Real trading patterns and prices in spot foreign exchange markets

J. Daníelsson, R. Payne *

Department of Accounting and Finance and Financial Markets Group, London School of Economics, Houghton Street, London WC2A 2AE, UK

Abstract

Most empirical FX microstructure research uses indicative quotes to proxy for firm, tradeable quotes. This paper presents comparison of the characteristics of one week of indicative DEM/USD data with contemporaneous, transactions-based data from an electronic FX brokerage. A high-frequency analysis yields the following results. Indicative returns are more volatile and more strongly autocorrelated than firm returns. Unlike firm spreads, indicative spreads contain no information on market liquidity. Indicative returns lag firm returns by around 3 minutes. These differences disappear with aggregation. Return properties are similar when sampled every 5 minutes and are essentially indistinguishable when sampled every 10 minutes. © 2002 Elsevier Science Ltd. All rights reserved.

JEL classification: C22; F31

Keywords: Exchange rates; Foreign exchange market microstructure; High-frequency data

1. Introduction

In recent years there has been a large increase in the amount of research devoted to the microstructure of foreign exchange markets and the behaviour of high-frequency exchange rate data. While some of this research has been driven by theoretical and statistical advances, the main impetus has been the increased availability of tick-by-tick exchange rate data to researchers. These data have been derived primarily from the FXFX, and later EFX, pages of Reuters’ information systems and has been exten-
sively employed in empirical studies. For example, FXFX midquotes are widely used as proxies for transaction prices in analysis of intra-day exchange rate volatility (Bailie and Bollerslev, 1991; Dacorogna et al., 1993; Andersen and Bollerslev, 1997; Payne, 1996), triangular arbitrage relationships (de Jong et al., 1998), and intra-day technical trading rule performance (Curcio et al., 1997). FXFX spreads have been used as measures of FX market liquidity in studies such as Bollerslev and Melvin (1994) and Hartmann (1999). Finally, FXFX quote frequency is used as a proxy for traded currency volumes in Bollerslev and Domowitz (1993) and Melvin and Yin (2000).

EFX data, however, have a number of shortcomings. First, and most importantly from a microstructure perspective, they contain no measure of traded currency volumes. This renders many interesting microstructure hypotheses untestable. Second, the bid and ask quotes derived from EFX screens are indicative rather than firm. This means that such quotes are not binding commitments to trade from the originator and hence they may not be accurate measures of tradeable exchange rates. Furthermore, while the EFX system gives a timestamp for the entry of a quote pair, no such timing is given for the exit of quotes. Hence there is no information on the effective lifetime of EFX quotes. Finally, each EFX bid and ask quote pair is input by a single dealer. As such, these quotes are likely to reflect dealer-specific characteristics (e.g. inventories or beliefs) and may be a poor representation of ‘market quotes’.

These shortcomings raise concerns about the validity of EFX data as a proxy for both transaction prices and firm quotes. Below, we present an empirical comparison of the features of data derived from both indicative EFX quotes and firm data drawn from Reuters D2000-2 (an electronic FX broking system), using tick-by-tick observations for the same five days. As such, this work builds upon, extends and clarifies results in Goodhart et al. (1996), who compare the statistical features of EFX data and data from D2000-2. However, their analysis is subject to a number of limitations. First, their data covered only a single day. Hence, for example, they could not examine recurrent intra-daily patterns in the data or provide accurate information on how the statistical features of the data varied with time of day. The second main limitation of their study is that their D2000-2 data are not timestamped. Hence, in order to construct approximately contemporaneous data sets, they match them by maximizing the correlation between EFX and D2000-2 midquotes. Clearly, this implies that Goodhart et al. (1996) cannot properly examine the lead–lag relationships between EFX and D2000-2 returns and volatility measures.

Our analysis is based on a newly available transactions-based data set on the DEM/USD drawn from Reuters D2000-2.1 The data covers the week from October 6 to October 10, 1997, and, in addition, we have EFX data from the same week. We examine the following statistical features of series drawn from the two data sources;

---

1 The data are available for academic research from the Financial Markets Group at the London School of Economics. See http://fmg.lse.ac.uk. To our knowledge, the data have not been used as yet in any published work.
1. Intra-day activity patterns.
2. Sample moments and dependence measures for returns.
3. Cross-correlations between firm and indicative returns.
4. The information content of each quotation series.
5. Volatility spillovers from D2000-2 to EFX returns and vice versa.
6. The manner in which the relationship between the data series depends on sampling frequency.

Our results indicate that EFX quotes are a poor proxy for firm quotes at very high sampling frequencies (20 seconds). The EFX midquote is far more volatile and more strongly autocorrelated than its D2000-2 counterpart. Further, EFX spreads contain little or no information on the pace of the market. D2000-2 midquote returns lead those on EFX by 2 to 3 minutes and the proportion of all information entering the exchange rate via the D2000-2 price is above 90% (during normal trading hours).

We surmise that EFX quotes tend to be used primarily as an indication of willingness to trade on a certain side of the market. Generally, an initiating trader wishes to either buy or sell, but not both. To signal his/her intentions, a standard, large spread is used to make only one side of the quote a keen price. Then, as the quotes contain the name and location of the originating institution, this trader can be easily contacted by potential counterparties. Such quote-setting behaviour will imply precisely the statistical data features indicated above.

As one might expect, as sampling frequency is reduced, many of the discrepancies between the indicative and firm data disappear (with the exception of the differences in spread behaviour). We find that at levels of aggregation of 5 minutes and above, EFX returns are a fairly good proxy for firm returns. At a sampling frequency of 10 minutes, the two return series are extremely similar.

The rest of the paper is set out as follows. The next section contains a brief description of the nature of the spot FX markets and how the D2000-2 and EFX systems fit into the market structure. Sections 2.1.1 and 2.1.2 give an overview of the two data sources and the structure of each data set. Section 2.2.1 provides our analysis of the activity patterns on EFX and D2000-2, mainly through a set of graphics. In Section 2.2.2 we describe the basic statistical nature of the two sets of mid-quote returns. Section 3 contains an analysis of the information content of each of the two return series and our bivariate GARCH results, and our analysis of the effects of temporal aggregation on return characteristics is contained in Section 4. Section 5 provides conclusions and presents ideas for further work.

2. The spot FX market

2.1. Spot FX market structure

The spot FX market has grown tremendously in recent years. According to the BIS surveys of FX activity in 1995 and 1998 (Bank for International Settlements 1996, 1998), total spot volume in all exchange rates has risen annually by over 4%
in recent years to a daily turnover figure of just above US$590bn. Of this global turnover the most commonly traded currency was the US dollar which was involved in 44% of transactions and the second largest was the German mark with 15%. The largest trading center was London, with 32% of global activity. The market for spot DEM/USD (analyzed here) was one of the most liquid in the world and DEM/USD activity can be expected to center around London trading hours.

Transaction activity in spot FX is heavily fragmented and largely opaque. Order flow can be divided into a number of segments:

1. Customer–dealer (approx. 25% of volume)
2. Inter-dealer (approx. 75% of volume)
   (a) Direct inter-dealer trade
      (i) Telephone based trade
      (ii) D2000-1 based trade
   (b) Indirect inter-dealer trade
      (i) Voice brokered trade
      (ii) Electronically brokered trade (EBS and D2000-2)

   Electronically brokered trade has been the major market innovation of the last few years and has grown tremendously in volume terms, mainly at the expense of voice brokered trade. One of the two main vendors of electronic brokerage is Reuters through its D2000-2 service.

2.1.1. EFX

A source of FX price information which is available to all market participants is the Reuters’ EFX page. This screen provides a continuously updated sequence of bid and ask exchange rate quotation pairs (timestamped to the nearest second) from individual institutions whose names and locations are also displayed. EFX does not contain any data on traded volumes but gives an evolving picture of the quotes available from other dealers. These quotes, however, are not firm but indicative, i.e. they do not present a binding commitment from the advertising institution to trade at these prices. It has been argued that reputation considerations would almost force those submitting EFX quotes to regard them as firm (see Bollerslev and Domowitz (1993), for example), but this is an assertion which has not been empirically validated. Information from EFX constitutes one component of our data set and during our sample period there were 32,121 quotes entries on the EFX system.

2.1.2. The Reuters D2000-2 dealing system and data

The second component of our data is the Reuters D2000-2 information, consisting of all entries to the D2000-2 system for the week of October 6 to October 10, 1997. The system display and its basic trading mechanism are as follows.

Reuters D2000-2 operates as an electronic limit order book with liquidity supply via limit order and liquidity demand via market order (and direct limit order crosses). A subscriber to D2000-2 sees the following items on the trading screen, for up to six exchange rates:
• Best limit buy and sell prices
• The quantities available for trade at the best prices
• An indicator of the characteristics of the last trade

These data items are available to us on a tick-by-tick basis. Furthermore, the data set contains information not available to market participants, specifically every away limit order on the D2000-2 order book at every point in time.\(^2\) Hence we can examine variations in liquidity supply which are unknown to those actually trading. The entry and exit times of all limit and market orders are supplied to the one hundredth of a second.

Limit orders are queued via price and then time priority. In general, market orders will hit the best outstanding limit order on a given side of the market. There are exceptions to this rule, however. In a small number of cases in the data set a market order failed to complete because of a lack of bilateral credit arrangements between the counterparties. As a result, at some point in time, the initiator of a market order may find the best priced limit order unavailable to him/her, implying that market orders may also trade outside the touch. This causes complications in processing the D2000-2 data set, but since such occurrences are rare, we feel that they will not affect the results in any way.

Trades also occur when bid and offer limit orders cross, i.e. the book contains a buy limit order with price greater than or equal to the best outstanding limit sell.\(^3\) These crosses occur automatically on the system and are straightforward to retrieve from the data set supplied. Again, in a few cases the best limit buy and sell do not cross for reasons outlined above, resulting in negative bid–ask spreads. Such observations are deleted before constructing the data used in this study.

For the five-day trading period included in the data set, there were 130,535 system entries of one of the following types:

1. Limit buy entry.
2. Limit sell entry.
3. Take: a market buy order.

In addition, there are a few other entry types, most of which can be reclassified within one of the four preceding categories.

2.1.3. Data processing and aggregation

The following basic D2000-2 variables were employed. First, the best bid and offer in the system at every observation point was used as our firm quote. From these, the midquote and bid–ask spread were constructed. Further, as measures of

\(^2\) By an away order we mean a limit sell with price above the current best or a limit buy with price below the best.

\(^3\) The price of such a transaction is that of the limit order entered earliest, i.e. the system treats the order entering latest similar to a market order.
FX market activity we constructed time series of the number and aggregate size (in US$m) of all limit orders outstanding as well as traded volume. Finally, the actual transaction prices are directly observed, and are used here as well. Midquote returns for the EFX data were constructed along with spread and quote frequency variables.  

In order to convert both EFX and D2000-2 data from event to calendar time, the following procedure was used. For EFX quotes, the final observation pair in each calendar time interval was recorded. The D2000-2 quotes used are the best limit bid and offer prices outstanding at the end of each interval. For EFX quote frequency and D2000-2 transaction frequency/volume, the number (or, in the case of volume, aggregate US$ quantity) of each event occurring in each interval was calculated. Finally, the number and total size of limit orders outstanding at the end of each calendar time interval were recorded. 

It has been suggested to us that our sampling convention of taking the final EFX observation in each interval biases our results against the EFX data due to their single dealer nature. The suggested alternative was to attempt to combine information from current and past EFX quotes in order to gain a set of bids and offers which better approximate at the market quotes. Our response to this is twofold. First, most work on EFX has used the quotes as we do. Second, implementation of this suggestion implies that one needs to use a necessarily ad hoc procedure for estimating the lifetimes of EFX quotes. There is no guarantee that such an ad hoc scheme will generate sensible output (e.g. non-negative spreads). 

2.2. Statistical features of D2000-2 and EFX data

Initially, we sample the data with a 20-second frequency. This frequency was chosen to balance two factors. On the one hand, one would like to sample very frequently in order to maximize the information content of the data, i.e. not omitting large numbers of observations in busy intervals. Second, however, sampling very finely leads to observation intervals which are empty (when markets are quiet) especially for the EFX data. The 20-second frequency represents a reasonable compromise between these two considerations. We investigate the impact of changing sampling frequency in Section 4.

2.2.1. Seasonal patterns

Fig. 1 clearly shows the existence of strong intra-day seasonal patterns in D2000-2 and EFX activity. Liquidity supply to D2000-2 is very light in the GMT evening and overnight period, while from 6 to 18 GMT the level of activity is very high.

---

4 More information on the structure of the raw data set and the processing of the data is available from the authors on request.

5 A very simple example would be to take the highest bid and lowest offer from the last 10 EFX quote pairs as the ‘market’ bid and offer at every observation time.

6 One exception to this is Bollerslev and Domowitz (1993).

7 For example, the minimum time between EFX quote arrivals was 2 seconds, a lower bound imposed by the processing constraints of the EFX system.
These hours correspond broadly to European and North American trading hours, reflecting the fact that we are examining the USD/DEM, the lack of impact of the D2000-2 system in Asia and the pre-eminence of London as FX trading center. EFX quote frequency correlates very strongly with D2000-2 transaction frequency. Both are high relative to their unconditional means in the periods covering 6 to 10 GMT and 12 to 16 GMT, while both measures are close to zero on average from 18 GMT to 6 GMT.

Certain differences between EFX and D2000-2 emerge in Figs 1(b) and 1(d). In the former, EFX quote frequency during the 10 to 12 GMT period stays relatively high while transaction frequency dips strongly on D2000-2. This is most likely to be due to the processing limits of the EFX system, i.e. the minimum time between quotes of 2 seconds, which will mask changes in activity at higher frequencies.8 Fig. 1(d) shows that there is essentially no intra-day seasonal pattern in the EFX bid–offer spread. D2000-2 spreads, on the other hand, vary widely across the GMT day.

8 The argument here is that the if the input to EFX is such that it is operating at maximum capacity, then marginal changes in quote inputs will not show up in the data which are displayed on screen.
During European and North American trading hours, D2000-2 spreads are an order of magnitude lower than those observed in the GMT evening and overnight period. As such, D2000-2 spreads seem to follow a similar intra-day pattern to the U-shape found for spreads observed on many major stock markets (see Foster and Viswanathan (1990) and Biais et al. (1995) for results from the NYSE and Paris Bourse). Hence EFX spreads are much greater than those on D2000-2 during peak D2000-2 trading hours and much lower outside of that period. This implies that using EFX spreads as a measure of liquidity underestimates true market liquidity throughout the hours from 6 to 18 GMT and overstates liquidity from 18 to 6 GMT.

Our analysis of activity patterns indicates that certain market activity statistics derived from the EFX data are likely to misrepresent actual FX market liquidity and trading activity. The EFX spread is particularly problematic. EFX quote frequency, on the other hand, shares a fairly similar seasonal pattern to aggregate liquidity demand measures on D2000-2 aside from the interval around midday.

2.2.2. Characterizing D2000-2 and EFX returns

A complementary picture emerges from a statistical analysis of the characteristics of the EFX midquote return, D2000-2 transaction price returns and D2000-2 midquote returns.9

Table 1 presents the first four moments of the three return series along with their first autocorrelation and a fifth-order Box–Ljung statistic. Given the evidence of strong seasonal patterns in activity from Fig. 1, the statistics are computed for seven non-overlapping subsamples of the entire data sample for each return series.10 It is immediately clear that EFX midquote returns are generally twice as volatile as D2000-2 returns. An exception to this result occurs for the overnight subsample where D2000-2 returns are much more volatile due to the effect of illiquidity. The volatility results are illustrated in Fig. 2. There seems to be little useful comparative information in the reported return skew and kurtosis figures. As one might expect, all three return series share similar patterns in the sign of skew across subsamples. Also all return series exhibit excess kurtosis.11

A great deal of prior work using EFX data has noted that returns contain a strong negative moving average component. Various explanations have been put forward for this phenomenon, e.g. the effect of idiosyncratic inventory positions of individual dealers, dealers working with different information sets, and noise in the EFX data. Analysis of our data demonstrates that the EFX data used here also displays negative first-order autocorrelation. Across the entire trading day the first return autocorre-

9 Goodhart et al. (1996) present comparative statistical results similar to those below. However, they compare the properties of tick-by-tick data from D2000-2 and EFX, rather than data series with common calendar time-sampling frequency. Hence, their results are not directly comparable with ours.
10 Each subsample corresponds to a given 2-hour interval from every trading day. Hence, each subsample contains four breaks (between the last observation in that period on day $k$ and the first on day $k+1$). When appropriate, these breaks are modeled in estimation using dummy variables.
11 Formal tests for the existence of the fourth moment of returns, using a procedure proposed by Danielsson and de Vries (2000), indicate that it is unbounded in all cases.
Table 1
Summary statistics for 20-second return subsamples

<table>
<thead>
<tr>
<th>Data</th>
<th>Mean</th>
<th>Variance</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>$\rho_1$</th>
<th>$Q(5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2000-2 midquote</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–8 GMT</td>
<td>0.00011</td>
<td>0.00014</td>
<td>-0.04</td>
<td>10.66</td>
<td>-0.14</td>
<td>63.28</td>
</tr>
<tr>
<td>8–10 GMT</td>
<td>-0.00001</td>
<td>0.00008</td>
<td>0.28</td>
<td>19.48</td>
<td>-0.10</td>
<td>37.65</td>
</tr>
<tr>
<td>10–12 GMT</td>
<td>0.00050</td>
<td>0.00019</td>
<td>-1.63</td>
<td>31.53</td>
<td>0.03</td>
<td>35.49</td>
</tr>
<tr>
<td>12–14 GMT</td>
<td>0.00011</td>
<td>0.00020</td>
<td>-0.82</td>
<td>20.66</td>
<td>-0.15</td>
<td>59.90</td>
</tr>
<tr>
<td>14–16 GMT</td>
<td>0.00028</td>
<td>0.00015</td>
<td>0.07</td>
<td>6.76</td>
<td>-0.19</td>
<td>76.75</td>
</tr>
<tr>
<td>16–18 GMT</td>
<td>-0.00024</td>
<td>0.00023</td>
<td>0.12</td>
<td>17.76</td>
<td>-0.33</td>
<td>205.65</td>
</tr>
<tr>
<td>18–6 GMT</td>
<td>0.00003</td>
<td>0.00075</td>
<td>-0.01</td>
<td>124.71</td>
<td>-0.26</td>
<td>742.01</td>
</tr>
<tr>
<td>EFX midquote</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–8 GMT</td>
<td>0.00006</td>
<td>0.00046</td>
<td>-0.52</td>
<td>14.29</td>
<td>-0.37</td>
<td>263.69</td>
</tr>
<tr>
<td>8–10 GMT</td>
<td>-0.00002</td>
<td>0.00034</td>
<td>0.06</td>
<td>6.46</td>
<td>-0.37</td>
<td>257.00</td>
</tr>
<tr>
<td>10–12 GMT</td>
<td>0.00050</td>
<td>0.00055</td>
<td>-1.42</td>
<td>34.45</td>
<td>-0.27</td>
<td>144.19</td>
</tr>
<tr>
<td>12–14 GMT</td>
<td>0.00013</td>
<td>0.00053</td>
<td>-0.13</td>
<td>6.75</td>
<td>-0.34</td>
<td>216.55</td>
</tr>
<tr>
<td>14–16 GMT</td>
<td>0.00029</td>
<td>0.00052</td>
<td>0.11</td>
<td>10.57</td>
<td>-0.41</td>
<td>318.37</td>
</tr>
<tr>
<td>16–18 GMT</td>
<td>-0.00022</td>
<td>0.00055</td>
<td>-0.24</td>
<td>14.06</td>
<td>-0.25</td>
<td>122.17</td>
</tr>
<tr>
<td>18–6 GMT</td>
<td>0.00002</td>
<td>0.00032</td>
<td>0.13</td>
<td>33.49</td>
<td>-0.15</td>
<td>442.98</td>
</tr>
<tr>
<td>D2000-2 prices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–8 GMT</td>
<td>0.00010</td>
<td>0.00014</td>
<td>-0.11</td>
<td>13.63</td>
<td>-0.06</td>
<td>17.24</td>
</tr>
<tr>
<td>8–10 GMT</td>
<td>-0.00001</td>
<td>0.00009</td>
<td>0.21</td>
<td>8.40</td>
<td>-0.09</td>
<td>38.39</td>
</tr>
<tr>
<td>10–12 GMT</td>
<td>0.00050</td>
<td>0.00024</td>
<td>-1.79</td>
<td>34.26</td>
<td>0.02</td>
<td>23.29</td>
</tr>
<tr>
<td>12–14 GMT</td>
<td>0.00010</td>
<td>0.00026</td>
<td>-0.34</td>
<td>13.83</td>
<td>-0.19</td>
<td>83.21</td>
</tr>
<tr>
<td>14–16 GMT</td>
<td>0.00029</td>
<td>0.00017</td>
<td>0.31</td>
<td>9.35</td>
<td>-0.07</td>
<td>19.93</td>
</tr>
<tr>
<td>16–18 GMT</td>
<td>0.00025</td>
<td>0.00012</td>
<td>-1.33</td>
<td>32.18</td>
<td>-0.06</td>
<td>22.78</td>
</tr>
<tr>
<td>18–6 GMT</td>
<td>0.00002</td>
<td>0.00037</td>
<td>2.03</td>
<td>727.06</td>
<td>-0.15</td>
<td>393.02</td>
</tr>
</tbody>
</table>

*The first four columns of the table give the sample mean, variance, skewness and kurtosis of returns. The next column gives the first-order return autocorrelation and the final column a fifth-order Box–Ljung statistic for return dependence. The 5% critical value for the Box–Ljung statistic is 11.07.

Furthermore, the Box–Ljung statistics indicate that the hypothesis of a lack of up to fifth-order autocorrelation can be rejected at the 5% level in all EFX return subsamples. A different picture emerges when we look at the autocorrelations for D2000-2 returns. These are generally between one-third and one-half the value for the corresponding EFX subsample. Furthermore, D2000-2 return autocorrelation displays a marked, inverted U-shape over the trading day, a pattern which does not appear in the EFX autocorrelations. This implies that D2000-2 return autocorrelation is inversely related to D2000-2 trading activity.

These patterns can be explained with our hypothesis regarding EFX quote-making behavior. Assume an agent who wishes to buy dollars. He submits a competitive bid to EFX and computes the ask by adding a large, standard spread. Another agent who, instead, wishes to sell dollars will enter a competitive offer and subtract the standard spread to gain his bid quote. Note that even if both of these agents agree on the competitive bid and offer, the midquote submitted by the first will exceed that submitted by the second. Then, if agents wanting to trade at the bid and ask...
Fig. 2. Intra-day pattern in return variances. These figures are based upon data sampled once every 20 seconds.

...arrive randomly to EFX, quote returns will contain (inflated) negative first-order autocorrelation. This is exactly the intuition used in the spread estimator of Roll (1984). Furthermore, this hypothesis delivers the result that EFX midquote returns should be much more volatile than the ‘true’ midquote returns.

A preliminary look at the relationships between the market activity measures derived from the D2000-2 and EFX data is given in Table 2 in the form of contemporaneous cross-correlations. These results mostly confirm the insights from the seasonal patterns. D2000-2 spreads are strongly negatively correlated with D2000-2 liquidity supply measures. D2000-2 transaction intensity correlates far less well with spreads, however. As expected, EFX spreads are extremely poorly correlated with all other series while EFX quote frequency is positively correlated with both D2000-2 liquidity and transaction activity. Hence, prior insights are corroborated in that...

Table 2
Cross-correlations of D2000-2 and EFX activity variables

<table>
<thead>
<tr>
<th>Data</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2 spread</td>
<td>1 ( \cdots ) ( \cdots ) ( \cdots ) ( \cdots ) ( \cdots ) ( \cdots ) ( \cdots ) ( \cdots )</td>
</tr>
<tr>
<td>D2 orders</td>
<td>-0.50</td>
</tr>
<tr>
<td>D2 size</td>
<td>-0.47</td>
</tr>
<tr>
<td>EFX spread</td>
<td>0.05</td>
</tr>
<tr>
<td>EFX quotes</td>
<td>-0.26</td>
</tr>
<tr>
<td>D2 deals</td>
<td>-0.13</td>
</tr>
<tr>
<td>D2 volume</td>
<td>-0.11</td>
</tr>
</tbody>
</table>

*S*preads are measured in percentage terms. Orders and size refer to the number of outstanding D2000-2 limit orders and their aggregate size, respectively. EFX quotes refers to the number of EFX quotes posted in each interval. Deals is a count of the number of transactions in each interval and Volume is the aggregate transacted volume in a given interval.
EFX spreads seem to convey little information regarding underlying spot market activity while EFX quote frequency can be regarded as a fair proxy for volumes and liquidity supply.\(^{12}\)

3. The inter-relationship between D2000-2 and EFX quotes

The previous section provides a general comparison of the patterns in and sample statistics derived from D2000-2 and EFX activity measures and returns. In this section, we focus solely on returns. Given the number of existing studies which use EFX returns as proxies for firm quote returns, a clear examination of their similarities and comovements would seem to be important. We analyze the dynamic relationship between the two quote return series using three techniques: return cross-correlations, a cointegrating vector autoregression (VAR) in returns and a bivariate GARCH specification. Hence we examine not only causality in returns but also return volatility.

3.1. Cross-correlation analysis

As a first pass at examining the lead–lag relationships in the two return series we calculated Box–Ljung \(Q\)-statistics for the cross-correlations between D2000-2 returns and leads and lags of EFX midquote returns.\(^{13}\) Fig. 3 plots the cross-correlations

![Cross-correlations between D2000-2 and EFX returns](image)

Fig. 3. Cross-correlations between D2000-2 and EFX returns (with 95% confidence band). The graph plots the correlation between D2000-2 returns and lagged EFX returns. A negative lagged EFX return is an EFX lead.

\(^{12}\) Using half-hour aggregates of the relevant series, Goodhart et al. (1996) also report a positive correlation between D2000-2 market order frequency and EFX quote frequency.

\(^{13}\) Table 1 indicates that both D2000-2 and EFX returns contain MA(1) components. We filter the MA(1) structure from each sample before constructing the cross-correlations. This prewhitening of returns implies that the estimated cross-correlations can be treated as independent and asymptotically normal with variance \(T^{-1}\) under the null of return independence.
between D2000-2 and EFX returns calculated using data from the period 6 to 18 GMT. It is immediately clear that there is strong asymmetry in cross-correlations with those for negative lags (i.e. leads) larger and more significant than those for positive lags. The implication is that D2000-2 returns tend to lead EFX returns (i.e. EFX returns are predictable with D2000-2 returns) while the converse is not true. Given the 20-second sampling of the data, the estimated cross-correlations imply that EFX returns are predictable between 2 and 3 minutes ahead using D2000-2 returns. As the D2000-2 midquote is formed from firm quotes and the EFX midquote is indicative, this is in line with intuition.

3.2. Cointegrating VAR analysis

The cross-correlation analysis in Section 3.1 gives rough answers to questions of information transmission between EFX and D2000-2 midquote returns. It ignores, however, the fact that the two midquote series are likely to be cointegrated. This cointegration comes from the fact that we are essentially examining two sources of price information for the same asset. In this section we present a bivariate vector error correction mechanism (ECM), suggested by Hasbrouck (1995), which takes account of the cointegration and allows one to conduct appropriate inference regarding the roles played by D2000-2 and EFX in price discovery.

3.2.1. VAR model

Hasbrouck (1995) proposes a cointegrating vector autoregression (VAR) procedure to determine the relative information content of two different price series for the same financial asset. The basic empirical model is

\[ r_t = \mu + \alpha z_{t-1} + \beta(L)r_{t-1} + \epsilon_t, \quad E(\epsilon_t\epsilon_t') = \Omega, \]

where \( r_t = (r_{tD2}, r_{tEFX})' \), is a vector containing the D2000-2 and EFX midquote returns, \( \mu = (\mu_1, \mu_2)' \), \( \alpha = (\alpha_1, \alpha_2)' \), \( \epsilon = (\epsilon_1, \epsilon_2)' \), and \( \beta(L) \) is a conformable polynomial in the lag operator. Equation (1) is just an error correction representation for the pair of return series where we assume that the difference between the two midquote variables (denoted \( z_t = q_t^{D2} - q_t^{EFX} \)) is I(0). This assumption guarantees that the two-price series cannot permanently diverge and is shown to be valid through a series of unit root tests on the midquotes and the difference between them.\(^{14}\) We refer to this difference as the pricing error from now on. In equilibrium, the EFX and D2000-2 midquotes are identical (i.e. \( z_t = 0 \)).

One measure of each system’s contribution to price discovery can be gained through comparison of the \( \alpha \) coefficients, i.e. how each series reacts to pricing errors. Given the way in which we constructed \( z_t \), one would expect \( \alpha_1 \) to be negative and \( \alpha_2 \) positive, but asymmetry in their sizes will indicate one system reacting more to

---

\(^{14}\) ADF test statistics for both quote series exceed \(-2.5\) and for both return series are less than \(-40\). Hence we conclude that the quote series are I(1). The ADF test statistic for the difference between the quote series is \(-17.1\), implying that the quote series are CI(1,1).
deviations from equilibrium than the other and hence an asymmetry in price discovery contributions (Harris et al., 1995). If one of the two quote series dominates completely in terms of information assimilation, the $\alpha$ coefficient for that return series is zero.

To calculate the information share of each source of USD/DEM quotes explicitly, we first compute the vector moving average representation of the system through inversion of Eq. (1):

$$r_t = \Phi(L)\epsilon_t.$$  \hspace{1cm} (2)

The long run impacts of the pair of innovations on both quote series are summarized in the value of $\Phi(1)$. Note that due to the cointegration between quotes, $\Phi(1)$ must contain two identical rows. Denote a row of $\Phi(1)$ by $\phi$.

By assuming that the VAR innovations are uncorrelated, the calculation of the information share of each price measure is straightforward. The information share for price measure $j$ is calculated as

$$S_j = \frac{\phi_j^2\Omega_{jj}}{\phi\Omega\phi}.$$  \hspace{1cm} (3)

However, the assumption that innovations are uncorrelated is unlikely to be satisfied in practice. Hence we must modify the preceding estimator. In order to bound the information share we use a Choleski decomposition of the variance–covariance (VCV) matrix. An upper bound on the information share of system $i$ is obtained from this decomposition when the innovation to equation $i$ is represented in the first row of the VCV matrix. A lower bound on $i$’s information share is obtained when system $i$ is represented in the second row of the VCV. Denoting the Choleski factor of $\Omega$ by $F$, the information share of series $j$ is then:

$$S_j = \frac{([\phi F]_j)^2}{\phi\Omega\phi},$$  \hspace{1cm} (4)

where $[\phi F]_j$ is the $j$th element of the given row vector. In the empirical results which follow, we present upper and lower bounds on the D2000-2 information share.

3.2.2. Empirical estimates

The properties of the pricing error ($z_t$) are presented in Table 3, broken down across intervals of the trading day. Several facts are immediately apparent. First, the difference between the two midquote series is small on average. Second, examination of Table 3 in conjunction with Fig. 1 shows that the variance of the pricing error covaries negatively with D2000-2 trading volume. In times of heavy D2000-2 activity, D2000-2 and EFX midquotes stay closer together on average. Finally, the dependence in pricing errors, measured by the first-order autocorrelation and a fifth-order Box–Ljung statistic, is also inversely related to D2000-2 volume. Hence, when trading volume is large, the pricing error is less persistent, implying that deviations from equilibrium are removed more speedily. Also, this final result suggests that the ECM structure in Eq. (1) will not be stable across the trading day and we therefore estimate the ECM separately for each of our trading day subsamples.
Table 3
Properties of pricing error between D2000-2 and EFX

<table>
<thead>
<tr>
<th>Subsample</th>
<th>$z$</th>
<th>$\sigma_z^2$</th>
<th>$\rho_{1z}$</th>
<th>$Q(5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6–8 GMT</td>
<td>−0.00088</td>
<td>0.00040</td>
<td>0.36*</td>
<td>386.8*</td>
</tr>
<tr>
<td>8–10 GMT</td>
<td>0.00030</td>
<td>0.00030</td>
<td>0.34*</td>
<td>405.9*</td>
</tr>
<tr>
<td>10–12 GMT</td>
<td>0.00134</td>
<td>0.00074</td>
<td>0.56*</td>
<td>1256.1*</td>
</tr>
<tr>
<td>12–14 GMT</td>
<td>0.00213</td>
<td>0.00040</td>
<td>0.23*</td>
<td>172.9*</td>
</tr>
<tr>
<td>14–16 GMT</td>
<td>−0.00141</td>
<td>0.00044</td>
<td>0.31*</td>
<td>391.0*</td>
</tr>
<tr>
<td>16–18 GMT</td>
<td>−0.00478</td>
<td>0.00093</td>
<td>0.59*</td>
<td>1445.8*</td>
</tr>
<tr>
<td>18–6 GMT</td>
<td>−0.00646</td>
<td>0.0069</td>
<td>0.92*</td>
<td>37706.3*</td>
</tr>
</tbody>
</table>

$z_t$ is the pricing error at $t$, $\rho_{1z}$ is the first autocorrelation of the pricing error. $Q(5)$ is the fifth-order Box–Ljung statistic for the pricing error. The 5% critical value for the Box–Ljung statistic is 11.07.

Table 4 contains the key ECM parameters for all seven trading day subsamples plus the estimated upper and lower bounds on the D2000-2 information share for each. The crucial ECM parameters are $\alpha_1$ and $\alpha_2$ and we expect the former to be negative while the latter should be positive. For all of the trading day subsamples, the sign of $\alpha_2$ is as expected, i.e. when the D2000-2 midquote exceeds the EFX midquote, the EFX midquote adjusts upwards. In only four of the six subsamples does $\alpha_1$ take the expected sign.

With regard to the relative size and significance of the parameters on the pricing error, a clear asymmetry is visible. For the subsamples covering 6 to 18 GMT, $\alpha_2$ is always an order of magnitude larger than $\alpha_1$ and is always highly significant. On the other hand, the lagged equilibrium error is only significant in three of six D2000-2 return regressions. Further, the $R^2$ for the EFX return equation is always an order of magnitude greater than that for the D2000-2 equation. These results imply that the majority of any disequilibrium in the system is removed through the adjustment of EFX quotes. D2000-2 quotes react very weakly to disequilibrium.

Table 4
Cointegrating VAR results for D2000-2 and EFX returns

<table>
<thead>
<tr>
<th>Subsample</th>
<th>Lags</th>
<th>$\alpha_1$</th>
<th>$t(\alpha_1)$</th>
<th>$R^2_{D2}$</th>
<th>$\alpha_2$</th>
<th>$t(\alpha_2)$</th>
<th>$R^2_{EFX}$</th>
<th>$D2_{\text{min}}$</th>
<th>$D2_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 a.m. to 8 a.m.</td>
<td>2</td>
<td>−0.05</td>
<td>−2.10</td>
<td>0.03</td>
<td>0.55</td>
<td>9.35</td>
<td>0.34</td>
<td>90.5</td>
<td>91.6</td>
</tr>
<tr>
<td>8 a.m. to 10 a.m.</td>
<td>1</td>
<td>−0.03</td>
<td>−1.77</td>
<td>0.02</td>
<td>0.56</td>
<td>16.26</td>
<td>0.33</td>
<td>94.7</td>
<td>95.8</td>
</tr>
<tr>
<td>10 a.m. to 12 p.m.</td>
<td>1</td>
<td>0.09</td>
<td>1.65</td>
<td>0.03</td>
<td>0.48</td>
<td>3.93</td>
<td>0.34</td>
<td>93.0</td>
<td>99.8</td>
</tr>
<tr>
<td>12 p.m. to 2 p.m.</td>
<td>1</td>
<td>0.06</td>
<td>0.20</td>
<td>0.02</td>
<td>0.72</td>
<td>22.48</td>
<td>0.41</td>
<td>92.9</td>
<td>1.00</td>
</tr>
<tr>
<td>2 p.m. to 4 p.m.</td>
<td>1</td>
<td>−0.05</td>
<td>−3.05</td>
<td>0.04</td>
<td>0.51</td>
<td>11.82</td>
<td>0.34</td>
<td>89.9</td>
<td>90.6</td>
</tr>
<tr>
<td>4 p.m. to 6 p.m.</td>
<td>2</td>
<td>−0.05</td>
<td>−2.43</td>
<td>0.14</td>
<td>0.27</td>
<td>8.40</td>
<td>0.17</td>
<td>86.0</td>
<td>88.8</td>
</tr>
<tr>
<td>6 p.m. to 6 a.m.</td>
<td>7</td>
<td>−0.03</td>
<td>−4.74</td>
<td>0.10</td>
<td>0.01</td>
<td>5.34</td>
<td>0.08</td>
<td>35.4</td>
<td>35.7</td>
</tr>
</tbody>
</table>

$\alpha_1$ and $\alpha_2$ are the coefficients on the lagged pricing error in the D2000-2 and EFX return equations, respectively. The columns immediately following give the $t$-values for these coefficients. The $R^2$ for the D2000-2 and EFX return equations are also presented and, finally, we present the lower and upper bounds on the D2000-2 information share.
Results are very different for the overnight subsample (6 p.m. to 6 a.m.). In this case the parameters on the lagged equilibrium error have similar size and significance level. This is likely to be due to the lack of activity on both D2000-2 and EFX during this period. A final point regarding the ECM results relates to those parameters which have been omitted from Table 4, specifically the coefficients on lagged EFX and D2000-2 returns in each equation.¹⁵ For the D2000-2 return equation, neither lagged own returns nor lagged EFX returns are significant in general. On the other hand, both sets of lagged returns are significant in the EFX equation with lagged D2000-2 returns generally having a positive effect and lagged EFX returns a negative impact on current EFX returns.¹⁶ Note that as equilibrium variation in the EFX quote occurs through the coefficient on \( z_t \), this predictability reflects inefficiency in the EFX quote-setting process. Finally, the last two columns of Table 4 present the estimated lower and upper bounds on the D2000-2 information share for each subsample. It is clear that for those hours of the day when D2000-2 is active (i.e. 6 a.m. to 6 p.m.), it is the dominant location for price discovery with a minimal information share in excess of 85%. For the overnight period, however, the lack of D2000-2 activity implies that its share drops to around 35%. Within the effective trading day, the information share on D2000-2 follows an inverted U-shape, such that the information share is minimized in the early GMT morning and early GMT evening.

### 3.3. Bivariate GARCH analysis

Below we examine the inter-relationships between D2000-2 and EFX return volatility, similar to our analysis of lead–lag relationships in returns, but using a bivariate GARCH specification. We use the residuals from the VAR specifications of Section 3.2.2 in estimation so as to remove any covariation between the series which arises from the previously estimated lead–lag relationships in returns themselves.

In their most general forms, multivariate GARCH models are notoriously hard to estimate and, for practical purposes, a restricted version can often be used without detriment to the application. Hence, we follow prior work and employ the BEKK specification introduced in Engle and Kroner (1995). In a bivariate setting, the structure of the BEKK model is as follows:

\[
\begin{align*}
\sigma_t = \sum_n \epsilon_n \sum_n = V'V + B'r_{t-1}'r_{t-1}B + A' \sum A, \\
\epsilon_t = (\epsilon_{t}^{D2000-2}, \epsilon_{t}^{EFX})', \text{ i.e. the vector of D2000-2 and EFX midquote returns}, \text{ } \epsilon_t \text{ is a } 2 \times 1 \text{ vector of NID}(0,1) \text{ innovations and } \Sigma_t \text{ is the conditional variance–covariance matrix at } t.
\end{align*}
\]

This specification has the advantage of being parsimonious, while capturing the relevant dynamics, and ensuring positive definiteness of \( \Sigma_t \). Each matrix, \( V, A, B \) is
2×2 and $V$ is restricted to be upper triangular. Hence we have 11 free parameters in the model.

Estimated parameters from the BEKK specification for residual returns are given in Table 5. From these parameters it is clear that both volatility series are strongly autocorrelated, a result which is in line with those from univariate GARCH specifications on intra-day FX data. The diagonal elements of $A$ and $B$ are all positive and strongly significant. An interesting feature of the results appears in the off-diagonal elements. The upper right coefficients of $A$ and $B$ (i.e. $a_{12}$ and $b_{12}$) are greater in magnitude and more significant than $a_{21}$ and $b_{21}$. This implies that D2000-2 volatility affects EFX return volatility to a greater degree than in the converse direction. A further result which agrees with those from univariate volatility models (and the prior kurtosis measures for returns) is that the coefficients in Table 5 imply that the unconditional variance–covariance matrix does not exist, a multivariate equivalent of IGARCH.

Hence, these results imply that volatility spillovers between the two series tend to be in the direction from real to indicative quotes.

### Table 5
Bivariate GARCH results: 20-second data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Raw data</th>
<th>Deseasonalized data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>$t$-value</td>
</tr>
<tr>
<td>$v_{11}$</td>
<td>$1.43 \times 10^{-5}$</td>
<td>38.37</td>
</tr>
<tr>
<td>$v_{12}$</td>
<td>$1.33 \times 10^{-5}$</td>
<td>6.02</td>
</tr>
<tr>
<td>$v_{22}$</td>
<td>$5.14 \times 10^{-5}$</td>
<td>52.95</td>
</tr>
<tr>
<td>$a_{11}$</td>
<td>0.96</td>
<td>1053.18</td>
</tr>
<tr>
<td>$a_{12}$</td>
<td>0.08</td>
<td>37.70</td>
</tr>
<tr>
<td>$a_{21}$</td>
<td>−0.02</td>
<td>−15.64</td>
</tr>
<tr>
<td>$a_{22}$</td>
<td>0.90</td>
<td>367.68</td>
</tr>
<tr>
<td>$b_{11}$</td>
<td>0.30</td>
<td>88.28</td>
</tr>
<tr>
<td>$b_{12}$</td>
<td>−0.14</td>
<td>−22.51</td>
</tr>
<tr>
<td>$b_{21}$</td>
<td>0.02</td>
<td>6.39</td>
</tr>
<tr>
<td>$b_{22}$</td>
<td>0.33</td>
<td>66.59</td>
</tr>
</tbody>
</table>

*Estimated coefficients from the bivariate GARCH model for residual return data and estimates of the parameters from the residual return data with deseasonalized volatility.

---

17 Previous research on intra-day FX volatility has shown that the intra-day seasonal patterns in volatility can bias estimated volatility process coefficients. Hence in Table 5 we also present estimated coefficients from the bivariate GARCH using residual returns with deseasonalized volatility. Results on volatility spillovers from the two tables are qualitatively similar.

18 The unconditional covariance matrix is given by $\Omega = (I - [A^\otimes A] - [B^\otimes B])^{-1} \text{vec}(VV)$. 

4. Temporal aggregation

The preceding analysis was based on a 20-second sampling of the data. However, much previous work on EFX has used much coarser sampling frequencies, e.g. 5 or 10 minutes. The reason for that may be suspicion regarding the quality of EFX quotes at extremely high frequencies. Below we examine how the statistical properties of D2000-2 and EFX returns alter as sampling frequency is varied from 20 seconds to 10 minutes. We consider the following comparative measures:

1. Return variance
2. Return autocorrelation
3. Cross-correlation of return series
4. $Q$-statistics for lead and lag cross-correlations
5. Pricing error autocorrelation

We report the results from these investigations in Fig. 4.

The autocorrelation evidence is presented in Fig. 4(a). The main features of this plot are that the D2000-2 autocorrelation is significant only for frequencies less than or equal to 180 seconds and is always closer to zero than the EFX coefficient. In addition, the EFX coefficient is significant until we reach a sampling frequency of 5 minutes.

Fig. 4(b) plots scaled variance measures for returns across sampling frequencies. Here we see that (scaled) D2000-2 volatility is fairly constant under aggregation while (scaled) EFX volatility decreases dramatically. At a 5-minute sampling frequency the difference between the two scaled variances is small, and this difference has almost disappeared at a 10-minute sampling frequency.

This evidence is reinforced in Figs 4(c) to 4(e) which show the contemporaneous cross-correlation of the two return series, lead and lag $Q$-statistics and pricing error autocorrelations across sampling frequencies. At very high frequencies (i.e. lower than or equal to 5 minutes), EFX returns are predictable with lagged D2000-2 returns and the pricing error is strongly autocorrelated. Such effects disappear with aggregation such that at a 10-minute sampling the first-order autocorrelation of the pricing error is less than 0.1 and the contemporaneous cross-correlation between D2000-2 and EFX returns exceeds 0.90.

The implication of this analysis is straightforward. The bias in EFX data is strongest at very high frequencies, but disappears with data aggregation. At a 10-minute sampling frequency, the indicative quotes have statistical properties much like those of firm quotes.

---

19 Throughout this section, data from the overnight period from 18 GMT to 6 GMT are omitted due to lack of activity on both systems.

20 The variance measures are scaled by the degree of aggregation such that we plot the 20-second return variance, one-third of the one-minute return variance, one-sixth of the 2-minute variance, etc.
5. Conclusion

In recent years, a large body of literature on empirical FX microstructure has emerged. Most of this work has been conducted using data derived from indicative EFX quotes because firm quotes and/or transaction prices have been unavailable at sufficiently high frequencies. Researchers have, however, remained skeptical of the
accuracy of EFX quotes. Perhaps for this reason, EFX data have generally been employed at relatively coarse sampling frequencies, e.g. 5 or 10 minutes. Our analysis vindicates this practice. Based on a 20-second data sampling we derive the following results:

1. EFX midquote returns are excessively volatile when compared to returns in the midquote derived from the best D2000-2 limit buy and sell prices.
2. While EFX returns are strongly negatively autocorrelated, D2000-2 returns have much lower negative autocorrelation.
3. Bid–ask spreads in D2000-2 data exhibit the expected intra-day U pattern, while EFX spreads do not.
4. Quote frequency on the EFX system and D2000-2 transaction frequency are fairly strongly positively correlated.
5. News is first incorporated into D2000-2 quotes, and impacts EFX quotes over the next 2 to 3 minutes. This implies that D2000-2 returns forecast EFX returns.
6. On average, over 90% of price relevant information enters the D2000-2 midquote first, and from there is impounded in the EFX midquote.
7. D2000-2 volatility impacts EFX volatility. The converse is not true.

These results can be explained by the following model of trader behavior. Most agents submitting a quote to EFX only really wish to deal on one side of the market, the sell side for example, and hence place an approximately competitive ask quote. To complete the quote a standard spread is subtracted to get the bid. This implies that spreads contain little or no information, quote returns will be excessively volatile and will display inflated first-order autocorrelations. By this hypothesis, traders consider the EFX screens as a forum for advertising, signaling willingness to buy/sell, while trades are conducted directly, possibly at different prices.

Our final results will be comforting to those that have used EFX return data with more coarse sampling frequencies. Measurement errors in EFX disappear with aggregation. Most anomalies in EFX returns have disappeared at a 5-minute sampling frequency and all are insignificant at a 10-minute sampling.

Acknowledgements

We would like to thank Casper De Vries, Charles Goodhart, Enrique Sentana, two anonymous referees plus seminar participants at LSE and CEMFI for helpful comments. All remaining errors are our own responsibility. We also thank Reuters Group PLC and Olsen and Associates for providing the data.

References


