Empir Econ (2015) 49:1431–1462
DOI 10.1007/s00181-015-0917-z

Informed traders' arrival in foreign exchange markets: Does geography matter?

Ramazan Gençay · Nikola Gradojevic · Richard Olsen · Faruk Selçuk

Received: 29 October 2012 / Accepted: 17 December 2014 / Published online: 18 February 2015

Abstract This article critically investigates the possibility that private information offering systematic profit opportunities exists in the spot foreign exchange market. Using a unique dataset with trader-specific limit and market order histories for more

Ramazan Gençay and Nikola Gradojevic gratefully acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada.
Faruk Selçuk passed away during this research.

R. Gençay
Department of Economics, Simon Fraser University, 8888 University Drive,
Burnaby, BC V5A 1S6, Canada
e-mail: rgencay@sfu.ca

R. Gençay · N. Gradojevic
The Rimini Centre for Economic Analysis, Rimini, Italy

N. Gradojevic (✉)
IÉSEG School of Management (LEM-CNRS), Lille Catholic University,
3, rue de la Digue, 59000 Lille, France
e-mail: ngradojevic@ieseg.fr

R. Olsen
Olsen Ltd., Seefeldstrasse 233, 8008 Zurich, Switzerland
e-mail: richardo@olsen.ch

R. Olsen
Centre for Computational Finance and Economic Agents (CCFEA),
University of Essex, Essex, UK

F. Selçuk
Department of Economics, Bilkent University,
06800 Bilkent, Ankara, Turkey

than 10,000 traders, we detect transaction behavior consistent with the informed trading hypothesis, where traders consistently make money. We then work within the theoretical framework of a high-frequency version of a structural microstructure trade model, which directly measures the market maker's beliefs. Both the estimates of the trade model parameters and our model-free analysis of the data suggest that the time-varying pattern of the probability of informed trading is rooted in the strategic arrival of informed traders on a particular day-of-week, hour-of-day, or geographic location (market).

Keywords Foreign exchange markets · Volume · Informed trading · Noise trading

1 Introduction

Information arrivals and the existence of informed and uninformed trades have received considerable attention in equity markets. Easley and O'Hara (1992) have introduced a sequential trade model in which a market maker learns from both trades and from the lack of trades. That is, a market maker's beliefs are continuously updated with new information that may or may not be reflected in a transaction price. As a result, the timing of the trade plays an important role in price formation. In a series of papers, Easley et al. (1996a, 1997a, b) expand this work by modeling an equity market in which a competitive market maker trades a risky asset with informed and uninformed traders. Easley et al. (2008) further extend these models to allow for the time-varying arrival rates of traders. They show that both informed and uninformed traders are highly persistent in equity markets. In contrast to this extensive strand of research concerned with equity markets, studies on the behavior of informed and uninformed traders in foreign exchange (FX) markets are scarce, mostly due to lack of transaction data for prices and trading volume.

The notable exceptions include Yao (1998), Lyons (1995), Payne (2003) and Marsh and O'Rourke (2005). Lyons (1995) and Payne (2003) use 1 week of trade-by-trade data while Marsh and O'Rourke (2005) use about one year of daily data. The current paper is based on a dataset from OANDA that contains tick-by-tick data over eight months of 2003–2004 or about 6.5 million data lines. The data are anonymized, but retain the specific trading histories for each of the traders included. To control for high-frequency noise effects and no-trade periods, we have aggregated the data to hourly data in the present paper.

This paper contributes to the study of informed trading effects in FX markets in several ways. First, using a structural model, we identify a geographic (time-of-day) component in the activity of informed and uninformed FX traders.[1] Moreover, from our unique high-frequency trading data for major FX rates, we track individual transactions to reveal trading behavior consistent with the informed trading hypothesis. Another important issue we address is the role of informed and uninformed traders in price

[1] As in Dacorogna et al. (2001) and, recently, Kaul and Sapp (2009), we use the following geographic regions to cover the 24-h trading day: 03:00–07:00 EST (Europe only), 07:00–11:00 EST (both Europe and North America), 11:00–15:00 EST (North America only), 15:00–19:00 EST (post-North America), and 19:00–03:00 EST (Asia).

discovery as well as their contribution to intraday price volatility. Considering the low-transparency feature of the FX market, these questions are of immense importance in understanding the market dynamics, which can lead to new empirical and theoretical additions to the emerging field of FX market microstructure. More importantly, we consider the possibility that informed traders strategically time their arrival in the FX market. Most FX microstructure models fail to account for this phenomenon and assume that the arrival of informed traders is directly related to the information flow.[2] We believe that such analysis ignores potentially strategic, intricate interactions that can occur in FX markets.[3]

We draw our conclusions from the data, as well as from a continuous time sequential microstructure trade model in the spirit of Easley et al. (1996b). This model measures market maker's beliefs and, thus, complements our direct evidence of informed trading. In our approach, informed trading has a broader meaning: it refers to the ability to consistently generate profit. Informed trading is interpreted as the ability to realize profits, where profits are the realization of private information. We find strong support for intraday geographic dependence in the arrival of informed traders, whereas the uninformed traders arrive uniformly. The dependence on the time of day pinpoints two target markets for informed traders: North America and Asia. It is worth noting that the above-average trading activity of informed traders coincides with low overall activity in both markets. This indicates that informed traders strategically attach the largest market weight to a particular regional location, which is similar to the findings of Goldstein et al. (2006) for equity markets. In addition, tracking the actual trading activity of the consistently profitable, thereby presumably informed traders, corroborates these findings. The observed geographic component in the activity of informed traders also constitutes evidence in favor of Covrig and Melvin (2002) who demonstrate the dominance of Japan in setting JPY/USD quotes.

A few very recent papers demonstrate the existence of geographic or time-of-day patterns in the FX market, but not from the perspective of informed and uninformed trading in a retail trading platform. Gençay and Gradojevic (2013) attribute the observed intraday trading activity patterns in the interbank electronic FX broking service mostly to region-specific private information. Concerning the regional activity and its impact on the FX rate movements, Breedon and Ranaldo (2013) and Ranaldo (2009) observe that domestic currencies appreciate (depreciate) systematically during foreign (domestic) working hours. This literature and the current paper are in agreement that FX traders in general trade mostly during their country's working hours. Our focus is, however, on the informational content of FX trading, while the latter two papers emphasize the role of (realized) domestic FX order flow in intraday price formation.

The fact that our findings provide both direct and structural evidence of informed trading confirms the validity of the modern FX microstructure literature that is built

[2] For example, in the framework by Lyons (2001), customers are the primary source of private information, but the implications of their strategic behavior are not considered.

[3] This line of reasoning has also been documented for equity markets. For example, Foster and Viswanathan (1994) find that the optimal strategy of better informed traders is to delay trading on his or her private information in the early rounds of trading, while trading very intensely on the common information.

on a premise that private information in FX markets exists. For example, Payne (2003) reveals substantial informed trading effects in an electronic FX market. Further empirical evidence on the existence of private information in spot FX markets can be found in Lyons (1995) and Yao (1998). A theoretical model involving asymmetric information in FX markets is provided in Vitale (2012). This recent paper shows how informed traders influence exchange rates by inventory management as well as through their private information. This work raises the important question of whether private information in the FX market is short-lived. A noteworthy finding is that long-lived private information enables traders to be more strategic in deciding how to optimally profit from their information advantage. This notion is exactly consistent with our results.

Several recent papers reveal evidence of private information in FX markets. Albuquerque et al. (2008) structurally identify marketwide private information from firm-specific information in an equity market and prove its relevance for both equity and FX markets. Osler and Vandrovych (2009) show that the trades by leveraged investors are consistently informative for the price. Using a general equilibrium framework, Evans and Lyons (2012) show that the information content in transaction activity (such as customer order flow) can predict exchange rates as well as macroeconomic fundamentals. Bjønnes et al. (2008) use trader size as a proxy for the degree of private information and document information asymmetries based on the interdealer transactions of three spot traders at a large Scandinavian bank. Our findings complement and expand this literature as we are able to track individual transactions of more than 10,000 traders and assess the profitability of potentially informed traders directly.

The model also reveals that the day-of-week effects represent a significant component of the trading by both types of traders. We show that the day-of-week arrival rates of informed traders are inversely related to the day-of-week probability of informed trading (PIN) values. In other words, when a high (low) arrival rate of informed traders is observed on a given day, the estimated PIN is low (high). Also, informed traders strategically follow the arrival rates of the uninformed traders. In contrast to our paper, Easley et al. (2008) do not find evidence of strategic behavior by informed traders. They document that uninformed traders seem to avoid informed traders by “herding.”

The second main contribution of the present study is to quantify the price impact of informed and uninformed trading. More specifically, as in Odders-White and Ready (2008), we acknowledge that the market maker’s expected loss from informed trading is a function of both the PIN and its likely impact on the price. We extract the information content of the estimates by measuring the price impact over time. We find that the estimates of some of the model’s probabilistic parameters can potentially be used to explain the fluctuations in daily FX returns. In addition, the hourly order imbalances have significant power in determining FX returns as predicted by Evans and Lyons (2002).

We also show that the arrival rates of informed and uninformed traders have significant power in determining hourly and daily FX rate volatilities. The impact of the uninformed traders’ arrival on daily volatility is about twice the magnitude of that for informed traders. On the other hand, the findings for the hourly data reveal persistent and dominant effects of the arrival of informed traders on FX volatility. In addition, there is no evidence of the link between the PIN and volatility. The PIN can also be interpreted as the proportion of total trading activity accounted for by informed trading (or the trade composition). The fact that it is uninformative for volatility means that after sufficient trading activity takes place, information is fully reflected in the price. More precisely, when the information is fully revealed, the trade composition (which by definition captures the degree of information discovery) does not matter for volatility.[4]

The paper is organized as follows. Section 2 describes the data and the estimates of the benchmark model. Section 3 examines the informativeness of the model estimates for the FX rate dynamics, and Sect. 4 summarizes the study's conclusions.

2 Data and estimation

The arrival of modern internet technology has had a big impact on the structure of financial markets and in particular the currency markets. The FX market is the largest financial market in the world with a daily turnover of 4.0 trillion USD, as estimated by the Bank of International Settlement.[5] The FX market is an over-the-counter market, where the bulk of transaction volume is handled directly by banks and not by an exchange. Customers trade their currencies with one or several banks, who are market makers quoting bid and ask prices. Banks internalize customer flow of buyers and sellers, and only when their exposure exceeds certain thresholds, they hedge the positions with other market makers. The big banks trade their excess position inventory over EBS and Reuters. With the arrival of the internet technology, there emerged new platforms, such as FXAll and Currenex,[6] that aggregate price feeds of banks, offering a streamlined access to the FX market. The arrival of these new platforms and the introduction of internet trading platforms operated by the banks increased the efficiency of the FX markets. Today, traders routinely screen the prices of the competing market markers searching for arbitrage opportunities, and any price inefficiencies are immediately exploited.

In parallel with this development, a set of market makers entered the space offering internet-trading platforms targeted for the retail traders. Retail trading is not the biggest segment of the FX market, but it grows quickly and is therefore interesting. For instance, King and Rime (2010) report that the increased FX turnover from 2007 to 2010 is, among other factors, driven by the emergence of retail investors (individuals and smaller institutions). Retail FX trading takes place over the internet via trading platforms or the so-called retail aggregators. These institutions are financial firms (i.e., FX intermediaries) that give households and individuals better access to FX trading. They aggregate bid-offer quotes from the top FX-dealing banks thus facilitating trades by retail investors. Retail investors are attracted to the FX market by its long trading hours, market liquidity, tight spreads, and the possibility of margin trading. There is no

4 In a related study, Easley et al. (2008) find that the trade composition does not forecast intraday volatility. Lei and Wu (2005) also examine the time series properties of the PIN for a panel of stocks, arguing that the Easley et al. (1996b) model should be extended with the time-varying PIN.

5 See Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity in April 2010: http://www.bis.org/publ/rpfxf10t.

6 www.fxall.com and www.currenex.com.

(or little) price discovery taking place at the retail level, but, given the increased trading volume, banks are keen on getting business from retail platforms. Instead, price discovery happens in the interdealer market, where geographic location and hours of operation are the key factors that determine the information content of FX trades (D'Souza 2008). Papers such as Hau (2001), Menkhoff and Schmeling (2008, 2010) and Peiers (1997) also report geographic sources of informational advantage in trading and local price leadership. A refinement of this literature by Moore and Payne (2011) reveals that information is concentrated among FX dealers that trade most frequently, are located on larger trading floors, and specialize their activity in a particular rate. At the retail level, information is obtained through social interactions that contribute to the growth of active strategies (Heimer and Simon 2012). In addition, Nolte and Nolte (2012) find that retail investors' future FX order flow is driven by information extracted from past price movements.[7]

A company that has proven to be particularly successful as an internet retail aggregator is OANDA.[8] The company has built a highly scalable trading platform designed for retail and institutional traders. The processing efficiency enables OANDA to offer the exactly same terms to retail traders and institutions. The traders do not pay a commission, the whole transaction cost is included in the spread. Traders can for example buy anywhere from one US Dollar (USD) up to 10 Million USD per transaction at the same spread, which can be as low as 0.9 basis points, equivalent to 0.009 %.[9]

Because OANDA is competitive in both the retail and institutional market, it has succeeded in building a truly international customer base that includes retail traders, professional traders, hedge funds, and a limited number of corporations. OANDA's customer base is a representative sample of the active trading community of the global FX market. The availability of anonymized transaction data is unique and no similar dataset exists to our knowledge. The data offers the opportunity to research the trading behavior of a large sample of traders that trade at exactly the same terms, which is important because transaction costs determine, when positions are opened and closed.

2.1 OANDA FXTrade and model-free analysis

Our dataset consists of tick-by-tick FX transaction prices and the corresponding volumes for several exchange rates from October 1, 2003 to May 14, 2004.[10] The number of active traders during this period is 4,983, and they mainly trade four major exchange rates.[11] The data show that the overall trading frequency increases from Monday to Wednesday (the peak) and falls from Thursday to Saturday. Using the trader's identity

7 We are grateful for this and other useful comments from an anonymous referee.

8 www.oanda.com.

9 One basis point is defined as 1/100th of a percentage point.

10 In 2003/2004, the platform of OANDA did not have many decision support tools, so the traders were not well connected to the professional trading community. If traders can generate profits in a highly liquid and efficient market, such as the FX market, without access to "inside" information, then this raises interesting questions of why and how this can be feasible.

11 By "active," we refer to traders that did not simply receive interest on their positions, but placed orders during this period. The market share of these traders is approximately 86.4 %.

(trader ID), we next investigate the number of currency pairs traded by investors. We find that about 33 % of the investors specialize in exactly one currency pair, about 11 % in two currency pairs, and about 9 % in three currency pairs. This decreasing trend leads to only 2–4 % of active traders who deal in between 10 and 13 currency pairs. Hence, traders appear to specialize in a small number of currency pairs.

Since the bulk of all transactions (approximately 40 %)[12] involve only Euro–U.S. Dollar (EUR–USD) trading, we focus on transactions involving only EUR–USD. In particular, we analyze all EUR–USD buy and sell transactions (market, limit order executed, margin call executed, stop-loss, and take-profit transactions). In addition to price and volume, we know the trader ID for each transaction, which ranges from 123 to 5,904.[13] The average number of EUR–USD transactions per trader is 512. Using the trader ID, we observe that a few traders transact very frequently in this currency pair (between 10,000 and 25,000 times) over the time period that spans the data (Fig. 1). We also observe that day-of-the-week trading patterns for the EUR–USD transactions (trading frequency and volume) are similar to those of other currency pairs.

Further investigation shows that about 98 % of the EUR–USD traders close their positions with a single trade and that about 90 % of those "round-trip" transactions are closed intraday. The average duration of round-trip transactions ($\bar{\tau}$) is about 4 h. Informed trading activities are assumed, if a trader makes consistent profits. We find excess round-trip profits in each month of our sample for 29 traders (and for 42 traders in the first four months of 2004). One interesting finding is that the trader with the highest excess return in the EUR–USD market is also the most successful one in the USD–CHF market. We also identified a single trader that dominates six different markets (AUD–JPY, AUD–USD, EUR–JPY, GBP–JPY, GBP–USD, and USD–CAD) and another that dominates three currency pairs (EUR-CHF, EUR-GBP, and GBP-CHF). Next, to verify the robustness of our findings, we apply a simple, model-free technique. We compute the price impact of the time t signed transaction on the change of mid-quote from time t to time $t + \tau$. The idea is to reveal whether (round-trip) trades "predict" mid-quote movement. On average, we find that the 29 consistently profitable traders[14] are on the correct side of the trade about 90 % of the time and that the total profit per trader is roughly \$16,000. In comparison, an average trader in this market is correct in about 50–60 % of his or her trades. Taken together, these results exhibit transaction behavior compatible with the informed trading hypothesis.

One may argue that over time, purely random trading will lose money via the spread, but over any fixed time period, there is a chance that random trades will produce a profit. Note, however, that we find profitability in each of the eight months. The probability that a trader will be consistently profitable (or not profitable) over the period of eight months is $(1/2)^8 = 0.0039$. The ratio of the profitable traders to the total active traders is $(29/4, 983) = 0.0058$, which is larger than 0.0039. Alternatively, based on a binomial distribution with parameters $n = 4, 983$ and $p = 0.0039$, the

[12] The next most active currency pairs are USD-CHF (7.88 % share), GBP-USD (7.81 % share), USD-JPY (6.42 % share), and AUD-USD (5.98 % share).

[13] Of the 10,000 registered traders, about 5,000 were active in the EUR–USD trading over the sample period.

[14] Trader ID is withheld to preserve data confidentiality.

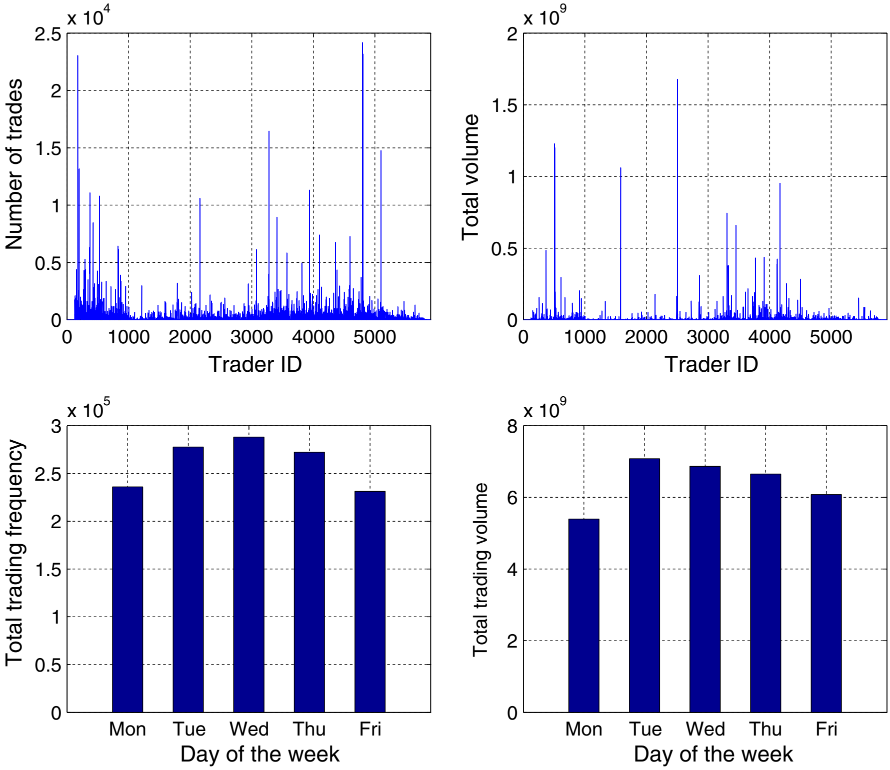

Fig. 1 *Top left* Total trading frequency (number of trades) per trader (trader ID). *Top right* Total trading volume (in the units of base currency—EUR) per trader (trader ID). *Bottom left* Total trading frequency (number of trades) for each day of the week. *Bottom right* Total trading volume (in the units of base currency—EUR) for each day of the week

probability of observing 29 or more successful traders in all eight months by pure chance is approximately 0.025 or 2.5 %.

It would be interesting to observe when the consistently profitable traders submit their orders. We find that, on average, they enter the market sometime between 09:00 EST and 10:00 EST. Of all the opened positions, about 95 % are closed by 19:00 EST, and only in about 5 % of the cases, they are left overnight. This indicates that these potentially informed traders are mostly active during the North American trading hours. In the following section, we will compare the model-free findings to the ones implied by the theoretical model.

Table 1 summarizes the institutional characteristics of the OANDA FXTrade trading platform. This platform is an electronic market making system (i.e., a market maker) that executes orders using bid/ask prices that are realistic and prevalent in the marketplace. The prices are determined by their private limit order book. The OANDA FXTrade policy is to offer the tightest possible bid/ask spread (e.g., 0.0009 % spread on the EUR–USD, regardless of the transaction size). Like most market makers, they profit from the spreads. Some of the other market features include continuous, second-by-second interest rate payments, no limit on the transaction size, no requirement for

Table 1 OANDA FXTrade institutional characteristics

Hours of operation	24 h/7 days per week
Number of currency pairs	30
Number of active traders	4,983
Number of trades (EUR–USD)	667,030 sell transactions
	666,133 buy transactions
Average number of trades per hour (EUR–USD)	192 sell transactions
	191 buy transactions
Total volume (EUR–USD)	32.6 billion USD
Average volume per day (EUR–USD)	224 million USD
Transaction types	Buy/sell market (open or close)
	Limit order buy/sell
	Cancel order (reason: bound violation, insuff. funds, none)
	Change order
	Change stop loss (sl) or take profit (tp)
	Sell/buy tp (close), sell/buy sl (close)
	Buy/sell limit order executed (open or close)
	Order expired
	Sell/buy margin called (close)
	Interest

minimum initial deposit, no charges for stop or limit orders, free quantitative research tools, and margin trading (maximum leverage of 50:1). Given these market characteristics, the OANDA FXTrade seems to be designed to attract small, uninformed traders. However, given the above findings, it is reasonable to assume that informed traders are also present in this market.

The theoretical model by Easley et al. (1996b) is developed in the context of equity markets. Adapting it to the FX market requires care. As opposed to the equity market, the FX market is open 24 h and is decentralized. Further, unlike the NYSE, it does not involve a so-called specialist responsible for maintaining fairness and order, with an insight into the limit order book. While the NYSE has recently introduced an open limit order book that provides a real-time view of the limit order book for all NYSE-traded securities, the FX market exhibits a low level of transparency. Finally, trading in the FX market is motivated by speculation, arbitrage, and, importantly, inventory management of currencies. Dealers in the FX market are generally quick to eliminate inventory positions (typically, they try to clear the inventory within seconds or minutes for large inventory imbalances). This process is sometimes referred to as "hot-potato-trading" (Evans and Lyons 2002; Bjønnes and Rime 2005). On the NYSE, however, inventory has an average half-life of over a week (Madhavan and Smidt 1993). Thus, inventory management is an important component of intraday FX trader activity.

The features of the OANDA FXTrade allow us to view it as a "special case" of the FX market that can be approached using the model by Easley et al. (1996b). First, as a market maker, the OANDA FXTrade promotes transparency: spreads are clearly visible, past spreads are published for public view and current open orders on major pairs are visible to all market participants. In regards to trader behavior, as we only focus on the informational aspect (i.e., informed vs. uninformed), market participants in the FX market can be treated in a fashion similar to those in equity markets. Section 3 suggests how intraday inventory management may relate to the findings.

Our preliminary analysis indicates that overall market activity was extremely low on certain days or during certain weeks. Therefore, we eliminate weekends, starting from every Friday 15:59:59 to Sunday 15:59:59 (all times are EST), including Christmas week (December 22–26), the first week of the year (December 29–January 2), and the week of Easter (April 5–9). This leaves us with 145 24-h periods. In order to avoid extremely high-frequency noise and no-activity periods in small time windows, we aggregated the data over 1-h intervals. Aggregating over trading intervals smaller than 1 h is not feasible, as this would not cover a sufficient number of buy and sell transactions for the model to be empirically applicable. On the other hand, longer trading intervals would "stretch" the assumptions of the model to a certain extent. For example, the news is assumed to arrive hourly (with probability α). It is unlikely that the flow of information would be less frequent, i.e., over longer time intervals. The final sample size is 3,480 hourly data points covering 145 business days, from October 5, 2003, 16:00 to May 14, 2004, 16:00 EST. There are 667,030 sell and 666,133 buy transactions in the sample period, with an average of approximately 6 transactions (3 buy and 3 sell) per minute. The transaction volume totals 32.6 billion USD. According to the BIS Triennial Survey for 2004, the daily average turnover in the EUR–USD currency pair was 501 billion USD. Hence, on average, our data represent about 0.045 % of the global daily EUR–USD trading volume. Nevertheless, it is one of the largest tick-by-tick FX datasets to be used in an academic study.

Since the trader's identity is known for each transaction (as stated before trader identities are anonymized), we are able to identify the number of unique traders in each 1-h window. For estimation purposes, the number of buy arrivals in each hour (B'_t) is defined as the number of unique traders involved in buy transactions in that hour. The number of sell arrivals in each hour (S'_t) is defined similarly. Therefore, the arrival of an individual trader who conducts several buy (sell) transactions in a given hour is counted as one buy (sell) arrival in that hour.

Figure 2 illustrates the number of hourly buy and sell arrivals (B', S') and the sample autocorrelation functions. We see strong daily time dependence and a time trend in both series. Therefore, we first estimate the linear time trend, $\hat{B}_t$ and $\hat{S}_t$, from the trend regression, which is free of temporal and irregular fluctuations. Assuming multiplicatively separable time dependence, we divide the original series by the trend estimates $\hat{B}_t$ and $\hat{S}_t$ to obtain an estimate of the time component

$$\tilde{s}_t^B = \frac{B'_t}{\hat{B}_t}, \quad \tilde{s}_t^S = \frac{S'_t}{\hat{S}_t}.$$

In order to estimate a time index for each hour of the day, we average the values of $\tilde{s}_t^B$ and $\tilde{s}_t^S$ corresponding to the same hour of the day across the sample and obtain

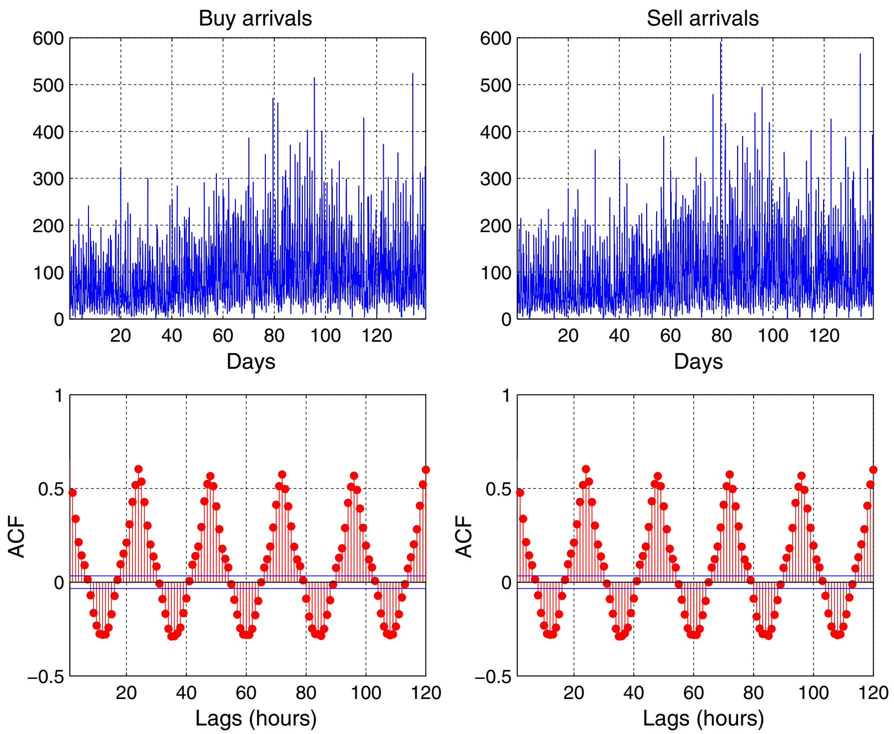

Fig. 2 The number of hourly buy (*top left*) and sell (*top right*) arrivals (B', S') and the sample autocorrelation functions at 120 hourly lags (5 days). The buy and sell arrivals in each hour are defined as the number of *unique* traders involved in buy or sell or both types of transactions in that hour. Sample period: October 5, 2003, 16:00—May 14, 2004, 15:59 (3,480 h, 145 business days)

the final hour-of-day indices s_i^B and s_i^S, $i = 1, 2, \ldots 24$ for the 24-h cycle. The 24-h adjusted number of buy and sell arrivals are obtained via

$$B_i = \frac{B_i'}{s_i^B}, \quad S_i = \frac{S_i'}{s_i^S}, \quad i = 1, 2, \ldots, 24$$

for each of the 145 days in the sample.

Figure 3 studies the final hour-of-day indices s_i^B and s_i^S, $i = 1, 2, \ldots 24$, for the number of unique buy and sell traders starting at midnight 00:00 EST. The average number of unique buy and sell traders increases after midnight, rising above the hourly average before the opening of the European market (at 03:00 EST). Similarly, we observe a sharp increase in the hours before the opening of the North American market (at 07:00 EST). The number of traders declines after 10:00, falling below the hourly averages after 14:00. They remain relatively low and stable in the subsequent hours until midnight. In the lower panel of Fig. 3, the sample autocorrelation functions of the diurnally adjusted number of buy and sell arrivals are studied at hourly lags. The removal of the daily temporal component reveals strong persistence in both series.

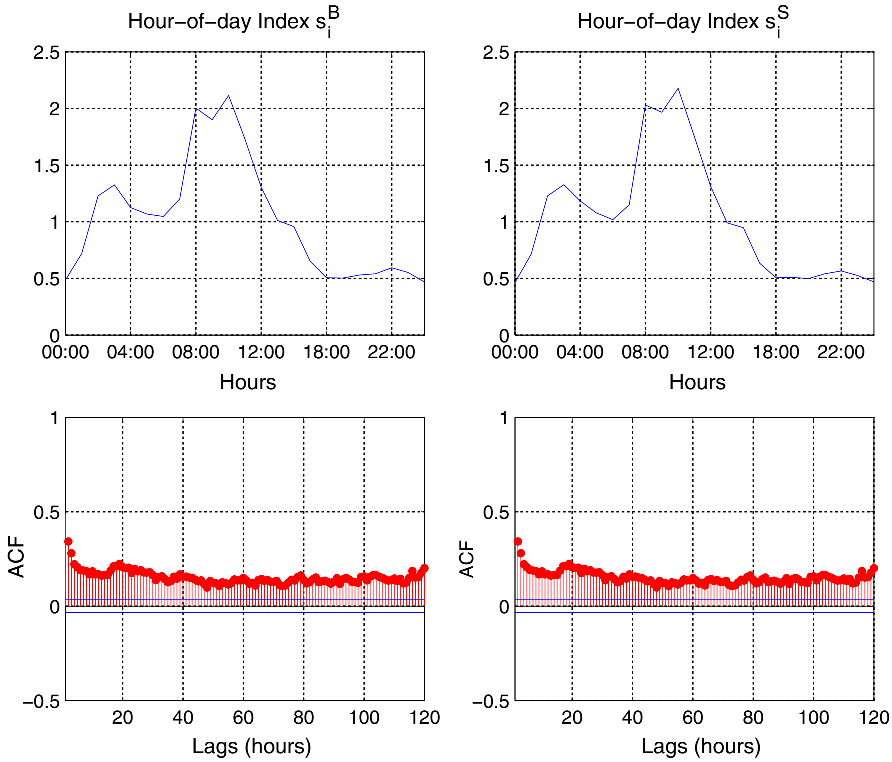

Fig. 3 Top: Hour-of-day indices s_i^B and s_i^S, $i = 1, 2, \ldots 24$, for the number of unique buy (*top left*) and sell (*top right*) traders starting at midnight 00:00 EST. *Bottom* Sample autocorrelation functions at 120 hourly lags (5 days) of the number of unique buy and sell traders

2.2 Informed and uninformed trades: When do they arrive?

According to the Easley et al. (1996b) model [please see Appendix 1: probability of hourly arrival of news (α), probability of bad news (δ), arrival rate of uninformed traders (ε) and arrival rate of informed traders(μ)], the expected value of the total number of trades per unit time, $E(TT) = E(S + B)$, is equal to the sum of the Poisson arrival rates of informed and uninformed trades:

$$E(TT) = \alpha(1-\delta)(\varepsilon + \mu + \varepsilon) + \alpha\delta(\mu + \varepsilon + \varepsilon) + (1-\alpha)(\varepsilon + \varepsilon) = \alpha\mu + 2\varepsilon$$

The expected value of the trade imbalance $E(K) = E(S - B)$ is given by

$$E(K) = \alpha\mu(2\delta - 1),$$

which provides information on the arrival of informed trades. When μ is large, the following approximate relation holds:

$$E(|K|) \simeq \alpha\mu.$$

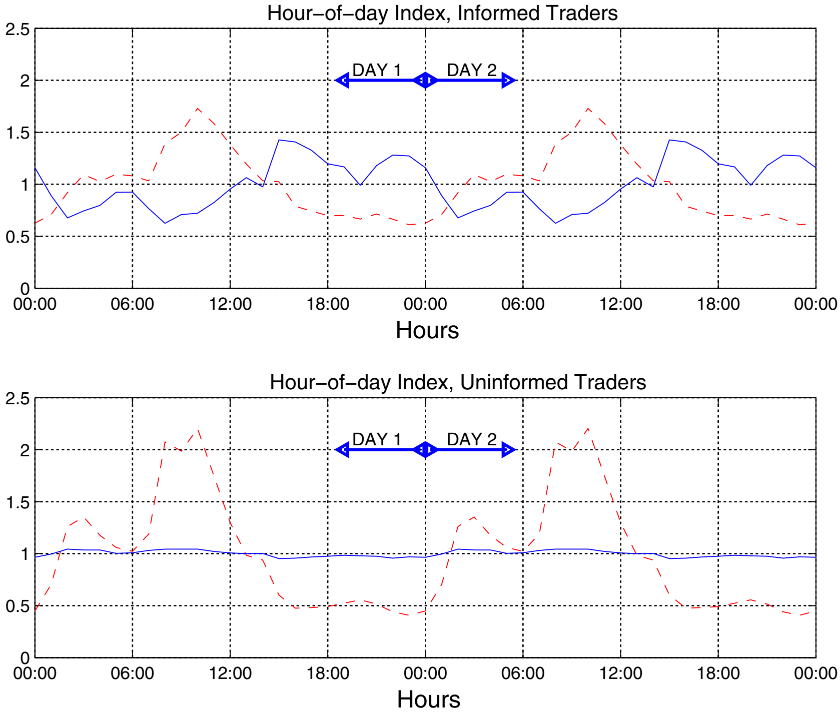

Fig. 4 Hour-of-day indices of informed (*top*) and uninformed (*bottom*) traders over 48 h, based on unbalanced traders ($|K|$) and balanced traders ($TT - |K|$). The solid line represents the data free of intraday fluctuations (the "de-seasonalized" data), while the dashed line represents the raw data. The hour-of-day index of uninformed traders is relatively stable for the de-seasonalized data, indicating a non-strategic, uniform arrival. The same index of informed traders fluctuates according to the observed regional dependence. For the raw data, the hour-of-day indices of both informed and uninformed traders exhibit similar patterns

Accordingly, the absolute trade imbalance $|K|$ provides information on the arrival of informed trades, $\alpha\mu$, while the difference between the total trade TT and the absolute trade imbalance $|K|$ contains information on the arrival of uninformed trades, ε. Note that our measure of the "number of buy and sell trades" in a given time period is the number of *unique* traders. Therefore, we can substitute the term "trader" for "trade" in the above expressions.

If we assume that the probability of information events α is constant, the hour-of-day average of the absolute trader imbalance $|K|$ provides information on the intraday time dependence of the orders from informed traders. In other words, since we know the number of unique individuals and corresponding trades at each hour of the day, we can obtain a measure of *the activity time* of the informed traders. We can similarly identify whether uninformed traders (liquidity traders) follow a distinct intraday pattern.

Figure 4 plots the hour-of-day indices of informed (top) and uninformed (bottom) traders based on unbalanced traders ($|K|$) and balanced traders ($TT - |K|$). Note that both $|K|$ and TT are calculated from B_t and S_t so that we do not expect any hour-of-day effect *a priori*.[15] However, the hourly activity of uninformed (liquidity) traders increases before the openings of the European, North American, and Asian markets (at 03:00, 07:00 and 19:00, respectively). Activity exceeds the hourly average from 01:00 until 14:00. Notice that the variation in the hour-of-day index of uninformed traders is relatively small, fluctuating between 0.95 and 1.04. Therefore, we may speculate that uninformed traders arrive uniformly at any time of the day. However, the hour-of-day index of informed traders is almost the opposite of that of uninformed traders. The volatility of informed traders during the day is high, with the index fluctuating between 0.63 and 1.43. The number of informed traders falls sharply after 01:00, before all three major markets open. It picks up around 03:00 at the opening of the European market, peaks and then dips down before the opening of the North American market. The number of informed traders increases sharply from the opening of the North American market until the market closes (at 15:00). This is followed by a decline until the opening of the Asian market (at 19:00). Above-average activity in Asia lasts until 01:00 and declines to the market close (at 03:00). The number of unique informed traders is well above the average after 15:00, i.e., it extends to the post-North American trading. This number remains above the average (except at the opening of the Asian market, at 19:00) until before the opening of the European market (at 03:00). To conclude, informed traders appear to primarily target the North American market and to a certain extent the Asian market. Recall from Fig. 2 that these are the hours with the fewest traders present in the market. The hour-of-day indices do not indicate the above-average arrival of informed traders during the European market trading hours.

2.3 Estimation of the Easley et al. (1996b) model

As mentioned above, the sample estimates of $E(TT)$ and $E(|K|)$ provide prior information about the parameters of the Easley et al. (1996b) model. The sample mean of total unique trades $\overline{TT}$ is 176.9 while $\overline{K} = -1.4$ and $\overline{|K|} = 19.4$. From the equation above,

$$\frac{E(K)}{E(|K|)} = \frac{\alpha\mu(2\delta - 1)}{\alpha\mu} = \frac{-1.4}{19.4} = -0.07$$

Accordingly, the implied probability that an event is bad news is 0.47 ($\bar{\delta} = (1 - 0.07)/2 = 0.47$). Uninformed buy and sell traders arrive at an hourly rate of ε. As

$$E(TT) = \alpha\mu + 2\varepsilon = 176.9 = 19.4 + 2\varepsilon,$$

[15] We also plot the hour-of-day indices for the raw data that use unadjusted B_t and S_t (dashed line). Similar arrival patterns are observed for the two types of traders. Thus, "hidden" hour-of-day patterns are present even after B_t and S_t are de-seasonalized. Since these effects are strong enough to persist even after adjusting for intraday time dependency, we will concentrate on the de-seasonalized data for the remainder of the paper.

Table 2 Day-of-week indices of estimated parameters and the PIN

Index	Monday	Tuesday	Wednesday	Thursday	Friday
SI_α	1.14	0.89	0.91	1.1	0.96
SI_δ	0.76	1.15	0.95	1.02	1.13
SI_ε	0.84	1.08	1.09	1.05	0.94
SI_μ	0.85	1.09	1.14	0.94	0.98
SI_{PIN}	1.16	0.92	0.91	1.01	1.00

The day-of-week indices are found using the ratio-to-moving average method

then $\hat{\varepsilon} = (176.9 - 19.4)/2 = 78.8$. Another measure based on the parameters of the Easley et al. (1996b) model is known as the PIN (see Appendix 2):

$$\text{PIN} = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} = 19.4/176.9 = 0.11,$$

for the market maker's initial belief. The estimated PIN in equity markets is usually between 0.15 and 0.25. This is relatively low, which may indicate that the number of informed traders is low (small μ), that the probability of the information event is low (small α), or both. In this particular case, the market maker's risk of informed trading is relatively low.

This estimation uses the above priors as the initial parameter set. That is, we set $\varepsilon = 78$, $\delta = 0.47$. Since we do not have a prior for α and μ separately, we assume $\alpha = 0.50$ so that $\mu = 38$. The log likelihood function in Eq. (11) from Appendix 1 is maximized daily ($T = 24$) for the entire sample period (145 days). As a result, we have 145 different estimates of each parameter. The two probability parameters α and δ are restricted to (0, 1) and the two arrival rates to (0, 500), since the maximum number of observed unique buy or sell traders in our sample is 474.

Table 2 reports the seasonal indices (day-of-week index) of the estimated parameters and the PIN.[16] The probability of an event α is higher on Mondays and Thursdays. Given that an event occurs, the probability that it is a bad event δ is lower than the average on Mondays and Wednesdays. Therefore, we speculate that Mondays were eventful, with good news for EUR–USD during the sample period. This may reflect that there is a flow of information over the weekend that is not impounded into prices during the weekend. Although it is possible to trade during the weekend, liquidity is so low that it is a different market and price discovery.

Figure 5 plots the daily estimates of the probability of an event α (top left) and the probability that an event is bad news δ (top right).[17] The estimated probability of an event $\hat{\alpha}$ fluctuates between 0.04 and 0.56, with an average of 0.30. This implies that there were no days without at least an event per hour. The lowest estimate is 0.04,

[16] The day-of-week indices, denoted by SI_i ($i \in \{\alpha, \delta, \varepsilon, \mu, \text{PIN}\}$), are found using the ratio-to-moving average method.

[17] The estimates over 145 days are stable with regard to the reasonable choice of their starting values. The only case in which the estimates begin to substantially change is when $\mu_0 > 200$ and $\varepsilon_0 > 200$.

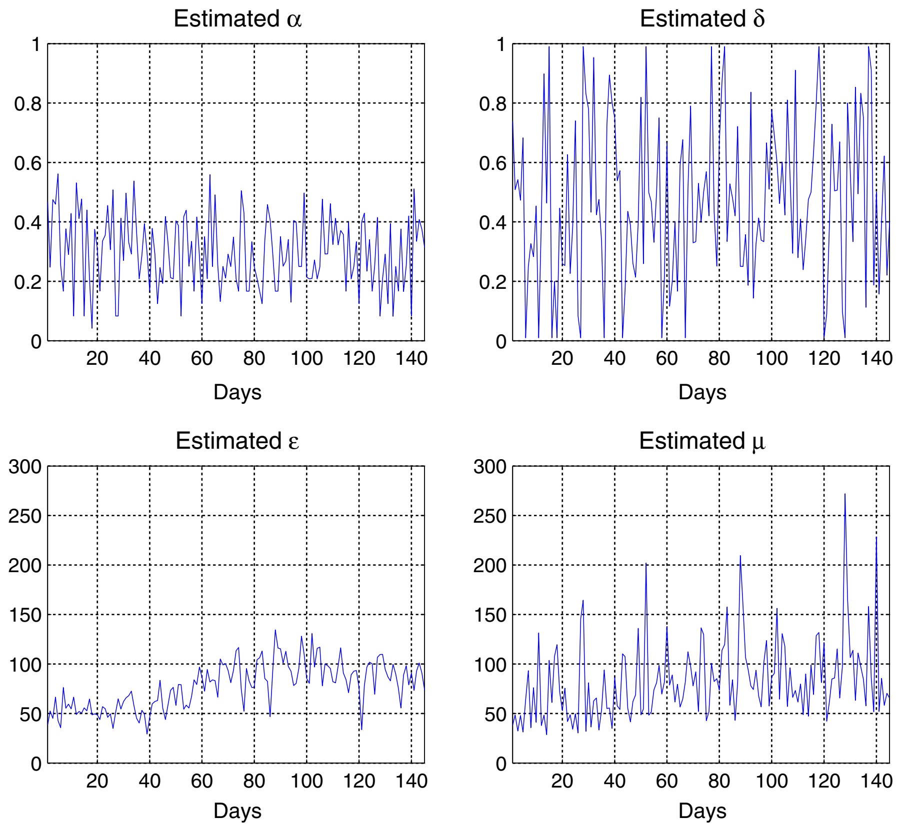

Fig. 5 Daily estimates of the probability of an event α (*top left*), the probability that an event is bad news δ (*top right*), the arrival rate of uninformed traders ε (*bottom left*), and the arrival rate of informed traders μ (*bottom right*)

which shows that there was a day with only one event in an hour ($0.04 \times 24 \approx 1$ day). Similarly, the highest estimate is 0.56, which shows that the most eventful day had 14 h with an event. The Shapiro-Wilk test (Shapiro and Wilk 1965) rejects the null hypothesis of normality at the 5 % significance level, as the p value is 0.035. Thus, for this sample, the market maker views the arrival of news as a non-normal process. The estimate that an event is bad news ($\hat{\delta}$) lies in between 0 and 1, with an average of 0.47 (our initial parameter). Note that $(1 - \delta)$ is the probability that an event is good news. For example, the 18th day of our sample covers the 24-h period from October 28, Wednesday 16:00 to October 29, Thursday 15:59. On this particular day, $\hat{\alpha} = 0.0412$ and $\hat{\delta} = 0.01$. This means that there was only 1 h with news (we do not know which hour) and the news was good ($1 - \hat{\delta} = 0.99$). According to the Shapiro-Wilk test, the estimate of δ is normally distributed, with p value=0.418. This result is expected, as there was no significant trend in EUR–USD prices during the sample period.

Figure 5 plots the daily estimates of the arrival rates. The estimated arrival rate of uninformed traders $\hat{\varepsilon}$ exhibits a sharp increase around the 60th day (in January 2004), from an average of approximately 50–80. The overall mean of this parameter is 78.8 (the same as the initial parameter). The estimate follows a normal distribution as confirmed by the Shapiro-Wilk test. The estimated arrival rate of informed traders $\hat{\mu}$ seems to be stable with occasional jumps. The Shapiro-Wilk test strongly rejects the null hypothesis of normality. The overall average of this parameter is 83.5. Table 2 shows that both informed and uninformed traders arrive less often than average on Mondays and Fridays. The highest arrival rates of both informed and uninformed traders are observed on Wednesdays. It is worth noting that the market maker attaches a non-normal component to the arrival of the informed traders, which goes against the assumption that informed traders are risk neutral. Another instance of empirics diverging from the market assumptions (dominance of uninformed traders) is the fact that informed and uninformed traders have similar arrival rates. This can be interpreted as an equal likelihood for the arrival of informed and uninformed traders, despite the fact that the institutional market characteristics encourage the participation of uninformed traders in particular.

Finally, the average estimated PIN is about 0.12, lower than that observed in equity markets. Over the 145 days in the sample, the PIN ranges between 0.04 and 0.21. The seasonal day-of-week indices for the PIN point to Monday as the only above-average day. The PIN is below average on Tuesdays and Wednesdays, and SI_{PIN} is close to unity on both Thursdays and Fridays. Therefore, although Mondays are viewed as eventful days with a relatively high PIN, this does not result in extraordinarily high arrival rates of informed traders. Rather, their activity appears to be more subtle, with most of their trading potentially taking place on days with lower-than-average news arrival and high arrival rates of uninformed traders. Hence, the PIN reveals that informed traders "conceal" their above-average activity on Tuesdays and Wednesdays as well as their below-average activity on Mondays and Fridays. In all, it appears that the arrival rates of informed traders are inversely related to the PIN values, i.e., high (low) arrival rates of informed traders imply low (high) PIN values. This can be explained by noting that on uneventful days with high arrival rates of uninformed traders, high arrival rate of informed traders does not necessarily imply high PIN, because there is not much to be informed about. Moreover, informed traders strategically match the arrival rates of the uninformed traders, thereby camouflaging their trading activity. For example, if there is a low arrival rate of uninformed traders and there is no significant news, the arrival rate of informed traders will also be low. If, however, there is a low arrival rate of uninformed traders and there is news, the arrival rate of informed traders may be strategically postponed and will also be low.

2.4 Independence of arrivals

The crucial underlying assumption in Easley et al. (1996b) is the hourly independence of information events in each 24-h sequence.[18] Thus, while deriving the log likelihood function, we assume that the arrival of traders in each hour, conditional on information events, is drawn from identical and independent distributions. Nevertheless, consider-

[18] Easley et al. (1997b) test for the independence assumption and find that information events in their dataset are independent.

ing evidence on the relationship between volatility clustering and trading volume (e.g., Gallant et al. 1992), it would be useful to test whether the independence assumption holds.

As a first step, we follow Easley et al. (1997b) and use a runs test for each day in the sample. The estimated $\hat{\alpha_i}$'s ($i = 1, \ldots, 145$) help us to classify hours into one category in which an event occurs or another in which no event occurs. As noted previously, *TT* is the total number of trades. On each day, we order hourly *TT* from the smallest to the largest and classify the upper $100 \times \hat{\alpha_i}$ percent as event hours. We then turn to the original *TT* sequence, classifying each event hour by one and each non-event hour by zero. The total number of event hours is denoted by e_i and non-event hours by n_i.[19] The results indicate that the null hypothesis cannot be rejected at the 5 % significance level for 72 days, although it is rejected for 73 days. This mixed evidence necessitates additional testing. We turn to the Ljung-Box portmanteau test (Ljung and Box 1978) for white noise next.[20]

We compute the Ljung-Box test statistic with up to the 10th-order serial correlation in levels of *S*, *B*, *TT* and *K* for each day. Hence, we compare 145 values for Q_L with the critical value χ^2_{10}. The null rejection frequencies at the 5 % significance level are: for *B* (frequency = 24, or 17 % of the days in the sample), for *S* (frequency = 21, or 14 % of the days in the sample), for *TT* (frequency = 23, or 16 % of the days in the sample), and for *K* (frequency = 18, or 12 % of the days in the sample). We conclude that the independence assumption is not disproven by this evidence, as our model agrees with the assumption on about 85 % of days. Easley et al. (1997b) argue that evidence of dependence does not affect the actual parameter estimates, but does affect their asymptotic standard errors. Since our inference is based on the mean values of the estimates and small standard errors (relative to the parameter values), we conjecture that not accounting for the dependence of information events does not have any major impact on our findings.

3 Price dynamics

3.1 Model estimates, returns and volatility

As a final test of the model's usefulness, we regress exchange rate returns (and squared returns) on the theoretical variables that comprise the model. Our regressions are of the following form

$$r_t = \alpha + \gamma M_t + v_t, \quad v_t \sim \mathrm{IID}(0, \sigma^2) \tag{1}$$

$$r_t^2 = \alpha + \beta r_{t-1}^2 + \gamma M_t + v_t, \quad v_t \sim \mathrm{IID}(0, \sigma^2) \tag{2}$$

[19] Under the null hypothesis of the randomness of information events across hours, the total number of runs r (sequences of ones or zeros) is normally distributed with $\bar{r} = \frac{2e_i n_i}{e_i + n_i} + 1$ and $\sigma_r^2 = \frac{(\bar{r}-1)(\bar{r}-2)}{(e_i+n_i)-1}$.

[20] For the null hypothesis of independence (or randomness) of information events over I=24 h, this test is based on the following statistic: $Q_L = I(I+2)\sum_{\tau=1}^{L} \frac{\hat{\rho}_\tau^2}{I-\tau}$, where L is typically chosen to be substantially smaller than I and $\hat{\rho}_\tau^2$ is the sample autocorrelation coefficient at lag τ.

Table 3 Regression results-returns

r_t	Constant	$M_t = \delta_t$
Coefficient	0.007***	−0.015***
Standard error	0.001	−0.002
t-statistic	6.26	−7.29
p value	0.000	0.000
Adjusted R^2	0.2675	

This table reports the results of the OLS regression: $r_t = \alpha + \gamma M_t + \upsilon_t,\ t = 1, \ldots, 145$. The dependent variable (r_t) is the return for the spot EUR/USD exchange rate at time t. The explanatory variable is the probability of a bad information event (δ_t). *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively

where $r_t = \ln(P_t) - \ln(P_{t-1})$, $M_t \in \{\alpha_t, \delta_t, \varepsilon_t, \mu_t\}$ and $t = 1, \ldots, 145$. By P_t we denote the daily close (EUR/USD) on day t, where the EUR is the base currency. Table 3 reports the results for the M_t variables found to be significant in the first regression. Since the estimate for δ_t is statistically significant with a relatively high adjusted R^2 value (0.2675), these results essentially confirm those from Easley et al. (1997b), which also find the probability of bad news informative for price determination. However, their estimated coefficient on the bad event probability variable is much larger than our $\hat{\gamma}$. More specifically, we find that a 1 exchange rate by about 1.5 cents, while the estimate in Easley et al. (1997b) is 0.61. This suggests that the equity market is more responsive than the FX market to the arrival of bad news.

We next discuss the results of regressing the squared FX returns, i.e., the measure of daily volatility, on the arrival rates (the only variables that are statistically significant in the second regression). We find that the impact of both arrival rates on volatility is positive and statistically significant. According to Table 4, the magnitude of the estimated γ for ε_t is about two times larger than that for μ_t. Since the PIN is insignificant in the second regression, we conclude that although volatility increases with the arrival rates of traders, it is independent of the trade composition.

Finally, to address the issue of the potentially strategic arrival of informed traders, we conduct Granger causality tests between the trader arrival rates. Essentially, the Granger causality test assesses the ability of one series to forecast another. The idea is that if the informed traders move strategically to match the activity of uninformed traders, we may be able to forecast their arrival.

Granger causality between ε_t and μ_t is estimated using a standard bivariate framework. In Table 5, we report the results of regressions estimated with two lags. The results are similar if we change to 1, 3, or more lags. We cannot reject the hypothesis that the arrival rates of informed traders do not Granger-cause the arrival rates of uninformed traders. However, we can reject the hypothesis that the arrival rates of uninformed traders do not cause the arrival rates of informed traders. This indicates that our conjecture of the strategic arrival of informed traders is valid. Combined with the findings from previous sections, this suggests that uninformed traders arrive first, while informed traders follow after strategically timing their arrival. Therefore, the

Table 4 Regression results-volatility

r_t^2	Constant	r_{t-1}^2	$M_t = \mu_t$
Coefficient	0.000***	−0.205**	3.25e−07*
Standard error	0.000	0.081	1.72e−07
t-statistic	2.70	−2.51	1.89
p value	0.008	0.013	0.061
Adjusted R^2	0.0509		
r_t^2	Constant	r_{t-1}^2	$M_t = \varepsilon_t$
Coefficient	0.000	−0.218***	7.03e−07**
Standard error	0.000	0.081	3.04e−07
t-statistic	0.73	−2.67	2.31
p value	0.466	0.008	0.022
Adjusted R^2	0.0626		

This table reports the results of the OLS regression: $r_t^2 = \alpha + \beta r_{t-1}^2 + \gamma M_t + v_t,\ t = 1, \ldots, 145$. The dependent variable (r_t^2) is the squared return for the spot EUR/USD exchange rate at time t. The explanatory variables are the lagged squared returns (r_{t-1}^2), the arrival rate of informed traders (μ_t), and the arrival rate of uninformed traders (ε_t). *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively

Table 5 Granger causality tests using daily observations

Null hypothesis	χ^2	Prob > χ^2
μ_t does not cause ε_t	2.51	0.28
ε_t does not cause μ_t	7.77	0.02

This table lists the probabilities from Granger causality tests on the estimates of the arrival rates for 145 days in the sample. The test statistic is distributed as $\chi^2(df = 2)$, with critical value χ^2_{cr}=5.991 for the 5 % significance level. The null hypothesis is stated in the first column. μ_t and ε_t ($t = 1, \ldots, 145$) denote the arrival rates of the informed and uninformed traders, respectively. The estimations are based on a standard bivariate framework. The figures in the third column are the probabilities of rejection

assumption that informed traders are risk neutral is not appropriate in the FX market context.[21]

3.2 Trading, returns, and volatility

Section 2 presented the hour-of-day indices of informed and uninformed traders based on unbalanced traders ($|K|$) and balanced traders ($TT - |K|$). By definition, K represents "trade imbalances" and can be interpreted as a variant of the market order flow

[21] One may expect the arrival of informed traders to be related only to the information flow. However, in this setting, they appear to use their private information strategically. Another possibility is that informed traders enter the market not only to establish speculative positions (information effects), but also to adjust their currency inventory (inventory effects), as previously mentioned.

Table 6 Trade imbalance regression results

r_t	Constant	K_t
Coefficient	4.12e−05*	−2.37e−05***
Standard error	2.39e−05	9.90e−07
t-statistic	1.72	−23.98
p value	0.085	0.000
Adjusted R^2	0.1417	

This table reports the results of the OLS regression: $r_t = \alpha + \beta K_t + v_t,\ \ t = 1, \ldots, 3479$. The dependent variable (r_t) is the hourly return for the spot EUR/USD exchange rate at time t. The explanatory variable is the trade imbalance (K_t) aggregated over a 1-hour period. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively

(Evans and Lyons 2002). We test whether this variable can "explain" FX returns on an hourly basis by running the following regression:

$$r_t = \alpha + \beta K_t + v_t,\quad v_t \sim \mathrm{IID}(0, \sigma^2),\quad t = 1, \ldots, 3479 \tag{3}$$

where r_t is defined as before and $K_t = S_t - B_t$ for each hour. Table 6 lists the estimates, demonstrating that trade imbalances significantly determine hourly returns. Essentially, a one-unit increase in the trade imbalance (i.e., one additional unique seller relative to the buyers of the EUR over the 1 h period) significantly decreases the EUR/USD exchange rate returns by 2.37e-05, i.e., the EUR depreciates. This confirms evidence from Evans and Lyons (2002) and many other authors who have documented (contemporaneous) microstructure effects in the FX market. We find it worthwhile to stress that price discovery typically does not occur in the retail market and our results should be interpreted with care. At best, the evidence shows that informed traders may operate and strategize in both the retail and interbank FX markets. The same note of caution applies to the remainder of the regressions in this subsection.

Since K_t is quite different from the order flow definition typically used in the FX microstructure literature (see, e.g., Lyons 2001; order flow is the difference between buyer-initiated and seller-initiated transactions), it would be useful to understand the relationship between the two measures more clearly. For that purpose, we construct hourly and daily market order flows based on the total buying and selling volumes. The OANDA FXTrade lists the transaction type and that allows us to aggregate buy/sell market orders. We follow this by regressing the exchange rate returns on order flow, as in Eq. (3). Table 7 shows the results of using both daily and hourly data. As expected, the impact of the order flow on FX returns is significant and positive in both cases: an increase in market order flow is related to the EUR appreciation against the USD. Surprisingly, the trade imbalance seems to be a more appropriate explanatory variable for the hourly data. Also, in line with other studies (Evans and Lyons 2005, Gradojevic 2007), the explanatory power of order flow increases with the aggregation to daily data, where the adjusted $R^2 = 0.09$. It is possible that the information content of order flow is obscured by high-frequency noise and that this effect becomes more pronounced through aggregation. On the other hand, by counting the number of *unique* buyers and

Table 7 Order flow regression results

r_t	Hourly		Daily	
	Constant	X_t	Constant	X_t
Coefficient	7.00e−06	5.67e−11***	0.0001396	1.33e−10***
Standard error	0.0000255	6.52e−12	0.0005237	3.20e−11
t-statistic	0.27	8.70	0.27	4.17
p value	0.784	0.000	0.790	0.000
Adjusted R^2	0.0210		0.0866	

This table reports the results of the OLS regression: $r_t = \alpha + \beta X_t + \upsilon_t$. The dependent variable (r_t) is the hourly or daily return for the spot EUR/USD exchange rate at time t. The explanatory variable is market order flow (X_t) aggregated over 1-h (Hourly) or 24-h (Daily) periods. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively

sellers, the trade imbalance variable reduces the impact of the returning, more frequent traders whose trades may be less informative.

Since $|K|$ and $TT - |K|$ contain information on the hourly arrival of informed and uninformed traders, we next examine their explanatory power with respect to hourly returns, using a regression similar to Eq. (1). Specifically, we run two regressions: one with $M_t = |K_t|$ and one with $M_t = TT_t - |K_t|$ $(t = 1, \ldots, 3479)$. To pin down the hourly impact of informed and uninformed trading on returns, we aggregate the information for each hour over 145 days (144 FX returns). This yields 24 regressions, each with 144 observations. The top two panels of Figure 6 show the absolute value of $\hat{\gamma}_i$ $(i = 0, \ldots, 23)$ for both types of traders. The top right panel models the series of regression coefficients using a cubic spline function.[22] This transformation is convenient because it provides smooth transition in the trader behavior over the 24-h cycle.

The price impact is time-varying, and when significant, is more pronounced among informed traders. Strong geographic dependence emerges again. The first region of activity commences at 03:00, the opening of the European market. This region ends after the opening of the North American market (at 07:00). We also observe a significant price impact before the closing of North American trading and during most of the hours of operation in Asia. These findings are in line with those for the hour-of-day index from Sect. 2. Therefore, in general, the above-average arrival of informed traders has a substantial effect on the exchange rate movements.

Our final objective is to assess the impact of the hourly arrival of informed and uninformed traders on FX volatility. We estimate Eq. (2) with hourly squared returns and, as before, use $M_t = |K_t|$ for informed traders and $M_t = TT_t - |K_t|$ for uninformed traders $(t = 1, \ldots, 3479)$. Here, we follow the same approach as for returns. The bottom two panels of Figure 6 show the estimated $\hat{\gamma}_i$ $(i = 0, \ldots, 23)$ for both types of traders.

[22] Cubic spline is an interpolation method that fits a curve by constructing piecewise third-order polynomials that pass through original data points (Burden and Faires 2004).

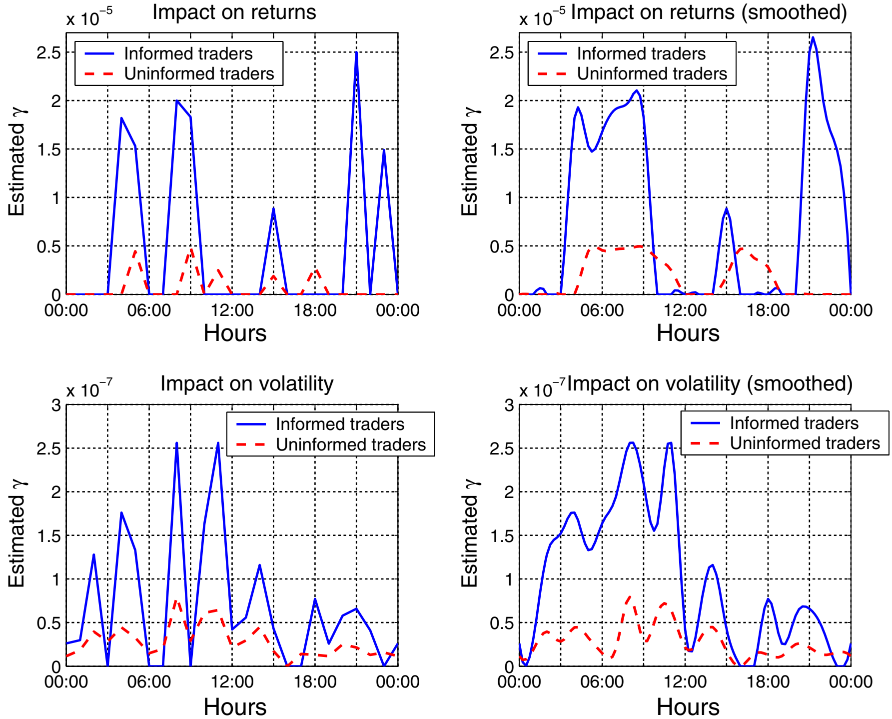

Fig. 6 *Top left* Geographic dependence of the impact of the traders on hourly FX returns. *Top right* Geographic dependence of the impact of the traders on hourly FX returns smoothed using the cubic spline interpolation method. *Bottom left* Geographic dependence of the impact of the traders on hourly FX volatility. *Bottom right* Geographic dependence of the impact of the traders on hourly FX volatility smoothed using the cubic spline interpolation method. The absolute value of $\hat{\gamma}_i$ (the impact of informed traders on FX returns or volatility) for each hour ($i = 0, \ldots, 23$) related to informed traders is given by a solid line. $\hat{\gamma}_i$'s (the impact of uninformed traders on FX returns or volatility) for uninformed traders are given by a dashed line. Sample period: October 5, 2003, 16:00—May 14, 2004, 15:59 (3,480 h, 145 business days)

We observe that informed traders dominate uninformed traders with regard to their influence on hourly FX volatility. The strongest geographic dependence effects take place when both North America and Europe are open, i.e., between roughly 07:00 and 11:00. During these particular hours, informed traders appear capable of driving FX volatility and, to a certain degree, FX returns, despite below-average arrival. Hence, this exercise extracts local information that was not obvious from the theoretical microstructure model.[23]

It is also of interest to explore the different patterns of FX volatility responses to daily and hourly arrival rates of informed/uninformed traders. Recall that the impact of ε on volatility is more than twice as much as the impact of μ. First, we compute the percentile of daily volatility operated in by informed and uninformed traders. This is done by comparing the absolute daily changes in volatility to the regression

[23] The Granger causality tests for the hourly arrivals yield findings similar to those for the daily arrival rates: $TT_t - |K_t|$ Granger-causes $|K_t|$, but not vice-versa.

coefficients (γ) on μ and ε in Eq. (2), listed in Table 4. Based on the estimated volatility percentiles, we find that informed traders operate in roughly the 1st percentile, while uninformed traders operate between the 1st and the 5th percentiles. We follow the same procedure for hourly data, finding that the range of γ from Figure 6 (bottom panel) falls into the 25th percentile for informed traders. Further, the range of γ for uninformed traders is more narrow. We find that they operate in roughly the 10th percentile of hourly volatility changes. As the correlation of the daily arrival rates for uninformed traders with volatility is more dominant at the lower frequency, one can conclude that uninformed traders are not risk seekers. Also, it appears that the impact of informed traders on hourly volatility becomes *averaged out* (more than it does for the uninformed traders) when it is translated into the daily data.

4 Conclusions

This paper utilizes a unique transaction-level dataset from OANDA to search for direct evidence of realized private information in terms of consistent profitability in an electronic spot FX market. To complement this analysis, we develop a high-frequency version of the structural model by Easley et al. (1996b) for the FX market and estimate parameters reflecting the market maker's beliefs about the exact intraday arrival of informed and uninformed traders. Using the model, we also estimate the impact of informed and uninformed currency order flows on the price and volatility.

Our model-free analysis reveals that some traders may possess profitable information in this market. We uncover two lines of evidence that support the informed trading hypothesis: consistent trading profitability and the ability to predict mid-quote movements. These intriguing findings are reinforced by the model estimates which indicate a strong strategic component in the activity of the informed traders that is not observed for the uninformed traders. This phenomenon operates at different levels, from the geographic (intraday) dependency to the day-of-week effects, and is substantiated by Granger causality tests.

What are the traders potentially informed about? An important feature of liquid markets is that price movements are subject to "dynamic effects." For example, if traders build up long positions rapidly and the market price fails to follow through and starts to consolidate, then traders sit on losing long positions, where a minor price reversal can trigger a cascade of margin calls. Traders who can spot this type of phenomenon can generate consistent profits.

We find that the estimates of some of the model's probabilistic parameters can potentially be used to explain the fluctuations in daily FX returns. In addition, the hourly order imbalances have significant power in determining FX returns as predicted by Evans and Lyons (2002). We also show that the arrival rates of informed and uninformed traders have significant power in determining hourly and daily FX rate volatilities.

An important advantage of our approach is that the nature of the OANDA FXtrade currency orders is directly observable, which attests to the accuracy of our results. This differs from equity markets, where trade classification data is often unavailable, and this meant that trade classification algorithms were not required to differentiate between buyer- and seller-initiated trades (see, e.g., Lee and Ready 1991). For example, Boehmer et al. (2007) find that trade misclassification results in a downward bias in the estimate of the PIN for the Easley et al. (1996b) framework.

The model assumes independence of information events across hours. However, we observe dependence for about 15 % of the days in our sample. This may bias the standard errors of our estimates. Though we conjecture this bias to be minimal, it is worth emphasizing that introducing dependency to the model is a key direction of future research. The impact of past transactions on current transactions could be modeled by a latent parameter measuring the degree of serial correlation. In this context, testing for inter-day dependency may offer broader insight into FX-trading patterns and strategies. Another future research avenue we envisage concerns the over-dispersion frequently found in transactions data, which reduces the usefulness of the Poisson distribution.[24] We hope that the evidence presented here will motivate research on new structural microstructure models applicable to both equity and FX markets.

Acknowledgments We are grateful to Olsen Ltd. (www.olsen.ch), Switzerland for providing the data.

Appendix 1: independent arrival model (Easley et al. 1996b)

The model consists of informed and uninformed traders and a risk-neutral competitive market maker. The traded asset is a foreign currency for the domestic currency. Similar to the portfolio shifts model (Evans and Lyons 2002), the trades and the governing price process are generated by the quotes of the market maker over a 24-h trading day. Within any trading hour, the market maker is expected to buy and sell currencies from his posted bid and ask prices. The price process is the expected value of the currency based on the market maker's information set at the time of the trade.

The hourly arrival of news occurs with the probability α. This represents bad news with probability δ and good news with $1-\delta$ probability. Let $\{p_i\}$ be the hourly price process over $i = 1, 2, \ldots, 24$ h. p_i is assumed to be correlated across hours and will reveal the intraday time dependence and intraday persistence of the price behavior across these two classes of traders. The lower and upper bounds for the price process should satisfy $p_i^b < p_i^n < p_i^g$ where p_i^b, p_i^n, and p_i^g are the prices conditional on bad, no news, and good news, respectively. Within each hour, time is continuous and indexed by $t \in [0, T]$.

In any trading hour, the arrivals of informed and uninformed traders are determined by independent Poisson processes. At each instant within an hour, uninformed buyers and sellers each arrive at a rate of ε. Informed traders only trade when there is news, and arrive at a rate of μ. All informed traders are assumed to be risk neutral and competitive and are therefore expected to maximize profits by buying when there is

[24] This point is also made in Wuensche (2007), who proposes mixed Poisson distributions to capture the characteristics of the trade data. Other potential weaknesses of the Easley et al. (1996b) model can be found in Venter and De Jongh (2004).

good news and selling otherwise.[25] For good news hours, the arrival rates are $\varepsilon + \mu$ for buy orders and ε for sell orders. For bad news hours, the arrival rates are ε for buy orders and $\varepsilon + \mu$ for sell orders. When no news exists, the buy and sell orders arrive at a rate of ε per hour.

The market maker is assumed to be a Bayesian who uses the arrival of trades and their intensity to determine whether a particular trading hour belongs to a no news, good news, or bad news category. Since the arrival of hourly news is assumed to be independent, the market maker's hourly decisions are analyzed independently from 1 h to the next. Let $P(t) = (P_n(t), P_b(t), P_g(t))$ be the market maker's prior beliefs with no news, bad news, and good news at time t. Accordingly, his or her prior beliefs before trading starts each day are $P(0) = (1 - \alpha, \alpha\delta, \alpha(1 - \delta))$.

Let S_t and B_t denote sell and buy orders at time t. The market maker updates the prior conditional on the arrival of an order of the relevant type. Let $P(t|S_t)$ be the market maker's updated belief conditional on a sell order arriving at t. $P_n(t|S_t)$ is the market maker's belief about no news conditional on a sell order arriving at t. Similarly, $P_b(t|S_t)$ is the market maker's belief about the occurrence of bad news events conditional on a sell order arriving at t, and $P_g(t|S_t)$ is the market maker's belief about the occurrence of good news conditional on a sell order arriving at t.

The probability that any trade occurring at time t is information based is (please see Appendix 2)

$$i(t) = \frac{\mu(1 - P_n(t))}{2\varepsilon + \mu(1 - P_n(t))} \tag{4}$$

Since each buy and sell order follows a Poisson Process at each trading hour and orders are independent, the likelihood of observing a sequence of orders containing B buys and S sells in a bad news hour of total time T is given by

$$L_b((B, S)|\theta) = L_b(B|\theta)L_b(S|\theta) = e^{-(\mu+2\varepsilon)T} \frac{\varepsilon^B(\mu + \varepsilon)^S T^{B+S}}{B!S!}, \tag{5}$$

where $\theta = (\alpha, \delta, \varepsilon, \mu)$.

Similarly, in a no-event hour, the likelihood of observing any sequence of orders that contains B buys and S sells is

$$L_n((B, S)|\theta) = L_n(B|\theta)L_n(S|\theta) = e^{-2\varepsilon T} \frac{\varepsilon^{B+S} T^{B+S}}{B!S!} \tag{6}$$

In a good-event hour, this likelihood is

$$L_g((B, S)|\theta) = L_g(B|\theta)L_g(S|\theta) = e^{-(\mu+2\varepsilon)T} \frac{\varepsilon^S(\mu + \varepsilon)^B T^{B+S}}{B!S!} \tag{7}$$

[25] This assumption may seem inappropriate, given that it rules out any strategic behavior. As shown in Sect. 2.3, informed traders have some tendency to trade strategically. Therefore, we concur that the assumption of risk neutrality needs defending, but we retain it for the sake of the model applicability.

The likelihood of observing B buys and S sells in an hour of unknown type is the weighted average of Eqs. (2), (3), and (4) using the probabilities of each type of hour occurring.

$$\begin{aligned} L((B,S)|\theta) &= (1-\alpha)L_n((B,S)|\theta) + \alpha\delta L_b((B,S)|\theta) + \alpha(1-\delta)L_g((B,S)|\theta) \\ &= (1-\alpha)e^{-2\varepsilon T}\frac{\varepsilon^{B+S}T^{B+S}}{B!S!} + \alpha\delta e^{-(\mu+2\varepsilon)T}\frac{\varepsilon^{B}(\mu+\varepsilon)^{S}T^{B+S}}{B!S!} \\ &\quad + \alpha(1-\delta)e^{-(\mu+2\varepsilon)T}\frac{\varepsilon^{S}(\mu+\varepsilon)^{B}T^{B+S}}{B!S!} \end{aligned} \tag{8}$$

Because hours are independent, the likelihood of observing the data $M = (B_i, S_i)_{i=1}^{I}$ over 24 h ($I = 24$) is the product of the hourly likelihoods,

$$\begin{aligned} L(M|\theta) = \prod_{i=1}^{I} L(\theta|B_i,S_i) = \prod_{i=1}^{I} \frac{e^{-2\varepsilon T}T^{B_i+S_i}}{B_i!S_i!} \times \\ \left[(1-\alpha)\varepsilon^{B_i+S_i} + \alpha\delta e^{-\mu T}\varepsilon^{B_i}(\mu+\varepsilon)^{S_i} + \alpha(1-\delta)e^{-\mu T}\varepsilon^{S_i}(\mu+\varepsilon)^{B_i}\right] \end{aligned} \tag{9}$$

The log likelihood function is

$$\begin{aligned} \ell(M|\theta) &= \sum_{i=1}^{I} \ell(\theta|B_i,S_i) \\ &= \sum_{i=1}^{I}[-2\varepsilon T + (B_i+S_i)\ln T] \\ &\quad + \sum_{i=1}^{I}\ln\left[(1-\alpha)\varepsilon^{B_i+S_i} + \alpha\delta e^{-\mu T}\varepsilon^{B_i}(\mu+\varepsilon)^{S_i} + \alpha(1-\delta)e^{-\mu T}\varepsilon^{S_i}(\mu+\varepsilon)^{B_i}\right] \\ &\quad - \sum_{i=1}^{I}(\ln B_i! + \ln S_i!) \end{aligned} \tag{10}$$

As in Easley et al. (2008), the log likelihood function, after dropping the constant and rearranging,[26] is given by

[26] To derive Eq. (11), the term $\ln[x^{M_i}(\mu+\varepsilon)^{B_i+S_i}]$ is simultaneously added to the first sum and subtracted from the second sum in Eq. (10). This is done to increase computational efficiency and to ensure convergence in the presence of a large numbers of buys and sells, as is the case in our dataset.

$$\ell(M|\theta) = \sum_{i=1}^{I} [-2\varepsilon + M_i \ln x + (B_i + S_i)\ln(\mu + \varepsilon)] + \sum_{i=1}^{I} \ln\left[\alpha(1-\delta)e^{-\mu}x^{S_i - M_i} + \alpha\delta e^{-\mu}x^{B_i - M_i} + (1+\alpha)x^{B_i + S_i - M_i}\right], \tag{11}$$

where $M_i \equiv \min(B_i, S_i) + \max(B_i, S_i)/2$, and $x = \frac{\varepsilon}{\varepsilon+\mu} \in [0, 1]$.

Appendix 2: derivation of the PIN

By Bayes' rule, the market maker's posterior probability with no news at time t, if an order to sell arrives at t, is

$$\begin{aligned} P_n(t|S_t) &= \frac{P_n(S_t|t)P_n(t)}{P(S_t)} \\ &= \frac{P_n(S_t|t)P_n(t)}{P_n(S_t|t)P_n(t) + P_g(S_t|t)P_g(t) + P_b(S_t|t)P_b(t)} \\ &= \frac{\varepsilon P_n(t)}{\varepsilon(1 - P_g(t) - P_b(t)) + \varepsilon P_g(t) + (\varepsilon + \mu)P_b(t)} \\ &= \frac{\varepsilon P_n(t)}{\varepsilon + \mu P_b(t)}, \end{aligned}$$

where $P_n(S_t|t)$ is the probability of the arrival of a sell order conditional on no news at time t, $P_g(S_t|t)$ is the probability of the arrival of a sell order conditional on good news at time t, and $P_b(S_t|t)$ is the probability of the arrival of a sell order conditional on bad news at time t.

Similarly, the posterior probability on bad news is

$$\begin{aligned} P_b(t|S_t) &= \frac{P_b(S_t|t)P_b(t)}{P(S_t)} \\ &= \frac{(\varepsilon + \mu)P_b(t)}{\varepsilon + \mu P_b(t)}, \end{aligned}$$

and the posterior probability on good news is

$$\begin{aligned} P_g(t|S_t) &= \frac{P_g(S_t|t)P_g(t)}{P(S_t)} \\ &= \frac{\varepsilon P_g(t)}{\varepsilon + \mu P_b(t)} \end{aligned}$$

The bid price, $b(t)$, conditional on S_t at time t at hour i is

$$\begin{aligned} b(t) &= P_n(t|S_t)p_i^n + P_b(t|S_t)p_i^b + P_g(t|S_t)p_i^g \\ &= \frac{\varepsilon P_n(t)p_i^n + (\varepsilon+\mu)P_b(t)p_i^b + \varepsilon P_g(t)p_i^g}{\varepsilon + \mu P_b(t)} \end{aligned}$$

Similarly, the ask price $a(t)$ is the market maker's expected value of the asset conditional on the history prior to t and on B_t.

Thus, the ask at time t at hour i is

$$\begin{aligned} a(t) &= P_n(t|B_t)p_i^n + P_b(t|B_t)p_i^b + P_g(t|B_t)p_i^g \\ &= \frac{\varepsilon P_n(t)p_i^n + \varepsilon P_b(t)p_i^b + (\varepsilon+\mu)P_g(t)p_i^g}{\varepsilon + \mu P_g(t)} \end{aligned}$$

The expected price conditional on t is

$$E\left[p_i|t\right] = P_n(t)p_i^n + P_b(t)p_i^b + P_g(t)p_i^g,$$

where $P_n(t)$, $P_b(t)$ and $P_g(t)$ are the prior beliefs of the market maker for no news, bad news, and good news at time t.

Substituting the expected price equation into the equations for bid and ask prices yields

$$\begin{aligned} b(t) &= \frac{\varepsilon P_n(t)p_i^n + \varepsilon P_b(t)p_i^b + \varepsilon P_g(t)p_i^g + \mu P_b(t)p_i^b}{\varepsilon + \mu P_b(t)} \\ &= \frac{\varepsilon E\left[p_i|t\right] + \mu P_b(t)p_i^b}{\varepsilon + \mu P_b(t)} \\ &= \left[1 - \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)}\right] E\left[p_i|t\right] + \frac{\mu P_b(t)p_i^b}{\varepsilon + \mu P_b(t)} \\ &= E\left[p_i|t\right] - \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)}\left(E\left[p_i|t\right] - p_i^b\right), \end{aligned}$$

and

$$\begin{aligned} a(t) &= \frac{\varepsilon P_n(t)p_i^n + \varepsilon P_b(t)p_i^b + \varepsilon P_g(t)p_i^g + \mu P_g(t)p_i^g}{\varepsilon + \mu P_g(t)} \\ &= \frac{\varepsilon E\left[p_i|t\right] + \mu P_g(t)p_i^g}{\varepsilon + \mu P_g(t)} \\ &= \left[1 - \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)}\right] E\left[p_i|t\right] + \frac{\mu P_g(t)p_i^g}{\varepsilon + \mu P_g(t)} \\ &= E\left[p_i|t\right] + \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)}\left(p_i^g - E\left[p_i|t\right]\right) \end{aligned}$$

Let $d(t) = a(t) - b(t)$ be the spread at time t.

$$\begin{aligned} d(t) &= E\left[p_i|t\right] + \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)}\left(p_i^g - E\left[p_i|t\right]\right) \\ &\quad - \left[E\left[p_i|t\right] - \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)}\left(E\left[p_i|t\right] - p_i^b\right)\right] \\ &= \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)}\left(p_i^g - E\left[p_i|t\right]\right) + \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)}\left(E\left[p_i|t\right] - p_i^b\right) \end{aligned}$$

The spread for the opening quotes is

$$\begin{aligned} d(0) &= \frac{\mu P_g(0)}{\varepsilon + \mu P_g(0)}\left(p_i^g - E\left[V_i\right]\right) + \frac{\mu P_b(0)}{\varepsilon + \mu P_b(0)}\left(E\left[p_i\right] - p_i^b\right) \\ &= \frac{\mu\alpha(1-\delta)}{\varepsilon + \mu\alpha(1-\delta)}\left(p_i^g - E\left[p_i|t\right]\right) + \frac{\mu\alpha\delta}{\varepsilon + \mu\alpha\delta}\left(E\left[p_i|t\right] - p_i^b\right) \end{aligned}$$

If good and bad events are equally likely, that is, if $\delta = 1 - \delta$, $\delta = 0.5$. Thus

$$d(0) = \frac{\mu\alpha}{2\varepsilon + \mu\alpha}(p_i^g - p_i^b)$$

The probability that any trade occurring at time t is information based is

$$\begin{aligned} i(t) &= \frac{P_b(t)\mu + P_g(t)\mu}{P(B_t, S_t)} \\ &= \frac{\mu(1 - P_n(t))}{P_n(t)P_n(B_t, S_t|t) + P_g(t)P_g(B_t, S_t|t) + P_b(t)P_b(B_t, S_t|t)} \\ &= \frac{\mu(1 - P_n(t))}{2\varepsilon + \mu(1 - P_n(t))} \end{aligned}$$

References

Albuquerque R, De Francisco E, Marques L (2008) Marketwide private information in stocks: forecasting currency returns. J Finance 63(5):2297–2343

Bjønnes G, Rime D (2005) Dealer behavior and trading systems in foreign exchange markets. J Finan Econ 75(3):571–605

Bjønnes G, Osler C, Rime D (2008) Asymmetric information in the interbank foreign exchange market, norges Bank Working Paper

Boehmer E, Grammig J, Theissen E (2007) Estimating the probability of informed trading—does trade misclassification matter? J Finan Mark 10(1):26–47

Breedon F, Ranaldo A (2013) Intraday patterns in FX returns and order flow. J Money Credit Bank 45(5):953–965

Burden R, Faires JD (2004) Numerical Analysis. Brooks Cole, Florence

Covrig V, Melvin M (2002) Asymmetric information and price discovery in the FX market: does Tokyo know more about the yen? J Empir Finance 9:271–285

Dacorogna M, Gençay R, Muller U, Olsen R, Pictet O (2001) An introduction to high-frequency finance. Academic Press, San Diego
D'Souza C (2008) Price discovery across geographic locations in the foreign exchange market. Bank of Canada Review pp 17–25
Easley D, O'Hara M (1992) Time and the process of security prices adjustment. J Finance 47:577–605
Easley D, Kiefer N, O'Hara M (1996a) Cream-skimming or profit-sharing? The curious role of purchased order flow. J Finance 51:811–833
Easley D, Kiefer N, O'Hara M, Paperman J (1996b) Liquidity, information and infrequently traded stocks. J Finance 51:1405–1436
Easley D, Kiefer N, O'Hara M (1997a) The information content of the trading process. J Empir Finance 4:159–185
Easley D, Kiefer N, O'Hara M (1997b) One day in the life of a very common stock. Rev Finan Stud 10:805–835
Easley D, Engle R, O'Hara M, Wu L (2008) Time-varying arrival rates of informed and uninformed trades. J Finan Econ 6(2):171–207
Evans M, Lyons R (2002) Order flow and exchange rate dynamics. J Political Econ 110:170–180
Evans M, Lyons R (2005) Meese-rogoff redux: micro-based exchange rate forecasting. Am Econ Rev 95(2):405–414
Evans M, Lyons R (2012) Exchange rate fundamentals and order flow. Q J Finance 2(4):1250,018
Foster F, Viswanathan S (1994) Strategic trading with asymmetrically informed traders and long-lived information. J Finan Quant Anal 29(4):499–518
Gallant A, Rossi P, Tauchen G (1992) Stock prices and volume. Rev Finan Stud 5(2):199–242
Gençay R, Gradojevic N (2013) Private information and its origins in an electronic foreign exchange market. Econ Model 33:86–93
Goldstein M, Van Ness BF, Van Ness RA (2006) The intraday probability of informed trading on the NYSE. Adv Quant Anal Financ Acc 3:139–158
Gradojevic N (2007) The microstructure of the Canada/U.S. dollar exchange rate: a robustness test. Econ Lett 94(3):426–432
Hau H (2001) Location matters: an examination of trading profits. J Finance 56(5):1959–1983
Heimer R, Simon D (2012) Facebook finance: how social interaction propagates active investing, brandeis University Working Paper
Kaul A, Sapp S (2009) Trading activity, dealer concentration and foreign exchange market quality. J Bank Finance 33(11):2122–2131
King M, Rime D (2010) The $4 trillion question: What explains FX growth since the 2007 survey? BIS Quarterly Review pp 27–42
Lee C, Ready M (1991) Inferring trade direction from intraday data. J Finance 46:733–746
Lei Q, Wu G (2005) Time-varying informed and uninformed trading activities. J Finan Mark 8(2):153–181
Ljung G, Box G (1978) On a measure of lack of fit in time series models. Biometrika 65:297–303
Lyons R (1995) Test of microstructural hypotheses in the foreign exchange market. J Finan Econ 39:321–351
Lyons R (2001) The microstructure approach to exchange rates. The MIT Press, Cambridge
Madhavan A, Smidt S (1993) An analysis of changes in specialist inventories and quotations. J Finance 48:1595–1628
Marsh IW, O'Rourke C (2005) Customer order flow and exchange rate movements: Is there really information content?, working Paper
Menkhoff L, Schmeling M (2008) Local information in foreign exchange markets. J Int Money Finance 27(8):1383–1406
Menkhoff L, Schmeling M (2010) Whose trades convey information? Evidence from a cross-section of traders. J Finan Mark 13(1):101–128
Moore M, Payne R (2011) On the sources of private information in FX markets. J Bank Finance 35(5):1250–1262
Nolte I, Nolte S (2012) How do individual investors trade? Eur J Finance 18(10):921–947
Odders-White ER, Ready MJ (2008) The probability and magnitude of information events. J Finan Econ 87(1):227–248
Osler C, Vandrovych V (2009) Hedge funds and the origins of private information in currency markets, working Paper No. 1484711
Payne R (2003) Informed trade in spot foreign exchange markets: an empirical investigation. J Int Econ 61:307–329

Peiers B (1997) Informed traders, intervention, and price leadership: a deeper view of the microstructure of the foreign exchange market. J Finance 52(4):1589–1614

Ranaldo A (2009) Segmentation and time-of-day patterns in foreign exchange markets. J Bank Finance 33(12):2199–2206

Shapiro S, Wilk M (1965) An analysis of variance test for normality (complete samples). Biometrika 52(3–4):591–611

Venter J, De Jongh D (2004) Extending the ekop model to estimate the probability of informed trading, working Paper

Vitale P (2012) Optimal informed trading in the foreign exchange market. Eur J Finance 18(10):989–1013

Wuensche O (2007) Using mixed Poisson distributions in sequential trade models, working Paper

Yao J (1998) Market making in the interbank foreign exchange market, working Paper No. S-98-3