistic for forecast improvement of the order flow model over the random walk is as given in Diebold(2001) (pp. 293).

		EUR/USD			USD/JPY			EUR/GBP			GBP/USD		
Freq	h	OF	RW	*t*-stats	OF	RW	*t*-stats	OF	RW	*t*-stats	OF	RW	*t*-stats
5m	1	0.05	0.06	-6.23	0.07	0.08	-0.52	0.05	0.05	-4.38	0.03	0.04	-4.28
	6	0.05	0.06	-6.23	0.07	0.08	-0.52	0.05	0.05	-4.38	0.03	0.04	-4.28
	12	0.05	0.06	-6.23	0.07	0.08	-0.52	0.05	0.05	-4.38	0.03	0.04	-4.28
15m	1	0.07	0.10	-7.82	0.09	0.10	-1.30	0.08	0.09	-3.44	0.06	0.07	-2.99
	6	0.07	0.10	-7.82	0.09	0.10	-1.30	0.08	0.09	-3.44	0.06	0.07	-2.98
	12	0.07	0.10	-7.82	0.09	0.10	-1.29	0.08	0.09	-3.43	0.06	0.07	-2.98
30m	1	0.10	0.13	-7.84	0.11	0.13	-1.72	0.12	0.13	-2.16	0.08	0.09	-2.32
	6	0.10	0.13	-7.84	0.11	0.13	-1.70	0.12	0.13	-2.15	0.08	0.09	-2.32
	12	0.10	0.13	-7.84	0.11	0.13	-1.70	0.12	0.13	-2.16	0.08	0.09	-2.31
1hr	1	0.16	0.20	-4.03	0.14	0.17	-2.57	0.18	0.19	-1.19	0.13	0.14	-1.29
	6	0.16	0.21	-4.03	0.14	0.17	-2.56	0.18	0.19	-1.20	0.13	0.14	-1.28
	12	0.16	0.21	-4.00	0.14	0.17	-2.55	0.18	0.19	-1.20	0.13	0.14	-1.28
4hr	1	0.33	0.42	-2.50	0.32	0.37	-0.94	0.36	0.37	-0.25	0.27	0.27	-0.04
	6	0.33	0.42	-2.46	0.33	0.38	-0.90	0.37	0.38	-0.26	0.27	0.27	-0.04
	12	0.33	0.41	-2.31	0.33	0.38	-0.89	0.37	0.38	-0.28	0.27	0.27	-0.05
6hr	1	0.39	0.50	-2.40	0.32	0.43	-2.97	0.48	0.48	-0.04	0.32	0.33	-0.08
	6	0.40	0.50	-2.38	0.32	0.43	-2.96	0.48	0.49	-0.02	0.32	0.33	-0.08
	12	0.40	0.49	-2.20	0.32	0.44	-2.97	0.48	0.48	-0.01	0.32	0.33	-0.08
12hr	1	0.54	0.66	-2.07	0.40	0.58	-2.63	0.67	0.66	0.07	0.46	0.47	-0.16
	6	0.53	0.65	-1.94	0.41	0.58	-2.47	0.69	0.68	0.12	0.47	0.47	-0.13
	12	0.54	0.66	-1.82	0.42	0.60	-2.47	0.68	0.67	0.16	0.47	0.47	-0.14
1wk	1	1.28	1.62	-1.10	0.76	1.20	-1.86	1.68	1.63	0.12	0.97	0.94	0.15
	6	1.39	1.76	-1.18	0.80	1.24	-1.51	1.85	1.84	0.04	0.97	0.94	0.12
	12	1.51	1.96	-1.16	0.62	1.18	-1.60	1.31	1.32	-0.02	0.98	1.01	-0.17

Table 2.6: Out-of-sample forecast experiments

$$\Delta P(k)_{i,t+1} = \alpha(k)_{i,t} + \beta(k)_{i,t} Q(k)_{i,t} + \varepsilon_{t+1}$$

where $\Delta(k)P_{i,t+1}$ is price change at sampling frequency k for exchange rate i at time $t+1$ and $Q(k)_{i,t}$ is order flow for the same exchange rate and same sampling frequency at t. The columns under *OF* and *RW* give the forecast RMSEs of the model above and random walk models and the t-statistic for the forecast improvement over random walk is reported in the last column of each panel. The order flow is scaled down 10^{-2}. ^{a,b,c} indicate the 1%, 5% or 10% significance level by using the Newey-West coefficient variance-covariance estimator.

	EUR/USD (a)					USD/JPY(b)				
Freq	$\hat{\beta}$	R^2	OF	RW	t-stats	$\hat{\beta}$	R^2	OF	RW	t-stats
5m	0.03^a	0.002	0.06	0.06	-0.02	0.09^b	0.000	0.09	0.09	0.00
15m	-0.01^c	0.000	0.10	0.10	0.00	-0.01	0.000	0.09	0.09	0.00
30m	-0.00	0.000	0.13	0.13	0.01	-0.12^a	0.003	0.13	0.13	0.03
1hr	0.01	0.001	0.20	0.20	0.00	0.02	0.000	0.17	0.17	0.00
4hr	0.01	0.000	0.42	0.42	0.02	0.09	0.002	0.37	0.37	0.01
6hr	0.00	0.000	0.50	0.50	0.02	0.03	0.000	0.43	0.43	0.04
12hr	-0.04	0.007	0.67	0.66	0.02	0.03	0.000	0.58	0.58	0.03
1wk	-0.10^b	0.041	1.62	1.62	-0.01	0.12	0.011	1.22	1.20	0.10

	EUR/GBP (c)					GBP/USD (d)				
Freq	$\hat{\beta}$	R^2	OF	RW	t-stats	$\hat{\beta}$	R^2	OF	RW	t-stats
5m	0.05^a	0.004	0.05	0.05	-0.04	0.02^a	0.001	0.04	0.04	0.00
15m	-0.01	0.000	0.08	0.08	0.02	-0.04	0.000	0.07	0.07	0.00
30m	-0.00	0.000	0.13	0.13	0.02	0.00	0.000	0.09	0.09	0.01
1hr	-0.00	0.000	0.19	0.19	0.02	0.00	0.000	0.14	0.14	0.01
4hr	-0.07^b	0.011	0.38	0.37	0.04	0.04^c	0.005	0.27	0.28	0.00
6hr	0.01	0.000	0.48	0.48	0.03	0.01	0.001	0.33	0.33	0.02
12hr	-0.01	0.000	0.67	0.66	0.12	0.00	0.000	0.47	0.47	0.06
1wk	-0.01	0.001	1.69	1.63	0.14	0.04	0.006	1.00	0.94	0.37

Table 2.7: Forecasting Order Flow Out-of-Sample

$$Q(k)_{i,t+1} = \alpha(k)_{i,t} + \sum_{j=1} \beta(k)_{j,i,t}\Delta P(k)_{i,t-j+1} + \sum_{l=1} \gamma(k)_{l,i,t}Q(k)_{t-l+1} + \varepsilon_{t+1}$$

where $\Delta(k)P_{i,t}$ is price change at sampling frequency k for exchange rate i at time t and $Q(k)_{i,t+1}$ is order flow for the same exchange rate and sampling frequency at time $t+1$. The columns under *OF* and *RW* give the forecast RMSEs of model (2.6) and random walk models and the t-statistic for the forecast improvement over RW is reported in the last column. ^{a,b,c} indicate 1%, 5% or 10% significance level by using the Newey-West variance-covariance estimator.

	k	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	R^2	OF	RW	t-stats
EUR/USD	5m	3.15^a	0.22	0.13^a	0.01	0.020	8.45	8.49	-0.10
	15m	2.53	-4.65^c	0.05^a	1.55	0.005	17.01	17.01	0.00
	30m	1.32	-2.77	0.04^c	0.72	0.002	25.97	25.95	0.01
	1hr	4.93	-3.66	0.04	5.44^c	0.003	31.39	31.34	0.03
	4hr	4.97	13.61	0.03	-7.44	0.007	70.46	70.23	0.04
	6hr	4.24	6.39	0.03	9.79	0.026	83.29	82.66	0.10
	12hr	37.13^b	14.59	0.01	-0.16	0.048	120.40	117.32	0.24
	1wk	-32.57	51.87	0.04	-27.96	0.072	331.68	278.70	1.10
USD/JPY	5m	2.44^a	0.79^a	0.19^a	0.04^a	0.064	1.72	1.79	-0.56
	15m	4.21^a	0.22	0.10^a	3.41^c	0.040	3.59	3.63	-0.09
	30m	5.25^a	1.29	0.03	9.20^a	0.032	5.96	6.03	-0.09
	1hr	5.86^a	0.99	0.09^b	0.88	0.039	8.35	8.55	-0.25
	4hr	-1.08	2.44	0.09^c	0.11	0.010	18.93	18.78	0.10
	6hr	1.82	-0.83	0.09	7.59	0.018	22.79	22.62	0.08
	12hr	9.59	14.30^b	-0.04	-0.11	0.050	35.25	35.39	-0.03
	1wk	-4.62	-50.06^b	0.34	0.42^c	0.161	126.06	108.08	0.71
EUR/GBP	5m	-5.70^a	-5.55^a	0.12^a	0.04^a	0.013	6.53	6.57	-0.11
	15m	-14.96^a	-10.00^a	0.10^a	0.04^b	0.016	0.00	0.00	0.00
	30m	-19.20^a	-4.07	0.09^a	-0.00	0.017	15.43	15.45	-0.03
	1hr	-18.91^a	-3.56	0.05^c	0.04	0.018	24.62	24.95	-0.19
	4hr	-3.39	-10.91	0.09^b	-0.04	0.017	53.35	53.09	0.06
	6hr	-8.26	3.94	0.02	0.06	0.008	68.27	67.30	0.16
	12hr	21.92^c	9.53	0.09	0.01	0.036	104.24	97.68	0.65
	1wk	-26.99	82.14^b	0.25^b	-0.08	0.249	273.88	262.86	0.21
GBP/USD	5m	-8.12^a	-11.83^a	0.07^a	0.04^a	0.007	7.03	7.08	-0.18
	15m	-23.80^a	-10.67^a	0.07^a	0.03^b	0.015	12.20	12.21	0.00
	30m	-27.30^a	-14.76^a	0.08^a	0.04^c	0.021	17.77	18.07	-0.29
	1hr	-31.37^a	-16.52^a	0.06^b	0.09^a	0.034	23.59	23.82	-0.18
	4hr	-14.35^b	7.23	0.13^a	0.03	0.021	49.42	49.46	-0.01
	6hr	-10.36	24.78^b	0.06	0.04	0.022	67.76	67.43	0.06
	12hr	18.99	4.51	0.09	0.05	0.027	90.09	88.26	0.23
	1wk	-13.88	41.59	0.09	0.05	0.044	271.59	244.34	0.80

Figure 2.1: Variation in R^2 of order flow model across sampling frequencies

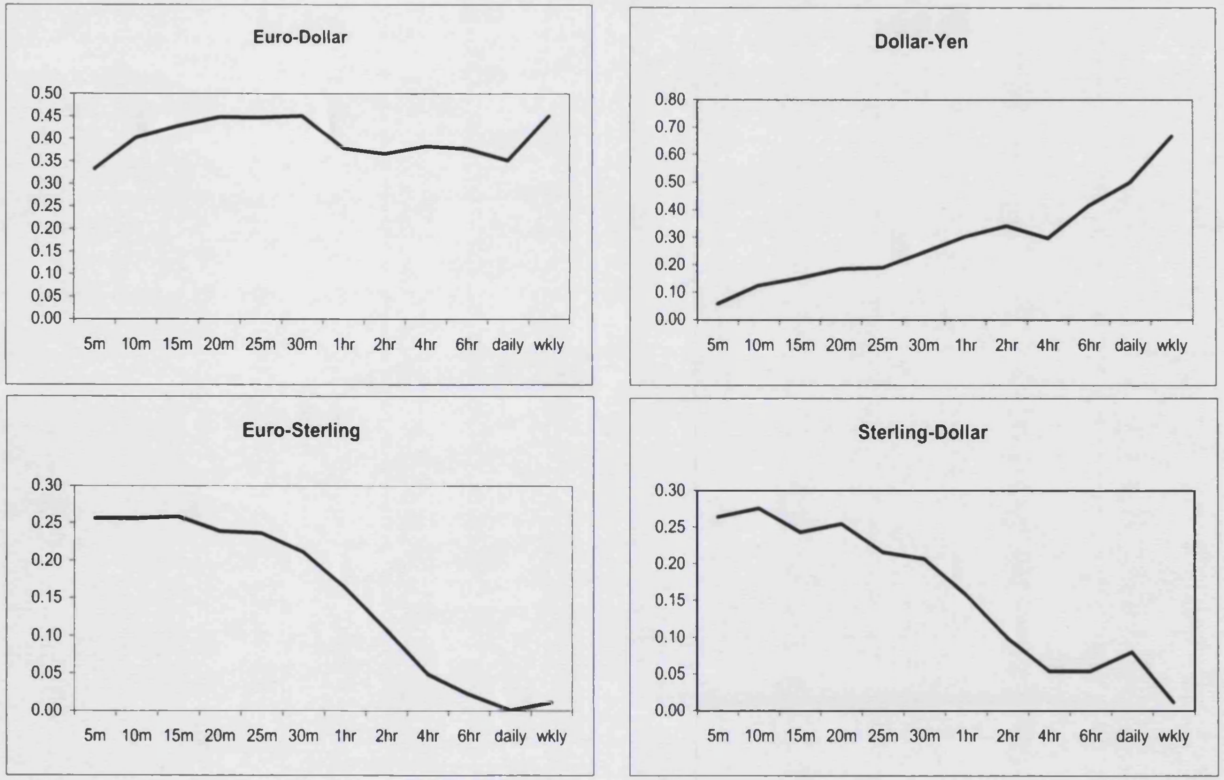
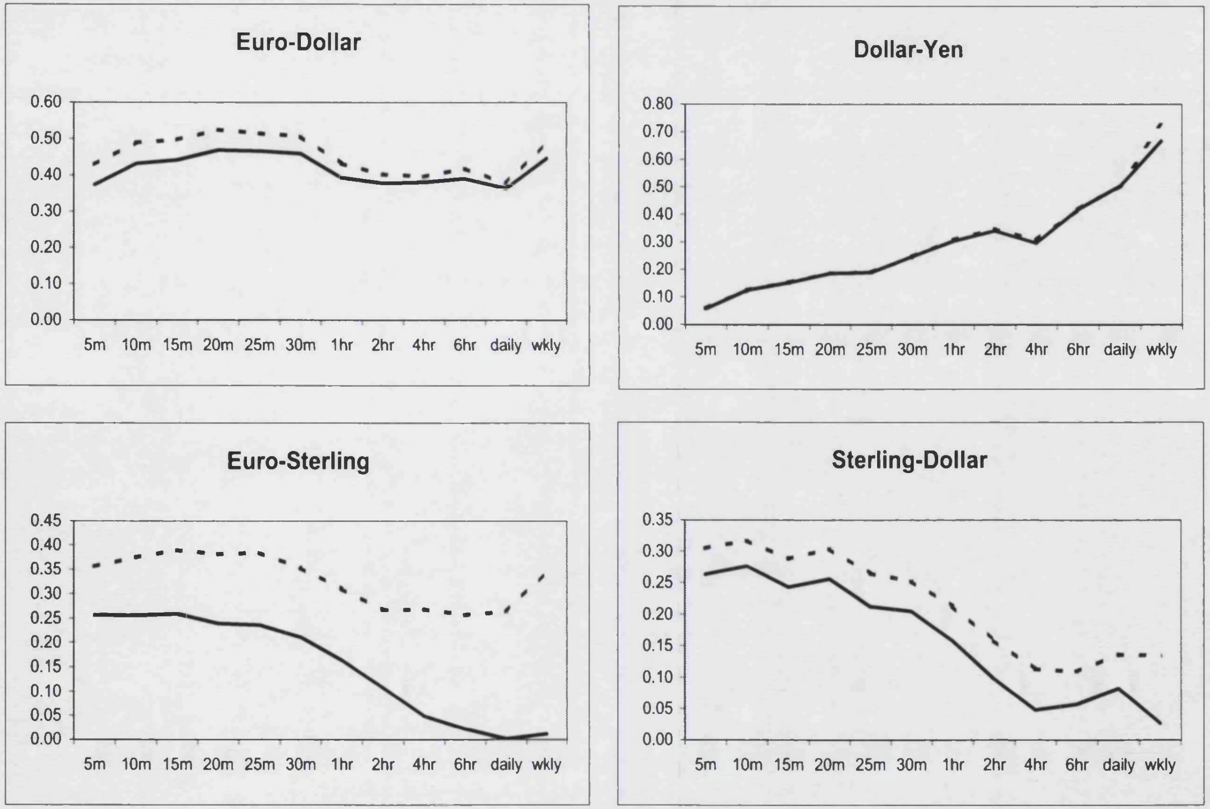


Figure 2.2: R^2 for univariate and multivariate order flow models



Chapter 3

Market conditions, order flow and exchange rates determination

3.1 Introduction

In macroeconomic models, foreign exchange rates are determined by the underlying fundamental factors in the economy. The information about these fundamental factors is public knowledge and rational agents all correctly understand the mapping from information to prices. In these models, there is no private information, trading activities play no role in exchange rates determination and price formation is straightforward and immediate. Unfortunately these models are associated with poor empirical performance. In general, their explanatory power is very low over short horizons. In their seminal papers Meese and Rogoff (1983a, 1983b) vigorously show that the proportion of monthly exchange rate movements that can be explained by macro models is virtually zero and their forecast ability is even worse than a random walk.[1]

In recent years, a new direction in exchange rate analysis, the market mi-

[1] For a recent survey of this literature, see Frankel and Rose (1995) and Isard (1995)

crostructure approach, has drawn growing attention. The microstructure approach assumes that the information structure in the market is asymmetric, i.e. some agents in the market have private information.[2] The informed traders can exploit their informational advantages by issuing orders to market makers. By observing order flow, the market maker makes inference about the private information and adjusts quotes accordingly. For example, if there is an incoming buy order, the market maker might increase the probability that the customer may have received 'good' news. If he saw a sell order, he will reduce this probability. In this way, private information is incorporated into the price and in this sense we say that order flow is informative.

This approach has recently been applied to foreign exchange markets and it has generated some promising results. Evans and Lyons (2002a) claim that order flow has substantial explanatory power for exchange rate changes on a daily basis and can explain 60 percent and 40 percent of return variations for DEM/USD and JPY/USD respectively. Rime (2000) shows that order flow can explain 30 percent of return variation for NOK/DEM on a weekly basis. Chapter two of this thesis works on high frequency data and shows that order flow has explanatory power for exchange rate changes for EUR/USD, GBP/USD, USD/JPY, and EUR/GBP at relatively high frequency levels even though the explanatory power varies across exchange rates and sampling frequencies. The central idea of this approach is that order flow carries information and thus has explanatory power for price movements.[3]

An implicit underlying assumption in the standard microstructure approach to exchange rates analysis is that the informativeness of order flow is constant even

[2]In this approach, private information is defined as any information that is not public and helps forecast the future price better than public information alone. See Lyons (2001)

[3]It is important to note that order flow is fundamentally different from volume. In the market microstructure literature, order flow is defined as the net of buyer-initiated orders and seller-initiated orders.

if the market experiences dramatically different situations. Consider the the following claim of the Reserve Bank of Australia: "$\cdots$ *highly leveraged institutions in mid-1998 deliberately traded during Sydney lunch time, or in the slow period between Sydney's wind-down and London's wind-up, in order to have maximum effects on the Australia dollar's exchange rate(Market Dynamics 2000, pp127-8).*"[4] This claim reflects the worries behind the following question: "Is the informativeness of order flow really constant?".

The objective of this chapter is to examine whether the above basic assumption is valid and extends the analysis of the role of order flow in information transmission by further investigating the relationship between market conditions and the informativeness of order flow. Even though the price impact of order flow is substantial, there is little consensus among the existing microstructure theories concerning under which market conditions order flow could be more informative. In the models of Admati and Pfleiderer (1988) and Subrahmanyam (1991), both the informed and the discretionary liquidity traders can strategically choose their trading time and such strategic behavior will impact on the informativeness of order flow. Admati and Pfleiderer (1988) predict that informativeness of order flow will be positively correlated with bid-ask spreads and negatively correlated with volatility and trading volume.

In the information uncertainty model of Easley and O'Hara (1992), however, the occurrence of trades is associated with information events. When a trade occurs, the market maker will update his/her belief towards a higher probability of the information events and adjust the quotes accordingly. The model predicts that order flow is more informative when spreads are high, volatility is high and trading volume is high.

This chapter develops a methodology to study the relationship between in-

[4]It is quoted from McCauley (2001), "Comments on 'Order flow and exchange rate dynamics'". pp.194

formativeness of order flow and market conditions and tests a set of hypotheses implied by classic market microstructure models. The market conditions studied in this paper are a set of important market statistics: bid-ask spread, volatility and trading volume. The data used in this study contains information of transactions and firm quotes generated by Reuters D2000-2 in the inter-dealer spot FX markets. It covers four major currency pairs (EUR/USD, EUR/GBP, GBP/USD and USD/JPY) and spans nine months on average between 1999 to 2000. The major findings of this study are:

1. The informativeness of order flow is not constant under different market conditions as was assumed in the standard microstructure approach. This result highlights the non-linearity in the relationship between order flow and price movement.

2. Order flow tends to be more informative when bid-ask spreads are high, volatility is high or trading volume is low. This relationship is significant and persistent across different sampling frequencies and exchange rates. Clearly these results are neither fully consistent with the predictions of Admati and Pfleiderer's model nor with those of Easley and O'Hara's information uncertainty model.

3. The relationship between the price impact of order flow and the market conditions is significantly captured by our Interaction model and Logistic Smooth Transition Regression (LSTR) model. In particular, the empirical results from the LSTR model also suggests that the price impact of order flow shifts within a relative small range of market conditions that the market is very likely to experience.

The remainder of the chapter is organized as follows: Section 3.2 discusses the theoretical background and generates hypotheses. Section 3.3 presents the methodology. Section 3.4 describes data. Model estimation and empirical analysis are presented in section 3.5. Section 3.6 presents two sensitive tests. Section 3.7 gives a brief discussion. Section 3.8 concludes.

3.2 The informativeness of order flow

Since the classic papers of Kyle (1985) and Glosten and Milgrom (1985), we have witnessed a large amount of work supporting the idea that order flow carries information and has impacts on the price formation process. Perraudin and Vitale [1996] and Evans and Lyons (2002a,b) explicitly model order flow as the means of information transmission in FX markets. In these models, dealers first receive information from their non-dealer customer order flow and then share or spread the information afterwards through inter-dealer trading. On the empirical side, French and Roll (1986), Hasbrouck (1991a,b), Ito, Lyons and Melvin (1998), Lyons (1995) and Payne (2003) show that order flow has a strong impact on price movements on both equity and FX markets. In particular, Evans and Lyons (2002a), Rime (2000) and the previous chapter claim that order flow has significant explanatory power for exchange rates movements in inter-dealer spot markets.

Since order flow, the information transmission medium, is one of the key steps to understanding security price behavior, characterizing this transmission mechanism under different market conditions will help people better understand the price formation process. This issue has been studied in a number of theoretical works by Admati and Pfleiderer (1988), Diamond and Verrecchia (1987), Subrahmanyam (1991), Easley and O'Hara (1992), Foster and Viswanathan (1990), etc. There is, however, little consensus among the theoretical models regarding the market conditions under which order flow could be more informative and indeed the results of some of these models contradict each other.

In the classic paper of Admati and Pfleiderer (1988), by introducing discretionary liquidity traders, the authors derive a model which endogenizes trading volume, volatility and the bid-ask spread. They show that in equilibrium discretionary traders will clump together. In order to camouflage their trading and minimize the price impact, the informed traders will also trade more heavily during the period when liquidity traders concentrate. The prediction of the model is that during the concentration: (1) volume will be higher because of increased trading activity for both informed and uninformed traders; (2) volatility could be higher because more informed trading occurs at that time; (3) spreads will be lower because competition of informed traders will decrease the bid-ask spread, leading to a better trading term for liquidity traders; (4) order flow will be less informative because of the clump of liquidity trading.

In the information uncertainty model developed by Easley and O'Hara (1992), high trading volume indicates a larger likelihood of an information event having occurred. Therefore during the period when volume is high, volatility will be high and order flow will be more informative. In their model, a small spread can be interpreted as a small likelihood of an information event. The authors argue that when there is a low probability of an information event the market maker will shrink the quotes toward "V^*, the unconditional expectation of V, and not toward the signal-based values of $\overline{V}$ or $\underline{V}$. $\cdots$ With trade now 'safer' the market maker reduces his spread."(pp.587). In other words, the model predicts that the informativeness of order flow is positively correlated with volume, volatility and spreads. Based on a different underlying information structure, the trading constraint model of Diamond and Verrecchia (1987) has a different implication. In the paper, the authors argue that due to trading constraints of informed traders, sparse trading could indicate a 'bad information event' rather than 'no information event'.

Paralleling the theoretical debate above, a fair amount of empirical work has been developed to test asymmetric information models (see, for example, Madhavan and Smidt (1991), Brock and Kleidon (1992), Bollerslev and Domowitz(1993), Hsieh and Kleidon (1996)). However, the focus of most of this literature is on the patterns in volume, volatility and spreads. Lyons (1996) is one of very few who addresses the above issue empirically. Lyons tests the event-uncertainty hypothesis against the *hot potato hypothesis*[5] by studying the relationship between informativeness of order flow and market activity intensity. The author finds evidence for the hot potato hypothesis if the market activity intensity is measured by trading intensity. He also finds evidence for event-uncertainty hypotheses if activity intensity is measured by quoting intensity. Yet it is unexplained why the different measurement of market pace leads to the opposite conclusions. One of the potential weaknesses of the paper comes from the data set, which covers only five working days for a single dealer in August 1992. As Mello (1996) argues, since markets might experience different conditions (turbulent vs calm periods) and dealers might have different characteristics (eg, capital size, trading strategies), one week data from a single dealer might not tell the whole story.

Given the order flow's significant impact on price and its increasing usage by practitioners in their decision making, characterizing the information transmission mechanism of order flow is important from the point of view of both academics and market participants. To characterize the informativeness of order flow under different market conditions, we first address the question related to the assumption of the standard microstructure approach: 'Is order flow equally informative under different market conditions?'. The assumption can be formalized in the following

[5]The *hot potato* is a metaphor used by foreign exchange dealers in referring to the repeated passage of idiosyncratic inventory imbalances from dealer to dealer following a customer order flow innovation. The *hot potato hypothesis* assumes that the trades are less informative when trading intensity is high.

set of hypotheses we will test in this study:

Hypothesis I: *Order flow is equally informative regardless of trading volume.*[6]

Hypothesis II: *Order flow is equally informative regardless of spread.*

Hypothesis III: *Order flow is equally informative regardless of market volatility.*

Nest we address the following question: 'if the informativeness of order flow changes under different market conditions, then what is the direction of the change?'. This question tries to characterize the informational aspect of price impact of order flow in security markets. In addressing this question, we test the predictions of market microstructure models.

3.3 Methodology

3.3.1 Core model

The asymmetric information pricing model of microstructure is based on a *learning process* faced by market intermediaries. Either in the sequential model of Glosten and Milgrom (1985) or in the batch trading model of Kyle (1985), a lot of attention has been given to the effect of asymmetric information on market prices. If a trader has superior information about the underlying value of the asset, his trades will reveal, at least partially, this private information about the value of the asset and will affect the behaviour of market prices.

The key to understanding the above information revealing process is Bayesian learning. Take the model of Glosten and Milgrom (1985) as an example, the market maker sets the ask price a_t to the expected value V of an asset after seeing a trader wishing to buy. a_t depends on the conditional probability that V is either

[6] The irrelevancy hypothesis implies that neither Admati and Pfleiderer (1988)'s prediction nor Easley and O'Hara (1992)'s prediction is valid.

lower ($V = \underline{V}$) or higher ($V = \overline{V}$) than his prior belief (V) given that a trader wishes to buy. The bid price b_t is defined similarly given that a trader wishes to sell. If the noise traders are assumed equally likely to buy or sell regardless of information, good news will result in an excess of buy orders and bad news will result in an excess of sell orders. In the model, the conditional probability incorporates the new information that the market maker learned from observing the order flow and is hence a posterior belief about the asset value V. The posterior will become a new prior in the next round of trading and the updating process continues.

The central idea of the above information extracting process is that the market maker adjusts the quoting prices by observing order flow which in turn is driven by information. In this sense we say that information, through order flow, drives price movements. This idea can be put in the following simple empirical model[7]:

$$\Delta P_t = \alpha + \beta * Q_t + \varepsilon_t \tag{3.1}$$

where ΔP_t is price change between the time $t-1$ and t. Q_t is aggregated order flow within time interval$[t-1,t]$. Generally speaking, the information effect of order flow will be captured by the regression coefficient β. In words, the more informative the order flow the larger the β, and vice versa.

Broadly speaking, β is a function of some structural parameter Z_t measuring market conditions, i.e $\beta = \beta(Z_t)$. The question whether the informativeness of order flow is constant under different market conditions can be addressed by testing whether $\beta(Z_t) = \beta_0$, where β_0 is constant. Z_t in this paper is a set of important market characteristic statistics: trading volume, return variance and bid-ask spreads. If $\beta(Z_t)$ is not constant, characterizing the function $\beta(Z_t)$ is a key aspect of studying price impact of order flow and will help people better understand the

[7] A similar modelling strategy has been extensively used in empirical work, such as Madhavan and Smidt (1991), Lyons (1995,2001), Foster and Viswanathan (1990)

price formation process.

3.3.2 Nonlinearity test

In this section we first use a quintile model, in which we study the price impact of order flow by dividing the whole sample into 5 sub-samples according to the market condition measuring variable, to test whether the regression coefficient β in the core model (3.1) is constant. If not, a recursive regression model is employed to further study the dynamics of β.

Quintile model

A very straightforward way to test model stability is to split the sample into different sub-samples and see whether the model is stable across sub-samples. In this paper, the sample is split into 5 sub-samples according to the variable Z_t, which measures market conditions. Division of the sample into 5 groups is based on the consideration of both the size of each group and the difference of market conditions of each group.[8]The market conditions of interest in this paper are market volatility, liquidity and volume. Take the case of volatility as an example, the observations are divided into 5 sub-samples by the magnitude of market volatility of the time interval from which the observation is drawn. Dummy variable S_{jt} is used to distinguish each sub-sample. With this constraint, the core model (3.1) can be expressed as

$$\Delta P_t = \alpha + \sum_{j=1}^{J} \beta_j * S_{jt} * Q_t + \varepsilon_t \tag{3.2}$$

where $S_{jt} = 1$ for the corresponding sub-sample j and 0 otherwise and J is equal to 5.

[8]However, it should be admitted that there is no absolute criterion on how many groups sample should be split. Indeed, we experiment with a moving window regression model and find that the results are similar to those of the Quintile model.

For the cases of market liquidity and volume, the quintile model is similarly constructed.

Under the null, i.e. the order flow is equally informative under different market conditions, all regression coefficients will be equal. This can be tested to see whether $\beta_i = \beta_j, \forall i, j$.

Recursive least squares regression model

Another way to investigate the variation of β is via a sorted recursive least squares regression model. In this model, the observations are sorted by the interesting variable Z_t and regression model (3.1) is run recursively on these re-sorted observations. This approach allows an analysis of the relationship between informativeness of order flow and market statistics with the least prior constraint. Since no model specification has been postulated on β and Z_t, we can obtain a graphical representation of the relationship between the informativeness of order flow and the market statistics.

For simplicity, we re-write the core model in a vector form:

$$\Delta P_t = \mathbf{x}_\mathbf{t}^{'}\mathbf{B} + u_t \tag{3.3}$$

where $x_t = [1, Q_t]^{'}$ and $\mathbf{B}$ is the coefficient vector $[\alpha, \beta]^{'}$ in the core model (3.1).

The recursive model is usually used to check whether the model structure varies for a time series. In this paper our purpose is slightly different. We aim to check whether the model structure varies along the third variable Z_t rather than along the time dimension. For this purpose the recursive model is constructed as in three steps:

Step one, Sort the observations x_t according to Z_t, which could be volatility, bid-ask spread or trading volume. After sorting, the observations are re-arranged by ascending value of Z_t.

Step two: Fit the model (3.1) to the first k ($k = 2$) observations and get the coefficient estimate b_k. Next use the first $k+1$ observations as regressor and compute the regression coefficient again. Proceed in this way, adding one observation at a time until the final regression coefficient, which is based on the all observations, is obtained. This process will generate a sequence of coefficient estimates, $\mathbf{b_k}, \mathbf{b_{k+1}}, ..., \mathbf{b_n}$. In general,

$$\mathbf{b_m} = (\mathbf{X}_\mathbf{m}^{'}\mathbf{X_m})^{-1}\mathbf{X}_\mathbf{m}^{'}\Delta\mathbf{P_m} \qquad m = k, k+1, ..., n \tag{3.4}$$

where $\mathbf{X_m}$ is the $m \times k$ matrix of regressors for the first m sample points, and $\Delta\mathbf{P_m}$ is the m-vector of the first m observations of the dependent variables.

Step three: The standard errors of the coefficients are calculated at each stage of the recursion (except the first one) and the evolution of the coefficients and their plus and minus two standard errors are graphed.

A visual inspection of the graph may suggest parameter constancy, or its reverse. A substantial vertical movement of a coefficient, to a level outside previously estimated confidence limits, is usually a result of the model trying to digest a structural change and may suggest parameter instability.

3.3.3 Nonlinearity modelling

In this section we try to answer the second question 'how does the informativeness of order flow change under different market conditions?' by modelling the nonlinearity in the order flow and price change relationship. We use simple interaction models and a more complex smooth transition regression model to characterize the informativeness of order flow under different market conditions.

Interaction model

Since so far there is little theoretical guide as to what specific form the relationship between the informativeness and market conditions should take, the interaction model simply conjectures that the informativeness of order flow has some linear relationship with the market condition measuring variable Z_t. Formally the following constraint is put on β in equation (3.1):

$$\beta = \beta_1 + \beta_2 * Z_t \tag{3.5}$$

Z_t is the measurement of market conditions of interest (it could be volume, volatility or spread). Inserting (3.5) back into (3.1) and rearranging it results in the following nonlinear regression model:

$$\Delta P_t = \alpha + \beta_1 * Q_t + \beta_2 * Z_t * Q_t + \varepsilon_t \tag{3.6}$$

In the interaction model, the regression coefficient β_2 captures the nonlinearity in the relationship between order flow and price change. A positive β_2 indicates order flow is more informative under conditions where Z_t is larger. In this sense, the interaction model can be used to test the predictions of various theoretical models about the relationship between the informativeness of order flow and market conditions in the market microstructure literature.

Logistic smooth transition regression: LSTR

An alternative approach to model the nonlinearity of order flow and price change is to relax the linear specification between the informativeness of order flow and market conditions and assume the relationship between β and Z_t is itself nonlinear.

In this section we choose the widely used logistic smooth transition regression (LSTR) to model the relationship between order flow and price movements. Formally the LSTR can be written as

$$\Delta P_t = \beta' \mathbf{x_t} + (\theta' \mathbf{x_t}) F(Z_t) + \varepsilon_t \tag{3.7}$$

where $\mathbf{x_t} = [1, Q_t]$, $\varepsilon_t \sim i.i.d(0, \sigma^2)$, $E[\mathbf{x_t}\varepsilon_t] = 0, \beta = (\beta_0, \beta_1)'$ and $\theta = (\theta_0, \theta_1)'$. $F(Z_t)$ is the logistic function and can be written as

$$F(Z_t) = (1 + exp\{-\gamma(Z_t - c)\})^{-1} - 1/2 \tag{3.8}$$

where the logistic parameter $\gamma > 0$ and c measures transition point.

The idea behind LSTR model is that the relationship between order flow and price movement changes gradually with the market condition measuring variable Z_t. The transitional feature is captured by the model parameters θ_1 and γ.

3.4 Data Description

3.4.1 The Data

The data set used in this Chapter is the same as that of Chapter two. The data set provides two types of tick level information: trades and firm quotes. In short, this data covers four currency pairs: EUR/USD, EUR/GBP, GBP/USD and USD/JPY. The samples for EUR/USD and GBP/USD cover a period of ten months from 28 September 1999 to 24 July 2000. EUR/GBP and USD/JPY samples cover a period of eight months from 1 December 1999 to 24 July 2000.

For the purpose of current study, this data set has significant advantages over foreign exchange data used in the past work (e.g. Bollerslev and Domowitz

(1993), and Lyons (1996)). The data used in Bollerslev and Domowitz (1993) are indicative quotes from Reuters FXFX.[9] The shortcoming of indicative quotes is that the return variance derived from them is far larger than that derived from the actual quotes or trades and the spread is less correlated with market activities.[10] The data used in Lyons (1996) is transaction data, but it covers only 5 working days for a single dealer. Our data set contains transaction and firm quote information, covers four major exchange rates and spans nine months on average. While the long sample period provides us with the opportunity to address our questions from the time aggregation angle without loss of statistical power, the multiple rates allow us to check the robustness of the estimation cross-sectionally.

3.4.2 Filtering and time aggregation

This chapter follows the same filtering and aggregation procedure of chapter two. Again, we exclude overnight periods, weekends, some world-wide public holidays and certain other dates where the feed from D2000-2 is very low. For a given time aggregation interval (e.g. 10 minutes), we record at every observation point the transaction price, order flow, total number of buys and sells, average bid-ask spread and volatility. Spreads are calculated as the percentage of trade price in basis point. Volatility is calculated as the return variance within the time interval. Within the generated time series, we further remove the observations that have "wrong" values of market condition variables, for example, zero or negative spreads.

To better capture the effects of the changing market conditions on the flow-return relationship, this research focuses on relatively high frequency intervals

[9]Reuters FXFX system is a screen system that is used to post quotes to attract customers. Unlike the quotes posted in the brokerage screens, the quotes displayed on FXFX screen are only indicative, not firm.

[10]Refer to Danielsson and Payne (2000) for a full discussion

because as sampling frequency declines, the feature of market conditions is attenuated. Specifically, we focus on 5 time aggregation levels: 5 minutes, 10 minutes, 20 minutes, 30 minutes and 1 hour.[11] The filtering and aggregation finally leaves us 20 databases (5 sampling frequencies $\times$ 4 exchange rates), each representing a different sampling frequency for an exchange rate. These aggregated databases are summarized in Table 3.1. As mentioned previously, the long covering periods is a valuable characteristic of our sample. After the filtration and aggregation we still have a reasonably large number of observations for each rate and each sampling frequencies in our generated databases.

3.5 Estimation and Analysis

In this section we use four major floating rates (EUR/USD, EUR/GBP, GBP/USD and USD/JPY) to estimate the models and present the main empirical results.

The following definitions will be used throughout the paper: ΔP_t is the log price change within the time interval $[t-1,t]$. Q_t is the order flow, defined as the difference between the number of buys and number of sells within time interval $[t-1,t]$, Z_t could be average spread, return variance or trading volume within time interval $[t-1,t]$[12].

3.5.1 Quintile model

In order to get a complete picture of variation in the price impact of order flow, we estimate the quintile model from two dimensions: market conditions and time ag-

[11] We have experimented with denser and longer time aggregation levels and the results do not alter the pattern we reported here but become less significant as aggregation level goes lower, e.g 12 hours.

[12] Since the size of each trade is not available in our data, the difference between buy initiated and sell initiated orders and the total number of trades are used as proxies for order flow and volume respectively in this paper. For more discussion of such proxies see Danielsson and Payne [2000].

gregations. For market conditions, we check the following three important market statistics: bid-ask spread, trading volume and market volatility. For time aggregation, we estimate the model for a series of sampling frequencies from 5 minutes to 1 hour.

In estimating the model, we divide the sample (for each sampling frequency database) into 5 equal-sized groups according to variable Z_t, which can be spread, volume or volatility. Take 10 minutes sampling frequency as an example, when Z_t represents spread, we divide the database into 5 groups by the magnitude of spread. So group 1 will have the 20% of the observations with the smallest spread and group 5 will have the 20% of the observations with the largest spread. The same idea applies to other sampling frequencies and market conditions.

The results of estimates of the quintile model are presented in Tables 3.2, 3.3 and 3.4. The most notable result is that the F-tests are significant in all three tables for most sampling frequencies and all currency pairs (except for JPY/USD and USD/GBP in Table 3.4). In particular, if the observations are arranged along bid-ask spread and volatility, F-tests are significant for almost all sampling frequencies from 5 minutes to 1 hour and all exchange rates. If the observations are arranged by volume, F-tests are significant for all sampling frequency for USD/EUR and GBP/EUR. But it is significant only for high sampling frequencies for JPY/USD and not significant for USD/GBP. Compared to the cases of spread and volatility, the results of volume market condition are less impressive, especially when GBP/USD is considered.[13] Nevertheless, the overall results from the quintile model as shown in Table 3.2, 3.3 and 3.4 suggest that the null hypothesis of order flow being equally informative under various market conditions, ie. $\beta_i = \beta_j, \forall i, j$, is overwhelmingly rejected in our samples.

Another interesting result is the changing pattern of β across groups. In Table

[13]This result can arise for various reasons, such as sample specific randomness or using proxy for volume. A much longer sample period is required to further investigate this issue.

3.2 and Table 3.3, β tends to increase from group 1 to group 5. In Table 3.4, β tends to decrease from group 1 to group 2 and then remains relatively stable (except for JPY/USD). For example, in Table 3.2 the β of EUR/USD based on the 10 minute sampling frequency increases from 0.0027 for group 1 (with smallest spreads) to 0.0077 (with largest spreads) for group 5. The pattern exhibited here is quite persistent for most sampling frequencies and exchange rates. The increasing β in Table 3.2 and Table 3.3 and the decreasing β in Table 3.4 indicates that β might be an increasing function of market spread and volatility and a decreasing function of trading volume (at least in some range of volume). In other words, order flow is more informative when market spread is large, volatility is high or trading volume is low.

3.5.2 Recursive least squares regression model

The purpose of using the recursive least squares regression model is to get a visual idea about the relationship between the informativeness of order flow and various market conditions without imposing other prior constraints. For this purpose we choose hourly frequency data as representative to estimate the model. We re-sort the observations by variable Z_t (which could be bid-ask spread, volatility or trading volume). In the first regression, we use the first 50 observations to void the possible excessive volatility early on in the recursive regression process. Then in each of the following recursive regression, we use additional 5 observations. The sequence of coefficients of β and their 2 standard deviations from the recursive regressions are drawn against Z_t in Figure 3.1, Figure 3.2 and Figure 3.3 for market conditions of spreads, volatility and volume respectively.

Clearly from Figure 3.1 and Figure 3.2 we can see that β increases with spread and volatility very sharply within a certain (small) range of spread and volatility and becomes more stable after that. In Figure 3.3 β decreases to some extent at the beginning (except for Dollar-Yen) and becomes stable (but still decreasing) after that. Since the number of observations become fewer for very large spreads, volatility and volume, the reduction of the additional number of observation can potentially explain the stability of β as spreads, volatility and volume increase. To clarify this ambiguity, in Figures 3.4, 3.5 and 3.6 we re-draw the sequence of β against the recursive regression process. For the purpose of studying the effect of market conditions on the price impact of order flow, we deliberately label the x-axis with the values of variable Z_t for those regressions rather than with the sequence of number representing the recursive process. Since 5 extra observations will be added to each of the regressions along the recursive process, the recursive process can also be viewed as a proxy for the number of observations used in the regression. For example, in the case of EUR/USD-spread (the first graph in Figure 3.4), the fifth point on the x-axis (labelled with 4.48) indicates that when 80% of observations enter the regression, the β will be about 0.0035 and the spread will increase to a level of 4.87 basis points. The advantage of this presenting method is that it allows us to study how β and Z_t change simultaneously along the recursive process.

In general, we can see a smoother change of β in all three Figures 3.4, 3.5 and 3.6. But it is still the case that the β increases or decreases more sharply for the first part of observations in all graphs.[14] In Figure 3.4 and Figure 3.5, β increases with bid-ask spread and market volatility respectively for a large proportion of the observations. In figure 3.6, β decreases with trading volume (except for Yen-Dollar) only for the first 30 percent of the observations and becomes more stable when observations with larger volume are added into the regression. It is important to note that in both Figure 3.4 and Figure 3.5, β completes the shift within a

[14]In recursive regression models, as more and more observations enter the regression the impact Z_t on β will attenuate and converges to the equilibrium pattern. As a complementary exercise, we experiment with moving window regressions and find the patterns confirm what reported in here.

fairly small range of market conditions and this range covers a large proportion of the total observations. For example, for USD/EUR-spread, the price impact of order flow increases with market spreads and this increasing trend covers 80 percent of the total observations. Even though the spread ranges from 1 basis point to 50 basis points for the total observations, the β finishes the shift within a fairly small range of spread from 1.4 basis points to 5 basis points (see USD/EUR in Figure 3.4). In other words, the results from the recursive model indicate that the shift of the price impact of order flow is not an extreme market condition phenomenon. Instead it shifts within a small range of market conditions which cover most of the cases the market is likely to experience.

The visual inspection indicates that the value of β varies substantially under different market conditions and moves outside previously estimated confidence limits in almost all graphs of Figure 3.4, 3.5 and 3.6. From an econometric point of view, this violation indicates that the structure imposed in the core model (3.1) is not stable as the market conditions (measured by Z_t) vary. The instability suggests a high possibility of the non-linear relationship between the order flow and price change.[15]

3.5.3 Interaction model

The interaction model (3.6) is evaluated along two dimensions: market conditions and time aggregations. The estimation is based on the 20 databases described in the Data Section. The estimates of model coefficients are presented in Tables 3.5 , 3.6 and 3.7.

In Table 3.5, spread is the market condition of interest. The most notable feature of this part is that β_2 is significantly away from zero for almost all the

[15]We exercise a group of CUSUM tests on the base of the regression residuals of the recursive model to check the model stability and the result confirms our speculation, i.e. the hypothesis of linear relationships between order flow and the price change is rejected.

sampling frequencies and all exchange rates. Another important feature we can see in this table is that β_2 is constantly positive for all time aggregation levels and all exchange rates. The significantly positive β_2 suggests that the order flow tends to be more informative when spreads are larger.

Volatility is considered in Table 3.6. Similar as in Table 3.5, the signs of the estimator of the regression coefficient β_2 are positive in most of the cases. Positive β_2 indicates that the order flow is more informative when market is more volatile. The β_2 is significant for USE/EUR and GBP/EUR but less significant for JPY/USD and USD/GBP.

Table 3.7 examines the market condition of trading volume. One notable feature of this part is the sign of the estimator of the regression coefficient. Contrary to Table 3.5 and Table 3.6, the estimates of the regression coefficients β_2 are negative in most of the cases(except for EUR/USD). Negative β_2 indicates that the order flow is less informative when trading volume is higher. The β_2, however, are not significant in most cases. The lack of significance might be due to the fact the informativeness of order flow doesn't have linear relationship with market volume (as seen in our quintile model).

Overall the results from the interaction model indicate that the linear relationship between the informativeness and market conditions imposed by equation (3.5) is positive and significant for a large spectrum of time aggregation levels when market condition is measured by spread and volatility. When market condition is measured by volume, the linear relationship is not significant for most cases.

3.5.4 LSTR model

In the LSTR model (3.7), we assume that the relationship between order flow (Q_t) and price change (ΔP_t) evolves smoothly with the market condition measuring variable Z_t. The coefficients (θ_1, γ) determine how the price impact of order flow varies as a function of the transition variable Z_t, which measures the market conditions. The logistic parameter γ determines the smoothness of transition, the sign of θ_1 determines the direction of the transition. A positive θ_1 suggests the informativeness of order flow is an increasing function of the transition variable Z_t and a negative θ_1 will suggest the opposite.

The Non-linear Least Squares approach is used to estimate the LSTR model and the estimates reported in Table 3.8 are based on hourly sampling frequency data[16]. In the top panel, where the transition variable is spread, the regression coefficient estimates of θ_1 are both positive and significant and the estimates of γ are also significant for all cases. In the middle panel, where the transition variable is volatility, the regression coefficient estimates of θ_1 are both positive and significant in three out of four cases and the estimates of γ are also significant for two out of four cases. In the bottom panel, where the transition variable is volume, we are not able to fit the model for GBP/USD. But for the other three exchange rates, the regression coefficient estimates of θ_1 are both negative and significant and the estimates of γ are also significant (except for USD/JPY).

The results reported in Table 3.8 indicate that the LSTR model can not only manifest the qualitative relationship between the informativeness of order flow and market conditions but also capture the transition feature of this relationship. To have a visual impression, the smooth transition relationships are drawn in Figures 3.7, 3.8 and 3.9. One of the most interesting points in Figure 3.7 is that the β in the core model (3.1) changes quickly within a specific range of bid-ask spreads. For example, the β triples if the spread increases from 2 basis points to 4 basis points for EUR/USD. For USD/JPY, the β shifts quickly from around zero

[16]Before estimating the LSTR, we evaluate an auxiliary regression as a test of linearity against Smooth Transition Regression model (see Granger and Tersvirta[1997] p.117): $\hat{u}_t = \alpha + \beta_1 Q_t + \beta_2 Q_t Z_t + \beta_3 Q_t Z_t^2 + \beta_4 Q_t Z_t^3 + \eta_t$, where $\hat{u}_t$ is the OLS residual from the core model (3.1). In 11 out 12(4 rates $\times$ 3 market conditions) cases, the results are in favor of STR model.

to 0.025 when spreads move from 10 basis points to 20 basis points. The transition feature shown in Figure 3.7 suggests that the price impact of order flow shifts quickly within a small range of market liquidity conditions.

In contrast to the graphs in Figure 8, the graphs in Figure 3.8 are smoother. The β increases smoothly as the volatility increases. It is important to note, however, that the β does not really shift a lot (except for JPY/USD, which doubles in the shift) as it does in Figure 3.7. β only increases from 0.0030 to 0.0045 for USD/EUR, from 0.0030 to 0.0034 for GBP/EUR and from 0.00200 to 0.00202 for USD/GBP. The small magnitude of shifts might suggest that the price impact of order flow is less sensitive to volatility than to bid-ask spreads.

In Figure 3.9, β shifts downward rapidly when trading volume increases. In particular, the shift occurs within a very small range around the lowest volume. The β finishes the shift before the volume reaches 20 for for USD/EUR and GBP/EUR and 40 for JPY/USD. The magnitude of the shift for different exchange rates is a mixture. While β decreases dramatically from 0.02 to less than 0.005 for EUR/USD, it decreases only a little for USD/JPY and even less for EUR/GBP.

Overall, the results from the LSTR model indicate that the relationship between order flow and price movement changes with the market conditions that FX market is very likely to experience.

3.6 Seasonality and Simultaneity Test

3.6.1 Intra-day Seasonality

It is well known that high frequency data usually displays some intra-day pattern of the market spread, volatility and trading volume. Such intraday regularity can introduce bias into our previous empirical analysis. In this subsection we examine the impact of intraday seasonality on our previous analysis.

A straightforward way to study the seasonality impact is to separate the market condition variable Z_t (which can be spread, volatility or volume) into an expected part Z_t^s and an unexpected part Z_t^u and explicitly model them separately. A simple but effective way to decompose the intra-day regularity is to use the intra-day pattern itself as a proxy of the expected part[17]. For instance, as far as the spread is concerned, the expected spread of a time interval can be proxied by the average spread of that specific time interval of all days over all sample period. The unexpected spread of that period can be proxied by the difference between the total spread Z_t and the expected spread Z_t^s:

$$Z_t^u = Z_t - Z_t^s \tag{3.9}$$

To examine the intraday impact, the decomposed market conditions can be used to re-estimate our empirical models. For simplicity, we only re-evaluate the interaction model with the de-seasonlised data. Plug (3.9) back into (3.6) resulting in:

$$\Delta P_t = \alpha + \beta_1 * Q_t + \beta_2 * Z_t^u * Q_t + \beta_3 * Z_t^s * Q_t + \varepsilon_t \tag{3.10}$$

where Z can be spread,volatility or volume. β_2 and β_3 measures unexpected and expected impact of market conditions on informativeness of order flow respectively. The seasonality impact can be evaluated by applying the F-test to the null hypothesis: $\beta_2 = \beta_3$.

The regression model (3.10) is estimated for 30 minute and 1 hour sampling frequencies and the results are reported in Panel A and Panel B respectively in Table 3.9. When $Z = Spread$, the impact of expected and unexpected spread, measured by $\hat{\beta}_2$ and $\hat{\beta}_3$ respectively, are at the same magnitude and most have expected signs. The p-value of F-test cannot reject the null that $\beta_2 = \beta_3$. This

[17]Goodhart, Love, Payne and Rime (2000) use the same method. An alternative approach is to use the ARMA model, eg. Hartmann (1999).

pattern is very consistent for different sampling frequencies and across different exchange rates. The results suggest that it is total rather than unexpected spreads that affect the informativeness of order flow. For volume, most cases of the expected part have the expected signs and are significant. Unexpected volume is less significant and 5 out 8 cases have the expected signs. The last column shows that in half of cases the F-test reject the null. Overall the results indicates that the expected volume plays a much bigger, significant role than does unexpected volume. Finally, when Z is volatility, we find that $\hat{\beta}_2$ and $\hat{\beta}_3$ are both positive and significant for most cases in both Panel A and Panel B. The magnitude of $\hat{\beta}_2$ and $\hat{\beta}_3$, however, are significantly different with $\hat{\beta}_3$ much larger than $\hat{\beta}_2$. This difference of magnitude is further reflected in the last column where F-test rejects null for most cases except USD/JPY. This evidence indicates that both components of volatility can affect the informativeness of order flow, but the seasonal component has a stronger impact. Note it is not a surprising result that expected components are significant in the regressions because the data used in this exercise hasn't been de-seasoned. The aim here is to test whether the results found in previous sections are due to pure seasonality effects. The results in Table 3.9 confirms that market conditions do impact on the informativeness of order flow.

3.6.2 Simultaneity Test

So far we have examined a set of market statistics separately and we found that find each of the market conditions variables (spreads, volatility and volume) can affect the informativeness of order flow. Since those statistics are three aspects of the same market, they are connected with each other[18]. For example, volatility will tend to be higher at the market 'opening' when bid-ask spread is larger. This

[18]See, for example, Hsieh and Kleidon (1996), Danielsson and Payne (2000) and Admati and Pfleiderer (1988).

overlapped trading pattern may bias our results based on individual factors. In this subsection we examine such simultaneity by putting all three factors simultaneously in our interaction model and evaluate the impact of each factor simultaneously. The formal regression model can be written as:

$$\Delta P_t = \alpha + \beta_1 Q_t + \beta_2 Z_t^{Spread} Q_t + \beta_3 Z_t^{Volume} Q_t + \beta_4 Z_t^{Volatility} Q_t + \varepsilon_t \qquad (3.11)$$

where ΔP_t is price change within $[t-1,t]$. Q_{jt} is order flow. Z_t^{Spread},Z_t^{Volume} and $Z_t^{Volatility}$ are the average spread, total volume and volatility within $[t-1,t]$ respectively.

The model is evaluated for 30 minutes and 1 hour sampling frequencies and the estimates are reported respectively in Panel A and Panel B in Table 3.10. In both Panel A and B, we find the β_2, which measures spreads effect, is positive and significant. The volatility effect is positive and significant for 6 out 8 cases. For volume, β_3 takes the correct sign for most cases but only half are significant. For the volume and volatility cases, the insignificance is mainly related with Sterling-Dollar and Dollar-Yen. To further test whether such insignificance is due to the particularly strong factor correlation of GBP/USD and USD/JPY, we calculated factor cross-autocorrelation for all currency pairs and report the results in Table 3.11. Comparing the cross-autocorrelations of GBP/USD and USD/JPY with those of EUR/USD and EUE/GBP, we find that the magnitude of cross-autocorrelations is either same or smaller in the cases of GBP/USD and USD/JPY. Also there is no particular lead-lag relationship among factors. Probably a more plausible explanation might be that the β does not take a simple linear relationship with volume and volatility. Overall, the results of this simple simultaneity analysis confirm our previous finding: each of the three factors, spreads, volatility and volume, can affect the informativeness of order flow.

3.7 Discussion

In this chapter, we attempted to characterize the informational aspects of order flow under different market conditions. The way we measure the informativeness of order flow, which is defined as the regression coefficient of price variation on order flow, is in the same spirit of that of Lyons (1995, 1996).[19] One point worth mention of this measurement is that the regression coefficient might just pick the 'liquidity' or 'digestion' rather than the information effect in the market. To mitigate the impact of such potential mis-measurement, this research bases the analysis on the use of aggregated order flow rather than order flow of individual transactions because noise trade can be largely cancelled out through the aggregation.[20]

To further support the robustness of our measurement, we conduct another experiment here which provides strong evidence that the order flow regression coefficients do capture information effects rather that liquidity effects. The study in chapter two and by Evans and Lyons (2002b) demonstrates that order flow of one market can be used to explain the price changes of inter-linked markets because order flow in one market could convey information about the valuation of the assets traded in other markets. Since this inter-market order flow effects cannot be attributed to liquidity effect, it provides an ideal framework to test the robustness of our results on the market conditions and order flow informativeness. In particular, we experiment with our interaction model in an inter-market framework. We study the impact of order flow of one rate (e.g. EUR/USD) on the variations of another exchange rate (e.g. EUR/GBP) under different market conditions of the

[19]Generally speaking, this measurement stems from the classic model of Kyle [1985]. In this model, equilibrium price is a linear function of total order flow. The coefficient of order flow, λ, equals to $2(\sigma_u^2/\Sigma_0)^{-1/2}$. When order flow is more informative, i.e. σ_u^2 is small, the coefficient will be larger.

[20]The study of Lyons (1995,1996) is more based on transaction level. The analysis of Hasbrouck (1991a,b) is also based transaction level.

first market (e.g.EUR/USD) and find that order flow tends to have a larger impact (i.e. larger β) on other rates when the market has larger spreads, higher volatility or lower volume.[21] In this inter-market framework, it can no longer be argued that the order flow effect picked by β is due to liquidity effects. Given the evidence from inter-market analysis, together with the considerable number of trades in our aggregation, we believe that the regression coefficient can be treated as a fair measurement of informativeness of order flow.

Using the order flow coefficient to measure its informativeness, we find that order flow tends to be more informative when market spreads are high, volatility is high or volume is low. The patterns are persistent across different exchange rates and over a wide range of sampling frequencies. The evidence highlights the non-linearity in the relationship between order flow and price movement and provides a potential to improve the methodology used in the current microstructure approach to FX analysis.

The relationships between order flow informativeness and market conditions found in this study are neither fully consistent with the predictions of the information uncertainty model of Easley and O'Hara (1992) nor with the predictions of the Admati and Pfleiderer (1988) model. In Easley and O'Hara's information uncertainty model, high trading volume is due to information event. But if at the same time there is substantial increase in liquidity trading, the information might be diluted and order flow become less informative. In the model of Admati and Pfleiderer (1988), during concentration period, the price volatility will be high because more informed trader will also choose to trade during concentration. However, if the information can be sufficiently diluted by noise orders,

[21] Take USD/EUR and GBP/EUR as example, we estimate regression model: $\Delta P_t^{ES} = \alpha + \beta_1 * Q_t^{ED} + \beta_2 * Z_t^{ED} * Q_t^{ED} + \varepsilon_t$. When Z is *Spread*, $\hat{\beta}_2$ is 0.0003 and significant at 1% level. When Z is *Volatility*, $\hat{\beta}_2$ is 0.0188 and significant at 5% level. When Z is *Volume*, $\hat{\beta}_2$ is -0.0001 and significant at 10% level. The R^2 is between 24% and 26%.

the order flow impact could be reduced and price variation could be lower. Since the empirical evidence provided in the current study cannot be fully explained by existing theory, further research to provide vigorous economic explanation is a worthwhile object of future study.

3.8 Conclusion

This chapter presents a methodology to characterize the price impact of order flow under different market conditions. In particular, research focuses on: (1) testing a set of hypotheses in the microstructure literature about the informativeness of order flow; (2) modelling the informativeness of order flow under different market conditions.

This chapter uses a high frequency data set of the inter-dealer FX spot market captured by Reuters 2000-2 to create a set of time series databases covering a wide range of sampling frequencies. We estimated the various models using these databases and find strong evidence to reject the hypothesis that the order flow is equally informative under different market conditions. More specifically, we find that order flow is more informative under such market conditions when spreads are high, volatility is high or volume is low. This pattern is persistent across different exchange rates and over a wide range of sampling frequencies. These findings are neither fully consistent with the prediction from the Easley and O'Hara's model nor with that from the Adamati and Pfleiderer's model.

It has been shown in this chapter that the relationship between order flow and price changes are non-linear. We model the nonlinearity between order flow and price changes by two alternative ways and they are both statistically significant. In particular, our LSTR model shows that the price impact of order flow can change within normal conditions that the FX market is very likely to experience.

Table 3.1: Summary of aggregated databases

In each panel of the table, Freq is sampling frequency. Obs is the total number of (derived) observations in that database. $\bar{r}$ is the average return for that sampling frequency. Return(ΔP_t) is defined as $(log(P_t) - log(P_{t-1}))100$. tno, qno, bno, sprd, std is the average number of trades, average number of quotes, average number of buys, average bid-ask spread and average standard deviation of return for that frequency.

EUR/USD(a)

Freq	obs	rtn	tno	qno	bno	sprd	std
5m	21155	-0.0011	22	68	11	3.17	0.0127
10m	10855	-0.0017	43	133	22	3.35	0.0139
20m	5522	-0.0044	84	262	42	3.50	0.0153
30m	3653	-0.0054	126	395	64	3.39	0.0159
1hr	1966	-0.0092	237	742	120	4.23	0.0185

USD/JPY(b)

freq	obs	rtn	tno	qno	bno	sprd	std
5m	6139	0.0010	4	15	2	18.89	0.0292
10m	3704	0.0041	6	28	3	19.65	0.0316
20m	2129	0.0014	11	51	5	19.39	0.0351
30m	1166	0.0018	17	83	9	18.24	0.0413
1hr	1056	-0.0009	24	127	12	22.85	0.0435

EUR/GBP(c)

Freq	obs	rtn	tno	qno	bno	sprd	std
5m	17008	-0.0003	19	45	10	3.50	0.0129
10m	8700	-0.0006	38	89	20	3.62	0.0135
20m	4420	-0.0018	74	175	39	3.69	0.0141
30m	2930	-0.0032	112	265	58	3.63	0.0146
1hr	1541	-0.0056	214	508	111	3.88	0.0152

GBP/USD(d)

freq	obs	rtn	tno	qno	bno	sprd	std
5m	22935	-0.0005	21	54	11	2.03	0.0079
10m	11730	-0.0010	42	106	21	2.05	0.0082
20m	5939	-0.0020	83	210	42	2.07	0.0086
30m	3973	-0.0030	124	314	63	2.07	0.0089
1hr	2038	-0.0060	243	615	124	2.08	0.0093

Table 3.2: Spread effect on price impact of order flow

The model to be estimated is:

$$\Delta P_t = \alpha + \Sigma_{j=1}^{J} \beta_j * S_{jt} * Q_t + \varepsilon_t$$

where ΔP_t is price change defined as $(log(P_t) - log(P_{t-1})) * 100$. Q_{jt} is order flow which is defined as the net of number of buys and number of sells within $[t-1, t]$. Observations are divided into 5 categories by spread. S_{jt} is indicator variable that takes on the value 1 if the spread belongs to the specific category and 0 otherwise. The informativeness of order flow of category j is measured by regression coefficients β_j. The last column is the p-value of F-test of the null that $\beta_i = \beta_j, \forall i, j$. The model was estimated for a spectrum of sampling frequencies.

EUR/USD

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0027	0.0034	0.0041	0.0053	0.0077	0.4730	< 0.01
10m	0.0027	0.0033	0.0040	0.0050	0.0077	0.4681	< 0.01
20m	0.0026	0.0031	0.0040	0.0052	0.0078	0.4759	< 0.01
30m	0.0026	0.0031	0.0039	0.0049	0.0073	0.4707	< 0.01
1hr	0.0025	0.0028	0.0040	0.0054	0.0070	0.4584	< 0.01

USD/JPY

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0064	0.0090	0.0143	0.0152	0.0211	0.0982	< 0.01
10m	0.0070	0.0099	0.0140	0.0145	0.0212	0.1209	< 0.01
20m	0.0070	0.0091	0.0127	0.0155	0.0202	0.1648	< 0.01
30m	0.0078	0.0103	0.0111	0.0137	0.0202	0.2858	< 0.01
1hr	0.0077	0.0112	0.0133	0.0141	0.0213	0.2735	< 0.01

EUR/GBP

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0029	0.0036	0.0044	0.0051	0.0069	0.3527	< 0.01
10m	0.0031	0.0037	0.0040	0.0049	0.0064	0.3256	< 0.01
20m	0.0028	0.0036	0.0035	0.0041	0.0060	0.3027	< 0.01
30m	0.0028	0.0031	0.0032	0.0037	0.0055	0.2479	< 0.01
1hr	0.0023	0.0029	0.0032	0.0031	0.0056	0.2123	< 0.01

GBP/USD

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0024	0.0029	0.0032	0.0034	0.0043	0.3551	< 0.01
10m	0.0024	0.0028	0.0030	0.0031	0.0041	0.3402	< 0.01
20m	0.0024	0.0027	0.0027	0.0028	0.0039	0.2963	< 0.01
30m	0.0023	0.0023	0.0026	0.0023	0.0035	0.2571	< 0.01
1hr	0.0022	0.0018	0.0021	0.0021	0.0034	0.1938	< 0.01

Table 3.3: Volatility effect on price impact of order flow

The model to be estimated is:

$$\Delta P_t = \alpha + \sum_{j=1}^{J} \beta_j * S_{jt} * Q_t + \varepsilon_t$$

where ΔP_t is price change defined as $(log(P_t) - log(P_{t-1})) * 100$. Q_{jt} is order flow which is defined as the net of number of buys and number of sells within $[t-1, t]$. Observations are divided into 5 categories by volatility which is defined as the standard deviation of returns within $[t-1, t]$. S_{jt} is indicator variable that takes on the value 1 if the volume belongs to the specific category and 0 otherwise. The informativeness of order flow of category j is measured by regression coefficients β_j. The last column is the p-value of F-test of the null that $\beta_i = \beta_j, \forall i, j$. The model was estimated for a spectrum of sampling frequencies.

EUR/USD

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0026	0.0032	0.0038	0.0046	0.0066	0.4650	< 0.01
10m	0.0025	0.0032	0.0037	0.0045	0.0059	0.4582	< 0.01
20m	0.0023	0.0030	0.0035	0.0044	0.0058	0.4674	< 0.01
30m	0.0024	0.0030	0.0036	0.0042	0.0056	0.4612	< 0.01
1hr	0.0022	0.0029	0.0034	0.0044	0.0052	0.4439	< 0.01

USD/JPY

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0048	0.0078	0.0113	0.0158	0.0241	0.1071	< 0.01
10m	0.0057	0.0087	0.0121	0.0148	0.0231	0.1272	< 0.01
20m	0.0067	0.0093	0.0121	0.0142	0.0192	0.1614	< 0.01
30m	0.0075	0.0102	0.0120	0.0143	0.0183	0.2775	< 0.01
1hr	0.0085	0.0100	0.0116	0.0144	0.0176	0.2603	< 0.01

EUR/GBP

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0033	0.0037	0.0041	0.0046	0.0062	0.3437	< 0.01
10m	0.0032	0.0035	0.0038	0.0045	0.0057	0.3217	< 0.01
20m	0.0031	0.0032	0.0035	0.0043	0.0047	0.2944	< 0.01
30m	0.0029	0.0032	0.0032	0.0040	0.0038	0.2384	< 0.05
1hr	0.0027	0.0028	0.0032	0.0033	0.0035	0.1979	> 0.10

GBP/USD

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0023	0.0028	0.0030	0.0033	0.0040	0.3545	< 0.01
10m	0.0022	0.0027	0.0029	0.0032	0.0035	0.3359	< 0.01
20m	0.0023	0.0023	0.0028	0.0029	0.0034	0.2947	< 0.01
30m	0.0021	0.0024	0.0026	0.0025	0.0028	0.2533	< 0.05
1hr	0.0020	0.0022	0.0024	0.0022	0.0021	0.1847	> 0.10

Table 3.4: Volume effect on price impact of order flow

The model to be estimated is:

$$\Delta P_t = \alpha + \sum_{j=1}^{J} \beta_j * S_{jt} * Q_t + \varepsilon_t$$

where ΔP_t is price change defined as $(log(P_t) - log(P_{t-1})) * 100$. Q_{jt} is order flow which is defined as the net of number of buys and number of sells within $[t-1, t]$. Observations are divided into 5 categories by volume which is defined as the total number of trades within $[t-1, t]$. S_{jt} is indicator variable that takes on the value 1 if the volume belongs to the specific category and 0 otherwise. The informativeness of order flow of category j is measured by regression coefficients β_j. The last column is the p-value of F-test of the null that $\beta_i = \beta_j, \forall i, j$. The model was estimated for a spectrum of sampling frequencies.

EUR/USD

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0070	0.0045	0.0040	0.0040	0.0042	0.4278	< 0.01
10m	0.0073	0.0044	0.0037	0.0037	0.0040	0.4303	< 0.01
20m	0.0076	0.0039	0.0037	0.0035	0.0040	0.4361	< 0.01
30m	0.0066	0.0033	0.0034	0.0034	0.0039	0.4356	< 0.01
1hr	0.0071	0.0042	0.0032	0.0034	0.0037	0.4206	< 0.01

USD/JPY

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0057	0.0141	0.0128	0.0146	0.0133	0.0871	< 0.05
10m	0.0089	0.0120	0.0193	0.0145	0.0126	0.1100	< 0.05
20m	0.0137	0.0158	0.0149	0.0131	0.0120	0.1473	> 0.10
30m	0.0144	0.0146	0.0120	0.0129	0.0128	0.2588	> 0.10
1hr	0.0103	0.0177	0.0141	0.0126	0.0130	0.2482	> 0.10

EUR/GBP

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0055	0.0046	0.0044	0.0043	0.0044	0.3275	< 0.01
10m	0.0058	0.0041	0.0041	0.0041	0.0042	0.3089	< 0.01
20m	0.0056	0.0041	0.0037	0.0041	0.0036	0.2901	< 0.01
30m	0.0053	0.0037	0.0036	0.0036	0.0032	0.2395	< 0.01
1hr	0.0064	0.0031	0.0033	0.0034	0.0027	0.2090	< 0.01

GBP/USD

Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	p-value
5m	0.0032	0.0031	0.0032	0.0031	0.0032	0.3434	> 0.10
10m	0.0032	0.0029	0.0029	0.0029	0.0031	0.3301	> 0.10
20m	0.0030	0.0028	0.0028	0.0028	0.0028	0.2887	> 0.10
30m	0.0029	0.0026	0.0026	0.0025	0.0025	0.2516	> 0.10
1hr	0.0027	0.0027	0.0024	0.0021	0.0020	0.1868	> 0.10

Table 3.5: Relationship between order flow informativeness and spread

The model to be estimated is:

$$\Delta P_t = \alpha + \beta_1 * Q_t + \beta_2 * Z_t * Q_t + \varepsilon_t$$

where ΔP_t is price change defined as $(log(P_{t+\Delta t}) - log(P_t)) * 100$. Q_t is order flow which is defined as the net of number of buys and number of sells within $[t-1, t]$. Z_t is average spread during the time interval $[t-1, t]$. The linear relationship between informativeness of order flow and spread will be captured by β_2. The reported t-value is based on Newey-West variance-covariance matrix estimate. The model was estimated for a spectrum of sampling frequencies.

	EUR/USD(a)					USD/JPY(b)				
Freq	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	R^2	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	R^2
5m	0.0025	5.45	0.0007	3.31	0.4547	0.0103	15.17	0.0002	4.34	0.0924
10m	0.0022	4.33	0.0008	3.47	0.4602	0.0102	13.22	0.0002	3.97	0.1171
20m	0.0022	3.24	0.0007	2.27	0.4593	0.0093	11.02	0.0002	3.49	0.1572
30m	0.0017	4.23	0.0008	4.41	0.4659	0.0092	10.35	0.0002	4.09	0.2815
1hr	0.0024	3.45	0.0005	1.64	0.4338	0.0098	9.09	0.0002	3.12	0.2612

	EUR/GBP(c)					GBP/USD(d)				
Freq	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	R^2	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	R^2
5m	0.0020	11.27	0.0008	13.37	0.3503	0.0013	8.33	0.0010	11.48	0.3559
10m	0.0019	7.36	0.0007	8.50	0.3271	0.0011	4.06	0.0010	7.13	0.3419
20m	0.0016	5.48	0.0007	7.41	0.3054	0.0008	2.09	0.0010	4.73	0.2987
30m	0.0013	3.65	0.0007	6.33	0.2532	0.0011	2.54	0.0007	3.25	0.2560
1hr	0.0010	1.87	0.0006	3.58	0.2123	0.0004	0.73	0.0009	2.86	0.1900

Table 3.6: Relationship between order flow informativeness and volatility

The model to be estimated is:

$$\Delta P_t = \alpha + \beta_1 * Q_t + \beta_2 * Z_t * Q_t + \varepsilon_t$$

where ΔP_t is price change defined as $(log(P_t) - log(P_{t-1})) * 100$. Q_t is order flow which is defined as the net of number of buys and number of sells within $[t-1,t]$. Z_t is volatility which is defined as the standard deviation of returns within $[t-1,t]$. The linear relationship between informativeness of order flow and volatility will be captured by β_2. The reported t-value is based on Newey-West variance-covariance matrix estimate. The model was estimated for a spectrum of sampling frequencies.

	EUR/USD(a)					USD/JPY(b)				
Freq	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	R^2	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	R^2
5m	0.0034	25.15	0.0569	4.99	0.4407	0.0124	12.67	0.0239	0.55	0.0865
10m	0.0035	28.10	0.0374	3.93	0.4335	0.0128	10.37	0.0194	0.40	0.1079
20m	0.0033	27.32	0.0390	4.72	0.4406	0.0118	15.68	0.0319	1.20	0.1501
30m	0.0033	20.51	0.0287	2.54	0.4347	0.0113	18.26	0.0374	2.14	0.2711
1hr	0.0032	14.93	0.0287	2.04	0.4194	0.0119	17.45	0.0278	2.34	0.2517

	EUR/GBP(c)					GBP/USD(d)				
Freq	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	R^2	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	R^2
5m	0.0030	16.00	0.1063	7.20	0.3400	0.0022	16.76	0.1104	6.47	0.3516
10m	0.0029	11.18	0.0899	4.61	0.3171	0.0024	7.01	0.0676	1.65	0.3340
20m	0.0031	9.18	0.0475	1.94	0.2900	0.0021	4.46	0.0751	1.36	0.2951
30m	0.0026	11.44	0.0605	3.77	0.2413	0.0025	7.57	0.0034	0.09	0.2512
1hr	0.0027	3.42	0.0299	0.54	0.1974	0.0017	4.57	0.0478	1.22	0.1876

Table 3.7: Relationship between order flow informativeness and volume

The model to be estimated is:

$$\Delta P_t = \alpha + \beta_1 * Q_t + \beta_2 * Z_t * Q_t + \varepsilon_t$$

where ΔP_t is price change defined as $(log(P_t) - log(P_{t-1})) * 100$. Q_t is order flow which is defined as the net of number of buys and number of sells within $[t-1,t]$. Z_t is volume (scaled by 10^{-2})which is defined as the total number of trades within $[t-1,t]$. The linear relationship between informativeness of order flow and volume will be captured by β_2. The reported t-value is based on Newey-West variance-covariance matrix estimate. The model was estimated for a spectrum of sampling frequencies.

	EUR/USD(a)					USD/JPY(b)				
Freq	β_1	$t(\beta_1)$	β_2	$t(\beta_2)$	R^2	β_1	$t(\beta_1)$	β_2	$t(\beta_2)$	R^2
5m	0.0039	35.28	0.0007	2.16	0.4245	0.0137	14.09	-0.0078	-0.64	0.0855
10m	0.0039	30.94	0.0001	0.80	0.4237	0.0150	11.98	-0.0138	-1.42	0.1074
20m	0.0037	17.96	0.0002	1.06	0.4275	0.0148	11.31	-0.0107	-1.84	0.1467
30m	0.0034	11.57	0.0002	1.10	0.4280	0.0139	9.91	-0.0038	-0.76	0.2584
1hr	0.0039	15.15	-0.0001	-0.96	0.4120	0.0147	10.41	-0.0037	-1.05	0.2459

	EUR/GBP(c)					GBP/USD(d)				
Freq	β_1	$t(\beta_1)$	β_2	$t(\beta_2)$	R^2	β_1	$t(\beta_1)$	β_2	$t(\beta_2)$	R^2
5m	0.0046	37.16	-0.0004	-1.07	0.3269	0.0030	36.36	0.0003	1.13	0.3436
10m	0.0045	22.63	-0.0004	-1.13	0.3075	0.0031	27.44	-0.0001	-0.36	0.3300
20m	0.0046	17.84	-0.0006	-2.62	0.2903	0.0028	15.43	0.0000	0.20	0.2886
30m	0.0050	12.81	-0.0009	-3.18	0.2481	0.0026	13.07	-0.0001	-0.43	0.2513
1hr	0.0045	9.74	-0.0004	-2.55	0.2066	0.0028	8.34	-0.0002	-1.60	0.1881

Table 3.8: Coefficients estimates for LSTR model

The model to be estimate is:

$$\Delta P_t = \beta_0 + \beta_1 * Q_t + (\theta_0 + \theta_1 * Q_t)F(Z_t) + \varepsilon_t$$

where $F(Z_t)$ is a logistic function that can be written as

$$F(Z_t) = (1 + exp\{-\gamma(Z_t - c)\})^{-1} - 1/2, \ \gamma > 0$$

and ΔP_t is price change defined as $(log(P_t) - log(P_{t-1})) * 100$. Q_t is order flow defined as the net of number of buys and number of sells within $[t-1, t]$. Z_t is the market conditions measuring variable. The model estimation (based on hourly sampling frequency) under different market conditions of spread, volatility and volume are reported in the following three panels respectively. The number in bracket is is t-value. '-' indicates that the model can not be fitted. The transition feature of the price impact of order flow is captured by the parameters γ and θ_1.

Spread

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\gamma}$	$\hat{c}$
EUR/USD	-0.0188	0.0042	0.0201	0.0048	8.5287	2.70
	(-5.30)	(35.42)	(1.87)	(5.78)	(3.27)	
USD/JPY	-0.0020	0.0113	-0.0458	0.0285	2.1476	14.00
	(-0.26)	(14.76)	(-1.11)	(3.86)	(2.48)	
EUR/GBP	-0.0427	0.0028	0.0906	0.0071	0.9801	2.70
	(-7.81)	(15.82)	(2.39)	(2.36)	(1.97)	
GBP/USD	-0.0086	0.0034	0.0200	0.0031	1.4006	3.00
	(-1.44)	(10.67)	(1.36)	(3.85)	(2.03)	

Volatility

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\gamma}$	$\hat{c}$
EUR/USD	-0.0165	0.0031	-0.0141	0.0046	5.7683	0.011
	(-4.13)	(28.20)	(-0.93)	(7.60)	(4.35)	
USD/JPY	0.0117	0.0067	-0.0902	0.0259	4.8007	0.006
	(0.83)	(3.17)	(-1.76)	(5.13)	(2.75)	
EUR/GBP	-0.0313	0.0030	-0.0187	0.0009	33.0713	0.012
	(-6.49)	(17.12)	(-1.70)	(2.13)	(0.48)	
GBP/USD	-0.0136	0.0020	-0.0900	0.0081	0.1736	0.006
	(-4.11)	(13.84)	(-0.41)	(0.42)	(0.36)	

Volume

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\gamma}$	$\hat{c}$
EUR/USD	0.0134	0.0130	-0.0741	-0.0189	4.1527	10
	(1.14)	(7.23)	(-2.87)	(-5.27)	(5.10)	
USD/JPY	-0.0098	0.0141	-0.0050	-0.0050	0.5887	18
	(-1.30)	(12.85)	(-0.08)	(-0.44)	(0.29)	
EUR/GBP	-0.0068	0.0081	-0.0724	-0.0107	1.9924	2
	(-0.43)	(4.40)	(-1.90)	(-2.92)	(3.73)	
GBP/USD	-	-	-	-	-	-
	-	-	-	-	-	-

Table 3.9: Seasonality impact analysis

The model to be estimated is:

$$\Delta P_t = \alpha + \beta_1 * Q_t + \beta_2 * Z_t^u * Q_t + \beta_3 * Z_t^s * Q_t + \varepsilon_t$$

where ΔP_t is price change within $[t-1,t]$. Q_{jt} is order flow which is defined as the net of number of buys and number of sells within $[t-1,t]$. Z_t^u and Z_t^s are the unexpected and expected part of market condition measuring variable (Z_t)of interest (Volume is scaled down by 10^3). The model is estimated for 30 minute and 1 hour sampling frequencies. The last column is the p-value of F-test of null hypothesis that $\beta_2 = \beta_3$

Panel A: Sampling Freq=30 minute							
$Z = Spread$							
CurrID	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	$\hat{\beta}_3$	$t(\hat{\beta}_3)$	p-value
EUR/USD	0.0019	(10.44)	0.0009	(11.74)	0.0008	(11.79)	> 0.10
EUR/GBP	0.0010	(2.63)	0.0006	(5.42)	0.0008	(6.55)	> 0.25
GBP/USD	0.0008	(1.34)	0.0007	(4.23)	0.0009	(2.96)	> 0.25
USD/JPY	0.0133	(4.80)	0.0002	(6.35)	0.0000	-(0.24)	> 0.10
$Z = Volume$							
CurrID	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	$\hat{\beta}_3$	$t(\hat{\beta}_3)$	p-value
EUR/USD	0.0061	(17.88)	0.0048	(5.58)	-0.0175	-(7.77)	< 0.01
EUR/GBP	0.0061	(12.38)	-0.0074	-(5.73)	-0.0180	-(4.67)	< 0.01
GBP/USD	0.0032	(11.92)	0.0002	(0.23)	-0.0052	-(2.75)	< 0.01
USD/JPY	0.0227	(3.63)	-0.0171	-(0.33)	-0.5529	-(1.53)	> 0.10
$Z = Volatility$							
CurrID	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	$\hat{\beta}_3$	$t(\hat{\beta}_3)$	p-value
EUR/USD	0.0004	(1.00)	0.0226	(5.47)	0.2317	(9.70)	< 0.01
EUR/GBP	-0.0009	-(0.89)	0.0471	(3.60)	0.3049	(4.33)	< 0.01
GBP/USD	-0.0002	-(0.18)	-0.0004	-(0.03)	0.3037	(2.56)	< 0.05
USD/JPY	0.0101	(3.22)	0.0371	(4.52)	0.0656	(0.88)	> 0.10

Panel B: Sampling Freq=1 hour							
$Z = Spread$							
CurrID	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	$\hat{\beta}_3$	$t(\hat{\beta}_3)$	p-value
EUR/USD	0.0025	(12.07)	0.0006	(5.80)	0.0004	(6.51)	> 0.25
EUR/GBP	0.0013	(2.54)	0.0008	(4.20)	0.0006	(4.24)	> 0.25
GBP/USD	-0.0002	-(0.22)	0.0008	(2.91)	0.0012	(2.73)	> 0.25
USD/JPY	0.0125	(5.94)	0.0002	(5.00)	0.0000	(0.51)	> 0.10
$Z = Volume$							
CurrID	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	$\hat{\beta}_3$	$t(\hat{\beta}_3)$	P-value
EUR/USD	0.0061	(13.64)	0.0013	(1.98)	-0.0089	-(5.84)	< 0.01
EUR/GBP	0.0054	(8.47)	-0.0036	-(3.62)	-0.0080	-(3.17)	< 0.10
GBP/USD	0.0032	(8.01)	-0.0015	-(2.41)	-0.0031	-(2.21)	> 0.10
USD/JPY	0.0206	(4.38)	-0.0279	-(0.77)	-0.2901	-(1.50)	> 0.10
$Z = Volatility$							
CurrID	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	$\hat{\beta}_3$	$t(\hat{\beta}_3)$	P-value
EUR/USD	0.0001	(0.25)	0.0246	(4.39)	0.2145	(7.41)	< 0.01
EUR/GBP	-0.0003	-(0.21)	0.0155	(0.82)	0.2275	(2.59)	< 0.05
GBP/USD	-0.0017	-(1.06)	0.0419	(2.54)	0.4058	(2.44)	< 0.05
USD/JPY	0.0099	(2.02)	0.0277	(3.04)	0.0729	(0.66)	> 0.25

Table 3.10: Simultaneity analysis

The model to be estimated is:

$$\Delta P_t = \alpha + \beta_1 * Q_t + \beta_2 * Z_t^{Spread} * Q_t + \beta_3 * Z_t^{Volume} * Q_t + \beta_4 * Z_t^{Volatility} * Q_t + \varepsilon_t$$

where ΔP_t is price change within $[t-1,t]$. Q_{jt} is order flow defined as the net of number of buys and number of sells within $[t-1,t]$. Z_t^{Spread},Z_t^{Volume} and $Z_t^{Volatility}$ are the average spread, total volume and volatility within $[t-1,t]$ respectively (Volume is scaled down by 10^3). The model is estimated for 30 minute and 1 hour sampling frequencies. The number reported in parentheses are t-value.

Panel A: Sampling Freq=30 minutes					
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	R^2
EUR/USD	0.0008	0.0008	0.0037	0.0093	0.4706
	(3.46)	(15.68)	(4.71)	(2.20)	
EUR/GBP	0.0027	0.0005	-0.0079	0.0429	0.2638
	(6.97)	(5.11)	-(6.32)	(3.06)	
GBP/USD	0.0010	0.0008	-0.0005	-0.0215	0.2565
	(2.85)	(5.22)	-(0.68)	-(1.55)	
USD/JPY	0.0091	0.0002	-0.0318	0.0303	0.2900
	(5.86)	(5.46)	-(0.66)	(3.70)	

Panel B: Sampling Freq=1 hour					
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	R^2
EUR/USD	0.0019	0.0005	0.0005	0.0200	0.4378
	(6.00)	(7.93)	(0.87)	(3.54)	
EUR/GBP	0.0024	0.0005	-0.0032	0.0069	0.2181
	(4.13)	(4.16)	-(3.35)	(0.35)	
GBP/USD	0.0013	0.0006	-0.0017	0.0446	0.1949
	(2.23)	(2.22)	-(2.89)	(2.57)	
USD/JPY	0.0101	0.0002	-0.0290	0.0226	0.2655
	(5.69)	(4.17)	-(0.80)	(2.44)	

Table 3.11: Factor Cross-autocorrelation

where *Vol*, *Std* and *Sprd* are volume, volatility and spread respectively. For each of the frequencies, first three lines are the correlations, the middle three lines are first order cross-autocorrelations and the bottom three lines are second order cross-autocorrelations.

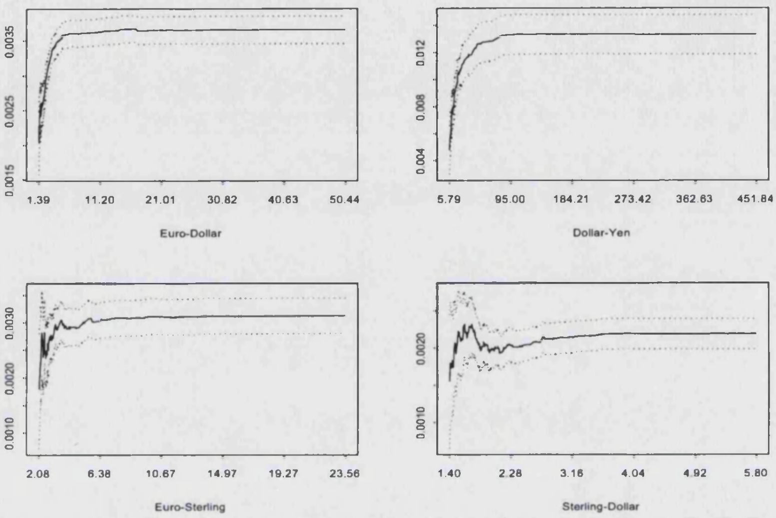
Freq=30 mins

	ED			ES			SD			YD		
	Vol	**Sprd**	**Std**	**Vol**	**Sprd**	**Std**	**Vol**	**Sprd**	**Std**	**Vol**	**Sprd**	**Std**
Vol	1			1			1			1		
Sprd	-0.48	1		-0.35	1		-0.32	1		-0.11	1	
Std	-0.05	0.32	1	0.02	0.28	1	0.08	0.23	1	0.03	0.16	1
Vol_{t-1}	0.6	-0.4	-0.11	0.63	-0.28	-0.01	0.59	-0.21	0.05	0.4	-0.18	-0.02
$Sprd_{t-1}$	-0.42	0.65	0.23	-0.31	0.58	0.18	-0.27	0.58	0.11	-0.11	0.26	0.03
Std_{t-1}	-0.07	0.22	0.13	-0.01	0.19	0.1	0.06	0.13	0.08	-0.02	0.02	0.03
Vol_{t-2}	0.34	-0.25	-0.05	0.41	-0.13	-0.03	0.36	-0.09	0.06	0.21	-0.1	-0.02
$Sprd_{t-2}$	-0.27	0.37	0.16	-0.2	0.37	0.14	-0.21	0.42	0.08	-0.11	0.12	0.04
Std_{t-2}	-0.07	0.16	0.09	-0.04	0.19	0.08	0.04	0.11	0.06	-0.04	0.02	0.04

Freq=1 hr

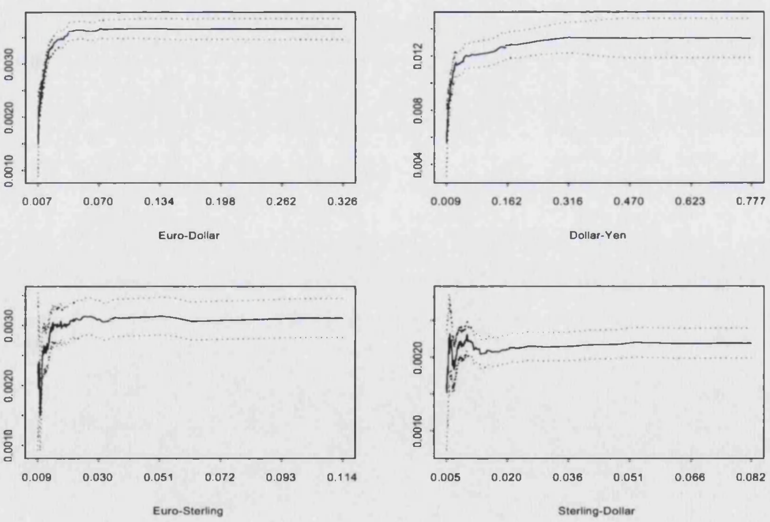
	ED			ES			SD			YD		
	Vol	**Sprd**	**Std**	**Vol**	**Sprd**	**Std**	**Vol**	**Sprd**	**Std**	**Vol**	**Sprd**	**Std**
Vol	1			1			1			1		
Sprd	-0.52	1		-0.46	1		-0.36	1		-0.18	1	
Std	-0.13	0.31	1	0.02	0.22	1	0.06	0.2	1	0.1	0.12	1
Vol_{t-1}	0.51	-0.36	-0.11	0.54	-0.23	0.03	0.47	-0.12	0.04	0.47	-0.19	0
$Sprd_{t-1}$	-0.34	0.36	0.17	-0.28	0.44	0.16	-0.27	0.51	0.09	-0.09	0.25	0.02
Std_{t-1}	-0.1	0.15	0.07	-0.02	0.21	0.1	0.03	0.13	0.05	-0.03	0.02	0.03
Vol_{t-2}	0.06	-0.06	-0.01	0.18	0.08	0.12	0.13	0.11	0.09	0.24	-0.11	0.04
$Sprd_{t-2}$	0.01	0.03	0.05	-0.05	0.11	0.03	-0.1	0.26	0.08	-0.06	0.04	0.01
Std_{t-2}	-0.02	0.07	0.04	0.02	0.08	0.08	0.03	0.12	0.06	0	0.02	0.05

Figure 3.1: Function $\beta(Z_t)$) from recursive model (Z_t is spread)



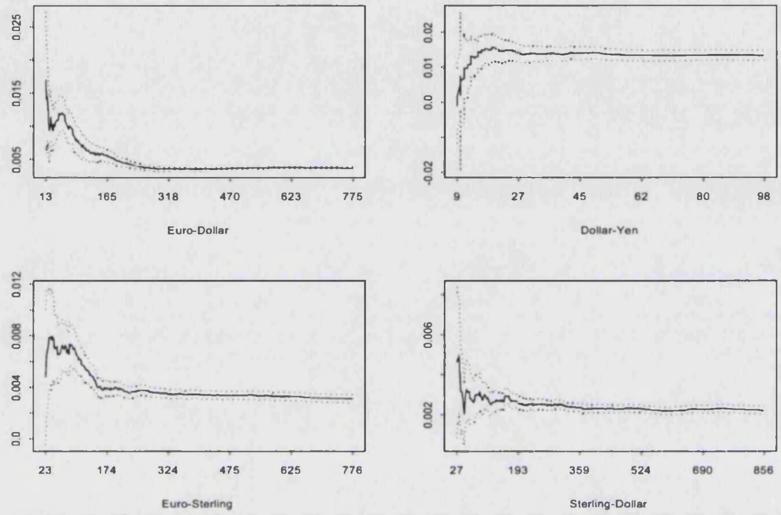
The regression coefficients of the recursive regression model are drawn against market spread. Spreads are measured in basis points. The dotted line is the two standard deviation confidence bounds.

Figure 3.2: Function $\beta(Z_t)$ from recursive model (Z_t is volatility)



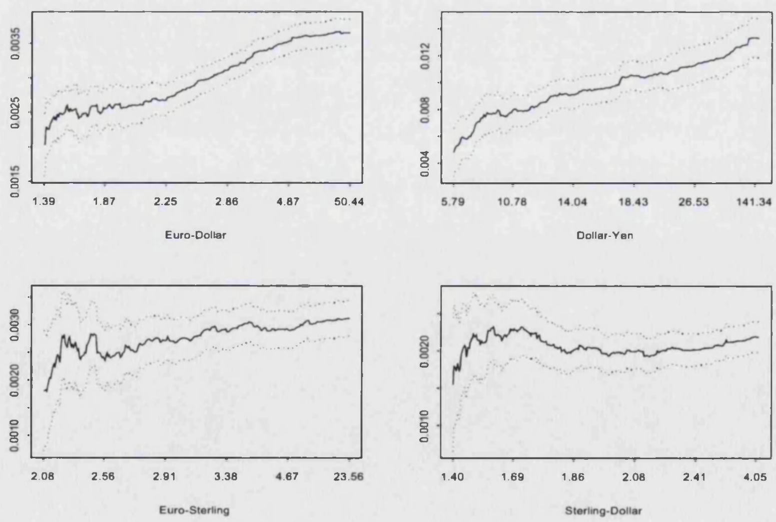
The regression coefficients of the recursive regression model are drawn against volatility. Volatility is computed as return standard deviation. The dotted line is the two standard deviation confidence bounds.

Figure 3.3: Function $\beta(Z_t)$ from recursive model (Z_t is volume)



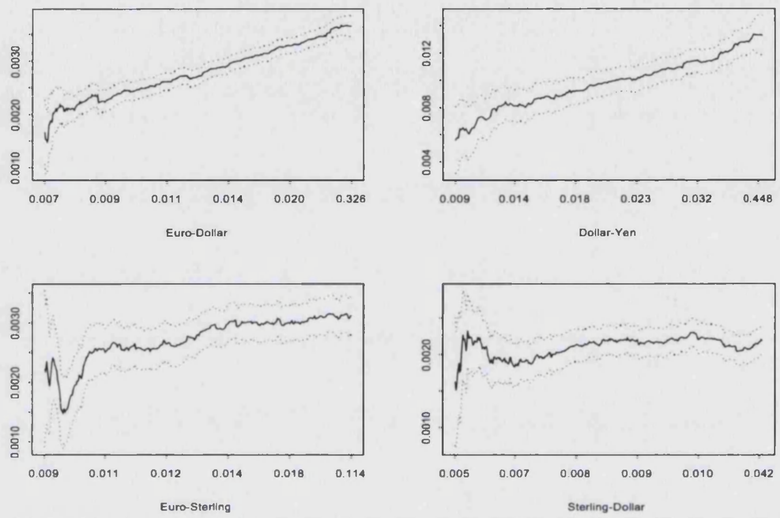
The regression coefficients of the recursive regression model are drawn against volume. Volume is proxied by the total number of trades. The dotted line is the two standard deviation confidence bounds.

Figure 3.4: Evolution of β along spread in recursive model



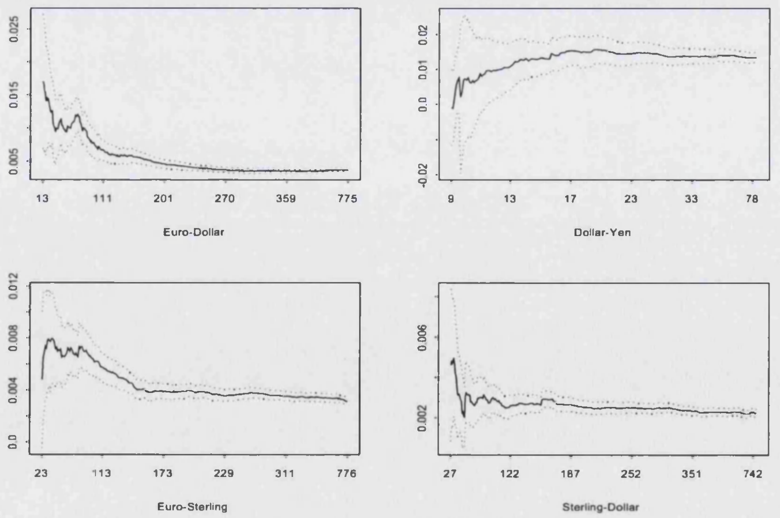
βs are drawn along the recursive regression process. Note: x-axis is deliberately labelled with the spread of the last observation in that regression rather than with the sequence number of regression. Spreads are measured in basis points. The dotted line is the two standard deviation confidence bounds.

Figure 3.5: Evolution of β along volatility in recursive model



βs are drawn along the recursive regression process. Note: x-axis is deliberately labelled with the volatility of the last observation in that regression rather than with the sequence number of regression. Volatility is computed as return standard deviation. The dotted line is the two standard deviation confidence bounds.

Figure 3.6: Evolution of β along volume in recursive model



βs are drawn along the recursive regression process. Note: x-axis is deliberately labelled with the volume of the last observation in that regression rather than with the sequence number of regression. Volume is proxied by the total number of trades. The dotted line is the two standard deviation confidence bounds.

Figure 3.7: Shift of price impact of order flow with spread

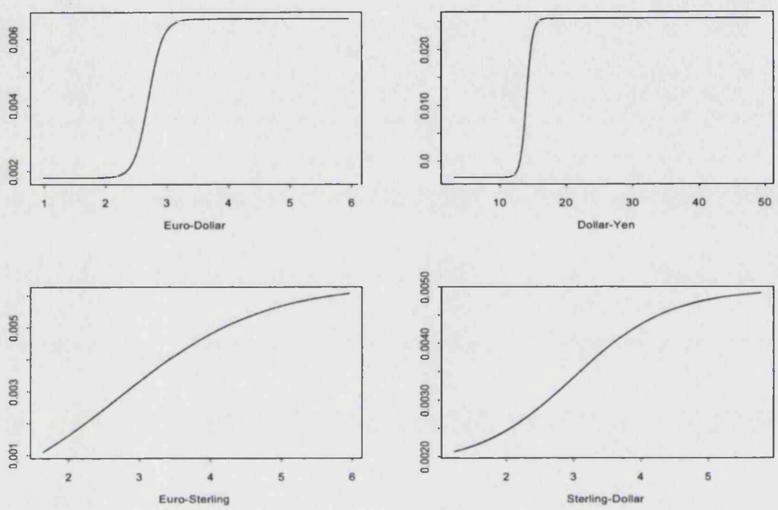


Fig.14 is a graphic representation of LSTR model when transition variable Z_t is spread. In this figure, $\beta_1 + \theta_1 * F(Z_t)$ in model (6) is drawn against Z_t with the estimated parameters $\hat{\beta}_1, \hat{\theta}_1, \hat{\gamma}, \hat{c}$.

Figure 3.8: Shift of price impact of order flow with volatility

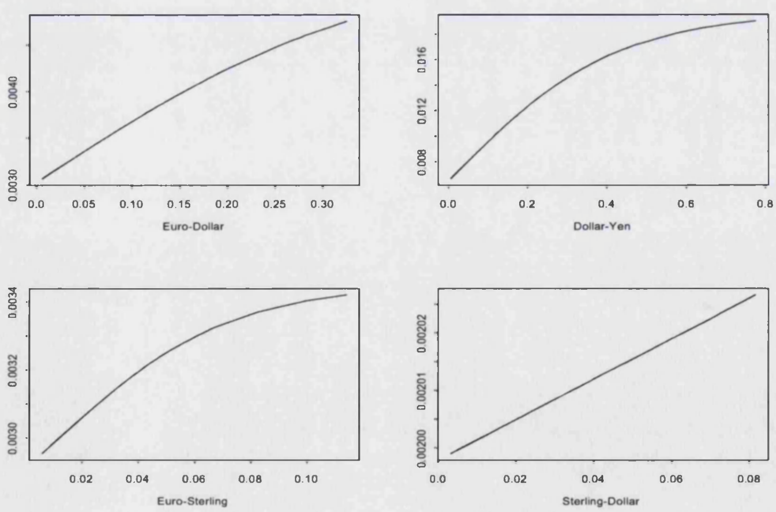


Fig.15 is a graphic representation of LSTR model when transition variable Z_t is volatility. In this figure, $\beta_1 + \theta_1 * F(Z_t)$ in model (6) is drawn against Z_t with the estimated parameters $\hat{\beta}_1, \hat{\theta}_1, \hat{\gamma}, \hat{c}$.

Figure 3.9: Shift of price impact of order flow with volume

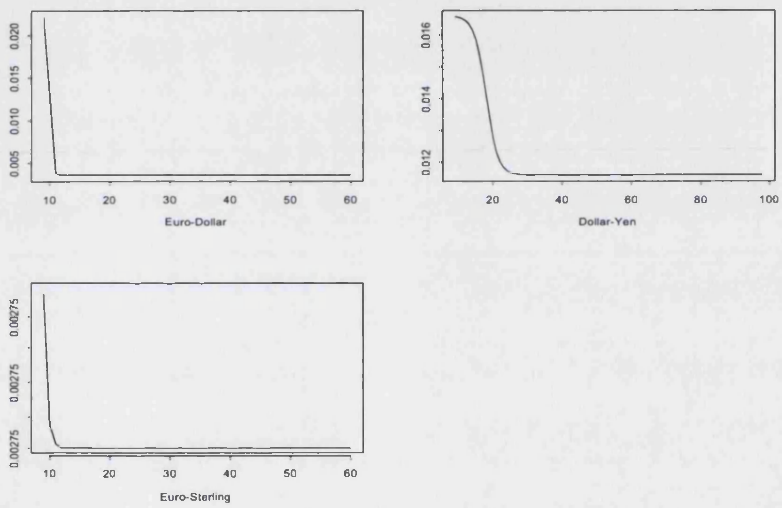


Fig.15 is a graphic representation of LSTR model when transition variable Z_t is volume. In this figure, $\beta_1 + \theta_1 * F(Z_t)$ in model (6) is drawn against Z_t with the estimated parameters $\hat{\beta}_1, \hat{\theta}_1, \hat{\gamma}, \hat{c}$.

Chapter 4

Limit Order Execution Probability: Evidence from the London Stock Exchange

4.1 Introduction

Interest in limit order trading has grown rapidly in recent years as more and more stock exchanges are adopting automatic order-driven systems. In order-driven systems, such as Tokyo Stock Exchange, Paris Bourse or Hong Kong Stock Exchange, there is no designed market maker as intermediary of trading. Supply and demand of liquidity is conducted by the natural buyers and sellers.

Comparing with market order, limit order has the benefit of price improvement and no price uncertainty. However, the benefit does not come without cost: execution is not secured. Such *implicit costs...associated with missed trading opportunities...are significant relative to explicit costs and realized portfolio returns.*[1] In a study by Handa and Schwartz (1996), the authors experiment with NYSE trans-

[1]Source: *Financial Times*, 23 July 2002, by Donald Keim and Ananth Madhavan

action data and find that limit order strategy is profitable conditional on execution but non-execution limit order is inferior to market order. Since non-execution cost could be very high, investors trading with limit order need to evaluate the execution probability of their submission strategies and take this into account in forming the investment decisions.

This chapter analyzes how such probability can be affected by various factors and attempts to provide a mechanism to evaluate the execution probability of different order submission strategies. The issue is clearly important to the market participants, especially to the trading desk of institutional investors to whom the trading cost is among the top concerns. It is also important from theoretical perspective because limit order placement has important impact on market liquidity and spread dynamics. In the classic market microstructure models, informed traders are overwhelmingly assumed to use market orders to exploit their information advantage. However, this prior is not necessarily true both in theory and reality. In theory, the optimal strategy of an informed investor doesn't necessarily lead to the use of market order. Using a market order, an investor has to pay bid-ask spread and price impact cost and these cost can reduce the potential benefit of the private information. If using a limit order, the probability of execution is less than one but transaction cost (including spread and negative price movement) is lower. Overall, the expected profit of using limit order is not necessarily less than that of using market order. When transaction cost is high and limit order execution probability is large enough, the informed investor can be better off by issuing limit order.[2] There is experimental evidence showing that informed investors do not restrict themselves to market orders. In their experiment study, Bloomfield, O'Hara and Saar (2002) find informed traders can use more limit orders than liq-

[2]When the execution probability is small, more market order will be used. The increasing usage of market order will increase the execution probability of limit order. In theory, there should be an equilibrium under which the expected profit is indifferent to the investors.

uidity traders do in certain market conditions. Clearly, the execution probability is a key determinant on the choice of order type. A better understanding of how such probability is determined — the objective of this study — will enhance our knowledge of price discovery process in limit order trading systems.

In recent years, limit order trading has drawn increasing attention in market microstructure literature. Cohen, Maier, Schwartz and Whitcomb (1981), Angel (1994), Glosten (1994), kumar and Seppi (1992), Seppi (1997), Handa and Schwartz (1996), Parlour (1998) and Foucault (1999) develop equilibrium models of limit order book. Chung, Ness and Ness (1999) examine the impact of limit order on the dynamics of bid-ask spreads in NYSE. Biais, Hillion and Spatt (1995) and Ahn, Bae and Chan (2001) empirically investigate the order flow dynamics in Paris Bourse and Hong Kong Stock Exchange respectively. These studies imply that various factors can impact on limit order execution, however, none of them focuses on the limit order execution itself. Instead the attention is mainly on the choice between market order and limit order and how such choice impacts on market equilibrium.

Works on limit order execution are relatively few and most concentrate on time-to-execution. Foucault, Kadan and Kandel (2001) develop an equilibrium model for time-to-execution of limit order. Under some simplified assumptions[3], the model predicts that time-to-execution depends on order aggressiveness and other market conditions. In another work, Lo, Mckinlay and Zhang (2002) compare three different econometric models for the time-to-execution of limit order and claim survival model is superior to other two models in fitting the data. Using limit orders via QuantEx of ITG, they also find the time-to-execution is sensitive to limit price and other explanatory variables. The current study focuses on the execution probability and can be seen as complementary study of the above

[3]In this model, if a trade submits a limit order, he/she must improve prevailing inside spread.

papers.

Research in this chapter is also related to the microstructure literature on profitability study of order submission strategies. Handa and Schwartz (1996) and Harris and Hasbrouck (1996) compare the profitability of different order submission strategies in NYSE. The performance in both paper is measured ex post, i.e. conditional on actual execution state of limit order. This chapter attempts to analyze how limit order execution probability is affected by various factors and evaluate such probability ex ante. Clearly this study can been seen as an extension of this literature because it provides the potential for further study of profitability of order submission strategies in real time.

Using a sample of FT30 stocks from the limit book system of the London Stock Exchange (LSE), we find three factors, *price aggressiveness, spread and potential market pressure*, have significant impact on limit order execution probability for all stocks. In particular, the more aggressive limit order the larger chance the order will get filled. For a given limit price, a wider bid-ask spread will reduce the execution probability. A limit order will have higher execution probability when market pressure from the opposite market is larger. Moreover, we also find liquidity has strong impact on limit order execution. Trading less liquid stock, a trader has to post a more aggressive limit price to get the same chance of execution as trading liquid stocks.[4] Impact of volatility is positive but weak.

Contrary to the intuition, order size and the time of the day do not have strong impact on limit order execution as expected. For order size, we find a U-shape of execution probability with middle-size order having the smallest execution chance. Many reasons could lead to such U-shape. For example, if large orders are pre-negotiated outside market, they will be easily executed. Since such information is not available, we cannot test this hypothesis in this study. However, we

[4]This result is consistent with the finding in NYSE by Chung, Ness and Ness (1999) that liquidity provided by specialist is most valuable with less liquid stocks.

do find that large orders usually are submitted more strategically. They cut spread aggressively when it is large. In this way they deviate themselves from the rest of limit orders to get a better execution chance.[5]

In evaluating the execution probability, we incorporating all factors analyzed in a probit model. We find the model is capable of capturing the major feature of the relationship between limit order execution probability and its determinant factors. Since the factors are measured conditional on information available at the moment of order submission, the model can be used to evaluate the execution chance of various order submission strategies in real time.

The rest of this chapter is organized as follows. The next section describes our empirical approach and defines hypotheses. Section 4.3 discusses market and data. Section 4.4 presents results on the impact of various factors. Section 4.5 models the limit order execution probability. Section 4.6 discusses and concludes.

4.2 Analysis Design

4.2.1 Factor selection and hypothesis specification

For the present purpose, a limit order strategy is defined as the combination of order choice (including limit price and order size) and market conditions chosen to place the order. Market conditions per se cannot be controlled by investors but investors can time their trading in more favorable conditions. In the current study, we analyze a set of factors might have impacts on limit order execution probabilities. The factors studied in this work can be roughly classified as theory-motivated and intuition-motivated factors. The former includes price aggressiveness, volatility and volume. The latter includes spread, order size, and potential market pres-

[5] We also find large orders tend to be submitted when market pressure from other side of market is large though this is only significant for sell orders.

sure. In the next subsections we define each factor and define hypotheses.

A. Price aggressiveness and Bid-Ask Spread

Foucault et al. (2001) derive a dynamic model of order driven market populated with discretionary liquidity traders. Traders differ by their impatience which is modelled as waiting cost. In general, impatient trades use market order while patient traders are more likely to use limit order. Under equilibrium, the competition among liquidity providers will force them to submit more aggressive orders to reduce the waiting cost. The model implies that the more aggressive is the limit order, the larger is the execution probability.

Actually the price aggressiveness is first introduced in an empirical study by Harris and Hasbrouck (1996). Denote the price of limit order, bid, ask and middle quotes prevailing at the time of limit order submission as p^l, q^{bid}, q^{ask}, and q^{mid} respectively. Price aggressiveness is defined as

$$L = \begin{cases} \frac{p^l - q^{bid}}{q^{mid}} & for\ a\ buy\ order, \\ \frac{q^{ask} - p^l}{q^{mid}} & for\ a\ sell\ order. \end{cases} \tag{4.1}$$

L measures the extent to which a limit order betters the existing quotes of the same side. A limit buy (sell) order posted at the current best bid (ask) will have $L = 0$. $L > 0$ indicates that a limit order cuts the current spread and $L < 0$ indicates a limit order is away from the market best quote.[6] In the current research, L is expressed in basis point relative to the middle quote.

[6]Foucault, Kadan and Kandel (2001) use J as price conservativeness measure. J is defined as the distance between the limit price and market price. For buy order, $J = q^{ask} - p^l$. For sell order, $J = p^l - q^{bid}$. Foucault et al. (2001) claim that the more conservative the limit order the longer will be the execution time. However, it should be noted that, the two measures, L and J, are equivalent for a given bid-ask spread because $L + J = Spread$.

It should be noted that L is a relative measure of price aggressiveness. For a given bid-ask spread, the larger is L the more aggressive is the limit order. But if spread is allowed to vary, as is the case when comparing the aggressiveness of orders submitted at different time when spread is unequal, this might not be true. The reason can be clearly seen in the Figure 4.1. In Panel A, spread for limit order (a) and (b) are the same. We say order (a) is more aggressive than order (b) since L of order (a) is larger. In Panel B, however, even if two orders has the same L-measure, we say order (a) is more aggressive than order (b) because it cut the larger proportion of prevailing spread. Clearly, price aggressiveness measure L is relative to spread. For a given L, a larger spread will be expected to reduce the execution probability of limit order. The above considerations lead to the following two hypotheses:

Hypothesis 1: *More aggressive limit order has larger execution probability for a given spread.*

Hypothesis 2: *For the same aggressiveness, limit order has larger execution probability when market spread is small.*

B. Volatility

Market volatility is a key factor determining liquidity dynamics in theoretical models. In Foucault's (1999) model, high volatility is caused by in-flow of information. In such circumstance, limit order traders will post their quotes more conservatively to compensate for the 'adverse selection' cost. This will make trading with market order more costly and investors will more likely use limit order rather than market order. Since more limit and less market orders are issued, the execution chance of limit order become smaller. The model predicts volatility has a negative relationship with limit order execution probability.

However, Handa and Schwartz (1996) argue that short term volatility is more likely caused by liquidity shock and can be seen as a profitable market condition for limit order trading. Thus volatility should have a positive relationship with limit order execution probability. However, it should also be noted that an increasing volatility will attract more limit orders into the market. Therefore volatility might be beneficial to the limit orders already in the market but might not for the limit orders submitted after observing the high volatilities because more limit order and less market order can reduce the execution probability of limit order.

From the implications of the models by Foucault (1999) and Handa and Schwartz (1996), we have the following null hypothesis regarding the relationship between volatility and limit order execution probability:

Hypothesis 3: *Limit order has larger execution probability when volatility is high, other conditions equal.*

In this study, we examine two volatility measures, *Pre-volatility* and *Post-volatility*. Pre-volatility is measured over 30 minutes immediately before the submission of a limit order and post-volatility is measured over 30 minutes immediately after the submission of a limit order.[7] If the high volatility can benefit the limit order already in the market, post-volatility will have a positive relation to the execution probability of limit order. But if the limit order is attracted to market after observing high pre-volatility, the impact of pre-volatility might be negative.

C. Volume

It is believed that how actively a stock is traded will have impact on the execution chance of limit orders. In his famous paper, Demsetz (1968) argued that " *The greater the frequency of transacting, the lower will be the cost of waiting in trading queue of a specified length.*" (pp.41). The Demsetz's model implies that high

[7] It should be noted that while *Pre-volatility* can be calculated but *Post-volatility* is not available at the time of order submission. We examine both in this study to better understand how volatilities impact on limit order trading environment.

trading volume will be associated with larger execution probability.

However, Foucault et al. (2001) have a different claim. They argue that when trading is sparse, the waiting cost will increase. The increased waiting cost will force the investor to cut the spread more aggressively and lead to a short waiting time. The model implies we might observe a higher execution chance when trading is sparse because investors post their limit orders more aggressively. Clearly two different models have different predictions regarding the relationship between trading volume and limit order execution probabilities. Thus the two model implications can be examined by testing the following null hypothesis,

Hypothesis 4: *Limit order has larger execution probability if trading volume is high.*

D. Order Size

Order size is one of the variables investor can control when submitting a limit order. Intuition suggests that order size should have a negative impact on the order execution probability (specially when completed execution is considered). In their study on the limit orders submitted to SuperDot to NSYSE, Harris and Hasbrouck (1996) find a negative relationship between the order execution probability and the size.[8] The discussion with market participants, however, indicates there is another possibility. When a investor issue a large order, he/she usually acts more strategically by either carefully choosing the market conditions or quoting limit price more aggressively. Such strategic behavior could lead to a higher execution probability for large orders. In this study, the size effect is investigated by testing the following hypothesis,

[8]Harris and Hasbrouck adopted a stock-insensitive classification rule, defining orders with less than 200 shares as tiny orders, orders with more than 200 shares but less than 500 shares as small orders, orders with more than 500 but less than 1000 shares as middle orders and those with more than 1000 shares as large orders.

Hypothesis 5: *Limit order execution probability negatively correlated with order size.*

To account for the characteristics of different stocks, we apply a stock-sensitive division rule to classify orders into different size groups. The limit orders of each stock are divided into four groups (tiny, small, middle and large) conditional on the characteristic of the stock.[9] By this way we can study execution probability of each size group. The incremental effect of order size is also examined in our probit model study in later part of this chapter.

E. Potential Market Pressure

The state of limit order book is an important factor investors take into account when making their order placement decisions.[10] In the market microstructure literature, investors who have inside information always use market order to exploit their information advantage. Thus market order is informative and is modelled to be the driving force of price movements. However, if transaction cost is high information trader might be better off by using limit order. This implies that limit order, like market order, can carry information about asset value and price movements.[11] When a large number of limit orders are built on one side of the order book, their information will predict opposite price movement and the limit orders on the other side will have larger chance of execution.

We use *potential market pressure* (PMP) to characterize the impact of book state on the limit order execution. The potential market pressure is defined as the difference between market depth' of buy side and sell side while accounting for

[9]Normal Market Size (NMS) can be used as size classification criteria. After careful study of each stock, we apply a finer classification criteria to account for the order distribution to avoid too few observations in any groups

[10]For example Parlour (1998) explicitly shows traders 'look at both sides of market' to determine the optimal order strategy.

[11]Unlike market order, limit order hasn't been modelled explicitly to transmit information in theory and the transmission mechanism is still not very clear.

the distance between the orders and the best market quotes;

$$PMP = \ln[\sum_{i}^{k} \frac{Size_i^{buy}}{1 + (q^{bid} - p_i^{buy})/d_i}] - \ln[\sum_{i}^{k} \frac{Size_i^{sell}}{1 + (p_i^{sell} - q^{ask})/d_i}] \tag{4.2}$$

where p_i and $Size_i$ are price and size of order i. d_i is the scaling factor (here d_i is the tick size applied). The PMP can be seen as a pre-cursor of 'realized market pressure' which is proxied by *order flow* in market microstructure literature. The empirical microstructure literature (e.g. Chapter Two of this thesis) has shown order flow can explain the price movements because information is incorporated into the price through order flow. Even though limit orders on the book cannot affect price directly because they are not true order flow, they can form the pressure on the market quotes which will eventually impacts on prices.

A positive *PMP* indicates more market pressure from buy side and will push the market quotes up while a negative *PMP* indicates larger pressure from sell side and can press market quotes down. Since the potential market pressure can affect quote prices it could impact on actual price movements indirectly. If a positive *PMP* forms upward pressure for price movement it will increase the execution probability of sell limit order and decrease the execution probability of buy limit order. This intuition can be formally examined by testing the following hypothesis;

Hypothesis 6: *A positive (negative) PMP is positively correlated with sell (buy) limit order execution probability.*

4.2.2 Empirical modelling

Our aim is to study the relationship between the limit order execution probability and a vector of explaining variables, $\mathbf{x}$. Let

$$Y = \begin{cases} 1 & if\ the\ limit\ order\ is\ executed, \\ 0 & if\ the\ limit\ order\ is\ not\ executed. \end{cases} \tag{4.3}$$

then relationship can be expressed as;

$$Prob\{Y = 1\} = F(\beta'\mathbf{x}) \quad and \quad Prob\{Y = 0\} = 1 - F(\beta'\mathbf{x}) \tag{4.4}$$

Usually function $F(\cdot)$ can be any continuous function subject to $F(\cdot) \in [0,1]$. In the current context, the probability itself is not observable. In stead we only observe the response variable Y which only takes value 0 or 1. This is typical set up for binary choice model. In this study, we choose a common probit model where $F(\cdot)$ takes a form of normal cumulative distribution function, $\Phi(\cdot)$, i.e.,

$$Prob\{Y = 1\} = \Phi(\beta'\mathbf{x}) \tag{4.5}$$

Using probit model (4.5), we can estimate the execution probability of a limit order strategy. Moreover, the simultaneous regression allows us to test the various hypotheses jointly. It is worth noting that the marginal effect of factors is not straightforward (as in linear regression) because probit model is non-linear. The marginal impact of explaining variable varies with $\mathbf{x}$ and is given by following formula

$$\frac{\partial E[y|\mathbf{x}]}{\partial x} = \{\frac{d\Phi(\beta'\mathbf{x})}{d(\beta'\mathbf{x})}\}\beta = \phi(\beta'\mathbf{x})\beta \tag{4.6}$$

4.3 Market and Data Description

4.3.1 Limit order trading in LSE

Prior to 1997, the London Stock Exchange was a pure dealership system. The major problems of the trading system are its pre- and post-trade opacity and over concentration of order flow.[12] To overcome these disadvantages, a limit order trading system, SETS, was introduced in October 1997. Since then, London stock market has become a hybrid market with a centralized limit order book system working in parallel with an off-book dealership market.[13] There are no forced interactions between either prices or quantities of book and off-book trades. The trades occurred in one system are not constrained in any way by market status of the other. Investors can choose to trade on any or both of the systems. Since the purpose of this paper is to examine the limit order execution condition we will focus our discussion more on the order book system.

The SETS system is becoming more and more important and is described as the 'main price formation mechanism' by the exchange. Unlike the dealership market, which is featured with low transparency,[14] SETS is among the most transparent of all limit order books available in major equity markets because the complete book state is visible to the member firms and transactions are publicized immediately to all participants. This feature is valuable to current study because investors trading in the book system can form their order placement strategies conditional on the rich information on book state and the transaction history easily.

[12] Roughly five major dealer firms handled more than two-thirds of order flow in actively traded stocks. See Friederich and Payne (2001) for detailed discussion.)

[13] In the dealership market, the market making is not a binding obligation. Dealers provide market liquidity voluntarily

[14] There is no pre-trade transparency as quotes are requested/provided on a purely bilateral basis but post-trade transparency is high because dealers are requested to report the trade within 3 minutes of its occurrence. In fact, as more and more dealers are adopting the electronic trading/recording system, the trades are usually reported immediately after it occurs.

The order book is a self execution system and no market maker is involved. There are two major types of order investors can submit to the book: market order and limit order.[15] Liquidity is provided by limit orders. Trade occurs when two limit order prices cross or a market order is submitted to trade against the existing limit orders. There is no pre-specified price for market order and it is executed against the most competitive existing limit orders. Limit order must be submitted with a pre-specified limit price and after submission it will sit in the order book and wait for execution against incoming orders.[16] The price priority and time priority are strictly maintained in SETS.

The book trading system in LSE starts with an opening auction from 07:50 to a random time between 08:00:00 and 08:00:30.[17] During the auction time, member firms are permitted to submit or delete limit orders and market orders. At the end of the auction period, the order book is frozen temporarily and an auction is called. The limit order and market will be executed at the auction price subject to price and time priorities. The remaining limit orders will sit on the book waiting for later execution. The price from the opening auction is the open price for the day. Once the auction matching process for a particular stock is complete, the continuous trading in that stock can begin.

The continuous trading is the main part of trading process in SETS and it lasts eight and half hours. During this period, investors submit market orders and limit orders to the book and transactions occur once a price cross is generated. Once order is submitted, it cannot be modified but can be deleted and re-submitted.

[15]In total there are five types of order in SETS system: market, limit, Execution and Eliminate, At Best and Fill or Kill order. The last three types of order only consist a tiny proportion of total orders and share very similar feature as market order in regard to the immediacy of execution. The minimum order size is 1 share and no limit on maximum order size.

[16]A limit order can be executed immediately if it generates price cross, i.e. the limit price is at least as good as the opposite best quote. The tick size is 0.25p, 0.50p and 1.00p for stocks with price below 500p, between 500p and 1000p and over 1000p respectively.

[17]The random time is due to the computer processing time for different stocks

The re-submitted order is treated as a new order. At 16:30, a closing auction period begins and it ends at a random time between 16:35:00 and 16:35:30. After the auction matching process, the limit trading system is officially closed.[18] The closing price for a stock is either the auction price or, in the event of no transactions from auction matching, the Volume Weighted Average Price ("VWAP") of all transactions from 16:20 to 16:30.

4.3.2 Data description

The data used in this study is tick data generated by SETS book system covering FT30 stocks. The sample covers randomly selected five working days in June.[19] For the purpose of this study, we exclude orders submitted and executed during the auction periods. The daily trading volume (number of shares), number of trades and names of the FT30 stocks are presented in Table 4.1. Clearly either by daily number of shares or by daily number trades, there is wide variation of how actively a stock is traded. In this research, we use daily volume as a proxy for liquidity of each stock to study volume impact on limit order execution.

Table 4.2 provide a sample of the original data used in this research. A major advantage of this data set is that it provides a complete image of whole book history. From this data, we can recover following information of each order entered into the limit order trading system:

- time of order entry
- whole book state at the time of order submission
- order type, direction, price (for limit order), and order size

[18] The dealership market has no formal closing time and trading activities could continue there.

[19] The five days are the 13th, 19th, 21th, 22th and 25th of June 2001 and cover all week days. We also experiment with another randomly chosen five days for robust check:1th, 7th, 13th, 19th and 25th. The results are similar to what reported here.

- the time of each subsequent event happened on the order, including (partial) execution, deletion or expiration
- prices and quantities of each transaction relevant to the order

Clearly the above information is critical for the current study because we can examine the market conditions at the time an order is submitted and track the life path of the order.

If a limit order is submitted with a price better than the opposite quote, it will generate a price cross and be executed immediately.[20] Because such orders behave like a market order, they are classified as market order in current study. The characteristics of different order type of each stock are summarized in Table 4.3. In general, the number of buy orders are slightly larger than the number of sell orders. Among the total orders, roughly 70 percent are limit orders. The proportion of limit buy order and limit sell order are generally balanced. The average execution probability of limit order is less than 50 percent.

4.3.3 Intraday pattern

Figure 4.2 plots the intraday pattern of order flow and limit order behavior of the SETS system. The upper panel is intraday dynamic of the order flow composition. It is clear that at the beginning and lunch time of a trading day, the total number of orders and the proportion of market order are relatively small. As US market starts to trade, more orders are attracted into the market. In the last 30 minutes, the proportion of market order increases a lot.

The lower panel plots the overall limit order execution probabilities and the price aggressiveness (in basis points). The most striking feature is the high correlation between the execution probability and the price aggressiveness of limit

[20] Such limit orders are called marketable limit order in the literature.

orders. Another interesting point is that investors' behavior has a clear intraday pattern. At the beginning of a trading day, investors are generally more conservative (they post their limit orders more conservatively by choosing a smaller L) because information uncertainty in the market could be high after overnight. As trading continues, more information is incorporated into the market and the investors become more confident. When market becomes more informative investors post their limit orders less conservatively and remain relatively stable for the most of the time. However, in the last 30 minutes, market becomes very aggressive, investors post the limit prices aggressively and the proportion of market order increases.

4.4 The Impact of Factors on Limit Order Execution Probability

As we have seen from the intraday pattern figure, investors' behaviour in the first and the last 30 minutes is quite different from other time. In this study, we exclude the new submissions of these two sub-periods from our analysis in order to focus on the continuous trading part and mitigate the impacts of the opening and closing auctions.

4.4.1 Volume Effect

To study the volume effect, we split the FT30 stocks equally into five groups according to how actively a stock is traded. Then we compare the execution probability of limit orders in the low, middle and high volume groups.[21] Since the evidence from intraday pattern study suggests that the execution probability of

[21] The low, middle and high volume groups are group 1, 3 and 5 respectively.

limit order could be highly correlated with the price aggressiveness, we also calculate the price aggressiveness (L-measure) of limit orders in different volume groups to account for this factor.

The main results are presented in Table 4.4. The first row of the table indicates there is a positive relationship between the limit order execution probability and how actively a stock is traded. The limit order execution probability of low volume group is smaller than that of the middle volume group which in turn is smaller than that of the largest volume group. The mean difference test indicates that such relationship is statistically significant[22]. The execution probability of limit orders of the least liquid and the most liquid stock groups are plotted in the upper panel of Figure 4.3.

The second row is the average price aggressiveness of limit orders of each volume groups. Surprisingly, we find a negative relationship between the stock volume and the limit order price aggressiveness. Limit orders trading on the most liquid stocks are the most conservative ones but such orders still enjoy a reasonable execution probability. This result suggests that trading on less liquid securities, investors have to quote their prices more aggressively to get a reasonable execution chance. The price aggressiveness (in basis points) of limit orders on the least liquid and most liquid stock groups are plotted in the lower panel of Figure 4.3.

In general, the evidence regarding the relationship between limit order execution probability and the stock volume is consistent with the prediction of Demsetz (1968). However, our results on price aggressiveness and volume do support the claim of Foucault et al. (2001) that investors trading on less liquid stocks will behave more aggressively.

[22] We use the paired t-test along the every 15 minutes interval to have smaller standard deviation.

4.4.2 Price aggressiveness and Spread

To study the impact of price aggressiveness and spread, we first calculate the price aggressiveness, L, of each limit order (by formula (4.1)) and the spread, S, at the time immediately before the order is submitted. Then for each stock within each day, we split the limit orders into five groups according to its L-measure. We group orders by their aggressiveness within each day because we believe that price aggressiveness is more relevant within a trading day than across days. For the similar reason, the limit orders are also split into five groups according to the spread at the time of submission[23]. After the grouping, each limit order is tagged with two labels L_i (price aggressiveness group i) and S_j (spread group j). Finally, we calculate the execution probabilities of each L-groups and S-groups for whole sample days. Table 4.5 reports the major results.[24]

The upper panel is the limit order execution probabilities of different price aggressiveness groups. The limit orders labelled with L_1 and L_5 are those most conservative and the most aggressive order groups respectively. The first row is the overall execution probabilities of the different L-groups. Not surprisingly, there is a monotonic relationship of the execution probability and the price aggressiveness. The more aggressive (when L is larger) , the larger the execution probability. This relationship is quite persistent when volume is controlled as shown in the next 5 rows of Panel A.

The lower panel is the limit order execution probabilities of different market spread conditions (S-groups). The first row is the overall unconditional execution probabilities of the limit orders under different spread conditions. Even though we find that the execution probability declines when spreads become very large (group S_4 and S_5), the general relationship between limit order execution proba-

[23] Spread grouping is independent of price aggressiveness grouping.

[24] We also studied price aggressiveness and spread impact on each of the 30 stocks. Since the patterns are similar to those reported here we skip them to save space.

bility and spread is not clear because investors can increase their order execution chance by cutting spread aggressively when it is large. To disentangle these two factors, we examine those orders that are submitted with the same price as the best market quote (which is labelled with L_3). When the price aggressiveness is controlled, a clear pattern emerges: the execution probability declines with spread monotonically. This result is consistent with the model prediction of Foucault, Kadan and Kandel (2001) that the distance between limit price and opposite market quote is negatively correlated with order execution probability.

4.4.3 Size effect

Following Harris and Hasbrouck (1996), we split the limit orders of each stock into four groups (tiny, small, middle and large) according to their order size. But we apply a stock-sensitive division criteria because order size is stock relevant in our sample. The limit order execution probability of each size group is calculated and reported in the first row of Table 4.6. Contrary to the intuition, we find a U-shape rather than a downward slope relationship between the execution probability and the order size: the middle size group has the smallest execution probability.

Many reasons could lead to this U-shape pattern. An obvious candidate could be that large orders tend to be more aggressive. To check this hypothesis, we re-calculate the execution probability of each size group conditional on the price aggressiveness, L_i. The result is reported in the lower panel of Table 4.6. Comparing with upper panel, we find that the U-shape is more flattened in many cases in the lower panel (for L_3, L_4 and L_5 groups) than in upper panel. This implies that price aggressiveness is part of the reason but not all because the U-shape is persistent when price aggressiveness is controlled.

Another potential explanation could be that investors with larger orders tend to act more strategically. When spreads are large, they can cut the spreads aggressively and deviate themselves away from the rest of limit orders in the book. To test this hypothesis, we calculate the price aggressiveness of each size groups conditional on spread. The results are reported in Table 4.7 and they confirms our conjecture. In all cases, L value of large orders is larger than that of middle orders. Specially when spreads are large. For example, in S_4 and S_5 spread groups, large orders have larger L and are much more aggressive than middle orders (also more aggressive than small orders).

In general, we find a U-shape relationship between the execution probability and the limit order size in our sample. This result is significantly different from that documented by Harris and Hasbrouck (1996) who found a negative correlation between the limit order size and the its execution probability. Our results indicate that the U-shape relationship found in this chapter is at least partially attributed to the strategic behaviour of the investors trading with larger orders.

4.4.4 Potential market pressure

In this study, we use formula (4.2) to calculate *PMP* for each limit order. We recovered from our data set the whole book state at any moment before a new order enters the book system and calculated the *PMP* for every limit order at the moment immediately before it is submitted. Like price aggressiveness, *PMP* is relevant more relevant within a trading day than acrose different trading days. We split the limit orders into three groups (Low, Middle and High) according to their associated *PMP* within each trading day for each stock. If, as argued in previous sections, limit order can also convey information, a positive *PMP*, which represents a market pressure from the buy side, will predict a upwards price movement and have a negative (positive) impact on the execution probabilities of buy (sell) orders. Similarly a negative *PMP*, which represents a market pressure from the sell side, will predict a downwards price movement and have a positive (negative) impact on the execution probabilities of buy (sell) orders. To test this hypothesis, we evaluate the effects of *PMP* on the limit buy and sell orders separately.

The average execution probability of all limit orders of each *PMP* group is calculated and reported in Table 4.8. The first row presents the unconditional execution probabilities of each *PMP* group. Clearly we find that *PMP* has impact on limit order execution probability. For buy orders, the larger the *PMP*, the smaller is the execution probability. For sell orders, the larger the *PMP*, the larger the execution probability. The mean difference test indicates that the potential market pressure derived from book state has a significant impact on the limit order execution probability. As it is shown in the next five rows, the pattern is consistent after control of volume groups.

A significant and persistent effect of *PMP* on the limit order execution is consistent with the hypothesis that limit order, like order flow, also conveys information. It should be noted, however, that information contained in *PMP* is information that hasn't been but is about to be incorporated into the prices. Thus, unlike *order flow*, which contemporaneously correlates with price, *PMP* can be used to predict future price changes. Ours results also confirm the theoretical argument that the book state has important implications for investors to form their order submission strategies (e.g. see Parlour (1998)).

4.4.5 Volatility

To study the effect of volatility on the limit order execution probability, we calculated two types of volatility: *pre-volatility* and *post-volatility*. The former is measured within 30 minutes immediately before an order is submitted. The latter is measured within 30 minutes immediately after an order is submitted. For each stock within each trading day, we split the limit orders into three groups (Low, Middle and High) according to their pre-volatility (and post-volatility).[25]

The execution probability of each volatility group is calculated and reported in Panel A (for post-volatility) and Panel B (for the pre-volatility) of Table 4.9. For the post-volatility, we find the evidence that volatility has a positive impact on the limit order execution probabilities (either volume is controlled or not). High volatilities do increase the execution chance of limit orders. This result rejects with the model prediction of Foucault (1999) and support the supports the hypothesis of Handa and Schwarts (1996) that volatility could be caused by liquidity reasons and is a favorable trading condition for limit orders.

When volatility is measured before the limit order submission (Panel B), its impact is much smaller and is insignificant in most cases. This result is consistent with the implication of Handa and Schwarts (1996) model and suggests that there is some equilibrium dynamics of order submission. However, it should be noted that even though volatility has been a major determinant of limit order execution probability in both models, the evidence found in this study suggests that the impact volatility is weak and economically insignificant, especially when it is compared with price aggressiveness, spreads and *PMP*.

4.5 Estimation of execution probability

4.5.1 Estimating the execution probability

In this section, we model the limit order execution probability with probit model (4.5) and evaluate the impacts of the various factors in the last section simultaneously. With the aim to evaluate the effects of these factors on the limit order submissions, we choose to measure all factors *ex ante* rather than *ex post*. More

[25]We experimented with volatilities generated by both market-wide trading and limit order trading and we found that the volatility generated from the limit order system has a stronger impact on the limit order execution probability.

specifically, we estimate model (4.5) with the following specification,

$$\mathbf{x}_i'\beta = \alpha + \beta_1 Spread_i + \beta_2 L_i + \beta_3 PMP_i + \beta_4 Volatility_i + \beta_5 tmID_i + \beta_6 Size_i \quad (4.7)$$

where i is the order index. $tmID_i = 2, 3, ..., 16$ is time zone within which order i is submitted,[26] $Size_i$ is the order size, $Spread_i$ is the percentage spread of middle quote (in basis points). L_i is the price aggressiveness defined by (4.1), PMP_i is the potential market pressure defined by (4.2). $Spread_i$, L_i and PMP_i are all measured at the time immediately before a limit order is submitted. $Volatility_i$ is the return variance of measured over the 30 minutes immediately before an order is submitted. Since our focus in this part is to estimate the model for each stock, volume is not included in the specification because it is not characteristics of an order but a stock.

As we have discussed in previous sections, potential market pressure has an opposite effect on buy and sell orders. We estimate the probit model for buy and sell orders separately. From our factor analysis, we would expect that β_1 is negative, β_2 and β_4 are positive for both buy and sell orders and β_3 is negative for buy orders and positive for sell orders. We use variable $tmID_i$ to test whether time of the day is an important factor because of the time priority rule. If the length of the remaining time is important, we would expect a negative β_5. Theoretically speaking, size itself should have non-positive effect, but if the large orders behave strategically, we might not be able to observe the negative coefficient of β_6 in this setup.

Using maximum likelihood method we estimate the probit model (4.7) for buy and sell orders for each stock and present the main results in Table 4.10 for buy

[26]The whole trading time from 8:00 to 16:30 is divided into 17 time zones with each zone has 30 minutes. Since the first and the last 30 minutes are excluded, *tmID* varies from 2 to 16. For example, if an order is submitted between 12:30 and 13:00, tmID=10.

orders and Table 4.11 for sell orders. As seen in formula (4.6), the marginal effect of explaining variables is $\phi(\beta'\mathbf{x})\beta$, the value of $\phi(\cdot)$ for 'an average observation' is calculated and reported as $\phi(\bar{\mathbf{x}})$ in the first column in both tables. The last columns of Table 4.10 and 4.11 are Likelihood Ratio Indices. The Likelihood Ratio (LR) test (column under LR-stat) highly rejected the null hypothesis that all coefficients except intercept are zero.

The most striking feature of Table 4.10 and 4.11 is the persistence and significance of the three factors we identified from our previous analysis: price aggressiveness, spread and and *PMP*. Under this setup where all factors are evaluated simultaneously, we find that spread has a significant negative impact on the limit order execution probability after accounting for the factor of price aggressiveness. The impact of the *PMP* is also of paramount. A negative (positive) β_3 for buy (sell) order implies that it is much harder to have a limit buy (sell) order executed when there is large market pressure from from the buy (sell) side, holding other market conditions equal.

For most of the cases, volatility has an positive impact on the limit order execution probability. However, compared to price aggressiveness, spreads and *PMP*, volatility has a much less significant effect. Only less than half of the cases have positive significant regression coefficients. It seems that time factor has stronger impact on sell orders than on buy orders. However, for both types of order more than half the cases either have a 'wrong' sign or are insignificant. Similarly, there are only a few cases where order size has a negative significant regression coefficient. These results implies that neither time of submission nor order size has an important impact on the limit order execution probability in our sample.

4.5.2 Predict the execution probability: An experiment

Using the probit model, we can evaluate the execution probability of different order submission strategies. To see the implication of the model, we randomly choose four stocks (BAE system, Barclays, HSBC Holdings and Shell) and evaluate the execution probability of various submission strategies in the following example. Consider the two orders:

*Order*1	*Order*2
$L = -5$ *basis points*	$L = -15$ *basis points*
average order size	*average order size*
$Spread = 15$ *basis points*	$Spread = 15$ *basis points*
$PMP \in [-1.5,\ 1.5]$	$PMP \in [-1.5,\ 1.5]$
Buy, *or Sell*	*Buy*, *or Sell*

When the *PMP* changes, the above two orders form two set of infinite number of submission strategies. The estimated execution probabilities are plotted in Figure 4.4. Note: in all graphs Order 1 is plotted with dotted line and Order 2 is plotted with solid line.

In all graphs we find that Order 1 has large execution probability than order 2 (because it is more aggressive). The execution chance of the buy orders decreases with *PMP* while that of that sell orders increases with *PMP*. For example, in the case of BARC buy order in Figure 4.4, when *PMP* varies from -1.5 to 1.5, the execution probability of order 1 decreases from 30% to less than 15%. For a given market condition, 10 basis points of return improvement (moving from order 1 to order 2) can reduce the execution probability dramatically. The decrease in execution probability, however, is not equal under different market conditions. When *PMP* is large, the decrease is less severe than when *PMP* is small. For example, when $PMP = 1.5$, the difference of execution probability between order

1 and order 2 is only 10%, but it will increase to 20% when *PMP* declines to -1.5. This result implies that the 'effort' (move from order 2 to order 1) is most valued when liquidity from the same side is most wanted.

4.6 Discussion and Conclusion

Execution probability of limit order is a critical determinant of order submission strategy which in turn has important impacts on price discovery process. In classic market microstructure literature, informed traders are uniformly modelled to use market order and demand liquidity while 'noise traders' provide liquidity and prevent market from breaking down. In this setup, market order, which is believed to carry information, is the driver of price discovery process. However, as argued in section 1.4, informed trader can be perfect rational to use limit order in reality where transaction cost is not negligible. In equilibrium, the expected profit of using different types of order should be equal, i.e. $\pi_m = \pi_l$ in equation (1.2). Clearly the expected profit of using a limit order (π_l) depends on the probability of execution of the order (Π), which in turn is a function of other microstructure factors.

Theory and intuition suggest that a number of factors can affect the limit order execution. This study empirically examines the impacts of these factors on the execution probability of limit orders. The factors investigated in this research include *Price Aggressiveness, Spread, Volatility, Volume, Potential Market Pressure (PMP)* and *Order Size.*

Using a sample of limit order segment of the London Stock Exchange, we find that *price aggressiveness, bid-ask spread* and *potential market pressure (PMP)* have the strongest impacts on the limit order execution probability for a given stock. Trading on a less liquid stock, investors have to post their orders more aggressively than trading on a more liquid stock. Volatility can increase limit order execution probability but its impact is weak. Contrary to the intuition, size itself doesn't have significant effect on limit order execution. In fact, we find a U-shape pattern of size effect in our sample. Such U-shape relationship between size and execution probability is at least partially attributed to the strategic behaviour of investors with large orders.

We evaluate the execution probability of different order submission strategies with a probit model which is capable of capturing the features regarding the relationship between limit order execution probability and its major determinants. The model provides a potential ground for future studies on the comparison of profitability of different order submission strategies.

The results from this study have important implications for investors to implement efficient trading strategies, especially for institutional investors who more and more depend on automatic trading algorithm. From theoretical microstructure perspective, this study provides clear evidence that a number of factors can impact on limit order execution probability and, therefore, can impact on order choice, liquidity dynamics and price formation process. For example, *PMP* effect indicates that limit order book state has predictive power for the direction of future price movements and can be viewed as an in-direct evidence for the hypothesis (which has been raised in the introduction section 1.4) that limit order, like market order, can convey information. Modelling the information incorporation mechanism and price formation process with limit order trading is outside the research agenda of this study, but the evidence presented in this chapter does suggest that future research in this direction, that we haven't seen so far, can enrich the market microstructure literature.

Table 4.1: Characteristics of FT30 Stocks

The table gives basic information on the components of FT30 index. *NMS* is 'Normal Market Size' is a measure of the average size (in number of shares) of an institutional trade in the stock concerned calculated by the London Stock Exchange. The last three columns are based on a sample of randomly chosen five working days in June 2001.,

Mnem	Name	NMS	Daily Volume(shrs)	Daily Trades
AL.	Alliance&Leicester	50000	746104	300
ANL	Abbey National	100000	1678846	533
ARM	Arm Holdings	100000	5999286	730
AZN	AstraZeneca	75000	2298534	852
BA.	BAE Systems	200000	4409697	471
BARC	Barclays	100000	1765250	919
BG.	BG Group	150000	7967388	522
BP.	BP PLC	200000	18933259	1208
BT.A	BT Group	200000	22980423	1299
CGNU	CGNU PLC (AVIVA)	100000	1919170	740
CTM	ColtTelecom	150000	2984904	715
CW.	Cable & Wireless	200000	8900083	912
DGE	Diageo	200000	4103698	740
GSK	GlaxoSmithKline	100000	7523137	1396
HFX	Halifax	75000	1509498	502
HSBA	HSBCHoldings	200000	6731891	901
ISYS	Invensys	200000	5686549	445
KGF	Kingfisher	100000	3242541	496
LLOY	Lloyds-TSB	200000	4817047	866
LOG	Logica	100000	2159368	745
MKS	Marks & Spencer	200000	4017295	447
PRU	Prudential	150000	3004526	678
PSON	Pearson	75000	1838736	614
RBOS	Royal Bank of Scot.	100000	2315341	800
RTR	Reuters Group	150000	2770629	870
SHEL	Shell Transport (Reg)	200000	17590314	1008
SPW	Scottish Power	100000	3083384	562
TSCO	Tesco	200000	9505741	722
ULVR	Unilever	200000	6346998	778
VOD	Vodafone Group	200000	123248459	1973

Table 4.2: Sample of SETs Data (AL. 1 June 2001)

The codes under column *Event* represent different types of event. N: new order entry, D: deletion of order, P: partial execution, M: complete execution. *Order Code* is the unique id of each order, *Time* is the event time stamp. 'LO', 'AA' in column *Type* represent limit order and market order respectively. 'S' and 'B' in column *Dir* represent sell and buy order respectively. The number under *Price* are limit order price or transaction price, depending on the Event code. *TSize* and *Size* are the transaction size and remaining size of the order. *TransCode* is the id of each transaction. *Counterpart* is the counterpart order code of the current transaction.

Event	Order Code	Time	Type	Dir	Price	TSize	Size	TransCode	Counterpart
N	901B54AI01	09:14:12	LO	S	761	0	13100		
N	601C6VER01	09:15:15	LO	B	757.5	0	1500		
N	A018KLFQ01	09:16:39	LO	S	759	0	4370		
N	501LRB0B01	09:17:15	LO	B	757	0	1500		
N	501LSANS01	09:27:25	AA	B	0	0	5000		
P	501LSANS01	09:27:25	AA	B	0	1081	3919	A018KLMI01	501LSANT01
D	301IOG3J01	09:47:48	LO	B	758.5	3480	0		
N	A018JJ3I01	09:51:25	LO	B	758.5	0	3480		
D	A018KO6R01	10:01:33	LO	B	763	1000	0		
N	101N37E801	10:05:16	LO	S	763.5	0	1000		
M	A018KOEX01	10:07:32	LO	B	764	13100	0	301IOGEL01	A018KOEY01
N	901B55B401	10:10:20	LO	S	766.5	0	1000		
N	601C6XQQ01	10:16:49	LO	B	764.5	0	8916		
N	701AEOOA01	10:16:57	LO	B	764.5	0	780		
D	601C6YM301	10:20:16	LO	S	769	7000	0		
M	C016Q5QZ01	10:20:19	LO	B	765	2000	0	201HJBOH01	201HJBOI01
N	B017AAHT01	10:24:16	LO	S	767	0	3000		
M	501LSEUL01	10:31:47	LO	S	767	3940	0	701AER2001	701AER2101
N	101N37VF01	10:31:48	LO	S	770	0	1970		
N	101N393T01	10:44:19	LO	B	765	0	1500		
N	201HJECH01	10:51:39	LO	S	767.5	0	1610		
N	501LSGHD01	10:53:43	LO	S	765	0	3915		
N	C016Q6J701	10:58:59	LO	B	766	0	6000		
N	701AESJR01	10:59:18	LO	S	769.5	0	1640		
M	F018XBV101	11:18:18	LO	B	768	387	0	E0194BJ501	F018XBV201
M	C016Q80A01	11:24:47	LO	B	768.5	621	0	501LSGKA01	C016Q80B01
P	301ION7U01	11:39:14	AA	B	0	15000	1000	601C6YI901	301ION7Z01
N	501LSKEX01	11:39:39	LO	S	770.5	0	2510		

Table 4.3: Summary statistics of limit orders of FT30 stocks

Total is the total number of limit and market (including marketable limit) orders. *Buy* and *Limit* are the proportion of buy and limit to total orders respectively. *Limit Buy* is the proportion of buy limit orders to total limit orders. *Exe.* is the execution probabilities of limit orders. The last two columns are the execution probabilities of limit buy and sell orders respectively. All statistics are daily averages computed based on a sample of randomly chosen five working days in June 2001.

Mnem	Total	Buy (%)	Limit (%)	Limit Buy (%)	Exe.	Exe. (Buy)	Exe. (sell)
AL.	4282	51.03	70.9	53.36	22.13	20.43	24.08
ANL	7293	55.61	73.12	58.45	27.62	28.04	27.03
ARM	7683	54	63.89	53.23	41.92	38.19	46.17
AZN	14740	52.36	78.73	51.71	21.98	23.51	20.34
BA.	5156	53.53	66.19	56.08	38.85	34.8	44.03
BARC	11579	59.06	71.97	65.32	30.58	28.86	33.81
BG.	5458	47.6	64.18	49.5	40.94	41.06	40.81
BP.	11133	54.24	63.45	56.17	52.32	52.72	51.81
BT.A	13659	51.89	66.59	54.9	40.27	43.53	36.3
CGNU	8471	54.5	66.62	58.67	34.34	34.22	34.52
CTM	8801	54.27	69.26	55.84	33.27	33.9	32.47
CW.	10053	54.92	66.38	54.47	36.63	37.8	35.22
DGE	9733	45.64	72.05	42.64	27.18	34.88	21.45
GSK	15091	52.02	66.73	55.63	39.28	40.95	37.18
HFX	5835	54.81	65.36	63.29	30.34	32.93	25.86
HSBA	9536	53.69	66.17	53.91	38.81	38.89	38.72
ISYS	4054	52.76	56.27	52.83	41.43	39.25	43.87
KGF	5350	49.44	64.79	46.83	38.29	36.29	40.04
LLOY	10243	53.31	65.97	54.25	29.91	29.43	30.48
LOG	8919	45.73	68.72	42.68	38.64	34.14	41.99
MKS	4778	54.02	65.24	56.4	36.8	37.26	36.2
PRU	7113	55.27	65.16	57.3	39.55	39.57	39.51
PSON	5981	54.72	58.87	53.54	41.89	43.93	39.55
RBOS	10091	58.25	71.19	58.74	32.41	28.72	37.65
RTR	10147	50.73	67.14	49.71	33.14	32.33	33.95
SHEL	22197	57.5	83.45	58.59	15.66	16.19	14.9
SPW	6056	56.61	63.72	60.51	36.9	37.17	36.48
TSCO	6723	54.48	61.49	57.5	43.57	43.25	44
ULVR	14095	36.13	79.83	31.97	17.06	30.8	10.61
VOD	15965	58.37	58.92	59.96	54.48	56.42	51.57

Table 4.4: Volume and Execution Probability

The FT30 stocks are equally divided into five groups (each group has 6 stocks) by their daily average trading volume. The first and the second rows under *Low, Middle and High* are the average execution probabilities and the price aggressiveness (defined by formula (4.1)) of limit orders of each volume groups respectively. *LMM* (Low minus Middle) and *MMH* (Middle minus High) are t-stat of the paired mean difference tests (paired by every fifteen minutes interval from 8:30 to 16:00). ^{a},^{b},^{c} indicate 10%, 5% and 1% significance levels.

	Low	Middle	High	LMM	MMH
Execution Probability	31.46	33.17	35.23	-1.72^{c}	-2.05^{b}
Price Aggressiveness (L-measure)	-4.97	-11.93	-21.2	6.97^{a}	9.27^{a}

Table 4.5: Price aggressiveness and bid-ask spread

The price aggressiveness is defined by formula (4.1). The orders of each stock are divided into five groups by their aggressiveness within each trading day (L_1 to L_5 with L_1 is the least aggressive group. Note that orders in L_3 group are defined to have $L = 0$.). The execution probabilities of each group (across stocks and trading days) are computed and reported in the first row of Panel A. The rest 5 rows of Panel A report the execution probabilities of each group controlled by trading volume group. In Panel B, the limit orders are grouped by their spreads (measured at the time of order entry) in a similar way and the execution probabilities are calculated for each spread group (S_1 represents the smallest spread group). The first and second rows of Panel B reports the execution probabilities of each spread group for all orders and orders in L_3 group respectively.

Panel A. Price Aggressiveness and Exe.Prob.

	L_1	L_2	L_3	L_4	L_5
All Orders	8	13	43	56	57
VolGrp1	7	12	39	48	51
VolGrp2	8	12	36	51	53
VolGrp3	7	11	41	57	59
VolGrp4	9	13	44	60	59
VolGrp5	10	15	50	63	66

Panel B. Spread and Exe.Prob.

	S_1	S_2	S_3	S_4	S_5
All Orders	36.66	37.85	37.37	35.46	31.19
L_3	54.41	48.97	42.28	33.76	24.85

Table 4.6: Execution Probability of Different Size Groups

Tiny orders on average have 200 shares or less. The remaining orders of each stock are equally divided into three groups: small, middle and large. The limit order execution probabilities of each size group are computed (across stocks) and are reported in the first row. The rest rows report the probabilities of each size group controlled for price aggressiveness of orders. L_1 and L_5 represent the least and the most aggressive groups respectively.

	Tiny	**Small**	**Middle**	**Large**
ALL	46.80	35.52	29.99	42.95
L_1	11.10	7.60	6.61	14.20
L_2	21.17	11.64	10.24	22.18
L_3	50.46	42.36	39.57	47.93
L_4	64.28	56.26	52.90	56.52
L_5	65.70	56.03	56.22	58.63

Table 4.7: Price Aggressiveness of Different size Groups

For each stock, its limit orders are ranked in in two dimensions simultaneously: size and the spreads. Orders are divided into tiny, small, middle and large groups by their size. Orders within each trading day are equally re-divided into five groups by their market spreads condition (measured at the time of order entry). In this way, all limit orders of each stock are split by a size-spread matrix. This table reports the price aggressiveness of each size-spread groups averaged across stocks.

Spread	**Tiny**	**Small**	**Middle**	**Large**
S_1	-5.38	-6.87	-11.77	-11.38
S_2	-2.69	-4.82	-9.52	-6.54
S_3	-0.22	-3.15	-6.9	-4.7
S_4	3.06	-1.66	-4.85	-1.58
S_5	5.82	0.95	-0.96	4.05

Table 4.8: PMP impacts and Execution Probability

The PMP (Potential Market Pressure) is first calculated for each limit order. Then the limit orders of each stock are equally divided into Low, Middle and High PMP groups within each trading day. The execution probabilities of limit buy (sell) orders of each PMP group (across stocks and trading days) are computed and reported in the first row of the left (right) panels of the table. *LMH* (Low minus High) is the execution probabilities difference between low and High PMP groups. *t*-stat is based on the mean difference test of Low and High PMP groups. The remaining rows are computed in the same way after controlling for trading volume group.

	BUY Order					**SELL Order**				
	Low	**Middle**	**High**	**LMH**	**t-stat**	**Low**	**Middle**	**High**	**LMH**	**t-stat**
Overall	41.5	36.2	32.4	9.0	22.18	31.4	34.5	37.8	-6.4	-14.54
VolGrp1	38.6	32.6	27.2	11.4	12.04	27.7	32.7	38.0	-10.3	-9.01
VolGrp2	35.7	30.1	28.2	7.5	8.93	29.9	32.6	35.9	-6.0	-6.84
VolGrp3	44.6	35.1	30.1	14.5	13.37	33.2	37.1	39.9	-6.7	-5.69
VolGrp4	44.7	41.2	37.0	7.8	7.89	28.4	31.2	34.4	-5.9	-6.57
VolGrp5	44.1	40.8	37.5	6.7	8.61	36.9	38.9	41.3	-4.3	-4.73

Table 4.9: Volatility and Execution Probability

Market volatility is computed as sum of squared return. The *pre-volatility* and *post-volatility* are measured over 30 minutes immediate before/after the order is submitted. For each stock within each trading day, the limit orders are equally divided into Low, Middle and High groups by their associated pre-volatility (and post-volatility). The execution probabilities of each volatility group are calculated and reported in the first rows of Panel A (post-volatility) and Panel B (pre-volatility). The probabilities in the remaining rows are similarly computed after controlling for trading volume groups. *LMH* (Low minus High) and *t*-stat are the difference of the execution probability between the Low and High volatility groups and the *t*-stats of the mean difference test.

Panel A. Post-volatility

	Low	Middle	High	LMH	t-stat
ALL	35.0	35.7	36.6	-1.6	-5.47
VolGrp1	31.8	33.4	33.2	-1.6	-2.00
VolGrp2	31.1	31.4	33.8	-0.2	-4.34
VolGrp3	35.2	36.6	37.9	-1.3	-3.27
VolGrp4	34.7	36.4	36.7	-1.6	-2.95
VolGrp5	40.4	39.7	40.5	0.6	-0.19

Panel B. Pre-Volatility

	Low	Middle	High	LMH	t-stat
ALL	35.5	36.2	35.7	-0.2	-0.76
VolGrp1	31.9	33.2	33.3	-1.4	-1.93
VolGrp2	31.4	33.2	31.7	-0.3	-0.51
VolGrp3	36.9	36.7	36.0	0.9	1.12
VolGrp4	35.5	35.8	36.5	-1.0	-1.47
VolGrp5	40.3	40.5	39.8	0.5	0.88

Table 4.10: Estimation the Execution Probability for Buy Limit Order

The model estimated is (4.5) where

$$\beta' \mathbf{x} = \alpha + \beta_1 Spread + \beta_2 L + \beta_3 PMP + \beta_4 Volatility + \beta_5 tmID + \beta_6 Size$$

L and *PMP* are price aggressiveness and potential market pressure defined by formula (4.1) and (4.2) and measured at the time of order entry. *Volatility* is computed as the sum of squared returns and measured over 30 minutes immediate before the order is submitted. *tmID* is ordered number from 2 to 16 for time periods from 08:30 to 16:00. The column headed $\phi(\bar{\mathbf{x}}'\beta)$ is scaling coefficients of marginal effect for an 'average order'. *LR*-stats is based the Likelihood Ratio test of the null that all coefficients (except intercept) are zero. *LRI* is the likelihood ratio index defined by Greene [1993] (pp.891).

Mnem	$\phi(\bar{\mathbf{x}}'\beta)$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	*LR*-stats	*LRI*
AL.	0.243	-0.01^a	0.04^a	-0.18^a	-0.17^b	-0.04^a	1.74^b	183	0.136
ANL	0.345	-0.02^a	0.04^a	-0.22^a	0.10^c	0.01^c	1.06^c	432	0.201
ARM	0.385	-0.01^a	0.02^a	-0.01	0.08^b	-0.01	1.82^a	246	0.113
AZN	0.230	-0.03^a	0.07^a	-0.14^a	0.18^a	-0.01^a	-0.44	1123	0.226
BA.	0.328	-0.03^a	0.07^a	-0.28^a	0.36^a	0.00	-0.08	389	0.238
BARC	0.308	-0.04^a	0.08^a	-0.17^a	0.30^a	0.01^b	-2.60^a	887	0.190
BG.	0.388	-0.01^a	0.05^a	-0.14^a	-0.19^a	0.02^a	-0.29^b	202	0.141
BP.	0.398	-0.01^a	0.04^a	-0.22^a	0.40^a	0.01^a	0.11	336	0.100
BT.A	0.394	-0.02^a	0.04^a	-0.10^a	0.21^a	0.02^a	0.00	505	0.120
CGNU	0.358	-0.02^a	0.04^a	-0.24^a	0.03	0.00	1.15^b	374	0.126
CTM	0.372	0.00^a	0.01^a	-0.02	0.01^a	-0.03^a	1.24^a	104	0.037
CW.	0.380	-0.01^a	0.02^a	-0.10^a	0.07^a	-0.01^c	0.29^b	258	0.086
DGE	0.383	-0.03^a	0.06^a	-0.17^a	0.45^a	0.00	0.56^b	321	0.141
GSK	0.386	-0.02^a	0.06^a	-0.16^a	0.37^a	0.01^a	-0.53^a	589	0.121
HFX	0.345	-0.02^a	0.05^a	-0.18^a	0.12	0.00	-1.37^b	345	0.168
HSBA	0.382	-0.03^a	0.06^a	-0.15^a	0.14	0.02^a	0.16	425	0.153
ISYS	0.382	-0.01^a	0.03^a	-0.33^a	-0.13^a	0.01	0.50^a	201	0.200
KGF	0.376	-0.01^a	0.03^a	-0.18^a	-0.03	-0.01	1.10^a	152	0.113
LLOY	0.293	-0.02^a	0.07^a	-0.24^a	0.11	-0.01^a	0.59^b	738	0.240
LOG	0.360	-0.01^a	0.03^a	-0.19^a	-0.01	0.00	1.15^b	246	0.115
MKS	0.381	-0.02^a	0.04^a	-0.31^a	0.07^b	0.01^a	0.16	215	0.142
PRU	0.379	-0.02^a	0.06^a	-0.19^a	-0.08	0.02^a	-0.07	416	0.180
PSON	0.391	-0.02^a	0.03^a	-0.19^a	0.10	0.01	1.10^a	236	0.145
RBOS	0.318	-0.0^{a}2	0.06^a	-0.10^a	-0.16^a	0.00	-0.57	553	0.151
RTR	0.367	-0.01^a	0.03^a	-0.09^a	0.10^b	-0.01^a	1.08^a	215	0.074
SHEL	0.066	-0.03^a	0.07^a	-0.17^a	-0.21^b	0.02^a	0.19^a	2943	0.409
SPW	0.363	-0.02^a	0.04^a	-0.21^a	-0.13^a	0.02^a	0.91^a	398	0.199
TSCO	0.393	-0.01^a	0.05^a	-0.28^a	-0.20	0.01	0.16^c	244	0.121
ULVR	0.362	-0.03^a	0.05^a	-0.20^a	0.04	0.01^a	0.87^a	465	0.179
VOD	0.392	0.00	0.02^a	-0.13^a	0.07^b	0.01^a	0.02^b	237	0.052

Table 4.11: Estimation the Execution Probability for Sell Limit Order

The model estimated is (4.5) where

$$\beta' \mathbf{x} = \alpha + \beta_1 Spread + \beta_2 L + \beta_3 PMP + \beta_4 Volatility + \beta_5 tmID + \beta_6 Size$$

L and *PMP* are price aggressiveness and potential market pressure defined by formula (4.1) and (4.2) and measured at the time of order entry. *Volatility* is computed as the sum of squared returns and measured over 30 minutes immediate before the order is submitted. *tmID* is ordered number from 2 to 16 for time periods from 08:30 to 16:00. The column headed $\phi(\bar{\mathbf{x}}'\beta)$ is scaling coefficients of marginal effect for an 'average order'. *LR*-stats is based the Likelihood Ratio test of the null that all coefficients (except intercept) are zero. *LRI* is the likelihood ratio index defined by Greene [1993] (pp.891).

Mnem	$\phi(\bar{\mathbf{x}}'\beta)$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	*LR*-stats	*LRI*
AL.	0.280	-0.02[a]	0.04[a]	0.26[a]	0.01	-0.02[a]	3.75[a]	243	0.200
ANL	0.287	-0.02[a]	0.04[a]	0.32[a]	0.05	0.00	-0.43	414	0.221
ARM	0.397	-0.01[a]	0.02[a]	0.13[a]	0.00	0.00	1.46[a]	193	0.102
AZN	0.183	-0.02[a]	0.07[a]	0.17[a]	0.26[a]	-0.02[a]	1.37[a]	1307	0.264
BA.	0.396	-0.01[a]	0.04[a]	0.14[a]	0.28[b]	0.00	-0.25	130	0.102
BARC	0.367	-0.02[a]	0.04[a]	0.24[a]	0.19	-0.02[a]	5.03[a]	270	0.114
BG.	0.387	-0.01[a]	0.03[a]	0.09[b]	0.02	-0.02[a]	0.45[a]	139	0.092
BP.	0.399	-0.01[a]	0.05[a]	0.05	0.63[a]	0.01[b]	-0.30[a]	303	0.116
BT.A	0.356	-0.02[a]	0.04[a]	0.06[b]	0.05	-0.01[c]	0.28[a]	604	0.180
CGNU	0.370	-0.01[a]	0.04[a]	0.16[a]	0.07[c]	-0.01[b]	0.66	229	0.114
CTM	0.365	-0.01[a]	0.01[a]	0.01	0.00	-0.01[a]	0.93[b]	126	0.058
CW.	0.372	-0.01[a]	0.03[a]	0.15[a]	0.02	-0.01[a]	0.22[b]	273	0.107
DGE	0.271	-0.02[a]	0.06[a]	0.17[a]	0.13	-0.01[a]	0.96[a]	682	0.257
GSK	0.373	-0.01[a]	0.04[a]	0.12[a]	0.03	0.00	-0.18	362	0.096
HFX	0.289	-0.02[a]	0.04[a]	0.06	0.09	-0.03[a]	1.01[b]	215	0.181
HSBA	0.373	-0.01[b]	0.05[a]	0.23[a]	0.14	-0.01[b]	-0.21	341	0.142
ISYS	0.397	-0.01[a]	0.03[a]	0.20[a]	-0.06	0.00	0.54[a]	243	0.273
KGF	0.379	-0.02[a]	0.03[a]	0.11[b]	0.13[a]	0.00	1.35[a]	191	0.122
LLOY	0.351	-0.02[a]	0.07[a]	0.21[a]	-0.09	0.02[a]	0.91[a]	639	0.290
LOG	0.391	-0.01[a]	0.02[a]	0.02	0.06[a]	-0.02[a]	5.00[a]	263	0.092
MKS	0.359	-0.01[a]	0.03[a]	0.43[a]	0.10[a]	-0.02[a]	1.09[a]	234	0.199
PRU	0.380	-0.02[a]	0.04[a]	0.22[a]	-0.14	0.01[a]	0.95[b]	278	0.161
PSON	0.381	-0.01[a]	0.03[a]	-0.04	0.16[b]	0.00	-0.01	175	0.123
RBOS	0.384	-0.02[a]	0.04[a]	0.14[a]	0.30[a]	0.00	1.18	306	0.127
RTR	0.367	-0.01[a]	0.03[a]	0.08[a]	0.07	-0.02[a]	0.68[c]	194	0.066
SHEL	0.121	-0.02[a]	0.07[a]	0.18[a]	0.12	-0.01[b]	-0.02	1497	0.345
SPW	0.361	-0.02[a]	0.05[a]	0.05	0.04	0.01	1.10[a]	257	0.192
TSCO	0.394	-0.02[a]	0.04[a]	0.22[a]	0.55[a]	0.00	0.37[a]	184	0.123
ULVR	0.031	-0.03[a]	0.08[a]	0.25[a]	0.36[a]	-0.02[a]	0.52[a]	1658	0.310
VOD	0.397	0.00	0.02[a]	0.09	0.21[a]	-0.01	-0.03[b]	241	0.082

Figure 4.1: L-measure and price aggressiveness

In Panel A, limit order (a) and (b) face the same market spreads. Order (a) has larger *L* and is therefore more aggressive than order (b). In Panel B, limit order (a) and (b) have the same price aggressiveness measure *L* but face different market spreads. Order (a) is more aggressive than order (b).

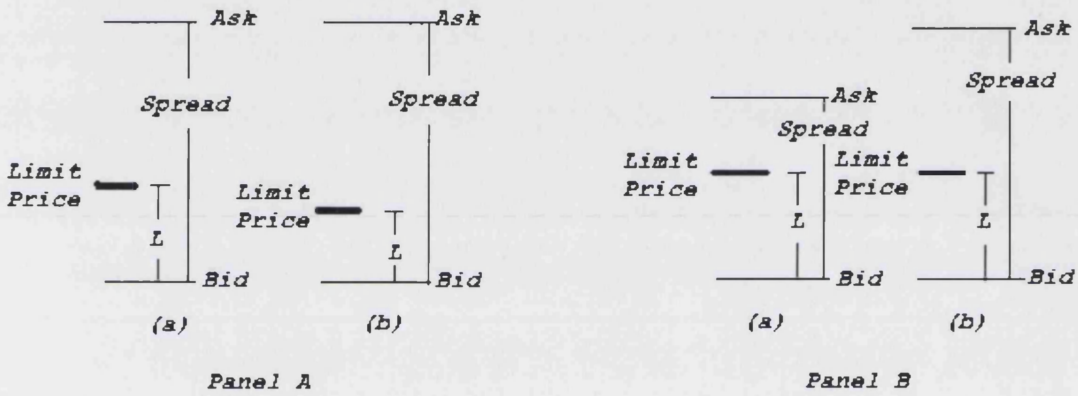


Figure 4.2: Intraday pattern of limit orders

Panel A is the number of total and limit orders (computed in every 15 minutes) of the FT30 stocks sumbitted to the SETS system in the LSE. Panel B plots the average execution probabilities and price aggressiveness of the limit orders (computed in every 15 minutes).

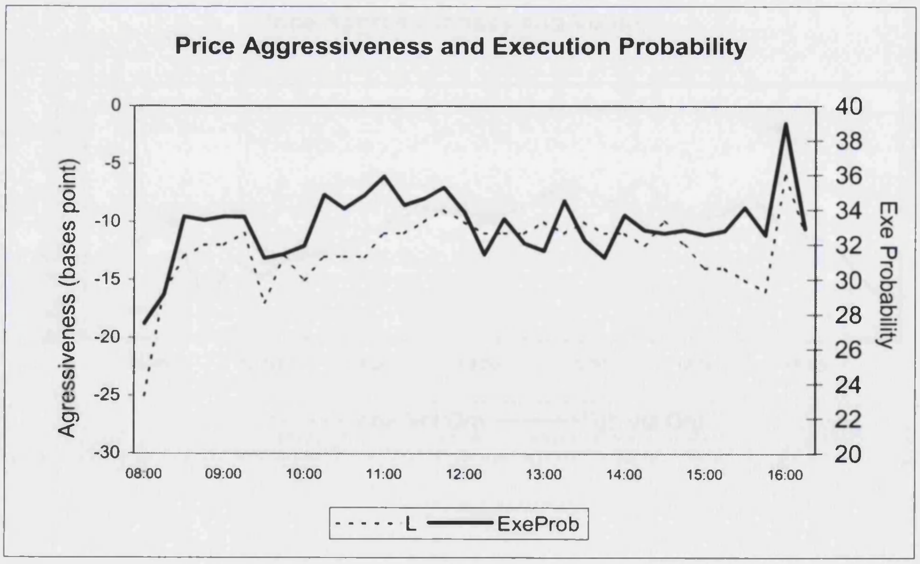


Figure 4.3: Volume Effect

The FT30 stocks are equally divided into five groups by their daily average trading volume. Panel A and B are the execution probabilities and price aggressiveness of the limit orders in the lowest and highest volume groups respectively.

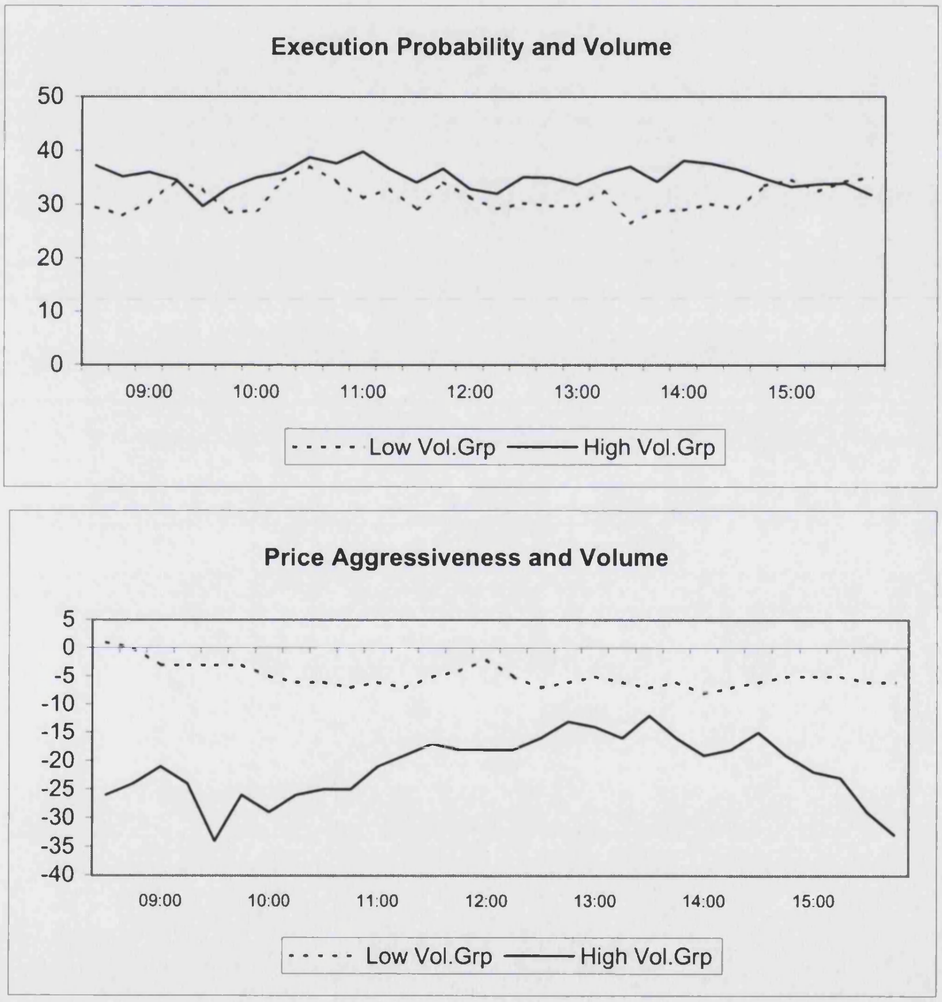
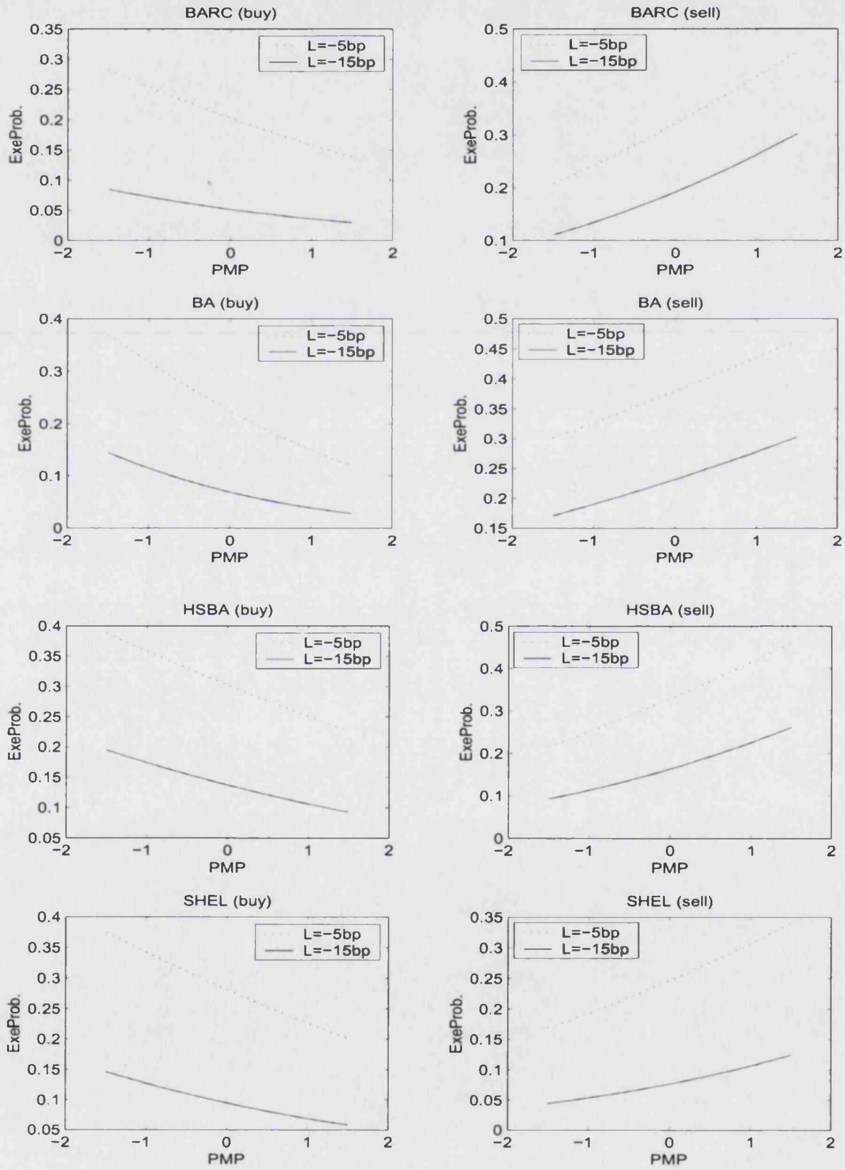


Figure 4.4: Predict Execution Probability: An Experiment

First type of orders for both sell and buy have L=-5 basis points and the second type of orders for sell and buy have L=-15 basis points. All orders take the average size of limit orders traded on Barclays, BAE, HSBC and Shell respectively. The PMP changes from -1.5 to 1.5. All orders are submitted between 12:30 to 13:00 (tmID=10).



Bibliography

[1] Admati, A.R., and P. Pfleiderer, 1988, A theory of intraday patters:Volume and price variability, Review of Financial Studies 1:3-40.

[2] Ahn, H.J., Bae, K.H., and Chan, K., 2001, Limit Orders, Depth, and Volatility: Evidence from the Stock Exchange of Hong Kong, Journal of Finance, 2:767-788.

[3] Angel, J., 1994, Limit vs. Market Orders, Working paper, No FINC-1377-10-293, School of Business Administration, Georgetown University.

[4] Bagehot, W., [Pseud], 1971, The Only Game in the Town, Financial Analysts Journal, 27:12-14,22.

[5] Bank of International Settlements, 2002, Foreign exchange markets turnover in April 2001, BIS, Basle.

[6] Biais, B. P.Hillion and C. Spatt, 1995, An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse, JOF, 50: 1655-1689

[7] Bloomfield, R., M. O'Hara and G. Saar, 2002, The 'Make or Take' Decision in an Electronic Market: Evidence on the Evolution of Liquidity, Unpublished paper, Cornell University.

[8] Blume, L.E., and D. Easley, 1990, Implications of Walrasian Expectations Equilibria, Journal of Economic Theory, 51:207-227.

[9] Bollerslev, T., and I. Domowitz, 1993, Trading patterns and prices in the interbank foreign exchange market. Journal of Finance 48:1421-43.

[10] Brock, W., and A. Kleidon, 1992, Periodic market closure and trading volume:A model of intraday bids and asks, Journal of Economic Dynamics and Control 16:451-489.

[11] Bronfman, C., Limit Order Placement and Execution Risk: An empirical Analysis, Working Paper, Commodity Futures Trading Commission.

[12] Caia, J., Yan-Leung Cheunga, Raymond S. K. Leea and Michael Melvin, 2001, Once-in-a-generation' yen volatility in 1998: fundamentals, intervention, and order flow, Journal of International Money and Finance, 20:327-347.

[13] Campell, J., and R. Shiller, 1988, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, Review of Financial Studies, 1:195-228.

[14] Chordia,T., R. Roll, and A. Subrahmanyam, 2002, Order imbalance, liquidity and market returns, Journal of Financial Economics, 65:111-130.

[15] Chung, Kee H., Bonnie F.Van Ness and Robert A. Van Ness, 1999, Limit orders and bid-ask spread, Journal of Financial Economics 53, 255-287.

[16] Clark, P., 1973, Asubordinated stochastic process model with finite variance for speculative prices, Econometica, 41:135-155.

[17] Cochrane, 2001, Asset pricing, Princeton University Press, Princeton

[18] Cohen, K., S.Maier, R. Schwartz, and D.Whitcomb, 1981, Transaction costs, Order Placement Strategy and the Existence of the Bid-Ask Spread, Journal of Political Economy, 14:71-100

[19] Cohen, B. H. and Shin, H. S., 2002, Positive feedback trading under stress: Evidence from the us treasury securities market, Working paper, Bank of International Settlements and the London School of Economics.

[20] Danielsson, J., and R. Payne, 2000, Real Trading Patterns and Prices in Spot Foreign Exchange Markets, Journal of International Money and Finance, forthcoming.

[21] Danielsson, J., R. Payne, and J. Luo, 2002, Exchange Rate Determination and Inter-market Order Flow Effects, Manuscript, FMG, London School of Economics.

[22] Danielsson Jon and Burak Saltoglu, 2003, Anatomy of a Market Crash: A Market Microstructure Analysis of the Turkish Overnight Liquidity Crisis, Working paper, FMG, London School of Economics.

[23] Demsetz, H., 1968, The Costs of Transacting, Quarterly Journal of Economics, 82:33-53.

[24] Diamond, D.W. and R.E. Verrecchia, 1987, Constraints on short-selling and asset price adjustment to private information, Journal of Financial Economics 18:277-311.

[25] Diebold, F., 2001, Elements of Forecasting, South-Western.

[26] Dominguez, K., 2003a, The market microstructure of central bank intervention, Journal of International Economics, 59:25-45.

[27] Dominguez, K., 2003b, When Do Central Bank Interventions Influcence Intra-daily and Longer-term Exchange Rate Movements, Discussion paper of 4th Empirical Finance Conference, London School of Economics, April, 2003.

[28] Easley, D., and M. O'Hara, 1992, Time and Process of Security Price Adjustment, Journal of Finance 47:577-605.

[29] Epps, T., and M. Epps, 1976, The stochastic dependece of security price changes and transaction volumes: Implication for the mixture-of-distributions hypothesis, Econometrica, 44:305-321.

[30] Evans, M., and R., Lyons, 2002a, Order flow and Exchange rate Dynamics, Journal of Political Economy, 110: 170-180.

[31] Evans, M., and R., Lyons, 2002b, Informational Integration and FX trading, Journal of International Money and Finance, 21:807-831.

[32] Evans, M., and R. Lyons, 2003, How is macro news transmitted to exchange rates?, NBER working paper, WP9433.

[33] Foster, D., and S. Viswanathan, 1990, A theory of intraday variations in volumes, variances and trading costs, Review of Financial Studies 3:593-624.

[34] Foucault, T., 1999, Order Flow Composition and Trading Costs in a Dynamic Limit Order Market, Journal of Financial Market, 2:99-134.

[35] Foucault, T., Ohad Kadan, and Eugene Kandel, 2001, Limit Order Book as a Market for Liquidity, HEC Working paper.

[36] Frankel, J., Giampaolo Galli, and Alberto Giovannini, 1996, The Mircostructure of Foreign Exchange Markets, The University of Chicago Press.

[37] Frankel, J., and A.K. Ross, 1995, Empirical Research on Nominal Exchange Rates, in G.Grossman and K.Rogoff(eds.), Handbook of International Economics,Elsevier Science: Amsterdam, pp.1689-1730

[38] French, K., and R. Roll, 1986, Stock return variance: The arrival of information and the reaction of traders, Journal of Financial Economics 17:99-117.

[39] Friederich, S. and R. Payne, 2001, Dealer liquidity in an auction market:evidence from the London Stock Exchange, FMG discussion paper.

[40] Froot, K., and Tarun Ramadorai, 2002, Currency Returns, Institutional Investor Flows and Exchange Rate Fundamentals, NBER working paper, 9101.

[41] Gehrig, T., and L. Menkhoff, 2002, The Use of Flow Analysis in Foreign Exchange: Explanatory Analysis, CEPR Working paper, DP 3221.

[42] Glosten, L., 1994, Is the electronic open limit order book inevitable? Journal of Finance, 49:1127-1161

[43] Glosten, L.R., and P.R. Milgrom, 1985, Bid,ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders, Journal of Financial Economics, 14:71-100.

[44] Goodhart, C., 1988, The Foreign Exchange Market: A Random Walk with a Dragging Anchor, Economica, 55:437-460

[45] Goodhart, C., R. Love, R. Payne, and D. Rime, 2002, Analysis of spreads in the Dollar/Euro and Deutsche Mark/Dollar foreign exchange markets. Economic Policy, 35, October 2002.

[46] Granger, C.W.J., and T. Tersvirta, 1997, Modelling Nonlinear Economic Relationships. Oxford University Press, Oxford, UK.

[47] Greene, William, 1993, Econometric Analysis, third edition. Prentice-Hall, Inc.

[48] Grossman, Sanford, 1976, On the Efficiency of Competetive Stock markets Where Traders have Diverse Information,Journal of Finance, 31:573-584.

[49] Grossman, Sanford and Joseph Stiglitz, 1980, On the Impossibility of Informationally Efficient Markets, American Economic Review, 70:393-408.

[50] Handa, p., and R. Schwartz, 1996, Limit Order Trading, Journal of Finance, 51:1835-1861

[51] Harris, L., and J. Hasbrouk, 1996, Market vs. Limit Orders: The SuperDOT Evidence on Order Submission Strategy, JFQA, 31: 213-231

[52] Hartmann, P., 1999, Trading volume and transaction costs in the foreign exchange market, Evidence from daily dollar-yen spot data. Journal of Banking and Finance.

[53] Hasbrouck, J., 1991a, Measuring the Information Content od Stock Trades, Journal of Finance, 46:179-206

[54] Hasbrouck, J., 1991b, The summary informativeness of Stock Trades: An Econometric Analysis, Review of Financial Studies, 4:571-595.

[55] Hasbrouck, J., and D. Seppi, 2001, Common factors in prices, order flow and liquidity, Journal of Financial Economics, 59:383-411.

[56] Hsieh, D.A., and A.W. Kleidon, 1996, Bid-Ask Spread in Foreign Exchange Markets:Implications for Models of Asymmetric Information. in J.Frankel et al (1996, eds.), The Microstructure of Foreign Exchange Markets, University of Chicago Press, Chicago,IL., p.41-65.

[57] Isard, P., 1995, Exchange Rate Economics,Cambridge University Press: Cambridge, UK.

[58] Ito, T., R. Lyons, and M. Melvin, 1998, Is there private information in the FX market? The Tokyo experiment, Journal of Finance 53:1111-1130.

[59] Karpoff, J., 1987, The relation between price changes and trading volume: A survey, Journal of Financial and Quantitative Analysis, 22:109-126.

[60] Kumar, P., and D.J. Seppi, 1992, Limit and Market Orders with Optimizing Traders, Working paper, Carnegie Mellon University.

[61] Kyle, A.S., 1985, Continuous Auctions and Insider Trading, Econometrica, 53:1315-1335.

[62] Lo, A., MacKinlay. C., and J.Zhang, 2002, Econometric Models of Limit-Order Executions, Journal of Financial Economics, 65:31-71.

[63] Love, R., and R. Payne, 2002, Macroeconomic news, order flows and exchange rates, Manuscript, FMG, London School of Economics.

[64] Lyons, R., 1995, Test of microstructure hypotheses in the foreign exchange market, Journal of Financial Economics. October 1995, 39,321-351.

[65] Lyons, R., 1996, Foreign Exchange Volume: Sounds or Furry Signifying Nothing?, in J.Frankel et al (1996, eds.), The Microstructure of Foreign Exchange Markets, University of Chicago Press, Chicago,IL., p.183-201.

[66] Lyons, R., 2001, The Microstructure Approach to Exchange Rates, MIT Press.

[67] Lyons, R., and M. Moore, 2003, An Information Approach to International Currencies, discussion paper of the 4th Empirical Finance Conference, April 2003, FMG, LSE.

[68] Madhavan, A and S. Smidt, 1991, A Bayesian model of intraday specialist pricing, Journal of Financial Economics, 30:99-134.

[69] Mark, C. Nelson, 1995 American Economic Review, 85:201-218.

[70] Matsuyama, K., N. Kiyotaki, and A.Matsui, 1993, Towards a theory of international currency, Review of Economic Studies, 60:283-320.

[71] Meese, R., and K. Rogoff, 1983a, Empirical exchange rate models of the seventies, Journal of International Economics, 14:3-24.

[72] Meese, R., and K. Rogoff, 1983b, The out-of-sample failure of empirical exchange rate models. in Jeffrey Frenkel(ed.), Exchange Rate and International Macroeconomics, vol.14, University of Chicago Press: Chicago.

[73] McCauley, R., 2001, Comments on "Order flow and exchange rate dynamics", Market Liquidity: proceedings of workshop help at BIS, April 2001. http://www. bis.org/publ/ bispap02.htm

[74] Mello, Antonio, 1996, Comments on "Foreign Exchange Volume: Sound and Fury Signifying Nothing?", in J.Frankel et al (1996, eds.), The Microstructure of Foreign Exchange Markets, University of Chicago Press, Chicago,IL., p.205-205.

[75] Parlour, C., 1998, Price dynamics in Limit Order Markets, Review of Financial Studies, 11:789-816

[76] Pastor, Lubos and Robert Stambaugh, 2003, Liquidity Risk and Expected Stock Returns, Journal of Political Economy, 111:642-685.

[77] Payne, R., 2003, Informed trade in spot foreign exchange markets: An empirical investigation, Journal of International Money and Finance, forthcoming.

[78] Payne, R., and P. Vitale, 2000, A Transaction Level Study of the Effects of Central Bank Intervention on Exchange Rates. FMG working paper, London School of Economics.

[79] Perradin W., and P. Vitale, 1996, Interdealer Trade and Information Flows in a decentralized Foreign Exchange Market, in J.Frankel et al (eds.), The Microstructure of Foreign Exchange Markets, University of Chicago Press, Chicago,IL., p.73-99.

[80] Radner, R., 1982, Equilibrium Under Uncertainty, in the Handbook of Mathematical Economics, Vol. 2, ed. by K.Arrow and M. Intrilligator, North-Holland, Amsterdam.

[81] Rey, H., 2001, International Trade and Currency Exchange, Review of Economic Studies, 68: 443-464.

[82] Rime, Dagfinn , 2000, Private or Public Information in Foreign Exchange Markets? An Empirical Analysis. Typescript, University of Oslo. July 2000.

[83] Seppi, D., 1997, Liquidity Provision with Limit Orders and a Strategic Specialist, Review of Financial Studies, 10:103-150

[84] Subrahmanyam, A., 1991 Risk aversion,market liquidity, and price efficiency, Review of Financial Studies 4:417-441.

[85] Tauchen, G., and Pitts, M., 1983, The price variability-volume relationship on speculative markets, Econometrica, 51:485-505.