Informed Trade in Spot Foreign Exchange Markets: an Empirical Investigation

Richard Payne*

January 11, 1999

Abstract

This paper extends the extant evidence on asymmetric information in inter-dealer spot FX trades. Using a new sample of USD/DEM data derived from an electronic FX brokerage and covering one trading week, we employ the VAR framework contained in Hasbrouck (1991a) and Hasbrouck (1991b) to test for the existence of private information effects of trading on prices. Our basic results confirm the existence of private information on FX markets, indicating that asymmetric information costs account for around 60% of the half-spread. Further, 40% of all permanent price variation is shown to be due to the information contained in order flow. We also document strong time-of-day effects in the price impact of trades. The price impact is shown to be inversely related to volume and liquidity as the model of Admati and Pfleiderer (1988) would suggest. Finally, we extend the linear VAR of Hasbrouck (1991a) to account for the impact of market activity on the response of quotes to trades. Estimates of the non-linear VAR framework indicate a decreasing, non-linear relationship between the price impact of trades and proxies for liquidity trader activity.

*I would like to acknowledge the financial support of the Financial Markets Group at LSE and the ESRC (UK). ...

1 Introduction

Spot markets for foreign exchange are widely characterised as the most efficient of asset markets. Transaction costs are small, volume is huge and information is impounded in prices extremely rapidly. Further, it is widely believed that information asymmetries in these markets are of minor importance. Bessembinder (1994), for example, argues that ”[w]hereas equity values depend on both macroeconomic and firm-specific information, currency valuation depends primarily on macroeconomic information, which may imply a reduced potential for market-maker losses to better-informed traders.”

In recent years, however, a few empirical papers have emerged which challenge the notion that spot FX markets are free from asymmetric information problems. Lyons (1995) and Yao (1997) use data from single FX dealers to demonstrate that spreads contain an asymmetric information component. Both use extensions of the empirical model presented in Madhavan and Smidt (1991). Furthermore, these papers provide strong evidence of inventory control by their dealers.[1] Lyons (1996) extends his prior work by examining the role of *time* in the relationship between trades and quotes. He finds that trades occurring in periods when the market is active convey less information than those when the market is quiet. This is interpreted as consistent with his ‘hot potato hypothesis’ by which high inter-dealer volumes are generated more by inventory rebalancing than exploitation of information. A final paper which tackles private information in FX markets is Ito, Lyons, and Melvin (1998). This study uses the variance ratio framework of French and Roll (1986) to characterise the effect of the removal of restrictions on lunch-time inter-dealing trading in Tokyo on USD/JPY volatility. The results obtained are consistent with the existence of private information in this market segment.

Theoretical models focussing on private information in inter-dealer spot FX markets can be found in Lyons (1995) and Perraudin and Vitale (1996). The mechanism emphasised in these studies relies heavily on the lack of transparency in actual spot FX markets.[2] In particular, trade between dealers and non-dealer customers is entirely opaque. Assuming that non-dealer order flow signals the future evolution of the ex-

[1] This evidence is far stronger than that obtained in equity market studies. Compare with Madhavan and Smidt (1991), for example.

[2] Similar mechanisms are employed in Vogler (1993) and the model of the London Stock Exchange presented in Naik, Neuberger, and Viswanathan (1994).

change rate, this implies that dealers observing a large portion of customer trade will be better informed than those with fewer customer contacts.[3] The better informed can then exploit their advantage through trading activity on the inter-dealer market. A first implication of this is that inter-dealer spreads will contain a component which covers asymmetric information costs as in the standard theoretical models of Glosten and Milgrom (1985) and Easley and O'Hara (1987). A further implication is that inter-dealer quotes will respond permanently to unexpected transaction activity, due to the possibility that a trade is information motivated. It is the latter implication which is the focus of the empirical methodology used in this analysis.

The current study provides a contribution to the literature on private information effects in spot FX trading. One of the main innovations of this work is the use of a new data set on inter-dealer USD/DEM trades, drawn from an electronic brokerage known as D2000-2.[4] Whilst earlier work in this area has employed data based on the activity of a single dealer, the data drawn from D2000-2 reflect the activity of multiple dealers. As such, these data provide a wider coverage of activity on the inter-dealer market and hence yield more broadly based results. Further, the D2000-2 data can be used to construct proxies for the liquidity of the FX market as a whole. D2000-2 operates as a closed electronic order book and every limit and market order entered onto the system are available from the data. Liquidity measures can be constructed from the subsidiary data and used as conditioning variables in the analysis of the effects of private information.

The specific research questions treated in our analysis are as follows. Primarily we examine whether trading activity on D2000-2 can be characterised as information based. This is achieved using the reduced form VAR model for trades and quote revisions introduced in Hasbrouck (1991a). Further, whilst earlier studies have demonstrated that at least some FX trades carry information, none has computed the aggregate impact of such information. Use of the variance decomposition presented in Hasbrouck (1991b) allows us to calculate the proportion of all information entering the quotation process via order flow and hence address this issue.

We also examine variations in the information content of trades across the GMT time-

[3]The most stark example of an informative customer trade would be a dealer trading with (or on behalf of) a central bank. See Peiers (1997). Unfortunately, the data used in this study contains no information on customer trades such that a direct test of this hypothesis is not available.

[4]We would like to thank Reuters for the provision of the D2000-2 data. Previous studies also using D2000-2 data are Goodhart, Ito, and Payne (1996) and Goodhart and Payne (1996).

of-day. Theoretical contributions such as Admati and Pfleiderer (1988) and Foster and Viswanathan (1993) predict correlations in the intra-day variation of transactions costs, volume and the intensity of informed trade which we evaluate via a subsample analysis of the D2000-2 data. Finally, we examine the stability of the VAR structure across different levels of market activity and liquidity via an arranged VAR technique. This leads to a fully parameterised non-linear VAR representation in which the impact of trades on quotes depends on the level of a weakly exogenous proxy for market activity.

Our main results are as follows. First, our full sample estimates imply that over 60% of the D2000-2 spread can be characterised as compensation for informed trade. An unexpected market buy, for example, leads to an upwards equilibrium quote revision of 1 basis point. Further, we estimate that around 40% of all information entering the quotation process does so through order flow, a figure which is comparable in magnitude to equivalent measures from equity market studies.

The impact of informed trade on D2000-2 quotes is shown to depend strongly on the level of market liquidity. When the D2000-2 order book is relatively thin and volume low (i.e. from the late GMT afternoon to the early GMT morning,) unexpected trading activity has a much larger effect on quotes than in peak trading periods. This result is in line with the empirical analysis of Lyons (1996) and the theoretical predictions of Admati and Pfleiderer (1988) who argue that this correlation should be observed due to the actions of discretionary liquidity traders. A more direct examination of the relationship between informed trade and market conditions is obtained from the arranged VAR results. These demonstrate that the impact of informed trade on quotes and the level of liquidity trader activity are negatively and non-linearly related, a result which is reinforced by the estimates from an explicitly non-linear VAR structure.

The rest of the paper is set out as follows. Section 2 introduces the basic features of the D2000-2 data set and Section 3 details the empirical methodology employed in the current study. Section 4 presents the empirical findings from the VAR estimations. Finally, Section 5 concludes and presents ideas for further work.

2 The Dealing System and the Data

The major proportion of all spot FX trade is inter-dealer.[5] Until recently, all of this segment of trade (including both direct inter-dealer and brokered trade) was carried out over the telephone. One implication of this was that, aside from the tri-annual BIS surveys of FX market activity, no consolidated source of FX volumes was available. Further, the order flow information available to dealers themselves was limited. Quote information was available via a number of screen based systems and reports of brokered trades were broadcast over intercom systems, but indications of market wide order flow were not available.

In the early 1990's, however, a shift away from telephone based trade occurred with the introduction of two electronic broking systems. The first of these is that run by the EBS consortium and the second the D2000-2 system operated by Reuters.[6] These systems have grown relatively quickly, driving the voice brokered portion of trade down considerably. In June 1997 EBS claimed to handle 37% of brokered trade in London, with Reuters' market share commonly assumed to be similar.

The data used in this study cover all DEM/USD trade on D2000-2 over the week 6th-10th October 1997. Around 30,000 transactions occurred during this time with total volume approaching $60bn.

2.1 The D2000-2 Dealing System

D2000-2 functions as an electronically governed continuous auction. Liquidity is supplied to the system via limit orders to buy and sell currency. Liquidity is drained from the system in two ways. First through market orders to buy and sell and second through the crossing of limit orders.[7]

Transaction consummation is governed by strict rules of price and time priority subject to one proviso. Participants in the system must bilaterally negotiate credit lines if they are to transact. This implies that a dealer entering a market buy, for example,

[5] In the BIS survey of foreign exchange market activity carried out in April 1995, the reported inter-dealer share is 64%.

[6] Reuters also operates a system called D2000-1 which allows electronic bilateral communication between individual dealers.

[7] By this we mean a situation in which the D2000-2 order book contains a limit buy with price greater than or equal to that of a limit sell. In this case the system automatically transacts the overlapping quantity.

may find the best priced limit sell unavailable to him as a credit channel has not been opened with the initiator of the limit order. Hence market orders sometimes execute outside the inside spread.

A subscriber to D2000-2 sees the following items on the trading display. First there is an indication of the most competitive limit buy and sell prices in the system along with quantity available at those prices. Data on limit orders outside the touch are not displayed. The screen also indicates the last transaction which occurred in a given currency pair, detailing price and volume. The above information is simultaneously available for up to 6 exchange rates.

From the DEM/USD data set supplied by Reuters, one can reconstruct all of the information available to subscribers. Additionally, the data set contains information on every subsidiary limit order entered on D2000-2. Hence, it is possible to compute the entire excess demand and supply schedules for foreign exchange at each point during the week long sample. The data on subsidiary prices are not employed in the current study and are analysed in Danielsson and Payne (1998).

2.2 The Data Set

The raw D2000-2 data feed consists of just over 130,000 data lines. Each line contains a 10 fields detailing the type of event to which it refers, timestamps with a one hundredth of a second granularity, price and quantities. Just over 100,000 of these data lines refer to limit order entries (with timestamps for entry and exit times, a buy/sell indicator, quantity available, quantity traded and price.) The rest of the data consists of market order entries. The market order lines give the quantity transacted and price of transaction, a timestamp and whether buyer or seller initiated.

For the purposes of this study, an event time data set including the following variables was constructed; the average of best bid and offer quotes (the midquote), a signed transaction indicator variable, signed volume, the inside spread, aggregate bid and offer order book size and the number of bid and offer orders outstanding.[8] Table 1 presents descriptive statistics for the market activity variables for seven non-overlapping subsamples of all trading days. The first six of these subsamples consist

[8] We define an event as a revision in the best bid or best offer or the occurrence of a transaction. Order book size is defined as the aggregate quantity outstanding across all bid/offer limit orders. Transactions indicators were constructed both including and excluding crosses. When included, crosses are signed by treating the latest entering limit order as the aggressor.

of observations from 2 hour segments of each day, covering the period from 6 to 18 GMT. The seventh subsample represents data from all GMT overnight periods, 18 to 6 GMT.

The main feature of the statistics presented in Table 1 is that all series have strong intra-day seasonal patterns.[9] The number of limit orders outstanding and aggregate size on the book both follow an inverted U-shape across the GMT trading day. The transaction frequency and volume data show a similar pattern, aside from a lull in activity in the period from 10 to 12 GMT. Hence, focussing on the hours between 6 and 18 GMT, the seasonal in D2000-2 transaction activity corresponds with those found in studies of equity market volumes.[10] Market liquidity as measured by the percentage spread follows the inverse pattern to transaction activity. It is particularly clear that D2000-2 is extremely illiquid between the hours of 18 and 6 GMT. Spreads are very high and book size extremely limited. Conversely, during the complementary portion of the trading day, spreads are very tight (with a model value of one basis point) and the D2000-2 order book is very deep.

A graphical representation of the seasonal patterns in D2000-2 activity is given in Figures 1 to 3 for limit orders outstanding, market order frequency and percentage spreads. These plots were constructed using a 20 second calendar time sampling of the data and reinforce the numerical evidence given in Table 1; D2000-2 is extremely active during European and North American trading hours but has low volume during Asian and Pacific trading.[11]

Table 2 contains summary statistics for the percentage midquote return for our seven subsamples. Again, the effects of the seasonal in D2000-2 activity are apparent. Examining the return variance across the seven subsamples, it is clear that D2000-2 volatility is *inversely* related to traded volume. An information flow model of the volume/volatility relationship (for example Clark (1973) or Tauchen and Pitts (1983)) would suggest a strong positive correlation. Our explanation for the observed correlation is based on the impact of D2000-2 liquidity on volatility. Whilst in peak trading periods more information is impounded into midquotes, in less active periods the thin-ness of the D2000-2 order book causes measured return volatility to be great

[9] Such patterns are also observed in the indicative data derived from EFX and FXFX. See, amongst others, Andersen and Bollerslev (1997) and Payne (1996).

[10] See the studies by .....

[11] The low D2000-2 volume during Asian trading hours is likely to be due to the dominance of the EBS system in this region. EBS now incorporates the Minex system which previously dealt with much of Asian inter-dealer trade.

as order removal or transactions will cause midquotes to jump a long way, The latter effect is dominant on D2000-2 leading to the negative correlation in the data.[12]

There is little evidence of any systematic patterns in the third and fourth moment of returns across the trading day. However, the measured kurtosis values suggest that the fourth moment of returns is unlikely to exist implying that these statistics should be viewed with some caution.

The final column of Table 2 presents first order autocorrelations for returns in each of the subsamples and indicates that dependence in returns is inversely related to liquidity and transaction activity as one would expect i.e. in busier market segments, the return process has less memory. Further results demonstrate that the first order effect is the only significant autocorrelation in returns and GARCH(1,1) estimations indicate the existence of strong conditional heteroskedasticity for all subsamples.

The observations from Figures 1 to 3 and results in Tables 1 and 2 imply that the return process behaves very differently in periods when D2000-2 activity is slow. Hence, in the empirical analysis presented in Section 4, observations from the 18 to 6 GMT period are omitted such that our estimations concentrate on European and North American trading hours only.

[12]Danielsson and Payne (1998) examine the intra-day variation in liquidity supply to D2000-2 in more depth. Constructing the excess supply and demand curves implied by subsidiary limit orders they show that D2000-2 depth covaries positively with volume and liquidity, supporting the interpretation given here.

3 Empirical Methodology

As discussed in the preivous section, D2000-2 is a multi-lateral order driven system. This implies that the empirical models used in Lyons (1995) and Yao (1997) are inapplicable here as they are based on an underlying quote-driven, single-dealer structure. Instead, we employ the reduced form VAR in trades and quote revisions developed in Hasbrouck (1991a) and Hasbrouck (1991b). This framework is not predicated on any particular underlying microstructure model and has been used in the analysis of order driven markets in de Jong, Nijman, and Röell (1995) and Hamao and Hasbrouck (1995). Informed trade is identified via a non-zero impulse response of quotations to transaction activity, in line with the theoretical argument presented in Section 1. Further, the variance decomposition presented in Hasbrouck (1991b) permits one to evaluate the proportion of all information entering the FX quotation process through transactions and hence the contribution of informed trade to price discovery.

There are two main assumptions which underlie the application of this framework to the current data set. The first of these is that informed agents exploit their advantage through the use of market orders rather than limit orders. Non-informed agents, on the other hand, submit either market or limit orders to D2000-2 depending on their desire for execution speed. This implies that private information can *only* influence prices through unexpected trading activity. The assumption may be justified by noting that an informed agent submitting a limit order is noisily advertising his beliefs and hence possibly eroding his advantage. Moreover, the analysis of informed order placement strategy presented in Harris (1995) suggests that informed agents should prefer to trade via market order in fast paced, liquid markets. The spot FX market certainly exhibits these features lending credence to our assumption.

The second assumption which is required is that public information is immediately reflected in quotes. If this was not the case, traders observing public information announcements could base profitable trading strategies upon their observations and hence transaction activity and public information would be correlated. However, given the levels of liquidity and competition in FX markets it would seem reasonable to assume semi-strong form efficiency.

3.1 The VAR model

Denote the percentage return in the midquote by r_t and let x_t represent a vector of transaction characteristics, where t is an event-time observation counter.

The basic VAR formulation used in our empirical work is as follows;

$$r_t = \sum_{i=1}^{p} \alpha_i r_{t-i} + \sum_{i=0}^{p} \beta_i x_{t-i} + \epsilon_{1t} \tag{1}$$

$$x_t = \sum_{i=1}^{p} \gamma_i r_{t-i} + \sum_{i=1}^{p} \delta_i x_{t-i} + \epsilon_{2t} \tag{2}$$

with the following restrictions placed on the innovations,

$$E(\epsilon_{1t}) = E(\epsilon_{2t}) = E(\epsilon_{1t}\epsilon_{2t}) = 0$$

$$E(\epsilon_{1t}^2) = \sigma^2, E(\epsilon_{2t}^2) = \Omega$$

$$E(\epsilon_{1t}\epsilon_{2s}) = E(\epsilon_{1t}\epsilon_{1s}) = E(\epsilon_{2t}\epsilon_{2s}) = 0, \quad \forall \quad t \neq s$$

Equations (1) and (2) form a general model of the dynamics of trades and quotes and the interactions between these variables. Note that the VAR is not entirely standard as the contemporaneous realisation of x_t enters the return equation. Hence trades logically precede quote revisions with the reverse causality prohibited. This, coupled with the fact that the innovations to the equations are uncorrelated, identifies the VAR.[13]

Given the two assumptions detailed above, the innovations in the VAR can be interpreted as follows. The innovation to the return equation may represent transitory quote returns or the effects of public information. The innovation to the trade equation represents unpredictable transaction activity and hence the possibility of information-based trade. Using this classification, the effects of private information on quotations are easily retrieved.

First, we rewrite the VAR in matrix form;

[13] This identification scheme is identical to a Sims orthogonalisation of an unidentified VAR system.

$$\begin{pmatrix} 1 & -\beta_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_t \\ x_t \end{pmatrix} = \begin{pmatrix} \alpha(L) & \beta(L) \\ \gamma(L) & \delta(L) \end{pmatrix} \begin{pmatrix} r_t \\ x_t \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \tag{3}$$

where the elements of the right hand side matrix which multiplies the vector of r_t and x_t are polynomials in the lag operator formed from the coefficients in the original representation (equations (1) and (2),) and

$$VCV \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \Omega \end{pmatrix}$$

Inverting the VAR representation we get the following VMA model;

$$\begin{pmatrix} r_t \\ x_t \end{pmatrix} = \begin{pmatrix} a(L) & b(L) \\ c(L) & d(L) \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \tag{4}$$

The coefficients in the VMA lag polynomials are precisely the impulse response functions implied by the VAR.[14] The coefficient a_k, for example, gives the effect of a unit return innovation on the midquote return at a k period horizon. The effects of private information will be revealed in the $b(L)$ polynomial, the impact of trade innovations on subsequent returns. The possibility of informed trade will imply that quotes respond permanently to trade innovations and hence $\sum_{i=0}^{\infty} b_k$ represents the information content of a trade .

Estimation of the preceding VAR and calculation of the implied impulse response functions will allow us to characterise the existence of information based trade on D2000-2. In the empirical analysis reported in Section 4, the lag length of the VAR was chosen via application of the Schwarz information criterion. The VAR equations were estimated by OLS and are reported with heteroskedasticity robust standard errors. Calculation of the VMA representation was performed by simulation.

3.2 The Variance Decomposition

Whilst the VAR model allows us to quantify the information content of a trade, it does not permit one to assess the overall importance of informed trade in deter-

[14]See, for example Hamilton (1994).

mining the evolution of the asset price in question. Hasbrouck (1991b) presents a variance decomposition for quote returns, based on the estimated VAR structure. This decomposition permits retrieval of the variance of the permanent component of midquotes, plus the proportion of permanent variation related to order flow. The overall influence of informed trade can be estimated with the latter measure.

Denoting the (logarithm) of the midquote as q_t, the following structure is assumed;

$$q_t = m_t + s_t \tag{5}$$

Equation (5) decomposes the midquote into a random walk component (m_t) and a stationary term (s_t). The following equations formalise these conditions;

$$m_t = m_{t-1} + \omega_t,$$

$$\omega_t \sim N(0, \sigma_w^2), \quad E(\omega_t \omega_s) = 0 \quad for \quad t \neq s,$$

$$\lim_{k \to \infty} E_t s_{t+k} = \lim_{k \to \infty} E s_{t+k} = 0,$$

From an economic perspective, m_t can be thought of as the efficient price process. The transitory component, s_t, represents that portion of the current midquote generated by any non-information based microstructure effect e.g price discreteness, digestion effects or inventory control.

The key parameter in the preceding formulation is σ_w^2 which measures variation in the permanent component. This can be estimated by equating the return representation implied by equation (5) and the return equation from the VMA. Furthermore, variation in the permanent component due to order flow alone can also be retrieved. These measures are calculated as;

$$\sigma_w^2 = \left(\sum_{i=0}^{\infty} b_i\right) \Omega \left(\sum_{i=0}^{\infty} b_i'\right) + \left(1 + \sum_{i=1}^{\infty} a_i\right)^2 \sigma^2 \tag{6}$$

$$\sigma_x^2 = \left(\sum_{i=0}^{\infty} b_i\right) \Omega \left(\sum_{i=0}^{\infty} b_i'\right) \tag{7}$$

Standard errors of measures based on these two variance computations are calculated via a residual based bootstrap of the estimated VAR system.

An economic interpretation of equations 6) and (7) is as follows. Public information events are incorporated into the exchange rate via the return innovation, ϵ_{1t}. The permanent effect on midquotes of a unit return innovation is given by unity (the contemporaneous impact) plus $\sum_{i=1}^{\infty} a_i$ and hence the variation in the permanent component implied by public information events is given by the second term on the right hand side of equation (6). Private information is impounded into the exchange rate via trade innovations with the permanent impact of an unexpected unit trade given by $\sum_{i=0}^{\infty} b_i$ in the case where x_t is scalar. For vector x_t, the variation in the permanent component driven by trade innovations is thus the first term on the right hand side of (6).

The crucial assumption in this variance decomposition is that the permanent component follows a random walk rather than a general I(1) process. In terms of the spectral representation of returns the assumption implies that the spectrum of w_t is constant and, due to the stationarity of s_t, the spectra of r_t and s_t coincide at the origin. Hence the variance of the permanent component is quite easily tied down. In a macroeconomic setting Quah (1992) criticises the random walk assumption as ad hoc and restrictive, although in the current context efficient markets theory provides a strong rationale for its validity.

3.3 Non-linear Effects in the Trade-Quote relationship

Whilst the VAR structure presented in Section 3.1 provides a fairly robust characterisation of the dynamics between trading activity and quote revisions, it restricts the relationship between these two variables to be invariant to changes in underlying market activity. Recent work, however, suggests that measures of the pace of market activity might impinge upon the response of quotes to trades. From a theoretical perspective, the work of Admati and Pfleiderer (1988) suggests that there should be a negative correlation between the information content of trades and overall volume. An empirical example of such work is Dufour and Engle (1998) who analyse the impact of the time between stock trades on the information content of trading activity. These authors find that as the time between trades widens, the permanent impact of trades on quotes falls and the adjustment process of quotes to their long run value is more protracted. Conversely, Lyons (1996) finds that FX trades are less informative when the time between trades is low.

We analyse the effect of variables describing the state of market activity on the VAR structure in a couple of ways. The first, and most crude, of these methods is through a subsampling of the data. In the analysis of time of day effects, for example, we split the entire D2000-2 data set into several non-overlapping subsamples representing certain hours of the day. The VAR structure in equations (1) and (2) are then estimated separately for each subsample.

A similar strategy can be employed in an analysis of the effects of liquidity measures on the VAR structure. Consider a candidate liquidity measure, z_t. We assume that this variable imparts certain non-linearities to the VAR structure. To examine this, one could sort the return and trade data by the associated values of z_t, split the sorted data into quintiles and estimate the VAR model separately for each subsample.

Such a subsampling of the data is, however, a crude way to get at the issue of non-lineairities. Hence, we progress by modifying the Hasbrouck (1991a) model and using an *arranged VAR* technique. This method also employs a sorting of the VAR data by the candidate variable z_t in the following manner. Denote by D_t the time-series of return and trade data along with their respective lags i.e. $D_t = \{r_t, r_{t-1}, ..., r_{t-P}, x_t, x_{t-1}, ..., x_{t-P}\}$. Associated with D_t we have the variable governing the non-linear behaviour, z_t, and we then rearrange the observations of D_t according an ascending order sort of z_t. Call the sorted data set D_s. We then run recursive least squares regressions equivalent to the VAR model in equations (1) and (2), yielding a series of VAR coefficients based on successively larger numbers of sorted observations.[15] Graphical examination of the recursively generated VAR coefficients may then indicate non-linear behaviour in the trade-quote relationship related to z_t.

Finally, based on the results of the arranged VAR technique, the basic VAR structure is modified to cope with the non-linearities observed. Specifically, the following two regressions are run;

$$r_t = g(z_t, \theta_1) \sum_{i=1}^{p} \alpha_i r_{t-i} + g(z_t, \theta_2) \sum_{i=0}^{p} \beta_i x_{t-i} + \epsilon_{1t} \tag{8}$$

[15]Such arranged regression techniques have previously been used in the testing for specific non-linear models, for example threshold AR processes. See Tsay (1989).

$$x_t = g(z_t, \theta_3) \sum_{i=1}^{p} \gamma_i r_{t-i} + g(z_t, \theta_4) \sum_{i=1}^{p} \delta_i x_{t-i} + \epsilon_{2t} \tag{9}$$

The function $g(z_t, \theta)$ governs the non-linearities in the VAR structure and is chosen after consideration of the arranged VAR results. Given an assumed functional form for $g(z_t, \theta)$, the equations are estimated separately using non-linear least squares.

4 Results

This section presents the results from estimations of the VAR structure given in Section 3.1 and the variance decomposition in Section 3.2. We begin with results for the entire D2000-2 trading day data set. Results for time-of-day based subsamples are discussed next and we finally present the VAR estimates from data sets constructed by conditioning on market liquidity variables.

The basic variables used in the VAR analysis were percentage midquote returns and a signed transaction indicator.[16] In the construction of the transaction indicator only market orders were used, limit order crosses were ignored. A change in this assumption affects the VAR results to a very small degree. Along the same lines, a signed volume variable was also constructed.

The first row of Table 3 gives a summary of the relevant VAR parameters estimated using all trading day observations. Some general comments on those parameters not presented are as follows. First, quote returns demonstrate significant negative autocorrelation at all included lags. This is line with findings from analysis of EFX data (see Baillie and Bollerslev (1991), Bollerslev and Domowitz (1993) and Danielsson and Payne (1998)). The transaction indicator displays strong positive autocorrelation. This indicates runs in buying and selling activity and may be due to dealers splitting large orders. Finally, the effects of lagged quote returns on transaction direction are mixed. Although marginally significant as a group, these coefficients were generally individually insignificant. The R^2 in the return equation was 0.25 whilst in the trade equation the R^2 was approximately 0.075.

From the current perspective, the key parameters in the VAR formulation are those labelled β_i in equation (1) i.e. the effects of trades on current and subsequent midquote returns. As shown in row one of Table 3 the sum of these coefficients is positive. Moreover, each individual coefficient is positive and all are statistically significant. Hence, a positive trade surprise increases quotes. Computing the VMA representation and calculating the equilibrium midquote impulse response shows that an unexpected market buy leads to an upward quote revision of around 0.005% on average.[17] Comparison of this figure with the mean half-spread on D2000-2 implies that over 60%

[16] This variable was constructed to take a value of unity for a market buy, zero for no trade, and minus one for a market sell.

[17] This translates to a 1 basis point quote increase approximately.

of the bid-offer spread is compensation for liquidity suppliers facing an asymmetric information problem.

It should be noted that we also experimented with the use of trade size variables in the trade description vector (x_t). One such experiment involved the use of signed volume and signed squared volume in addition to trade direction and another used a signed indicator variable identifying the largest 5% of trades. In contrast to the results of Yao (1997), in no case were the size variables significant. This is likely to be due to the fact that there is very little variation in trade size on D2000-2. Over 90% of D2000-2 trades are for less than $5m. such that the data may not contain the power needed to identify a size effect.

Our first result confirms the findings of previous research in that a portion trade on D2000-2 can be characterised as information motivated. A question which has not been addressed in the literature, however, is the extent to which such asymmetries alter over the trading day. Microstructure theory relevant to intra-day variations in the intensity of informed trade and liquidity can be found in Admati and Pfleiderer (1988) and Foster and Viswanathan (1990), inter alia. The endogenous information acquisition model of Admati and Pfleiderer (1988), for example, predicts that high volume periods should be characterised by relatively small price impacts from trade. This is due to the clustering of discretionary liquidity trade and increased competition between informed traders in equilibrium. The model of Foster and Viswanathan (1990) has a similar setup except that the information advantage of insiders is assumed to be eroded over time.

We examine these issues by estimating the VAR separately for the six previously defined non-overlapping subsamples of the trading day. Results are given in rows 2 to 7 of Table 3.

The results correspond broadly with the predictions made by Admati and Pfleiderer (1988). First note that Table 1 shows that the segments of the trading day with greatest volume are those with lowest percentage spreads. Admati and Pfleiderer (1988) present a model based on the batch trading analysis of Kyle (1985) and, hence, does not incorporate a spread. However, given that their framework predicts more liquid markets in high volume intervals this result can be seen as broadly consistent with their analysis. Further, Table 3 indicates that the information content of trading activity follows an intra-day pattern which is the inverse of that followed by volume. Thus, in high volume/liquidity periods, the price response to a trade is relatively low. This is also consistent with the 'hot-potato' hypothesis of Lyons (1996). A further point to note is that, while the size of the asymmetric information effect is inversely related to volume, the significance of the trade variables is greater in high-volume periods. These results contrast with those of Foster and Viswanathan (1993) who demonstrate a positive relationship between adverse selection costs and trading volume for a sample of NYSE equities and hence reject the models of Admati and Pfleiderer (1988) and Foster and Viswanathan (1990).

Hence, our results conform with standard theoretical predictions for intra-day variation in aymmetric information costs, trading costs and volume. High volume periods are characterised by a concentration of liquidity trades and increased competition between informed agents, with these effects reducing the price impact of a trade. Figures 4 to 6 graphically illustrate the VAR results, plotting the quote impulse response functions for the entire trading day data set plus the 8 to 10 GMT and 16 to 18 GMT subsamples. These two subsamples were chosen as they are those with the lowest and highest average spread respectively. The figures also demonstrate the variation in D2000-2 order book depth across the trading day through variation in the *immediate* response of the midquote to a trade. In Figure 6 this is about five times as large as the corresponding value from Figure 5. The implication is that liquid periods are characterised by a greater clustering of limit orders around the inside spread.

Based on the VAR results presented in Table 3, the return variance decompositions are presented in Table 4. Across the entire trading day, the VAR estimates imply that about 40% of the permanent return variance is attributable to order flow. Comparable figures are contained in Hasbrouck (1991b) who reports an average value of 33% for a sample of U.S. equities and de Jong, Nijman, and Röell (1995) who analyse a sample of French stocks traded on the Paris Bourse and report an average trade correlated component of 40%. Hence our results imply that the information content of USD/DEM order flow on D2000-2 is of the same magnitude as that on equities markets. Examination of the last two columns of Table 4 shows that the permanent component accounts for only one quarter of all return variation such that order flow contributes one tenth of total return variance.

The variance decompositions for the six trading day subsamples are also presented in Table 4. Whilst the results in Table 3 showed that the information content of a single trade was inversely related to volume, the variance decompositions show that the information contained in order flow as a whole is positively related to volume.

The ratio of the trade-correlated component to the variance of total permanent quote variation ($\sigma_x^2/\sigma_\omega^2$) is greatest during high volume episodes. Furthermore, the proportion of all return variation explained by order flow and the size of the permanent component have a positive correlation with volume.

Again, these results can be reconciled with the predictions from theoretical models of intra-day variation in transaction activity and price formation. Admati and Pfleiderer (1988) implies that in high-trading intervals return variance should be high and prices more informative. In Section 2 we discussed the *negative* correlation between volume and volatility arguing that the lack of order book depth in low trading periods was behind this result such that the impact of information on volatility was obscured. However, Table 4 indicates that, despite permanent return variation increasing with volume, the information entering prices through order flow is increased also. Aggregate trading activity contributes a greater proportion of all information in peak trading periods in line with Admati and Pfleiderer (1988).

The prior results present evidence that the information content of a trade and that of aggregate order flow are strongly linked to market activity with this evidence derived from a set of non-overlapping time-of-day subsamples of the D2000-2 data. Our final set of estimations refine this analysis. The relevant theoretical models predict such patterns due to clustering of liquidity trading in equilibrium. Hence, empirical identification of times of concentration in uninformed activity should corroborate the prior results.

Given the assumptions on order placement strategy made in Section 3, a concentration of non-information based trade should be associated with a high level of limit order placement (the informed preferring to trade by market order due to the execution certainty it yields.) Data from such episodes should hence display small price impacts from trades. This is examined using the arranged VAR technique and non-linear VAR estimations discussed in Section 3.3 with both the first lag of the number and aggregate quantity of all limit orders outstanding used as the variable governing the non-linearities in the VAR (z_t).[18]

Figures 7 and 8 present the recursive OLS estimates of the asymmetric information coefficients from the VAR (i.e. $\sum_{i=0}^{P} \beta_i$) and the implied quote impulse response after arranging the data by the aggregate limit quantity outstanding. Figures 9 and 10 present the same measures using the total number of limits outstanding as the

[18] We used 500 observations to initialise the recursive OLS procedure in all cases.

arranging variable. These estimates are based on the entire D2000-2 sample and the x-axis in these figures represents the number of observations used in the recursive OLS estimation.

All of these figures strongly suggest that the asymmetric information problem is lower in times of intense liquidity trader activity. When either the number or size of outstanding limits is low, the equilibrium impact of an unexpected trade on the midquote is high. The relationship is clearly non-linear, however, with the equilibrium quote impact dropping off sharply over the first 20000 observations or so and much less rapidly thereafter. Nonetheless, the figures provide evidence in favour of the Admati and Pfleiderer (1988) model and the 'hot-potato' hypothesis of Lyons (1996).

The evidence from Figures 7 to 10 indicates that the linear VAR model is not well specified as it fails to capture the effects of liquidity on the quote response to trades. Hence, a fully parametric non-linear VAR model is estimated using the structure presented in equations (8) and (9). As the figures suggest that the price impact declines non-linearly with liquidity trader activity the following function is used to model the relationship;

$$g(z_t, \theta_i) = 1 + \frac{\theta_i}{z_t} \tag{10}$$

where z_t is the chosen proxy for liquidity trader activity. Estimates of equations (8) and (9) using the transition function in equation (10) and a lag length of 5 are given in Tables 5 and 6.

Both tables imply similar patterns of serial and cross correlations of trades and returns to those obtained for the linear full-sample VAR. Returns demonstrate significant negative autocorrelation, trades positive correlation and there is strong causality running from trades to quotes. The θ_i parameters are those which govern the non-linear behaviour in the VAR. Estimates of these parameters indicate significant non-linearities in the return equation (i.e. θ_1 and θ_2 significantly different from zero) regardless of the proxy used for liquidity trader activity whilst the evidence for non-linear behaviour in the trade equation is weaker. In terms of size and sign, θ_1 and θ_2 are consistently positive and θ_3 and θ_4 negative, whilst θ_2 is an order of magnitude larger than the others.

Interpretation of the prior results is as follows. Both private information effects and return autocorrelation are smaller in more liquid periods. The first of these results is particularly pronounced, conforming with prior observations on time-of-day effects and is consistent with Admati and Pfleiderer (1988) and Lyons (1996). The second may be seen as greater informational efficiency in more active periods. Non-linearities in the trade equation are much weaker, with marginal evidence that both trade sign autocorrelation and the impact of past returns on trade direction are increasing with liquidity trader activity.

Figures 11 and 12 give a graphical representation of these estimates. They plot the equilibrium quote response to an unexpected market buy across varying levels of limit orders and limit quantity outstanding. The impulse responses were constructed via simulation of the estimated VAR structure assuming a fixed level of z_t throughout each individual simulation experiment. The figures clearly demonstrate that the non-linear VAR structure captures the impact of liquidity trader activity on the price impact of trades very well. Both show a very similar pattern to that observed in Figures 8 and 10.

5 Conclusion

This paper analyses trading behaviour in the inter-dealer spot FX market. Specifically, we examine whether the inter-dealer market can be characterised as subject to information asymmetries. Using the technology introduced by Hasbrouck (1991a) and Hasbrouck (1991b) we find that FX trades *do* carry information. Roughly 60% of the bid-offer spread in our data can be related to the asymmetric information problem . Further, around 40% of the variation in the efficient price is shown to be contributed by order flow, a proportion which is comparable to those found in studies of equity markets.

We also uncover significant intra-day variation in the information content of a trade and the total information content of order flow. In line with the theoretical predictions contained in Admati and Pfleiderer (1988), in high volume/liquidity intervals the information content of a single trade is low whilst the share of information entering the midquote through order flow is high.

A more refined analysis of the relationship between the price impact of trades and market activity is also presented. In line with the time-of-day estimations, results from an arranged VAR technique indicate that the asymmetric information coefficients are not stable across different levels of market liquidity. We then propose a simple non-linear VAR structure which relates the coefficient instability to variation in proxies for liquidity trade intensity. Results demonstrate significant non-linearity in the return equation of the VAR which again conforms with the predictions of Admati and Pfleiderer (1988) and the 'hot-potato' hypothesis of Lyons (1996).

There are several possible extensions to the current study. First, whilst the non-linear VAR used in Section 4 is useful in the current context, it is a restricted version of a much more general model. Modelling coefficient variation due to multiple conditioning variables and with more general functional forms may prove to be a fruitful exercise. Another interesting area of research would be to examine how, amongst other factors, transaction activity affects subsequent liquidity supply through analysis of the entire excess demand and supply schedules for currency. Furthermore, econometric analysis of individual limit order entry and execution would provide stronger evidence on the order placement strategy of dealers and how it reacts to volume, liquidity and quote volatility. Such issues are currently under investigation.

References

Admati, A., and P. Pfleiderer (1988): "A Theory of Intra-Day Patterns: Volume and Price Variability," *Review of Financial Studies*, 1, 3–40.

Andersen, T., and T. Bollerslev (1997): "Intraday Seaonality and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115–158.

Baillie, R., and T. Bollerslev (1991): "Intra-Day and Inter-Market Volatility in Foreign Exchange Markets," *Review of Economic Studies*, 58, 565–586.

Bessembinder, H. (1994): "Bid-Ask Spreads in the Interbank Foreign Exchange Market," *Journal of Financial Economics*, 35, 317–348.

Bollerslev, T., and I. Domowitz (1993): "Trading Patterns and Prices in the Interbank Foreign Exchange Market," *Journal of Finance*, 48, 1421–1444.

Clark, P. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135–155.

Danielsson, J., and R. Payne (1998): "The Information in Traded and Indicative Exchange Rate Data: a Comparison of Reuters' D2000-2 and EFX Systems," Working Paper, Financial Markets Group, London School of Economics.

de Jong, F., T. Nijman, and A. Röell (1995): "Price Effects of Trading and Components of the Bid-Ask Spread on the Paris Bourse," Discussion Paper 207, Financial Markets Group, London School of Economics.

Dufour, A., and R. Engle (1998): "Time and the Price Impact of a Trade," Working Paper, University of California, San Diego.

Easley, D., and M. O'Hara (1987): "Price, Trade Size and Information in Securities Markets," *Journal of Financial Economics*, 19, 69–90.

Foster, D., and S. Viswanathan (1990): "A Theory of the Interday Variations in Volume, Variance and Trading Costs in Securities Markets," *Review of Financial Studies*, 3, 593–624.

——— (1993): "Variations in Trading Volume, Return Volatility and Trading Costs: Evidence on Recent Price Formation Models," *Journal of Finance*, 48(1), 187–211.

French, K., and R. Roll (1986): "Stock Return Variances: The Arrival of Information and the Reaction of Traders," *Journal of Financial Economics*, 17, 5–26.

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

Goodhart, C., T. Ito, and R. Payne (1996): "One Day in June, 1993: A Study of the Workings of Reuters D2000-2 Electronic Foreign Exchange Trading System," in *The Microstructure of Foreign Exchange Markets*, ed. by J. Frankel, G. Galli, and A. Giovannini. University of Chicago Press.

Goodhart, C., and R. Payne (1996): "Microstructural Dynamics in a Foreign Exchange Electronic Broking System," *Journal of International Money and Finance*, 15(6), 829–852.

Hamao, Y., and J. Hasbrouck (1995): "Securities Trading in the Absence of Dealers: Trades and Quotes on the Tokyo Stock Exchange," *Review of Financial Studies*, 8(3), 849–878.

Hamilton, J. (1994): *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Harris, L. (1995): "Optimal Dynamic Order Submission Strategies in some Stylized Trading Problems," Working Paper, University of Southern California.

Hasbrouck, J. (1991a): "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46(1), 179–206.

——— (1991b): "The Summary Informativeness of Stock Trades: An Econometric Analysis," *Review of Financial Studies*, 4(3), 571–595.

Ito, T., R. Lyons, and M. Melvin (1998): "Is There Private Information in the FX market: The Tokyo Experiment," *Journal of Finance*, 53(3).

Kyle, A. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.

Lyons, R. (1995): "Tests of Microstructural Hypotheses in the Foreign Exchange Market," *Journal of Financial Economics*, 39, 321–351.

Lyons, R. (1996): "Foreign Exchange Volume: Sound and Fury Signifying Nothing?," in *The Microstructure of Foreign Exchange Markets*, ed. by J. Frankel, G. Galli, and A. Giovannini. University of Chicago Press.

Madhavan, A., and S. Smidt (1991): "A Bayesian Model of Intraday Specialist Pricing," *Journal of Financial Economics*, 30, 99–134.

Naik, N., A. Neuberger, and S. Viswanathan (1994): "Disclosure Regulation in Competitive Dealer Markets: Analysis of the London Stock Exchange," Working Paper.

Payne, R. (1996): "Announcement effects and seasonality in the intra-day foreign exchange market," Discussion Paper 238, Financial Markets Group, London School of Economics.

Peiers, B. (1997): "Informed Traders, Intervention, and Price Leadership: a Deeper View of the Microstructure of the Foreign Exchange Market," *Journal of Finance*, 52(4), 1589–1614.

Perraudin, W., and P. Vitale (1996): "Interdealer Trade and Information Flows in the Foreign Exchange Market," in *The Microstructure of Foreign Exchange Markets*, ed. by J. Frankel, G. Galli, and A. Giovannini. University of Chicago Press.

Quah, D. (1992): "The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds," *Econometrica*, 60(1), 107–118.

Tauchen, G., and M. Pitts (1983): "The price variability-volume relationship on speculative markets," *Econometrica*, 51, 485–505.

Tsay, R. (1989): "Testing and Modelling Threshold Autoregressive Processes," *Journal of the American Statistical Association*, 84, 231–240.

Vogler, K.-H. (1993): "Inter-Dealer Trading," Discussion Paper 174, Financial Markets Group, London School of Economics.

Yao, J. (1997): "Market Making in the Interbank Foreign Exchange Market," Working Paper, New York University.

Table 1: Summary Statistics for D2000-2 Order Book

Sample	Obs	Bids	Offers	Q_b	Q_o	$\bar{s}$	Deal	Vol	Size
6 to 8 GMT	10816	31.14	31.25	55.24	58.68	0.015	4921	8742	1.78
8 to 10 GMT	11587	46.63	47.02	91.39	89.82	0.011	5751	11036	1.92
10 to 12 GMT	10446	47.74	42.15	99.84	79.29	0.017	4654	8600	1.85
12 to 14 GMT	17339	50.22	46.77	95.31	80.79	0.014	7519	13888	1.85
14 to 16 GMT	12088	41.65	33.31	86.59	57.37	0.018	4101	7302	1.78
16 to 18 GMT	3333	20.04	14.46	54.89	22.04	0.042	671	1080	1.61
18 to 6 GMT	4938	9.33	7.43	25.45	15.06	0.076	766	1155	1.51

Note: The above table gives summary statistics for the basic order book variables. The columns denoted *Bids* and *Offers* give the average number of orders outstanding on that side of the book during the given data sample. Correspondingly, the following two columns give average aggregate quantity outstanding on both sides. The column labelled $\bar{s}$ gives the average spread in a given subsample of data. The final three columns give the total number of deals, total volume and average volume in each sample.

Table 2: Summary Statistics for Returns in all subsamples

Sample	$\bar{r}$	σ_r^2	Skew	Kurtosis	ρ_1
6 to 8 GMT	0.00002	0.00006	0.40	92.5	-0.38
8 to 10 GMT	-0.00000	0.00003	0.10	34.2	-0.35
10 to 12 GMT	-0.00009	0.00013	-0.38	127.2	-0.39
12 to 14 GMT	0.00001	0.00005	0.04	55.9	-0.34
14 to 16 GMT	0.00004	0.00007	-0.05	180.9	-0.35
16 to 18 GMT	-0.00013	0.00036	0.11	18.9	-0.42
18 to 6 GMT	0.00001	0.00230	0.05	34.5	-0.50

Note: The above table gives summary statistics for the midquote return variable in the given subsamples of data. The first four columns of the table give the first four sample moments of the return series. The final column gives the first order return autocorrelation.

Table 3: Summary of VAR results for Trading Day and Subsamples

Subsample	Lag	SIC	$\sum_i \beta_i$	$\chi^2(\sum_i \beta_i)$	QIR	Q(10)
6 to 18 GMT	8	89877	0.00668	4178.7	0.00518	36.1
6 to 8 GMT	5	-8425	0.00565	837.8	0.00502	42.4
8 to 10 GMT	5	-16964	0.00406	1060.1	0.00404	19.9
10 to 12 GMT	5	-106	0.00758	561.5	0.00583	70.0
12 to 14 GMT	4	-7213	0.00548	1344.1	0.00490	55.6
14 to 16 GMT	9	-4220	0.00750	804.6	0.00564	11.2
16 to 18 GMT	5	-406	0.01853	247.1	0.01192	56.7

Note: This table summarises the VAR results over the entire trading day and across the 6 trading day subsamples. SIC is the value of the Schwarz information criterion. $\sum_i \beta_i$ is the sum of the asymmetric information coefficients from the VAR and the following column gives a Wald test statistics for the null that all are zero. QIR denotes the final quote impulse response implies by the VAR and $Q(10)$ gives a Box-Ljung statistic for the null that the residuals in the return equation are uncorrelated up to order 10.

Table 4: Variance Decomposition for Trading Day and Subsamples

Subsample	$\sigma_x^2/\sigma_\omega^2$	$\sigma_\omega^2/\sigma_r^2$	σ_x^2/σ_r^2
6 to 18 GMT	0.41 (0.015)	0.27 (0.009)	0.11
6 to 8 GMT	0.43 (0.030)	0.32 (0.021)	0.14
8 to 10 GMT	0.47 (0.027)	0.42 (0.022)	0.20
10 to 12 GMT	0.32 (0.034)	0.25 (0.019)	0.08
12 to 14 GMT	0.36 (0.020)	0.39 (0.018)	0.14
14 to 16 GMT	0.41 (0.041)	0.23 (0.021)	0.10
16 to 18 GMT	0.35 (0.046)	0.17 (0.014)	0.06

Note: $\sigma_x^2/\sigma_\omega^2$ is the ratio of the trade correlated component to permanent midquote variance. $\sigma_\omega^2/\sigma_r^2$ is the size of the permanent component. σ_x^2/σ_r^2 is the ratio of the trade correlated component to overall return variation. Numbers in parentheses in the first two columns are bootstrap standard errors for the ratios presented. These standard errors are based on 500 bootstrap replications.

Table 5: Non-linear VAR estimation using Limit Quantity Outstanding

Quantity	Value	Quantity	Value
θ_1	1.3458^*	θ_3	-2.8951
θ_2	38.3577^*	θ_4	-2.0150
$\sum_{i=1}^{5} \alpha_i$	-1.4364^*	$\sum_{i=1}^{5} \gamma_i$	-0.1906^*
$\sum_{i=0}^{5} \beta_i$	0.00582^*	$\sum_{i=1}^{5} \delta_i$	0.4401^*
R^2	0.3143	R^2	0.0714
$D.W.$	2.0107	$D.W.$	2.0024

Note: The table summarises the estimates of the non-linear VAR structure presented in equations (8) and (9). $D.W.$ is the Durbin-Watson statistic and * denotes significance from zero at a 5% level.

Table 6: Non-linear VAR estimation using Orders Outstanding

Quantity	Value	Quantity	Value
θ_1	1.6232*	θ_3	-2.6194*
θ_2	30.7756*	θ_4	-1.2569
$\sum_{i=1}^{5}\alpha_i$	-1.3444*	$\sum_{i=1}^{5}\gamma_i$	-0.0055*
$\sum_{i=0}^{5}\beta_i$	0.00498*	$\sum_{i=1}^{5}\delta_i$	0.4411*
R^2	0.3217	R^2	0.0714
$D.W.$	2.0157	$D.W.$	2.0024

Note: The table summarises the estimates of the non-linear VAR structure presented in equations (8) and (9). $D.W.$ is the Durbin-Watson statistic and * denotes significance from zero at a 5% level.

Figure 1: Intra-day Seasonal Aggregate D2000-2 Order Book Depth

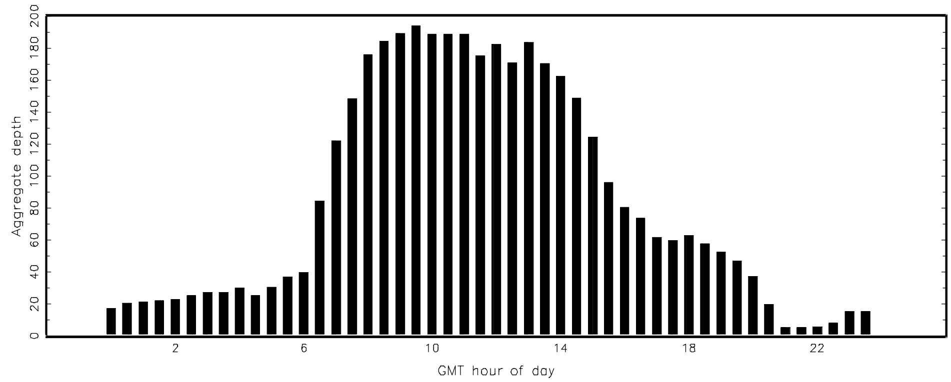

Figure 2: Intra-day Seasonal in D2000-2 Transaction Frequency

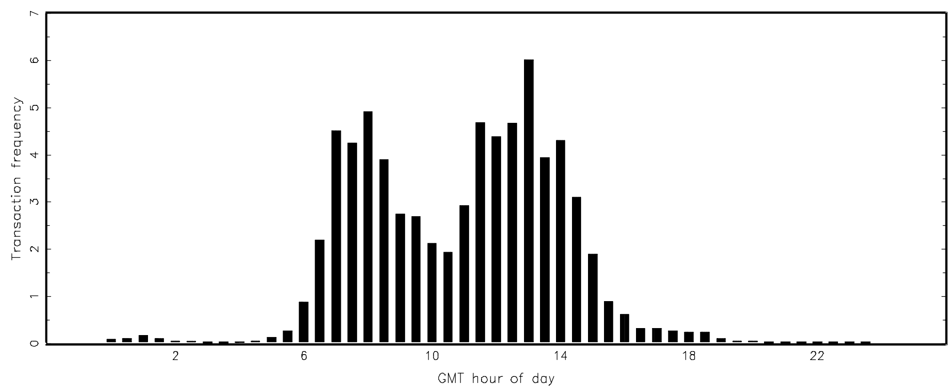

Figure 3: Intra-day Seasonal in D2000-2 Percentage Spread

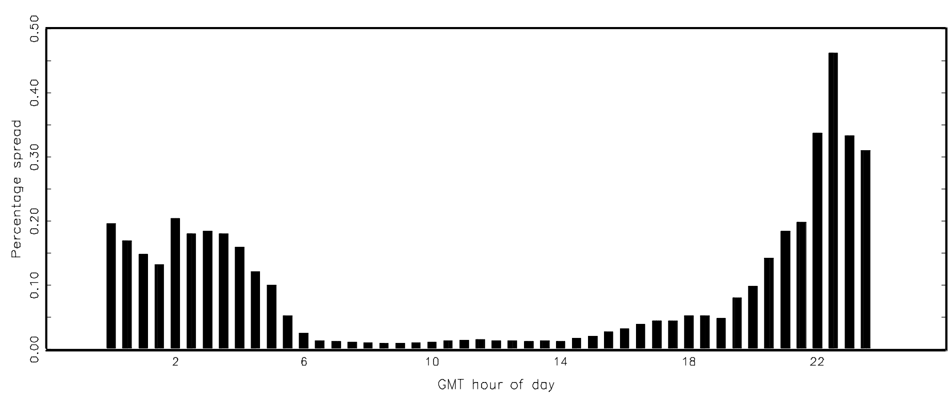

Figure 4: Basic Midquote IRF, all data

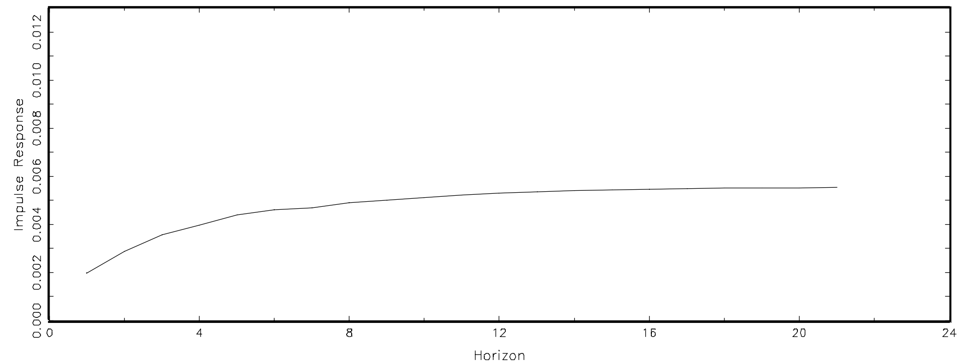

Figure 5: Basic Midquote IRF, subsample 2

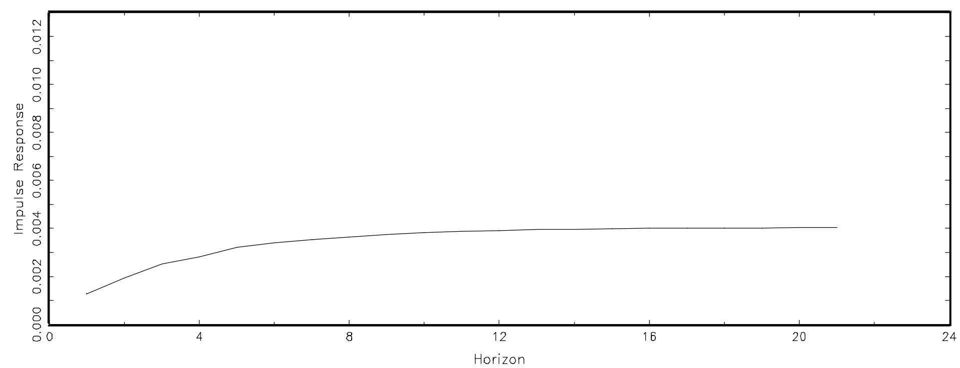

Figure 6: Basic Midquote IRF, subsample 6

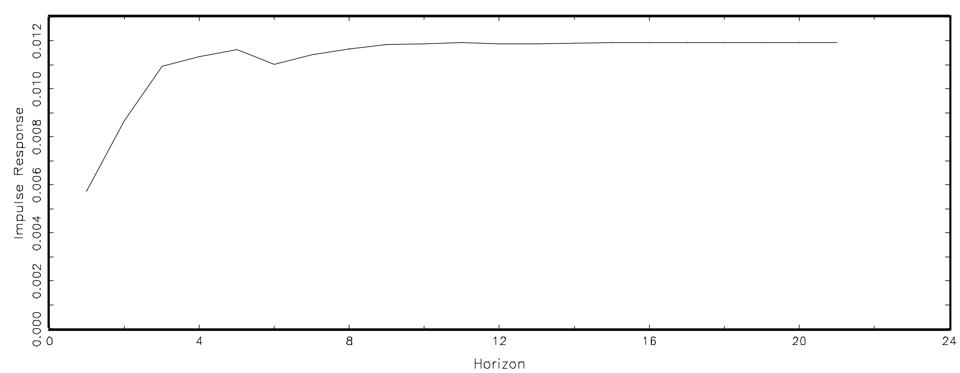

Figure 7: AI coefficients arranged by Size

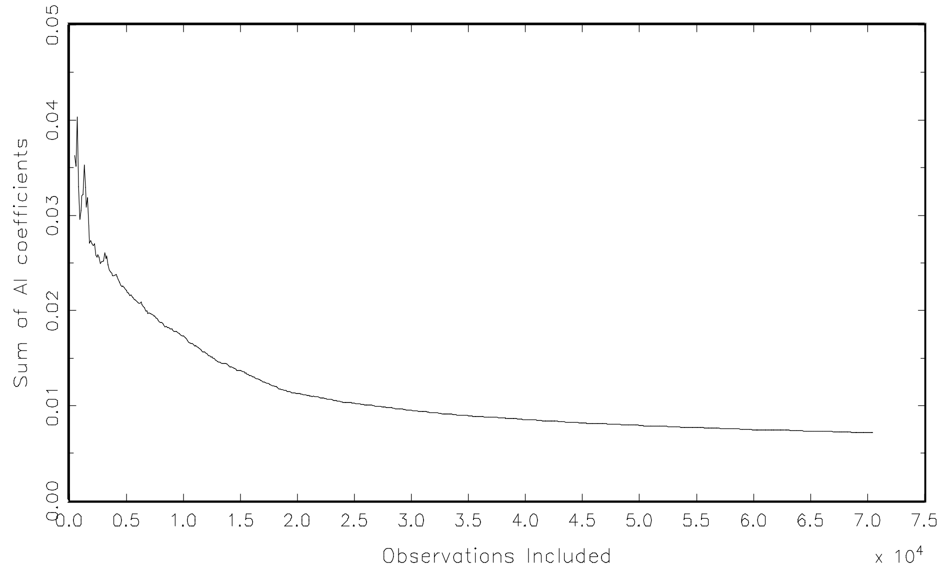

Figure 8: Quote Impulse Responses arranged by Size

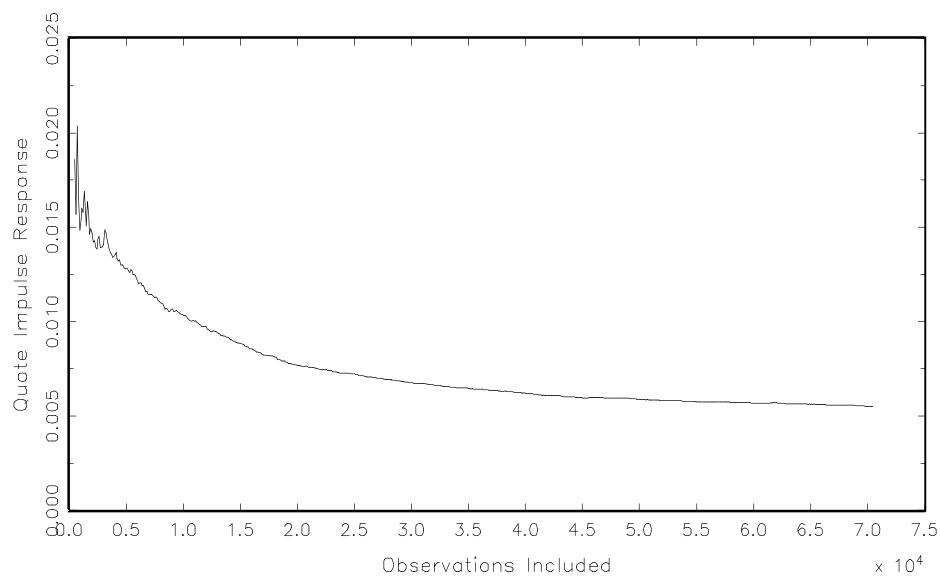

Figure 9: AI coefficients arranged by Orders

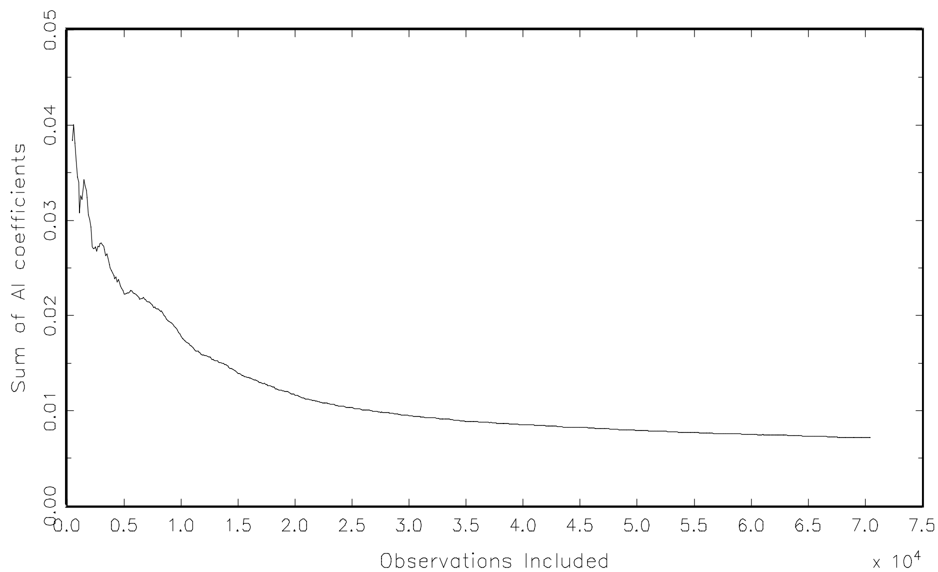

Figure 10: Quote Impulse Responses arranged by Orders

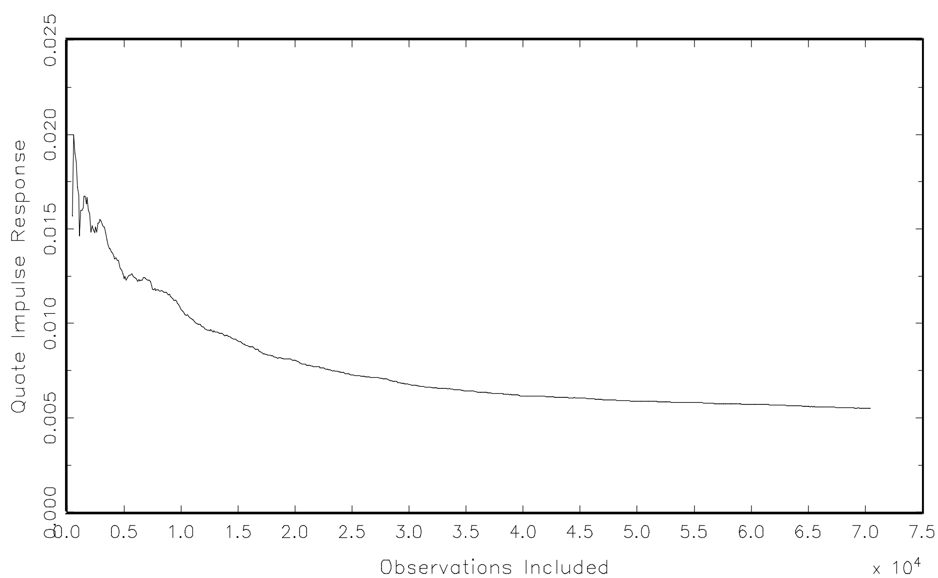

Figure 11: Price Impact Variation across Orders Outstanding

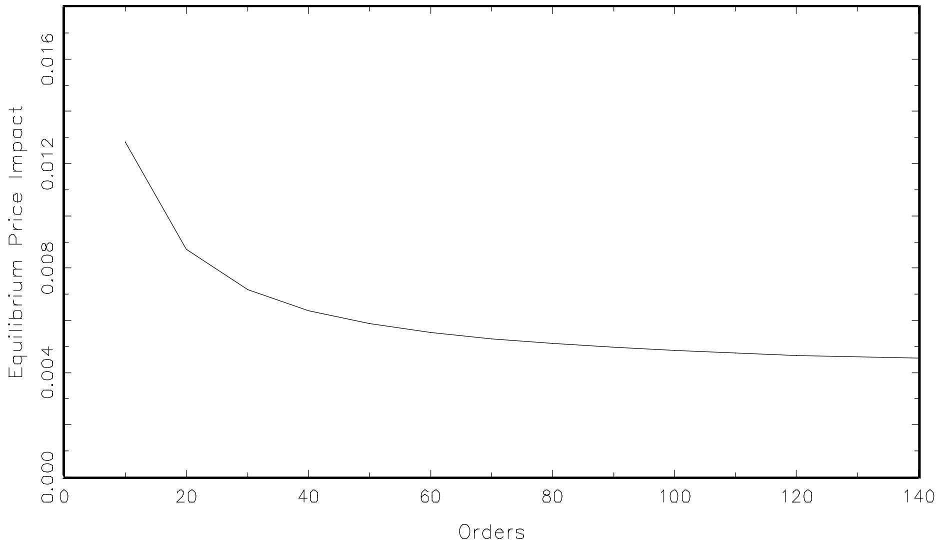

Figure 12: Price Impact Variation across Limit Quantity