

Market conditions, order flow and exchange rates determination

Jinhui Luo*

Financial Market Group
Department of Accounting and Finance
London School of Economics

December 29, 2001 Second draft

Abstract

In comparison to macroeconomic models of nominal exchange rates, the market microstructure approach performs better in explaining exchange rate changes over short time horizons. The microstructure approach implies that information in order flow drives the dynamic processes of price evolution. This paper studies the informativeness of order flow under different market conditions in the foreign exchange market. We find that order flow tends to be more informative when the market experiences large bid-ask spreads, high volatility or trading volumes. We also identify nonlinearities in the relationship between order flow and foreign exchange rate changes and find that these nonlinearities can be captured by the Interaction model and the Logistic Smooth Transition Regression (LSTR) model .

JEL Classification: D82,F31,G15

Keywords: Market microstructure, Foreign exchange market, Order flow

*I would like to thank Richard Payne and Jon Danielsson for their helpful comments and suggestions and the time they spent discussing my ideas. I also want to thank FMG director David C.Webb for his support and Charles A.E. Goodhart, Paolo Vitale, Ryan Love and seminar participants at London School of Economics for valuable comments. Any errors remain my own. For correspondence: Financial Market Group, London School of Economics, Houghton Street, London, WC2A 2AE, UK; tel: ++44-(0)20-79556511; email: J.Luo@lse.ac.uk.

1 Introduction

In macroeconomic models, foreign exchange rates are determined by the underlying fundamental factors in 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 rate 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 famous papers Meese and Rogoff [1983a, 1983b] vigorously showed 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. The empirical evidence of the 1980s drove economists to the conclusion that the most critical determinants of exchange rate volatility may not be macroeconomic¹.

In recent years, a new direction in exchange rate analysis, the market microstructure 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 information². When the market is not fully efficient, 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 order flow is informative. Indeed dealers in foreign exchange markets claim that trading with their customers is one of the most important sources of information (Cheung and Wong [2000], Yao [1998a] and Goodhart [1988]). From the point of view of asset pricing, the crucial idea here is that information flow, through order flow, drives price movements³.

This approach has recently been applied to foreign exchange markets and generated some promising results. Evens and Lyons [2001] claim that net order flow has substantial explanatory power for exchange rate changes on a daily basis and can

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

²Private information is not necessarily about fundamentals. In this approach, private information is defined as any information that is not public and help forecast the future price better than public information alone. This definition is adopted from Lyons [2000], *The Microstructure Approach to Exchange Rates*, forthcoming, MIT Press.

³It is important to note that order flow is fundamentally different from volume. In market microstructure literature, order flow is defined as the net of buyer-initiated orders and seller-initiated orders and is a measure of market buying/selling pressure.

explain 60 percent and 40 percent of return variation 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. Payne, Luo and Danielsson [2001] work on very high frequency data and show that order flow has explanatory power for exchange rate changes for USD/EUR, USD/GBP, JPY/USD, and GBP/EUR at a relatively high frequency level even though the explanatory power varies across exchange rates and sampling frequencies. All these results indicate that order flow is informative. More accurately, order flow plays the role of a channel through which information is impounded into price.

An implicit underlying assumption in all this research is that the informativeness of order flow is constant even if the market experiences dramatically different situations. Consider the the following claim of the Reserve Bank of Australia:

*"... 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

The claim of the Reserve Bank of Australia reflects the worries behind the following question people would like to ask: "Is the informativeness of order flow really constant?". In this paper we test whether this basic assumption is true and further investigate the question of under what market conditions order flows tend to be more informative. Since the price impact of order flow is empirically important, characterizing the information transfer mechanism of order flow would be an important step to better understand exchange rate volatility and has important implications for central banks and practitioners in the foreign exchange market. This direction of investigation, however, has hardly been touched in the literature⁵.

This paper develops a methodology to study variation in the price impacts of order flow and provides empirical evidence to characterize the information transfer mechanism in the inter-dealer spot FX markets via the analysis of a transaction data set recorded in the Reuters D2000-2 dealing system. We test a set of hypotheses regarding whether order flow is equally informative under various market conditions and we also model the relationship between the price impact of order flow and market condition measuring variables. The main conclusions are:

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

⁵Lyons [1996] is an exception. In this paper Lyons tested a hypothesis regarding the relationship between the order flow informativeness and market intensity and concluded that the order flow is more informative when trading intensity is high and less informative when quoting intensity is high.

- The informativeness of order flow is not constant under different market conditions as was assumed in the literature. This result highlights the non-linearity in the relationship between 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 strong at relatively high sampling frequencies and becomes weaker as the sampling frequencies get lower.
- 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 Interaction model and LSTR model also suggest that the price impact of order flow is more sensitive to bid-ask spreads than to market volatility and trading volume.

The remainder of the paper is organized as follows: Section 2 gives the theoretical background and motivation. Section 3 presents the methodology. Section 4 describes the FX market background and the data. Model estimation and empirical analysis are presented in section 5. Section 6 concludes.

2 Theoretical issue and motivation

In contrast to standard exchange rate models where information affecting rates and the mapping from information to prices are known to all, the market microstructure approach to exchange rate determination explicitly assumes that information that affects prices could be private. In the classic market microstructure frameworks by Kyle [1985] and Glosten and Milgrom [1985], informed traders exploit their informational advantage by placing their orders against the market maker. The market maker adjusts price by observing the order flow. In this way the private information is transmitted into the market. Here the core is order flow which acts as the vehicle of information transmission.

In the last decade, we have witnessed a large amount of empirical work that provides evidence supporting the above idea in both equity and FX markets. French and Roll [1986] study the volatility of returns from Tuesday to Thursday on the New York Stock Exchange (NYSE) and conclude that informative order flow arriving at the market during non-halt-Wednesday periods causes an increase in volatility over halt-Wednesday periods when there is no such order flow. Ito, Lyons and Melvin [1998] conduct the analogous work in the Tokyo FX market. They find the volatility of JPY/USD doubled over the lunch period during which

the only change was that a trading rule, under which all banks were restricted from trading in Tokyo over lunch time, was lifted. The authors conclude that the increase in volatility was, at least partially, due to the informative order flows that arrived during the lunch time after the policy was changed.

Another approach to test the informative role of order flow is to investigate the persistent effect of order flow on prices and this approach has been extensively used in studying the inter-bank foreign exchange markets⁶. Payne [2000] uses a VAR model introduced by Hasbrouck [1991] to study the long-run effect of transactions on price in the inter-dealer FX market. After studying the USD/DEM data recorded in Reuters D2002⁷, the author shows that 40% of the permanent price variation can be attributed to information contained in transactions. A slightly different way to study the long-run effect of order flow on prices is to use time-aggregated order flow to explain price movements⁸. The papers along this line include Evans and Lyons [2001], Rime [2000] and Payne, Luo and Danielsson [2001]. Even though the research is conducted on different time aggregation levels⁹, all authors show that the time-aggregated order flow has substantial explanatory power for exchange rate changes and claim that the evidence presented supports the hypothesis that order flow conveys information.

The theoretical models that focus on asymmetric information of inter-dealer FX markets include Lyons [1995, 2001] and Perraudin and Vitale [1996]. In these models dealers first receive order flow from their non-dealer customers, which is believed to contain some private information. The information is then shared among the dealers through inter-dealer trading. In particular, Perraudin and Vitale [1996] show that to be able to exploit the surplus from liquidity traders the dealers are willing to share the information. The authors argue that information sharing is achieved by maintaining a certain level of inter-dealer trading. An implication of these models is that order flow can be employed to transfer information in inter-dealer spot FX markets and therefore affect exchange rates¹⁰.

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

⁶Lyons [2000] provides a full discussion of the approaches to test the informativeness of order flow.

⁷Reuters D2000-2 is a brokered inter-dealer trading system. A detailed introduction of the D2000-2 system is also provided in Payne [2000]. Reuters also operates a system called D2000-1, which is a direct inter-dealer system.

⁸This method is first introduced in an earlier version of Evans and Lyons [2001].

⁹Evans and Lyons [2000] work on daily bases, Rime [2000] works on weekly bases and Payne, Luo and Danielsson [2001] work on a range of higher intra-day frequencies.

¹⁰A full discussion of the order flow impact on foreign exchange rate is covered in Lyons [2000]

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, no consensus on the implications of the theoretical models, and indeed some appear to be contradictory.

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 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. However, Subrahmanyam [1991] argues that the conclusion of the Admati and Pfleiderer [1988] model relies on the assumption of risk neutrality of informed traders and demonstrates that if the informed traders are risk averse then, during the concentration, the increased informative trading will lower the market liquidity and increase the spread. The additional assumption adopted by the author that more individuals are informed traders during the concentration (the beginning and the end of the day) implies that the order flow is more informative in this period.

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 \bar{V} or \underline{V} With trade now 'safer' the market maker reduces his spread." (pp.587). In words, the model predicts that the informativeness of order flow is positively correlated with volume, volatility and spread. 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 checking the patterns in volume, volatility and spreads. The topic of the relations between the informativeness of order flow and the different market conditions is little explored. One of the few papers that address this issue is Lyons [1996]. In the paper, Lyons tests the event-uncertainty hypothesis against *hot potato hypothesis*¹¹ by empirically 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. The contradictory evidence drives the author to the conclusion that ‘the results highlight the potential complementarity between these seemingly polar views’(pp.199). Yet why the different measurement of market pace has led to the opposite conclusion is unexplained. 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 the market could undergo different periods (eg. turbulent vs calm, announcement vs non-announcement) and different dealers might have different characteristics (eg, capital size, trading strategies), the observations from a single dealer for one week might not tell the true story. Another concern, as Mello[1996] points out, is that “the results seems to be highly dependent on the definition of *short time*, a metric that must be endogenous and dependent on the prevailing market conditions”(p.206).

The aim of this paper is to characterize the information transmission mechanism of order flow under different market conditions, namely the informativeness of order flow under different regimes of three most important market statistics: volume, volatility and bid-ask spread.

Here the first basic question we would like to ask is that “Does informativeness of order flow change under different market conditions?”. This question can be put formally as the following three hypotheses:

Hypothesis I: *Order flow is equally informative regardless of trading volume.*¹²

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

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

A further question, also a more challenging one, is “if the informativeness of order flow changes in different market conditions, how does it change?”. This

¹¹The *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.

¹²The irrelevancy hypothesis implies that neither Admati and Pfleiderer[1988]’s prediction nor Easley and O’Hara[1992]’s prediction is valid.

question tries to characterize the informational aspect of price impact of order flow in security markets. Investigation of the above issues can't only help clarify the theoretical dispute in the market microstructure literature but also help people better understand the price formation process. The latter point has crucial implications for issues such as central bank intervention and practitioner's trading strategy design.

3 Methodology

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 [1984,1985], great 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 lower ($V = \underline{V}$) or higher ($V = \bar{V}$) 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 whatever the 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 are 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 model¹³:

$$\Delta P_t = \alpha + \beta * Q_t + \epsilon_t \quad (1)$$

¹³A similar modelling strategy has been extensively used in empirical work, such as Madhavan and Smidt[1991], Lyons[1995,2001], Foster and Viswanathan[1990]

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 of 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 price formation process.

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 (1) is constant. If it is not constant under different market conditions, a recursive regression model and a window regression model are employed to further study the dynamics of β .

3.2.1 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 interesting variable Z_t , which measures market conditions. The market conditions of interest in this paper are market volatility, liquidity and volume. Take the case of volatility as an example, the whole 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 (1) can be expressed as

$$\Delta P_t = \alpha + \sum_{j=1}^J \beta_j * S_{jt} * Q_{jt} + \varepsilon_t \quad (2)$$

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

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 $\beta_i = \beta_j, \forall i, j$.

3.2.2 Recursive least squares regression model

A simple 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 (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:

$$\Delta P_t = \mathbf{x}_t' \mathbf{B} + u_t \quad (3)$$

where $x_t = [1, Q_t]'$ and \mathbf{B} is the coefficient vector $[\alpha, \beta]'$ in the core model (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 follows: Step one, sort the observations x_t according to Z_t , which could be volatility, bid-ask spreads and trading volume. After sorting, the observations are re-arranged by ascending value of Z_t .

Step two, fit the model (1) to the first k ($k = 2$) observations and get the coefficient estimate b_k . Next use the first $k + 1$ data points as regressor and computer the regression coefficient again. Proceed in this way, adding one observations at a time, until the final regression coefficient is obtained, based on the all observations. This process will generate a sequence of coefficient estimates, $\mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n$. In general,

$$\mathbf{b}_m = (\mathbf{X}_m' \mathbf{X}_m)^{-1} \mathbf{X}_m' \Delta \mathbf{P}_m \quad m = k, k + 1, \dots, n \quad (4)$$

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 on the dependent variables.

Step three, the standard errors of various coefficients are calculated at each stage of the recursion (except the first one) and the evolution of the various 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 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 the parameter inconstancy.

3.2.3 Window regression model

A window regression model is proposed here as an alternative to the recursive model. In recursive models the effect of variable Z_t (measuring market conditions) will attenuate as more and more observations enter the regression and the β will eventually converge to the equilibrium value for the whole sample (without the effect of Z). In window regression models, the observations are also sorted by the interesting variable Z_t and regression model (1) is run repeatedly on a set of non-overlapping fixed-length windows. Comparison of the regression coefficient of each window can provide us with a closer look at the β dynamics. The window regression model can be viewed as a finer version of quintile models.

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.

3.3.1 Interaction model

Since so far there is no unanimous theoretical indication 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 (1):

$$\beta = \beta_1 + \beta_2 * Z_t \quad (5)$$

Z_t is the measurement of market conditions of interest (it could be volume, volatility or spread). Inserting (5) back into (1) and rearranging it results in the following nonlinear regression model:

$$\Delta P_t = \alpha + \beta_1 * Q_t + \beta_2 * Z_t * Q_t + \varepsilon_t \quad (6)$$

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 prediction of various theoretical models about the relationship between informativeness of order flow and market conditions in the market microstructure literature.

3.3.2 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 movement. Formally the LSTR can be written as

$$\Delta P_t = \beta' \mathbf{x}_t + (\theta' \mathbf{x}_t) F(Z_t) + \varepsilon_t \quad (7)$$

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

$$F(Z_t) = (1 + \exp\{-\gamma(Z_t - c)\})^{-1} - 1/2 \quad (8)$$

where the logistic parameter $\gamma > 0$.

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 .

4 The FX market background and the Data

4.1 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. In contrast to most of the equity markets, the FX market neither has a physical location or system where all quotes are posted and trades are excised and reported nor has a disclosure requirement. Banks can enter their quotes in a number of screen-based electronic systems, but until fairly recently, most of the trades have been carried out over the telephone and therefore not been observable¹⁴. Further, the transactions between dealers and their non-bank customers largely remain private information to the dealers themselves. This fragmentation feature makes the data on FX transactions very elusive and market transparency very low. Another feature that is different from the equity markets is the FX market has no daily open-close procedure. The FX market operates continuously on a 24-hour basis. 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.

The enormous trading volume is probably one of the most prominent features of the FX market. The daily trading volume in the spot FX is \$600 billion in 1998 according to the 1998 survey of the Bank for International Settlements¹⁵. Following Lyons [2000], 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 market and that information will be shared among dealers within the brokered and the direct inter-dealer segments later. It is believed that the inter-dealer market (including the last two segments) accounts for 80% of the spot FX trade (see Payne [2000]) and is the most liquid part of the market. In 1998 the brokered inter-dealer segment accounts for roughly 40% for the whole inter-dealer market volume. Currently Reuters D2000-1 is an electronic trading system operating in the direct inter-dealer segment¹⁶. EBS and Reuters D2000-2 are the dominant electronic broker systems operating in the brokered inter-dealer market¹⁷.

¹⁴Trading technology in the FX markets was altered by the introduction of several electronic dealing systems such as EBS, Reuters D2000-1 and Reuters D2000-2. EBS, Reuters D2000-2 are brokered inter-dealer systems and Reuters D2000-1 is a bilateral direct inter-dealer system.

¹⁵The usually quoted daily trading volume of \$1.5 trillion applies to a broader definition of market which includes derivatives market and spot market. For the purpose of this paper, the focus is the spot market since most of the FX derivative instruments have no order flow consequences. See Lyons [2000] for detail discussion in this aspect.

¹⁶Lyons [1995, 1996, 2001] analyze the transactions data from this system.

¹⁷EBS claimed to handle 37% of the brokered trade in London and it is believed that Reuters has the same share. See Payne [2000].

4.2 The Data

The data¹⁸ used in the paper are the tick data generated by Reuters D2000-2. This data set has significant advantages over foreign exchange data used in the past work (see, e.g., Bollerslev and Domowitz[1993], and Lyons[1996]). The data used in Bollerslev and Domowitz[1993] are indicative quotes from Reuters FXFX¹⁹. The shortcoming of indicative quotes is that the return variance derived from them is far larger than that derived from actual quotes or trades and the spread is less correlated with market activity²⁰. The data used in Lyons[1996] is real trade data, but it covers only 5 working days for a single dealer. As mentioned previously, the data generated by Reuters 2000-2 dealing system is real trade and quote data of the inter-dealer market. More over, the data set on average covers nine months and four major floating exchange rates.

4.2.1 Trades and quotes

The data used in this study is composed of two types of transaction level information: trades and firm quotes. Each type covers four currency pairs: Euro-Dollar, Euro-Sterling, Sterling-Dollar and Dollar-Yen²¹. Each piece of the trade information contains the time stamp of the transaction, buy/sell indicator (from market maker's point of view) and transaction price (accurate to the fourth decimal). Each piece of the quote information contains time stamp, bid and ask²². The samples for Euro-Dollar and Sterling-Dollar cover a period of ten months from 28 September 1999 to 24 July 2000. Samples for Euro-Sterling and Dollar-Yen cover a period of eight months from 1 December 1999 to 24 July 2000.

For the purpose of the study in this paper, one possible drawback of the data set is the lack of information about the size of each trade. This drawback calls for additional attention to the interpretation of our results. Nevertheless this high frequency data set has two valuable characteristics: long sample periods and multiple exchange rates. The long sample period provides us with the opportunity to study the above issues from the time aggregation angle without loss of statistical power. The broad currency scope provides a platform to check the robustness of estimation cross sectionally on major floating exchange rates.

¹⁸The data set used in this paper is from the Foreign Exchange Project of the Financial Market Group, London School of Economics

¹⁹Reuters 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.

²⁰Please refer to Danielsson and Payne[2000] for full discussion

²¹For the purpose of simplicity, the shortcut "ED", "ES", "SD", "DY" will be used to represent Euro-Dollar, Euro-Sterling, Sterling-Dollar and Dollar-Yen currency pairs respectively.

²²The stamp used in this paper is standard Daylight Saving Time (DST) of London.

4.3 Filtering and time aggregation

As shown in Figure 1, the market behavior overnight is dramatically different from that during the daytime²³. In particular, the volume (either proxied by number of trades or by number of quotes) is extremely low and spreads (calculated as percentage of trading prices and measured in basis point) are extremely high during the overnight period. For the purpose of this paper, we exclude all overnight periods, weekends and holidays from our sample before further procession. In this way, we concentrate more on the issues of informativeness of order flow under various market conditions.

Payne, Luo and Danielsson [2001] show that the relationship between order flow and exchange rate changes can vary widely for different time aggregation levels. To get a picture of the dynamics of price impact of order flow under different market conditions, we need to aggregate the transactions and quotes through a spectrum of sampling frequencies. In this paper, we choose 8 different time aggregation levels: 5 minutes, 10 minutes, 20 minutes, 30 minutes, 1 hour, 2 hours, 6 hour and 12 hours²⁴. The time aggregation is done as follows. First, we scan the sample along calendar time minute by minute. If no trade or quote occurs within one minute, an artificial trade or quote with the same trade price or quote as the last true trade or quote will be inserted into the time series for that minute²⁵. Second, for each sampling frequency we record the last transaction price, total number of buys and sells, average bid-ask spread and volatility within every time interval for that sampling frequency. Spreads are calculated as the percentage of trade price in basis point. Volatility is calculated as the return variance within the time interval. Third, we delete the outliers that have the extreme value of volume, spreads or volatility from the time series produced by the second step²⁶.

²³In this paper we define the overnight as a period from 18:00 to 6:00 DST next day. It should be noted this definition is only proper for the traders in London and New York but not for the traders in Asian markets. In our data set, JPY/USD is most likely affected by this definition since the daytime period of Tokyo and Hong Kong markets are exactly the overnight period by DST of London. After checking the data in our sample, we found that the market for JPY/USD stays at a reasonably active level during the period defined as overnight by DST of London. The possible reason could be that, in order to trade with their counterparts in European and American markets, a sizable proportion of traders in Asia markets remains active during the time when London and New York markets are “open”.

²⁴We have experimented with denser time aggregation levels and the results do not alter the pattern we reported in this paper.

²⁵These artificial trades and quotes are specially flagged so that they will not be counted when we calculate the number of trades and quotes.

²⁶These outliers are usually generated by key-board error or extreme market conditions such as when trading volume is extremely low or spreads and volatility is extremely high. The experiments have also been done to check the impact of this exclusion and we find that the exclusion doesn't change the pattern we show in this paper.

Through the filtering and aggregation, we generate 32 databases (8 sampling frequencies \times 4 exchange rates), each representing a different sampling frequency for an exchange rate. These aggregated databases are the foundation for our model estimations and their properties are summarized in Table 1. As mentioned previously, the long covering periods is a valuable characteristic of our sample. After the filtration and aggregation we still have about 150 observations for JPY/USD and GBP/EUR and 190 observations for USD/EUR and USD/GBP for the lowest sampling frequency (12-hour) in our generated databases. It is interesting to note that the periods our sample covers is coincidental with the time when foreign exchange markets experienced a depreciation of the euro against the US dollar and the sterling pound, a depreciation of the pound against the dollar and a depreciation of the Yen against the dollar. These market trends are reflected in the *rtn* columns of each panel in Table 1. Comparing panel (b) with the other three panels, we find that the number of trades and quotes are far less and the spreads and volatility are much higher. This is probably partially because the Reuters D2000-2 is not the most popular dealing system for JPY/USD and partially because the trading activity of JPY/USD is less concentrated during the periods we defined as DST daytime.

5 Estimation and Analysis

In this section we use four major floating rates (USD/EUR, GBP/EUR, USD/GBP and JPY/USD) 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 $[t-1, t]$ ²⁷, Z_t could be average spread (in basis point sense), return variance or trading volume²⁸ within time interval $[t-1, t]$.

²⁷In the standard microstructure model, order flow is the difference between buy initiated orders and sell initiated orders, i.e. the signed volume. Since we don't have the size of each trade in our time series, we take the same strategy as that in Evans and Lyons [2001] by defining order flow as the difference between the number of buys and number of sells.

²⁸Since the size of each trade is not observable in our data set, the total number of trades or quotes will be used as proxy of volume. For justification of this proxy please see Danielsson and Payne[2000].

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 aggregations. For market conditions, we check the following three important market statistics: bid-ask spread, trading volume²⁹ and market volatility. For time aggregation, we estimate the model for a series of sampling frequencies from 5 minutes to 12 hours.

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 table 2,3 and 4. The most notable result is that F -tests are significant in all three tables for most of the high sampling frequencies and all currency pairs (except for JPY/USD and USD/GBP in Table 4). In particular, if the observations are arranged along bid-ask spread (Table 2), F -tests are significant for sampling frequencies from 5 minutes to 6 hours on average. If the observations are arranged by volatility (Table 3), F -tests are significant for sampling frequencies from 5 minutes to at least 30 minutes. If the observations are arranged by volume (Table 4), F -tests are only significant for USD/EUR and GBP/EUR for some sampling frequencies and not significant for JPY/USD and USD/GBP for most of the sampling frequencies. Overall, the results from the quintile model suggest that the null hypothesis (order flow is 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 2 and Table 3, β tends to increase from group 1 to group 5. In Table 4, β tends to decrease from group 1 to group 2 and then remains relatively stable (except for JPY/USD). For example, in Table 2 the β of Euro-Dollar based on the 10 minute sampling frequency increases from 0.0027 for group 1 (with smallest spread) to 0.0077 (with highest spread) for group 5. The pattern exhibited here is quite persistent for relatively high sampling frequencies. The increasing β in Table 2 and Table 3 and the decreasing β in Table 4 indicates that β might be an increasing function of market spread and volatility and decreasing function of trading

²⁹In this paper we only report the results for number of trades since results from number of quote are more or less the same.

volume (at least in some domains of volume). In words, order flow is more informative when market spread is large, volatility is high or trading volume is low. The interesting point here is that our results are neither completely 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 since the information uncertainty model predicts a positive relation between informativeness of order flow and all three market conditions and the model of Admati and Pfleiderer [1988] predicts that informativeness of order flow should be negatively correlated with volume and volatility and positively correlated with market spreads.

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 relationships 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 resort the observations by variable Z_t (which could be bid-ask spread, volatility or trading volume). To avoid being too volatile at the beginning part of the recursive regression, we slightly alter the standard estimation procedure of the recursive least squares regression model. Specifically, we start with the first 50 observations in the first regression. We record the regression coefficient β_1 and the value of variable Z_t of the last observation³⁰. Then we add the next 5 observations into the regressor, run the model again on its first 55 observations and record β_2 and the value of variable Z_t of the last observation again. The core model is run recursively with 5 additional observations entering the regressor each time. In this way we obtain a sequence of coefficients and a sequence of Z_t . We draw the sequence of β s and its 2 standard deviations against Z_t in Figure 2, Figure 3 and Figure 4 for market conditions of spreads, volatility and volume respectively.

Clearly from Figure 2 and Figure 3 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 (but still increasing) after that. In Figure 4 β decreases to some extent at the beginning (except for Yen-Dollar) and becomes stable (but still decreasing) after that. Since the observations become fewer for very large spreads, volatility and volume, the stable part could be caused by the lack of observations. To clarify this ambiguity, in figure 5, figure 6 and figure 7 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

³⁰The starting number of observations are roughly equal to those of one week.

will be added into the 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. Take USD/EUR-spread (the first graph in Figure 5) as an example, 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 at the same time spread will increase to a level of 4.87 basis points. The advantage of this presenting strategy is that it allows us to study how β varies along the recursive process and how Z_t changes along this process simultaneously.

In general, we can see a smoother change of β in all three figures (Figures 5, 6 and 7). But it is still the case that the β increases or decreases more sharply for the first part of observations in all graphs. In figure 5 and figure 6, β increases with bid-ask spread and market volatility respectively for a large proportion of the observations. In figure 7, 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 5 and Figure 6 β completes the shift within a 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. For the total observations, the spread range is from 1 basis point to 50 basis points. The β , however, finishes the shift within a fairly small range of spread from 1.4 basis points to 5 basis points (see USD/EUR in Figure 5). In words, the results from 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 the most of the cases the market is likely to experience.

The visual inspection indicates that the value of β varies substantially and moves outside previously estimated confidence limits in almost all graphs of Figure 5, Figure 6 and Figure 7. From an econometric point of view, this violation indicates the structure imposed in the core model (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 change³¹.

One possible distortion in recursive models is that the price impact of order flow attenuates and converges to the equilibrium pattern as more and more observations enter the regression. In this case, the distortion could blind the real picture of the order flow effect on price at the later stage when spread, volatility or volume is higher. This drawback will be largely overcome by the following window regression model.

³¹We 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.

5.3 Window regression model

In implementation of the window regression, we use the same data set as in the recursive regression model estimation, i.e. the hourly frequency data. For each of the exchange rates, we re-sort the observations by bid-ask spread, volatility or trading volume. In this way 12 databases are generated (each for one exchange rate and one market condition). Then we divide each of the 12 databases into 15 equal-sized non-overlapping windows (sub-samples)³². Take Euro-Dollar/spread as an example, the hourly frequency data for Euro-Dollar are sorted according to the market spread. The re-sorted sample is divided into 15 non-overlapping windows with roughly 131 observations in each window. The observations in the first window have the smallest spread and the observations in the 15th window has the largest spread. Running core model (1) on each window will generate 15 regression coefficients of β . This procedure is repeated for other market measures and other exchange rates³³.

The window regression coefficients of the β s and their 2 standard deviations are drawn against the sequence of regression windows in figure 8, figure 9 and figure 10³⁴. For the purpose of studying the effect of market condition on the price impact of order flow, we deliberately labelled the x-axis with the average values of variable Z_t for those regression windows rather than with the window sequence number. The advantage of this labelling method is to allow us to study how β evolves along the window regression sequence and at the same time to have some idea of what the average values of Z_t are for those windows.

In figure 8 β increases with spread for all exchange rates. In figure 9, β increases with volatility for USD/EUR and JPY/USD but not very much for USD/GBP and GBP/EUR. In figure 10, β decreases with trading volume (except for Dollar-Yen). The results shown in figures 8, 9 and 10 are consistent with the general conclusions from recursive and quintile models: order flow is more informative when spread and volatility is high and is less informative when volume is high. Another interesting result from the window regression is that the β does not change "more sharply" within the first small range of observations as misleadingly shown in recursive model (Figures 5, 6 and 7). On the contrary, it changes even more widely for later observations in the cases of spread (Figure 8) and volatility (Figure 9). In the case of volume (Figure 10), β does become more stable for later observations as shown in the recursive model. This latter point might suggest that price impact

³²With the fixed number of windows for all exchange rates, the window length will be different for different rate since the number of observations for each rate, as shown in Table 1, is different.

³³We also experiment with the overlapped moving window regression and find the results are qualitatively similar as those we reported here.

³⁴In figure 8, 9 and 10 the interesting market conditions are bid-ask spreads, volatility and volume respectively.

of order flow changes only when trading volume is very low and it will remain stable when volume reaches some level.

In order to have a visual impression of the function $\beta(Z)$, β is drawn against spread, volatility and volume in figure 11, figure 12 and figure 13 respectively. We can see clearly that β is an increasing function of spread (Figure 11) and volatility (Figure 12) and a decreasing function of volume (Figure 13) in most cases, even though not in a strict monotonic sense. On further inspection of the graphs in Figures 11, 12 and 13 it is clear that the function $\beta(Z)$ is far from a straight line. This curvature suggests the relationships between β and market condition measuring variable Z_t themselves might be non-linear.

5.4 Interaction model

The interaction model (6) will be evaluated along two dimensions: market conditions and time aggregations. For market conditions, we check the following three important market statistics: bid-ask spread, trading volume³⁵ and market volatility. For time aggregation, we estimate the model for a series of sampling frequencies from 5 minutes to 12 hours. The estimation is based on the 32 databases described in the Data Section. The estimates of model coefficients are presented in table 5, table 6 and table 7³⁶.

In Table 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 most of the sampling frequencies for all four exchange rates (except for a few cases for EUR/USD and GBP/EUR at low frequencies). The rejection of the null is equivalent to the rejection of linearity between order flow and price change. Another important feature we can see in this table is that β_2 is constantly positive for almost all time aggregation levels and all exchange rates. The significant positive β_2 suggests that the order flow tends to be more informative when spread is big and less informative when spread is small. It should be noted that the informativeness pattern of order flow under different market liquidity conditions persists in a fairly large range of time aggregation horizons.

Volatility is considered in Table 6. Similar as in Table 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.

³⁵As in the quintile model we only report the results for number of trades since results from number of quote are more or less the similar.

³⁶The t-value reported in tables 4, 5 and 6 are based on the Newey-West estimator of the coefficient variance-covariance matrix.

The β_2 s, however, are not significant for JPY/USD and USD/GBP and are significant for USD/EUR and GBP/EUR only at relatively high frequencies. This result contrasts with those reported in Table 5 and might suggest that the relationship between order flow informativeness and market volatility is only a EUR phenomenon.

Table 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 5 and Table 6, the estimates of the regression coefficients β_2 are negative in most of the cases (except for USD/EUR at relatively high frequency levels). Negative β_2 indicates that the order flow is less informative when trading volume is higher. Another different point, if compared with Table 5 and Table 6, is that the t-values of β_2 are not significant in most cases in Table 7. The lack of significance might suggest that the informativeness of order flow is less sensitive to market volume than is assumed in the market microstructure literature³⁷.

Overall the results from the interaction model indicate that the linear relationship between the informativeness and market conditions imposed by equation (4) is positive and significant for a large spectrum of time aggregation levels when market condition is measured by spread and is positive and significant only for USD/EUR and GBP/EUR at high frequency level when market condition is measured by volatility. When market condition is measured by volume, the linear relationship between informativeness and volume is not statistically significant for most cases³⁸.

5.5 LSTR model

In the LSTR model, we assume the relationship between order flow (Q_t) and price change (ΔP_t) is not constant but evolves smoothly with the market condition measuring variable Z_t .

Before we go ahead to model estimation, we use the following auxiliary regression as a test of linearity against Smooth Transition Regression model (see Granger and Tersvirta [1997] p.117):

$$\hat{u}_t = \alpha + \beta_1 Q_t + \beta_2 Q_t Z_t + \beta_3 Q_t Z_t^2 + \beta_4 Q_t Z_t^3 + \eta_t \quad (9)$$

where \hat{u}_t is the OLS residual from the core model (1). The null here is $H_0 : \beta_2 = \beta_3 = \beta_4 = 0$.

³⁷A lot of attention has been given to volume in Easley and O'Hara [1992], Admati and Pfleiderer [1988] and Lyons [1996].

³⁸We are aware of precision of this later point because the volume used in this paper is total number of trades, not real trading volume but a proxy for it

Using the hourly frequency data (same as in the recursive model and window regression model) we estimate the above auxiliary model and report the results in Table 8. In nearly all cases (except for JPY/USD-volume) the null is rejected. The high $F - stats$ in Table 8 are in favor of the Smooth Transition Regression (STR) model.

In the LSTR model (7), regression coefficients (θ_1, γ) will 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 method is used to estimate the LSTR model. The estimates based on hourly frequency data are reported in Table 9. 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 USD/GBP. 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 9 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, specially how the relationship evolves with market liquidity, volatility and intensity. To have a visual impression, the smooth transition relationships are drawn in Figures 14, 15 and 16. One of the most interesting points in Figure 14 is that the β in the core model (1) changes quickly within a specific range of bid-ask spreads. For example, the β triples if the spread increases from 2 basis point to 4 basis point for USD/EUR. For JPY/USD, the β shifts quickly from around zero to 0.025 when spreads move from 10 basis points to 20 basis points. The transition feature shown in Figure 14 suggests that the price impact of order flow shifts quickly within a small range of market liquidity conditions.

In contrast to graphs in Figure 14, the graphs in Figure 15 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 14. 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 16, 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 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 USD/EUR, it decreases only a little for JPY/USD and even less for GBP/EUR.

Overall, the transitional features displayed in Figure 14, 15 and 16 of the LSTR model indicate that the relationship between order flow and price movement changes with the market conditions, the relationship shifts rapidly within a small specific range of bid-ask spreads and volume and change more smoothly with volatility. Further, the order flow impact on price is more sensitive to bid-ask spreads than to volume and volatility.

6 Conclusion

This paper presents a systematic methodology to characterize the informational aspects of the price impact of order flow. In particular, we focus 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. The methodology also explores the time-variation of order flow impact with time aggregation.

We use the high frequency data of the FX 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 evidence is neither fully consistent with the prediction from the Easley and O'Hara model nor with that from the Adamati and Pfleiderer model since the former predicts a positive relation between informativeness of order flow and all three market conditions and the latter predicts that informativeness of order flow should be negatively correlated with volatility and volume and positively correlated with market spreads.

We show in the paper that the relationship between order flow and price changes are highly non-linear. The detected relationships of informativeness of order flow and market conditions are stronger for higher sampling frequencies than for lower frequencies and become weak at low sampling frequencies such as 12 hours. We modelled the nonlinearity between order flow and price changes by two alternative ways and they are both statistically significant. Our interaction model can

express the general relationship between the price impact of order flow and the market conditions and the smooth transition model can further capture the transitional feature of price impact of order flow. Finally the results from our non-linear models suggests that the price impact of order flow is more sensitive to bid-ask spreads than to volume and volatility.

Reference

- Admati,A.R.,and P.Pfleiderer,1988, A theory of intraday patters:Volume and price variability, *Review of Financial Studies* 1:3-40.
- Bollerslev,T., and I.Domowitz, 1993, Trading patterns and prices in the interbank foreign exchange market. *Journal of Finance* 48:1421-43.
- 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.
- Cheung,Y., and C.Wong, 2000, A survey study of market practitioners' view on exchange rate dynamics, *Journal of International Economics*, 51:379-400.
- Dagfinn Rime, 2000, Private or Public Information in Foreign Exchange Markets? An Empirical Analysis. Typescript, University of Oslo. July 2000.
- Danielsson,J., and Payne,R., 2000, Real Trading Patterns and Prices in Spot Foreign Exchange Markets, *Journal of International Money and Finance*, forthcoming.
- 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.
- DLP, 2001, the draft Danielsson,,J, Luo,J., and Payne,R. are working on.
- Easley,D., and M.O'Hara, 1992, Time and Process of Security Price Adjustment, *Journal of Finance* 47:577-605.
- Foster,D., and S.Viswanathan, 1990, A theory of intraday variations in volumes, variances and trading costs, *Review of Financial Studies* 3:593-624.
- French,K., and Roll,R., 1986, Stock return variance: The arrival of information and the reaction of traders, *Journal of Financial Economics* 17:99-117.
- Frankel,J., and Ross,A.K., 1995, Empirical Research on Nominal Exchange Rates, in G.Grossman and K.Rogoff(eds.), *Handbook of International Economics*,Elsevier Science: Amsterdam, pp.1689-1730
- 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.
- Goodhart,C., 1988, The Foreign Exchange Market: A Random Walk with a Dragging Anchor, *Economica*, 55:437-460

- Granger,C.W.J., and Tersvirta,T., 1997, Modelling Nonlinear Economic Relationships. Oxford University Press, Oxford, UK.
- Hasbrouck, J., 1991, Measuring the Information Content of Stock Trades, *Journal of Finance*, 46:179-206
- Hsieh,D.A., and Kleidon,A.W., 1996, Bid-Ask Spread in Foreign Exchange Markets:Implications for Models of Asymmetric Information. in J.Frankel et al (eds.), *The Microstructure of Foreign Exchange Markets*, University of Chicago Press, Chicago,IL., p.41-65.
- Isard,P., 1995, *Exchange Rate Economics*,Cambridge University Press: Cambridge, UK.
- Ito,T., Lyons,R., and Melvin,M., 1998, Is there private information in the FX market? The Tokyo experiment, *Journal of Finance* 53:1111-1130.
- Kyle,A.S.,1985, Continuous Auctions and Insider Trading, *Econometrica*, 53:1315-1335.
- Lyons,R., 1995, Test of microstructure hypotheses in the foreign exchange market, *Journal of Financial Economics*. October 1995, 39,321-351.
- Lyons,R., 1996, Foreign Exchange Volume: Sounds or Furry Signifying Nothing?, in J.Frankel et al (eds.), *The Microstructure of Foreign Exchange Markets*, University of Chicago Press, Chicago,IL., p.183-201.
- Lyons, R., 2000, *The Microstructure Approach to Exchange Rates*, *forthcoming*, MIT Press.
- Lyons, R., 2001, Order flow and Exchange rate Dynamics, *Journal of Political Economy*, *forthcoming*.
- Madhavan, A and S.Smidt, 1991, A Bayesian model of intraday specialist pricing, *Journal of Financial Economics*, 30:99-134.
- Meese, R., and K.Rogoff, 1983a, Empirical exchange rate models of the seventies, *Journal of International Economics*, 14:3-24.
- 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.
- Mello,A. 1996, Comments on “Foreign Exchange Volume: Sounds or Furry Signifying Nothing?”, in J.Frankel et al (eds.), *The Microstructure of Foreign Exchange Markets*, University of Chicago Press, Chicago,IL., p.205-206.

McCauley,R., 2001, Comments on “Order flow and exchange rate dynamics”, Market Liquidity: proceedings of workshop held at BIS, April 2001. <http://www.bis.org/publ/bispap02.htm>

Payne, R.1999 Informed trade in spot foreign exchange markets: An empirical investigation, typescript, London School of Economics. January.

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

Yao,J. 1998a, Spread components and dealer profits in the interbank foreign exchange market, New York University, Salomon Centre working paper #S-98-4.

Yao,J. 1998b, Market making in the interbank foreign exchange market, New York University, Salomon Centre working paper #S-98-3.

Table 1: **Summary of aggregated databases**

In each panel of Tab. 1, freq is sampling frequency. obs is the total number of (derived) observations in that database. rtn 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.

USD/EUR(a)								JPY/USD(b)							
freq	obs	rtn	tno	qno	bno	sprd	std	freq	obs	rtn	tno	qno	bno	sprd	std
5m	21155	-0.0011	22	68	11	3.17	0.0127	5m	6139	0.0010	4	15	2	18.89	0.0292
10m	10855	-0.0017	43	133	22	3.35	0.0139	10m	3704	0.0041	6	28	3	19.65	0.0316
20m	5522	-0.0044	84	262	42	3.50	0.0153	20m	2129	0.0014	11	51	5	19.39	0.0351
30m	3653	-0.0054	126	395	64	3.39	0.0159	30m	1166	0.0018	17	83	9	18.24	0.0413
1hr	1966	-0.0092	237	742	120	4.23	0.0185	1hr	1056	-0.0009	24	127	12	22.85	0.0435
2hr	1038	-0.0150	448	1406	226	4.10	0.0213	2hr	716	0.0088	37	210	19	24.76	0.0454
6hr	371	-0.0268	1264	3959	638	3.52	0.0250	6hr	282	0.0296	97	549	50	29.06	0.0569
12hr	190	-0.0424	2443	7642	1234	3.54	0.0267	12hr	147	0.0200	178	1014	91	29.31	0.0634

GBP/EUR(c)								USD/GBP(d)							
freq	obs	rtn	tno	qno	bno	sprd	std	freq	obs	rtn	tno	qno	bno	sprd	std
5m	17008	-0.0003	19	45	10	3.50	0.0129	5m	22935	-0.0005	21	54	11	2.03	0.0079
10m	8700	-0.0006	38	89	20	3.62	0.0135	10m	11730	-0.0010	42	106	21	2.05	0.0082
20m	4420	-0.0018	74	175	39	3.69	0.0141	20m	5939	-0.0020	83	210	42	2.07	0.0086
30m	2930	-0.0032	112	265	58	3.63	0.0146	30m	3973	-0.0030	124	314	63	2.07	0.0089
1hr	1541	-0.0056	214	508	111	3.88	0.0152	1hr	2038	-0.0060	243	615	124	2.08	0.0093
2hr	823	-0.0078	402	955	209	4.06	0.0158	2hr	1046	-0.0113	474	1202	241	2.06	0.0100
6hr	292	-0.0170	1126	2678	587	3.68	0.0161	6hr	370	-0.0304	1345	3400	685	2.06	0.0108
12hr	150	-0.0127	2165	5144	1128	3.64	0.0167	12hr	190	-0.0552	2595	6573	1322	2.06	0.0112

Table 2: Spread effect on price impact of order flow

The model to be estimated is:

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

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 model was estimated for a spectrum of sampling frequencies

Euro-Dollar								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0027	0.0034	0.0041	0.0053	0.0077	0.4730	491.32	2.37
10m	0.0027	0.0033	0.0040	0.0050	0.0077	0.4681	226.64	2.37
20m	0.0026	0.0031	0.0040	0.0052	0.0078	0.4759	128.30	2.37
30m	0.0026	0.0031	0.0039	0.0049	0.0073	0.4707	74.88	2.37
1hr	0.0025	0.0028	0.0040	0.0054	0.0070	0.4584	42.30	2.37
2hr	0.0022	0.0030	0.0039	0.0055	0.0053	0.4160	18.15	2.37
6hr	0.0033	0.0022	0.0039	0.0032	0.0047	0.3173	2.16	2.39
12hr	0.0025	0.0027	0.0030	0.0054	0.0030	0.3104	1.84	2.42
Dollar-Yen								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0064	0.0090	0.0143	0.0152	0.0211	0.0982	21.56	2.37
10m	0.0070	0.0099	0.0140	0.0145	0.0212	0.1209	14.63	2.37
20m	0.0070	0.0091	0.0127	0.0155	0.0202	0.1648	12.01	2.37
30m	0.0078	0.0103	0.0111	0.0137	0.0202	0.2858	11.29	2.37
1hr	0.0077	0.0112	0.0133	0.0141	0.0213	0.2735	10.25	2.37
2hr	0.0096	0.0118	0.0156	0.0132	0.0211	0.2886	5.00	2.38
6hr	0.0091	0.0137	0.0134	0.0166	0.0198	0.4523	3.40	2.39
12hr	0.0082	0.0146	0.0097	0.0092	0.0187	0.5245	3.66	2.43
Euro-Sterling								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0029	0.0036	0.0044	0.0051	0.0069	0.3527	170.09	2.37
10m	0.0031	0.0037	0.0040	0.0049	0.0064	0.3256	59.39	2.37
20m	0.0028	0.0036	0.0035	0.0041	0.0060	0.3027	25.33	2.37
30m	0.0028	0.0031	0.0032	0.0037	0.0055	0.2479	12.32	2.37
1hr	0.0023	0.0029	0.0032	0.0031	0.0056	0.2123	7.94	2.37
2hr	0.0018	0.0016	0.0031	0.0026	0.0052	0.1261	4.72	2.38
6hr	0.0020	0.0005	0.0032	0.0026	0.0004	0.0823	2.76	2.39
12hr	0.0002	0.0000	-0.0003	0.0014	0.0001	0.0156	0.52	2.43
Sterling-Dollar								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0024	0.0029	0.0032	0.0034	0.0043	0.3551	103.62	2.37
10m	0.0024	0.0028	0.0030	0.0031	0.0041	0.3402	45.50	2.37
20m	0.0024	0.0027	0.0027	0.0028	0.0039	0.2963	16.09	2.37
30m	0.0023	0.0023	0.0026	0.0023	0.0035	0.2571	7.87	2.37
1hr	0.0022	0.0018	0.0021	0.0021	0.0034	0.1938	6.06	2.37
2hr	0.0016	0.0018	0.0020	0.0016	0.0034	0.1430	3.55	2.37
6hr	0.0022	0.0015	0.0011	0.0011	-0.0002	0.0864	2.78	2.39
12hr	0.0023	0.0013	0.0008	0.0023	0.0004	0.1198	1.76	2.42

Table 3: Volatility effect on price impact of order flow

The model to be estimated is:

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

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 model was estimated for a spectrum of sampling frequencies

Euro-Dollar								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0026	0.0032	0.0038	0.0046	0.0066	0.4650	405.04	2.37
10m	0.0025	0.0032	0.0037	0.0045	0.0059	0.4582	173.23	2.37
20m	0.0023	0.0030	0.0035	0.0044	0.0058	0.4674	104.22	2.37
30m	0.0024	0.0030	0.0036	0.0042	0.0056	0.4612	57.59	2.37
1hr	0.0022	0.0029	0.0034	0.0044	0.0052	0.4439	28.44	2.37
2hr	0.0022	0.0030	0.0037	0.0044	0.0040	0.3929	7.66	2.37
6hr	0.0028	0.0035	0.0034	0.0036	0.0033	0.3029	0.24	2.39
12hr	0.0028	0.0029	0.0036	0.0036	0.0022	0.2909	0.53	2.42
Dollar-yen								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0048	0.0078	0.0113	0.0158	0.0241	0.1071	37.16	2.37
10m	0.0057	0.0087	0.0121	0.0148	0.0231	0.1272	21.45	2.37
20m	0.0067	0.0093	0.0121	0.0142	0.0192	0.1614	9.83	2.37
30m	0.0075	0.0102	0.0120	0.0143	0.0183	0.2775	7.83	2.37
1hr	0.0085	0.0100	0.0116	0.0144	0.0176	0.2603	5.39	2.37
2hr	0.0084	0.0116	0.0134	0.0138	0.0163	0.2769	2.06	2.38
6hr	0.0112	0.0105	0.0163	0.0143	0.0135	0.4341	1.07	2.39
12hr	0.0077	0.0130	0.0105	0.0145	0.0114	0.4894	0.98	2.43
Euro-Sterling								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0033	0.0037	0.0041	0.0046	0.0062	0.3437	109.82	2.37
10m	0.0032	0.0035	0.0038	0.0045	0.0057	0.3217	46.86	2.37
20m	0.0031	0.0032	0.0035	0.0043	0.0047	0.2944	11.98	2.37
30m	0.0029	0.0032	0.0032	0.0040	0.0038	0.2384	2.98	2.37
1hr	0.0027	0.0028	0.0032	0.0033	0.0035	0.1979	0.90	2.37
2hr	0.0023	0.0023	0.0026	0.0026	0.0023	0.1064	0.11	2.38
6hr	0.0019	0.0016	0.0004	0.0012	0.0031	0.0706	1.83	2.39
12hr	0.0000	0.0000	0.0019	-0.0016	0.0015	0.0604	2.26	2.43
Sterling-Dollar								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0023	0.0028	0.0030	0.0033	0.0040	0.3545	98.48	2.37
10m	0.0022	0.0027	0.0029	0.0032	0.0035	0.3359	26.43	2.37
20m	0.0023	0.0023	0.0028	0.0029	0.0034	0.2947	12.71	2.37
30m	0.0021	0.0024	0.0026	0.0025	0.0028	0.2533	2.86	2.37
1hr	0.0020	0.0022	0.0024	0.0022	0.0021	0.1847	0.35	2.37
2hr	0.0019	0.0023	0.0017	0.0023	0.0015	0.1353	1.19	2.37
6hr	0.0018	0.0018	0.0016	0.0005	0.0006	0.0743	1.56	2.39
12hr	0.0015	0.0018	0.0016	0.0012	0.0012	0.0894	0.17	2.42

Table 4: Volume effect on price impact of order flow

The model to be estimated is:

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

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 model was estimated for a spectrum of sampling frequencies

Euro-Dollar								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0070	0.0045	0.0040	0.0040	0.0042	0.4278	34.75	2.37
10m	0.0073	0.0044	0.0037	0.0037	0.0040	0.4303	31.65	2.37
20m	0.0076	0.0039	0.0037	0.0035	0.0040	0.4361	21.89	2.37
30m	0.0066	0.0033	0.0034	0.0034	0.0039	0.4356	13.47	2.37
1hr	0.0071	0.0042	0.0032	0.0034	0.0037	0.4206	7.57	2.37
2hr	0.0054	0.0041	0.0030	0.0034	0.0035	0.3795	1.91	2.37
6hr	0.0022	0.0043	0.0046	0.0032	0.0026	0.3186	2.34	2.39
12hr	0.0040	0.0019	0.0044	0.0036	0.0022	0.3113	1.90	2.42
Yen-Dollar								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0057	0.0141	0.0128	0.0146	0.0133	0.0871	2.76	2.37
10m	0.0089	0.0120	0.0193	0.0145	0.0126	0.1100	3.14	2.37
20m	0.0137	0.0158	0.0149	0.0131	0.0120	0.1473	0.86	2.37
30m	0.0144	0.0146	0.0120	0.0129	0.0128	0.2588	0.31	2.37
1hr	0.0103	0.0177	0.0141	0.0126	0.0130	0.2482	1.06	2.37
2hr	0.0093	0.0161	0.0131	0.0188	0.0111	0.2849	4.06	2.38
6hr	0.0137	0.0154	0.0151	0.0118	0.0132	0.4290	0.45	2.39
12hr	0.0099	0.0115	0.0128	0.0154	0.0094	0.4957	1.43	2.43
Euro-Sterling								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0055	0.0046	0.0044	0.0043	0.0044	0.3275	4.48	2.37
10m	0.0058	0.0041	0.0041	0.0041	0.0042	0.3089	5.59	2.37
20m	0.0056	0.0041	0.0037	0.0041	0.0036	0.2901	5.27	2.37
30m	0.0053	0.0037	0.0036	0.0036	0.0032	0.2395	4.12	2.37
1hr	0.0064	0.0031	0.0033	0.0034	0.0027	0.2090	6.34	2.37
2hr	0.0034	0.0041	0.0025	0.0031	0.0017	0.1195	3.16	2.38
6hr	0.0058	0.0023	0.0017	0.0020	0.0008	0.0862	3.07	2.39
12hr	0.0027	0.0008	0.0010	0.0001	0.0000	0.0207	0.71	2.43
Sterling-Dollar								
Freq	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	R^2	F-stats	F-Critical
5m	0.0032	0.0031	0.0032	0.0031	0.0032	0.3434	0.18	2.37
10m	0.0032	0.0029	0.0029	0.0029	0.0031	0.3301	0.62	2.37
20m	0.0030	0.0028	0.0028	0.0028	0.0028	0.2887	0.23	2.37
30m	0.0029	0.0026	0.0026	0.0025	0.0025	0.2516	0.57	2.37
1hr	0.0027	0.0027	0.0024	0.0021	0.0020	0.1868	1.68	2.37
2hr	0.0029	0.0021	0.0019	0.0020	0.0017	0.1342	0.87	2.37
6hr	0.0029	0.0017	0.0005	0.0012	0.0012	0.0710	1.23	2.39
12hr	-0.0008	0.0014	0.0012	0.0023	0.0016	0.1104	1.26	2.42

Table 5: Linear relationship between the informativeness of order flow 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.

Freq	USD/EUR(a)					JPY/USD(b)				
	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	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
2hr	0.0022	5.31	0.0005	2.58	0.3979	0.0104	8.33	0.0001	3.04	0.2815
6hr	0.0025	4.05	0.0002	1.33	0.3041	0.0084	5.27	0.0002	3.10	0.4565
12hr	0.0026	2.69	0.0002	0.62	0.2840	0.0081	4.24	0.0001	2.29	0.4925

Freq	GBP/EUR(c)					USD/GBP(d)				
	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	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
2hr	0.0004	0.56	0.0006	2.90	0.1205	0.0001	0.18	0.0009	2.25	0.1364
6hr	0.0007	0.32	0.0003	0.39	0.0480	0.0015	4.15	-0.0001	-2.09	0.0604
12hr	-0.0024	-1.33	0.0008	1.39	0.0113	0.0027	6.15	-0.0005	-6.28	0.1033

Table 6: **Linear relationship between the informativeness of order flow 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.

Freq	USD/EUR(a)					JPY/USD(b)				
	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	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
2hr	0.0034	19.81	0.0045	0.59	0.3753	0.0130	15.19	0.0086	0.51	0.2691
6hr	0.0038	11.06	-0.0180	-1.07	0.3064	0.0135	9.64	-0.0009	-0.07	0.4253
12hr	0.0042	4.52	-0.0462	-1.00	0.3025	0.0110	6.51	0.0122	0.54	0.4768

Freq	GBP/EUR(c)					USD/GBP(d)				
	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	$\hat{\beta}_2$	$t(\hat{\beta}_2)$	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
2hr	0.0040	3.73	-0.0978	-1.33	0.1187	0.0018	5.02	0.0101	0.27	0.1315
6hr	-0.0011	-0.62	0.1565	1.44	0.0947	0.0017	2.38	-0.0506	-0.64	0.0621
12hr	-0.0017	-0.61	0.1140	0.66	0.0209	0.0029	4.68	-0.1259	-2.38	0.1149

Table 7: **Linear relationship between the informativeness of order flow 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 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.

USD/EUR(a)						JPY/USD(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	
2hr	0.0037	7.16	0.0000	-0.45	0.3751	0.0164	9.04	-0.0046	-2.16	0.2723	
6hr	0.0044	5.10	-0.0001	-1.34	0.3033	0.0158	8.55	-0.0016	-1.42	0.4294	
12hr	0.0034	3.24	0.0000	-0.24	0.2829	0.0143	6.46	-0.0010	-1.19	0.4799	

GBP/EUR(c)						USD/GBP(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	
2hr	0.0050	6.77	-0.0004	-2.81	0.1328	0.0026	6.77	-0.0001	-1.49	0.1350	
6hr	0.0031	1.70	-0.0001	-0.81	0.0552	0.0022	2.69	-0.0001	-1.14	0.0660	
12hr	0.0017	0.76	0.0000	-0.64	0.0068	0.0009	0.72	0.0000	0.38	0.0872	

Table 8: **Auxiliary regression: a test of linearity against STR model**

The model to be estimated is:

$$\hat{u}_t = \alpha + \beta_1 Q_t + \beta_2 Q_t Z_t + \beta_3 Q_t Z_t^2 + \beta_4 Q_t Z_t^3 + \eta_t$$

where \hat{u}_t is the OLS residual from the core model (1). 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 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 null is $H_0 : \beta_2 = \beta_3 = \beta_4 = 0$. The estimation is based on hourly sampling frequency. The number in the bracket is t -value and critical F-value for 5% significance level is 2.6.

Spread					
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	F-value(2.6)
Euro-Dollar	-4.2546E-03 (-11.52)	2.1426E-03 (11.2)	-1.4928E-04 (-8.23)	2.6452E-06 (6.18)	55.07
Dollar-Yen	-9.0723E-03 (-5.91)	5.6617E-04 (5.97)	-3.7044E-06 (-3.57)	5.3571E-09 (2.51)	17.18
Euro-Sterling	-2.4889E-03 (-1.86)	8.3053E-04 (1.12)	-1.6139E-05 (-0.14)	1.6867E-09 (0.00)	10.70
Sterling-Dollar	9.5424E-03 (2.22)	-1.3347E-02 (-2.45)	5.6365E-03 (2.54)	-6.9521E-04 (-2.42)	7.30
Volatility					
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	F-value(2.6)
Euro-Dollar	-2.6498E-03 (-9.44)	2.2821E-01 (9.48)	-2.8724E+00 (-7.7)	8.0851E+00 (6.78)	35.82
Dollar-Yen	-4.4801E-03 (-3.37)	1.4906E-01 (3.93)	-4.7964E-01 (-2.56)	3.7775E-01 (1.78)	8.35
Euro-Sterling	-3.9574E-03 (-3.89)	4.9041E-01 (4.40)	-1.5911E+01 (-5.12)	1.3829E+02 (5.86)	14.55
Sterling-Dollar	1.5112E-03 (2.15)	-3.2970E-01 (-2.69)	1.8164E+01 (3.23)	-2.2846E+02 (-3.20)	6.35
Volume					
	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	F-value(2.6)
Euro-Dollar	5.9606E-03 (6.59)	-5.3178E-05 (-6.89)	1.3700E-07 (6.79)	-1.0475E-10 (-6.58)	15.97
Dollar-Yen	2.2311E-03 (0.43)	-1.0742E-04 (-0.26)	1.6258E-06 (0.18)	-1.0640E-08 (-0.17)	0.37
Euro-Sterling	4.9224E-03 (4.24)	-4.2224E-05 (-3.79)	1.1338E-07 (3.52)	-9.6821E-11 (-3.51)	11.07
Sterling-Dollar	2.6911E-03 (3.58)	-2.3029E-05 (-3.56)	5.9027E-08 (3.59)	-4.6005E-11 (-3.73)	8.15

Table 9: **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}
Euro-Dollar	-0.0188 (-5.30)	0.0042 (35.42)	0.0201 (1.87)	0.0048 (5.78)	8.5287 (3.27)	2.70
Dollar-Yen	-0.0020 (-0.26)	0.0113 (14.76)	-0.0458 (-1.11)	0.0285 (3.86)	2.1476 (2.48)	14.00
Euro-Sterling	-0.0427 (-7.81)	0.0028 (15.82)	0.0906 (2.39)	0.0071 (2.36)	0.9801 (1.97)	2.70
Sterling-Dollar	-0.0086 (-1.44)	0.0034 (10.67)	0.0200 (1.36)	0.0031 (3.85)	1.4006 (2.03)	3.00
Volatility						
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\gamma}$	\hat{c}
Euro-Dollar	-0.0165 (-4.13)	0.0031 (28.20)	-0.0141 (-0.93)	0.0046 (7.60)	5.7683 (4.35)	0.011
Dollar-Yen	0.0117 (0.83)	0.0067 (3.17)	-0.0902 (-1.76)	0.0259 (5.13)	4.8007 (2.75)	0.006
Euro-Sterling	-0.0313 (-6.49)	0.0030 (17.12)	-0.0187 (-1.70)	0.0009 (2.13)	33.0713 (0.48)	0.012
Sterling-Dollar	-0.0136 (-4.11)	0.0020 (13.84)	-0.0900 (-0.41)	0.0081 (0.42)	0.1736 (0.36)	0.006
Volume						
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\gamma}$	\hat{c}
Euro-Dollar	0.0134 (1.14)	0.0130 (7.23)	-0.0741 (-2.87)	-0.0189 (-5.27)	4.1527 (5.10)	10
Dollar-Yen	-0.0098 (-1.30)	0.0141 (12.85)	-0.0050 (-0.08)	-0.0050 (-0.44)	0.5887 (0.29)	18
Euro-Sterling	-0.0068 (-0.43)	0.0081 (4.40)	-0.0724 (-1.90)	-0.0107 (-2.92)	1.9924 (3.73)	2
Sterling-Dollar	-	-	-	-	-	-

Figure 1: Intraday trading pattern for spread, volume and volatility

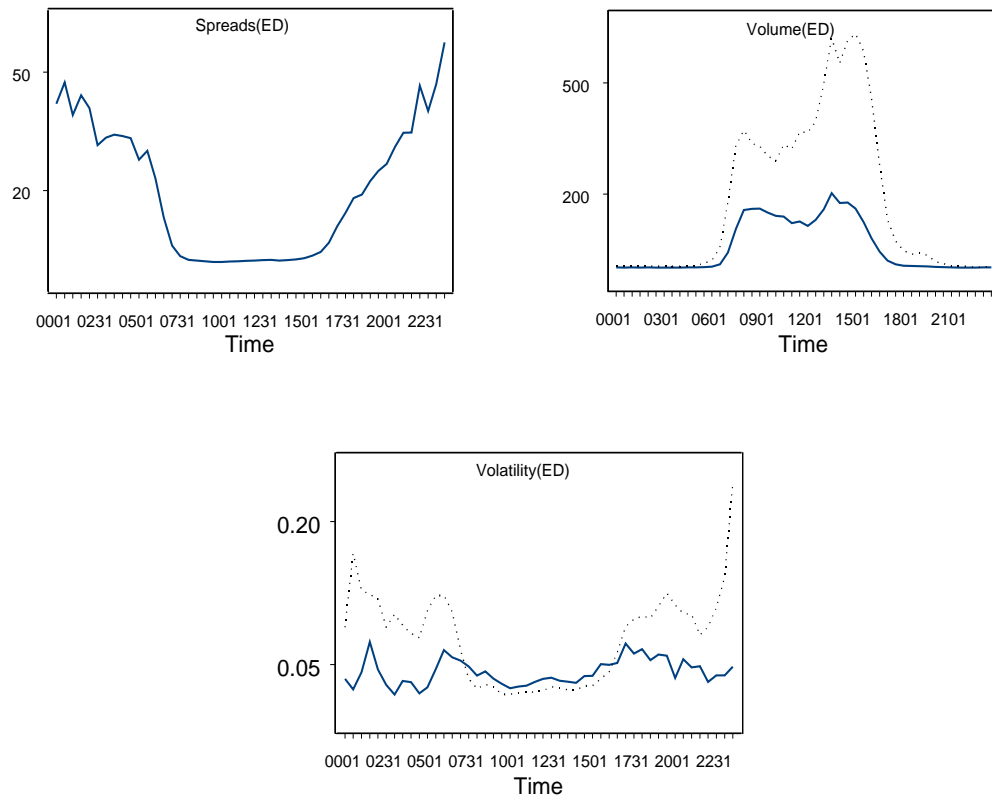
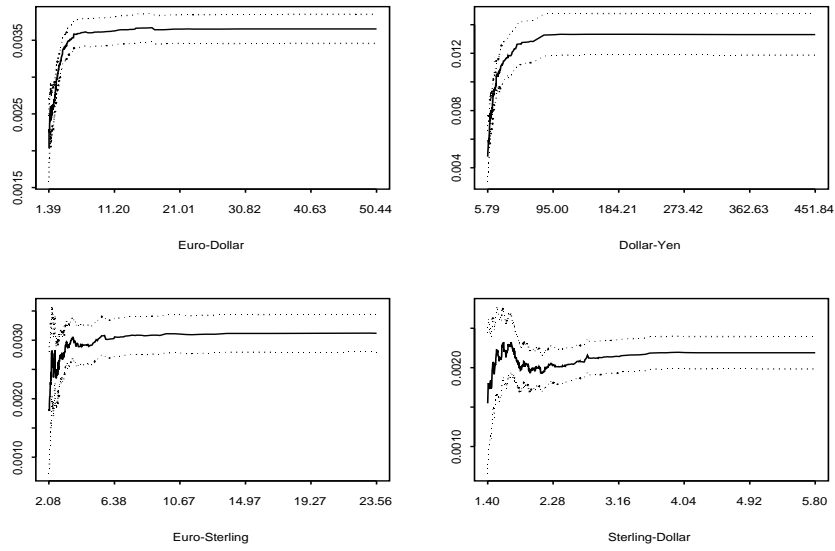


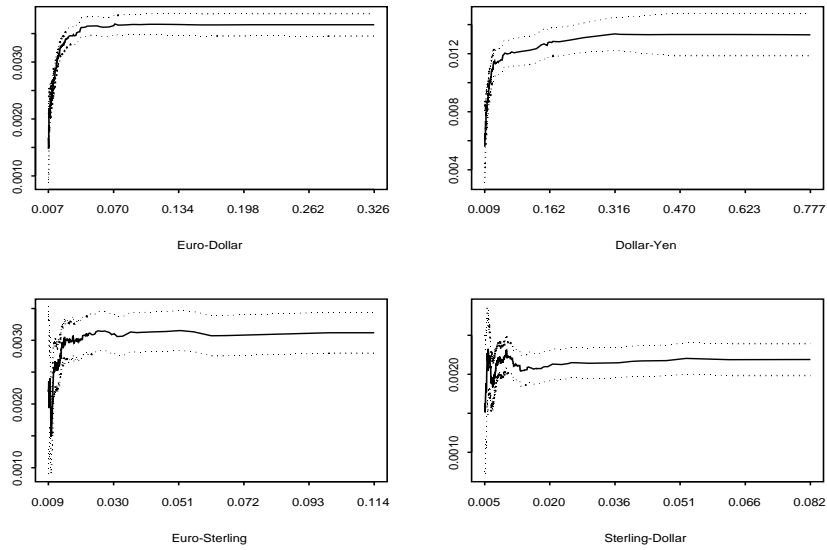
Figure 1 is drawn on Euro-Dollar. In volume graph, the solid line is for total number of trades. The dotted line is total number of quotes. In volatility graph, the solid line is for return volatility calculated from trade prices and the dotted line is return volatility calculated from mid-quote prices.

Figure 2: Function $\beta(Z_t = \text{spread})$ from recursive model



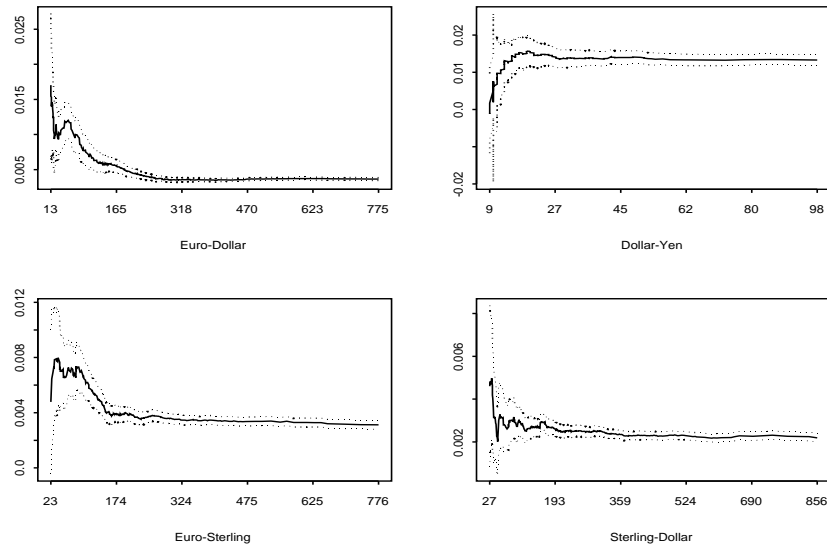
The regression coefficients of the recursive regression model are drawn against market spread. Spreads are measured in basis points. The dotted line is the two standard deviation confidence bounds.

Figure 3: Function $\beta(Z_t = \text{volatility})$ from recursive model



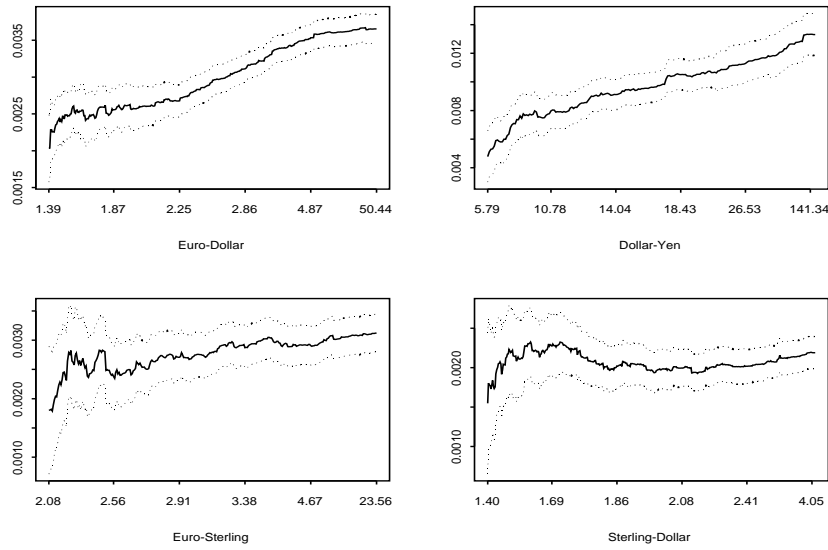
The regression coefficients of the recursive regression model are drawn against volatility. Volatility is computed as return standard deviation. The dotted line is the two standard deviation confidence bounds.

Figure 4: Function $\beta(Z_t = volume)$ from recursive model



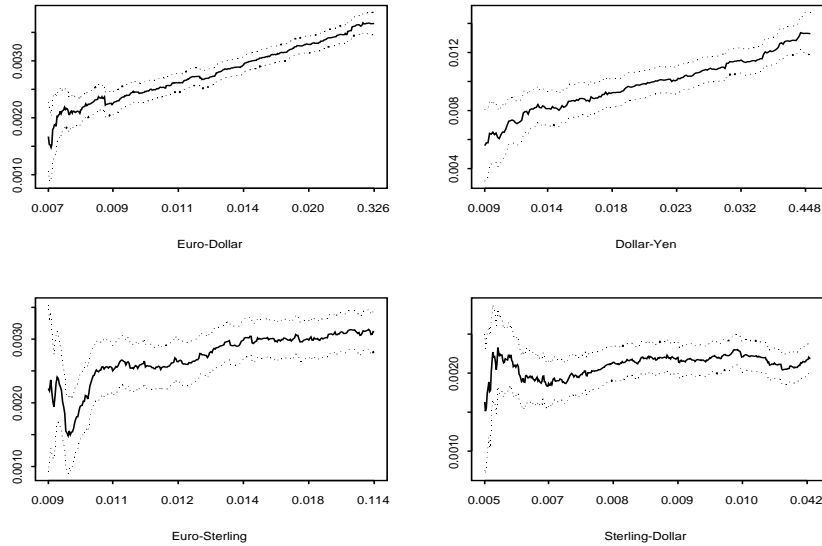
The regression coefficients of the recursive regression model are drawn against volume. Volume is proxied by the total number of trades. The dotted line is the two standard deviation confidence bounds.

Figure 5: Evolution of β along spread in recursive model



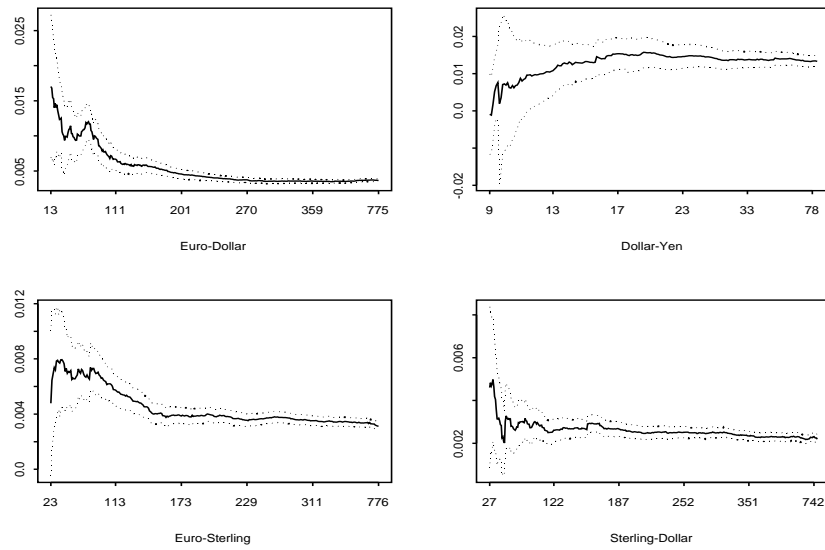
β s are drawn along the recursive regression process. Note: x-axis is deliberately labelled with the spread of the last observation in that regression rather than with the sequence number of regression. Spreads are measured in basis points. The dotted line is the two standard deviation confidence bounds.

Figure 6: Evolution of β along volatility in recursive model



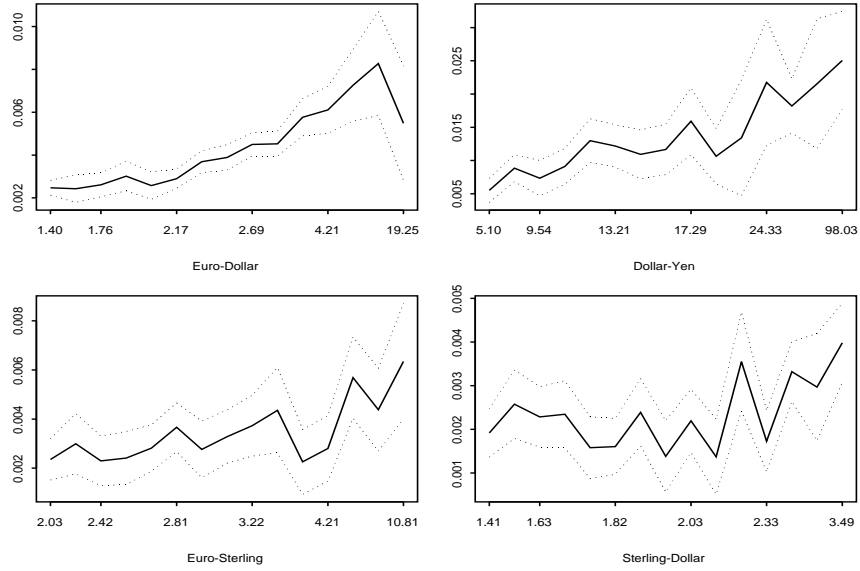
β s are drawn along the recursive regression process. Note: x-axis is deliberately labelled with the volatility of the last observation in that regression rather than with the sequence number of regression. Volatility is computed as return standard deviation. The dotted line is the two standard deviation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

Figure 7: Evolution of β along volume in recursive model



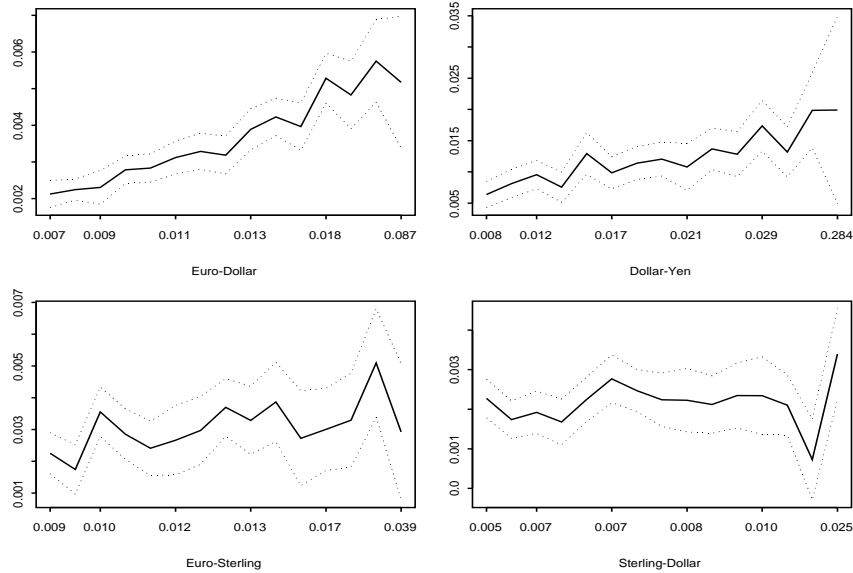
β s are drawn along the recursive regression process. Note: x-axis is deliberately labelled with the volume of the last observation in that regression rather than with the sequence number of regression. Volume is proxied by the total number of trades. The dotted line is the two standard deviation confidence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

Figure 8: Evolution of β along spread in window regression model



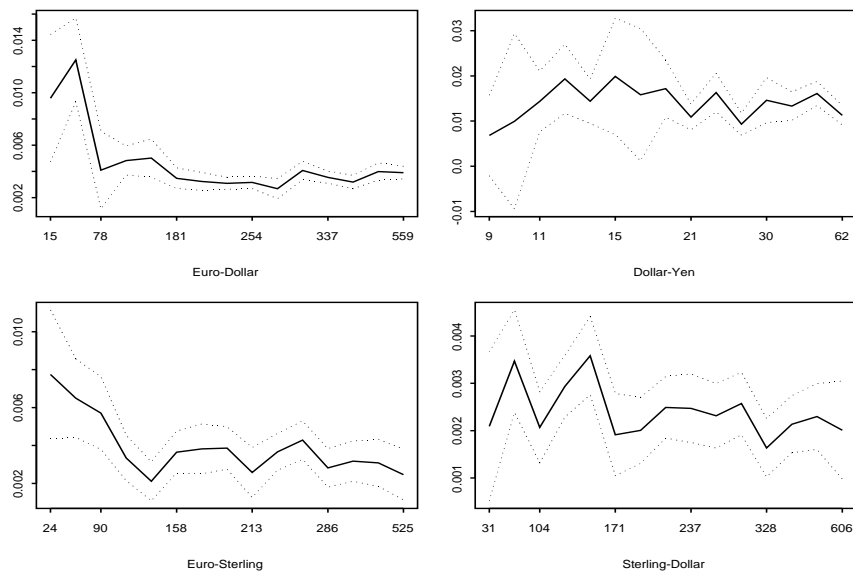
β s are drawn along the window regression process. Note: x-axis is deliberately labelled with the average spread for that window rather than with the sequence number of the window. Spreads are measured in basis points. The dotted line is the two standard deviation confidence bounds.

Figure 9: Evolution of β along volatility in window regression model



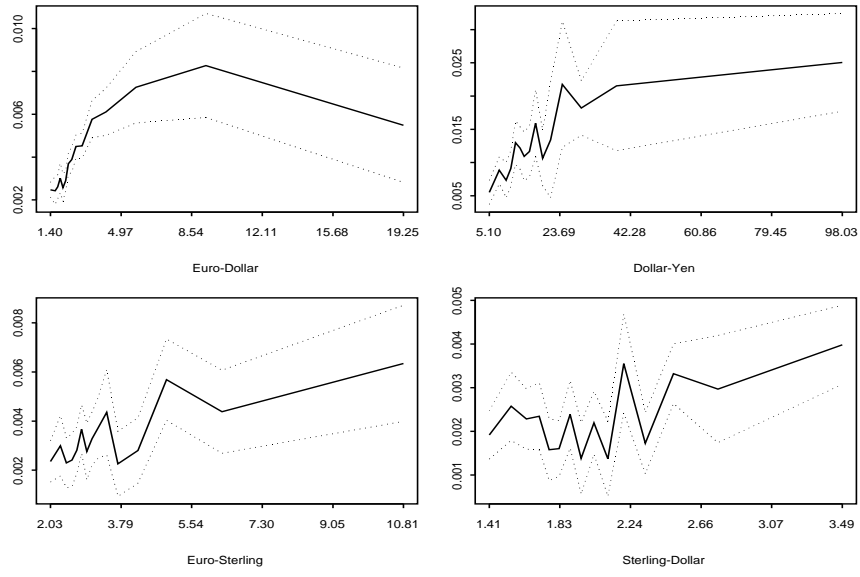
β s are drawn along the window regression process. Note: x-axis is deliberately labelled with the average volatility for that window rather than with the sequence number of the window. Volatility is computed as return standard deviation. The dotted line is the two standard deviation confidence bounds.

Figure 10: Evolution of β along volume in window regression model



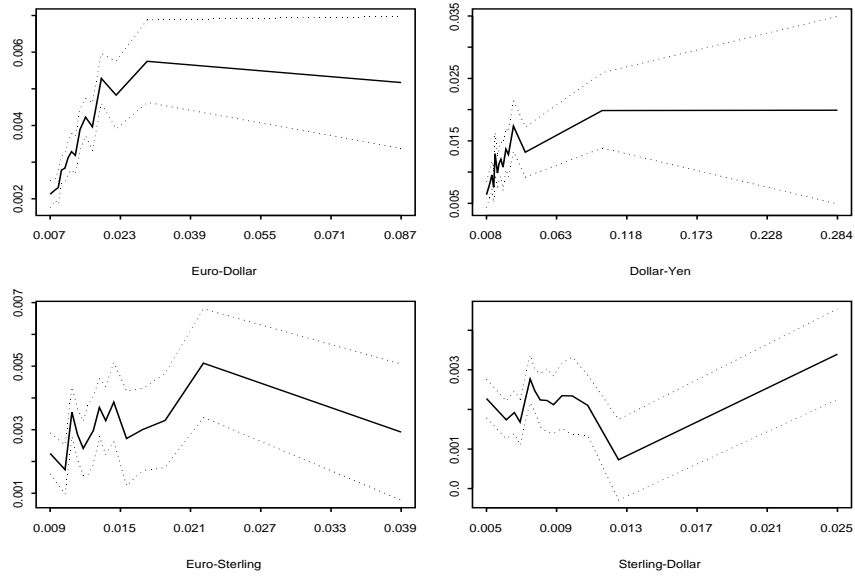
β s are drawn along the window regression process. Note: x-axis is deliberately labelled with the average volume for that window rather than with the sequence number of the window. Volume is proxied by the total number of trades. The dotted line is the two standard deviation confidence bounds.

Figure 11: Function $\beta(Z_t = \text{spread})$ from window regression model



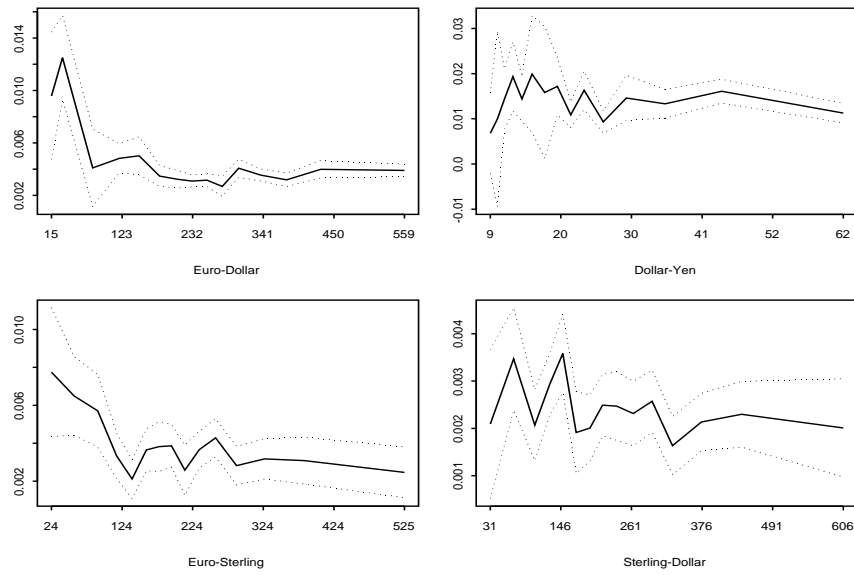
The regression coefficients of the window regression model are drawn against market spread. Spreads are measured in basis points. The dotted line is the two standard deviation confidence bounds.

Figure 12: Function $\beta(Z_t = \text{volatility})$ from window regression model



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

Figure 13: Function $\beta(Z_t = volume)$ from window regression model



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

Figure 14: Shift of price impact of order flow with spread

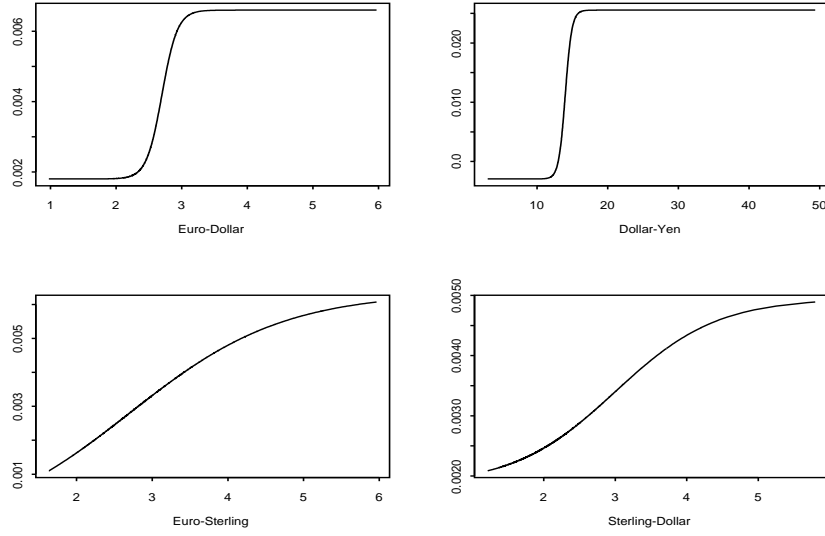


Fig.14 is a graphic representation of LSTR model when transition variable Z_t is spread. In this figure, $\beta_1 + \theta_1 * F(Z_t)$ in model (6) is drawn against Z_t with the estimated parameters $\hat{\beta}_1, \hat{\theta}_1, \hat{\gamma}, \hat{c}$.

Figure 15: Shift of price impact of order flow with volatility

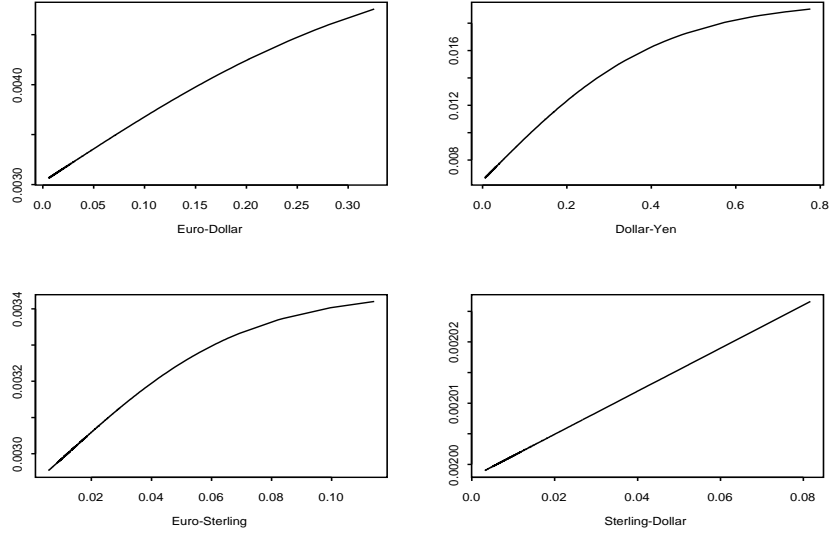


Fig.15 is a graphic representation of LSTR model when transition variable Z_t is volatility. In this figure, $\beta_1 + \theta_1 * F(Z_t)$ in model (6) is drawn against Z_t with the estimated parameters $\hat{\beta}_1, \hat{\theta}_1, \hat{\gamma}, \hat{c}$.

Figure 16: Shift of price impact of order flow with volume

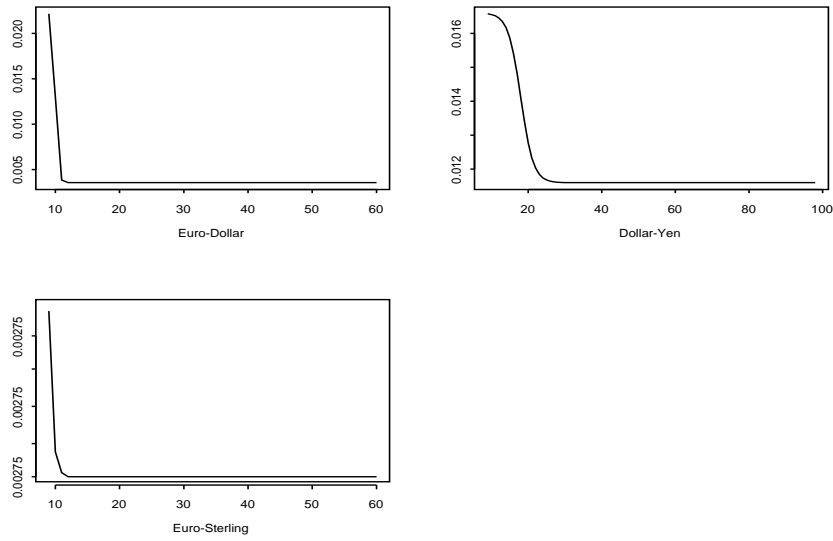


Fig.15 is a graphic representation of LSTR model when transition variable Z_t is volume. In this figure, $\beta_1 + \theta_1 * F(Z_t)$ in model (6) is drawn against Z_t with the estimated parameters $\hat{\beta}_1, \hat{\theta}_1, \hat{\gamma}, \hat{c}$.