Informativeness of trade size in foreign exchange markets

Nikola Gradojevic[a,b,*], Deniz Erdemlioglu[b], Ramazan Gençay[c]

[a] *Department of Economics and Finance, University of Guelph, Canada.*
[b] *IÉSEG School of Management (LEM-CNRS), Lille, France.*
[c] *Department of Economics, Simon Fraser University, Burnaby, Canada.*

This Version: April 25, 2016

Abstract

This article investigates a trading strategy that relies on private information in an electronic spot foreign exchange market. In a structural microstructure model extended for high-frequency data, our analysis links the informational content of trading activity to order size. We find that large currency orders are likely to be placed by informed traders, which generally results in increased price volatility. We further examine the reverse causality and show that trade size responds to price volatility in a nonlinear manner: the probability of observing large trades increases with volatility, but this probability decreases when volatility becomes relatively high. In addition, the data suggest that excess kurtosis in exchange rate returns (corresponding to large price-contingent trades) is significantly lower than that in small trades. We discuss the implications of such findings for studying extreme events and jumps in asset prices.

Keywords: Foreign exchange markets, Volume, Trade size, Volatility, Informed trading, Noise trading, Market microstructure

JEL: F3, F31, G0, G1

*University of Guelph, College of Business and Economics, Department of Economics and Finance, 50 Stone Road East, Guelph, Ontario, N1G 2W1, Canada. Phone: (+1) 519-824-4120/ ext. 52357.

Email addresses: ngradoje@uoguelph.ca (Nikola Gradojevic), d.erdemlioglu@ieseg.fr (Deniz Erdemlioglu), rgencay@sfu.ca (Ramazan Gençay)

1. Introduction

Do heterogeneously informed currency traders differ in their use of information? If so, how does private information impact their trade size? What is the relationship of trade size to foreign exchange (FX) rate volatility? This paper seeks answers to these questions by linking the information process to order size patterns while relying on a trading strategy designed in an electronic FX market. Extending Easley et al. (1997b) and Easley and O'Hara (1987), our high-frequency trading setup allows informed and uninformed traders to place orders sequentially in continuous time.[1] To test the predictions of the strategies, we derive tractable likelihood functions that identify the variation in trade size associated with the orders of informed and uninformed FX traders.

Based on a retail electronic trading platform data set, our empirical analysis reveals several notable findings. First, we empirically show that large orders are likely to be executed by informed traders rather than uninformed traders. This evidence is particularly pronounced for buy orders and remains strong regardless of the choice of size thresholds. These results highlight the importance of the information content of trade size (i.e., *informativeness*) in characterizing currency transaction data. More broadly, a direct implication of this analysis is that order flow size could be informative by itself even in the absence of information shocks.[2] Second, an estimated logit model suggests that large trade size appears to be an *endogenous* factor that depends (nonlinearly) on price volatility. This finding supports the *intraday trading invariance principle* proposed by Andersen et al. (2015). Third, when we reverse the causality direction in the logit model, we also reveal that large trades result in a statistically significant increase in price volatility. This effect, however, is not as pronounced as the impacts of volatility on large trades, which is sensible for an FX retail trading platform. Finally, we assess the distributional characteristics of price increments

[1] While informed traders utilize information surprises as principal motivation for their trading, uninformed traders consider non-news factors, such as liquidity or trade-driven shocks. See also Osler and Savaser (2011) and Osler (2005), who empirically show that extreme FX price movements could result from stop-loss orders even in the absence of any macroeconomic news announcements.

[2] Prior research on currency markets has focused on the link between expectations and shocks hitting currency markets (Evans, 2002; Evans and Lyons, 2002a). While information-driven shocks change expectations and often increase market volatility (see, e.g., Jiang et al., 2011; Ederington and Lee, 1995, 1993), price-contingent trading could also trigger large FX market swings even in the absence of any news arrivals (Osler, 2005; Osler and Savaser, 2011). Non-information shocks may include, for instance, liquidity (or trading-based) shocks (Caballero and Krishnamurthy, 2008), real shocks (i.e., innovations of preferences as in Allen and Gale, 2000) and structural shocks (see, e.g., Dungey et al., 2010).

and show that excess kurtosis in exchange rate data, corresponding to large price-contingent trades, is significantly lower than that in small trades. Our motivation for this assessment directly builds on the argument of Osler and Savaser (2011) and Osler (2005), who provide evidence that price-contingent trading could solely explain the extreme price cascades in the transactions of an FX dealer. We emphasize that the source of extreme events could be attributed to the informativeness of trade size: uninformed traders tend to place small orders as a range of extreme stop-loss and take-profit trades. This may result in jump cascades or excess kurtosis observed in transaction data. In addition to quotation bursts in equities (Gençay et al., 2016), the size and informational content of trades could thus be additional drivers of currency jumps.

There is an extensive body of microstructure literature on the role of high-frequency trading and information in price discovery. Recent examples of empirical studies include the works of Brogaard et al. (2014, 2015a) and Brogaard et al. (2015b). Using NASDAQ data, Brogaard et al. (2014) show that the direction of trading of high-frequency traders (HFTs) is linked to *public information*, such as macro news announcements, and that high-frequency trading activities facilitate the price discovery process. In the same strand, Brogaard et al. (2015a) find evidence that HFTs are liquidity suppliers, whereas non-HFTs are liquidity demanders. Building on these two works and using Canadian equity data instead, Brogaard et al. (2015b) reveal that price discovery is primarily driven by limit orders of HFTs (60-80%) and is linked to public information (as in Brogaard et al., 2015a). Considering a theoretical setup, Duffie and Zhu (2015) develop an equilibrium market microstructure model and show that price discovery, when combined with *size discovery*, improves (allocative) efficiency in markets with *private information*. Alternatively, Gençay et al. (2016) examine the dynamics of the quotation process and relate the stress cycles of trading activity to stock price swings. For Gençay et al. (2016), *burst* episodes in data signal the presence of asymmetric information, while average order size tends to decline surrounding bursts in quotes. In contrast to these studies investigating equity market microstructure, we analyze the properties of an electronic FX trading platform. Considering the role of private information in markets, we further characterize a trading strategy that permits us to test for the presence of private information in this context.

Our paper is also related to studies that link heterogeneous information to trading activity and market structure. As a seminal paper, Kyle (1985) derives a continuous auction model and shows that prices react gradually to the private information of informed traders. Glosten and Milgrom (1985) characterize adverse selection between end users and dealers. The key prediction of Glosten and Milgrom (1985) is that (equilibrium) bid-ask spreads reflect the information in trades and that such spreads should widen when traders (i.e., end users) are informed. Interactions between informed and liquidity stock traders have been studied from a theoretical perspective by Admati and Pfleiderer (1988). In line with the conclusions of Kyle (1985), Admati and Pfleiderer (1988) document that (monopolistic) informed traders can *camouflage* their trading activity by splitting large trades into small trades.[3] Moreover, as Glosten (1994) suggests, the frequency of trading could also vary with the information set of investors: in equilibrium, large traders with extreme news trade twice, whereas small traders with less extreme news trade only once.[4]

The above studies develop a theoretical foundation and concern the equity market, whereas the current paper is an empirical analysis of a retail FX market. Notably, the regularities and activity in an equity market microstructure may substantially deviate from those in an FX market microstructure. For instance, the adverse selection paradigm in the equity literature posits that spreads are typically wider when informed traders (end users) undertake large trades (see, e.g., Glosten and Milgrom, 1985, Glosten, 1989). Contrary to this dominant view in equity market microstructure, spreads are not primarily driven by adverse selection in currencies (Osler et al., 2011). In FX data, customer spreads are negatively related to trade size, and they are typically tight for financial customers. In light of these results from the currency market microstructure literature, we examine the role of heterogeneous information in explaining trade size variation. We conduct an empirical exercise that relies on a trading model for heterogeneously informed investors who can submit large or small orders depending on their access to news arrivals.

A parallel yet separate strand of research investigates the interactions among trade size, volatility and market impact (see, e.g., Madhavan and Smidt, 1991, Kalok and Fong, 2000 and Andersen

[3]The price impact of trade could be either linear (see, e.g., Kyle, 1985) or nonlinear (see, e.g., Dufour and Engle, 2000, Zhang et al., 2001 and Hautsch and Huang, 2012). In this context, Banerjee and Green (2015) further show that if investors do not know whether others are informed or noise traders, then the reaction of asset prices to news becomes nonlinear.

[4]See, e.g., Baruch and Glosten (2013), who show that equilibria exist in a market with noise and informed traders.

et al., 2015). Our analysis is linked to this extant literature in several ways. For instance, Madhavan and Smidt (1991) theoretically show that trade size is crucial because it carries information by influencing the spread. Examining NYSE/NASDAQ-listed stocks, Kalok and Fong (2000) further analyze how trade size (as an *exogenous* factor) impacts price volatility. The main empirical findings in Kalok and Fong (2000) suggest that volatility increases (decreases) with small to medium (medium to large) changes in size, and hence, the relationship could be nonlinear rather than linear. In contrast to Kalok and Fong (2000), we consider trade size as an *endogenous* variable. The underlying intuition behind our consideration is relatively straightforward: if large orders belong to traders with private information, then it is natural to expect that informed traders monitor market conditions and adjust orders in response to time-varying volatility or uncertainty. Kalok and Fong (2000) consider absolute *daily* stock returns as a volatility proxy, while we use absolute *intraday* currency returns (i.e., instantaneous price volatility). In this vein, Andersen et al. (2015) empirically show that trade size in the S&P 500 futures market declines as price volatility (per transaction) rises, consistent with the *intraday trading invariance principle*. We test this hypothesis in our study with FX data and find supporting evidence.

In the context of currency trading, our paper particularly extends and complements the FX market microstructure literature. In contrast to Easley et al. (1997a), who develop a stock market model and show that large buy-sell orders carry almost equal information content, our analysis reveals that informed buy orders in a retail FX market convey more information than uninformed buy orders. Our findings also complement Menkhoff and Schmeling (2010), as we examine a trading strategy that explains the trading behavior and market regularities observed in that paper. Chakravarty (2001) and Anand and Chakravarty (2007) link price discovery to trade size in stock and option markets, respectively. Our results support the conclusions of these studies: medium-sized trades have the most significant information content in currencies, similar to stocks and options.[5] As King et al. (2013) argue, private information has significant power to amplify short-

[5]See also Goodhart and Payne (1996), who study the interactions among quote revision, spreads and transactions: trades drive spreads and quote revisions. Furthermore, that information could be contagious, and the order flow of major FX rates could determine the returns on other exchange rates (Evans and Lyons, 2002b). Exploring the role of order flow, Menkhoff and Schmeling (2008) show that the price impact of currency order flow depends on regions, which in turn supports the local information hypothesis.

term currency fluctuations. Consistent with this view, we empirically show that private information leads to the execution of large orders that potentially create price impacts.

The remainder of the paper is organized as follows. Section 2 introduces the model and outlines the trading environment. Section 3 describes our data. In Section 4, we present and discuss the empirical results. Section 5 concludes the paper.

2. The Model

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, 2002a), the trades and the governing price process are generated by the quotes of the market maker over a 24-hour trading day. Within any trading hour, the market maker is expected to buy and sell currencies from his posted bid and ask prices.[6] The price process is the expected value of the currency based on the market maker's information set at the time of the trade.

2.1. Arrivals of news, traders and orders

The hourly arrival of news occurs with the probability α. This represents bad news with probability δ and good news with $1-\delta$ probability. We define the price process as follows.

Definition 1. *Let $\{p_i\}$ be the hourly price process over $i = 1, 2, \ldots, 24$ hours. 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 news, 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 trade only when there is news, arriving at a rate of μ.

[6]For ease of notation and exposition, we present our model based on hourly time scales. The predictions of the trading strategy remain the same at higher frequencies, such as 5 minutes or 10 minutes.

We also assume that all informed traders are risk neutral and competitive, and we therefore expect them to maximize profits by buying when there is good news and selling otherwise. 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.

2.2. The market maker and measuring the likelihood of orders

The market maker is assumed to be Bayesian, using the arrival of trades and their intensity to determine whether a particular trading hour belongs in the category of no news, good news or bad news. Because the arrival of hourly news is assumed to be independent, the market maker's hourly decisions are analyzed independently from one hour to the next.

Definition 2. *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, the prior beliefs before trading starts each day are $P(0) = (1 - \alpha, \alpha\delta, \alpha(1 - \delta))$.*

Given the definition above, let S_t and B_t further 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 occurs at time t (based on information) is then

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

Because 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

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

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

$$L_n((B,S)|\theta) = L_n(B|\theta)L_n(S|\theta) = e^{-2\varepsilon T}\frac{\varepsilon^{B+S}T^{B+S}}{B!S!}, \tag{3}$$

and in a good-event hour, this likelihood becomes

$$L_g((B,S)|\theta) = L_g(B|\theta)L_g(S|\theta) = e^{-(\mu+2\varepsilon)T}\frac{\varepsilon^{S}(\mu+\varepsilon)^{B}T^{B+S}}{B!S!}. \tag{4}$$

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

$$\begin{aligned} L((B,S)|\theta) &= (1-\alpha)L_n((B,S)|\theta) + \alpha\delta L_b((B,S)|\theta) + \alpha(1-\delta)L_g((B,S)|\theta) \\ &= (1-\alpha)e^{-2\varepsilon T}\frac{\varepsilon^{B+S}T^{B+S}}{B!S!} + \alpha\delta e^{-(\mu+2\varepsilon)T}\frac{\varepsilon^{B}(\mu+\varepsilon)^{S}T^{B+S}}{B!S!} \\ &\quad +\alpha(1-\delta)e^{-(\mu+2\varepsilon)T}\frac{\varepsilon^{S}(\mu+\varepsilon)^{B}T^{B+S}}{B!S!}. \end{aligned} \tag{5}$$

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

$$\begin{aligned} L(M|\theta) &= \prod_{i=1}^{I} L(\theta|B_i,S_i) = \prod_{i=1}^{I}\frac{e^{-2\varepsilon T}T^{B_i+S_i}}{B_i!S_i!} \times \\ &\quad \left[(1-\alpha)\varepsilon^{B_i+S_i} + \alpha\delta e^{-\mu T}\varepsilon^{B_i}(\mu+\varepsilon)^{S_i} + \alpha(1-\delta)e^{-\mu T}\varepsilon^{S_i}(\mu+\varepsilon)^{B_i}\right], \end{aligned} \tag{6}$$

and the log likelihood function is

$$
\begin{aligned}
\ell(M|\theta) &= \sum_{i=1}^{I} \ell(\theta|B_i, S_i) \\
&= \sum_{i=1}^{I} [-2\varepsilon T + (B_i + S_i)\ln T] \\
&\quad + \sum_{i=1}^{I} \ln\left[(1-\alpha)\varepsilon^{B_i+S_i} + \alpha\delta e^{-\mu T}\varepsilon^{B_i}(\mu+\varepsilon)^{S_i} + \alpha(1-\delta)e^{-\mu T}\varepsilon^{S_i}(\mu+\varepsilon)^{B_i}\right] \\
&\quad - \sum_{i=1}^{I} (\ln B_i! + \ln S_i!). \qquad (7)
\end{aligned}
$$

As in Easley et al. (2008), the log likelihood function, after dropping the constant and rearranging[7], is given by

$$
\begin{aligned}
\ell(M|\theta) &= \sum_{i=1}^{I} [-2\varepsilon + M_i \ln x + (B_i + S_i)\ln(\mu + \varepsilon)] \qquad (8) \\
&\quad + \sum_{i=1}^{I} \ln\left[\alpha(1-\delta)e^{-\mu}x^{S_i - M_i} + \alpha\delta e^{-\mu}x^{B_i - M_i} + (1+\alpha)x^{B_i+S_i-M_i}\right],
\end{aligned}
$$

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

2.3. Heterogeneous information and orders with different sizes

Given the (buy-sell) likelihoods presented in the previous section, we now utilize a procedure similar to Easley et al. (1997b), theoretically outlined in Easley and O'Hara (1987). This approach allows informed and uninformed traders to place both large and small orders. The extended model relies on the number of unique large buy (LB), small buy (SB), large sell (LS) and small sell (SS) trades that represent the set of possible trade outcomes.[8] This approach introduces two new parameters: ϕ (the probability that an uninformed trader trades a large amount) and ω (the probability that an informed trader trades a large amount). Naturally, $(1 - \phi)$ denotes the

[7] To derive equation (8), 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 equation (7). This approach increases computational efficiency and ensures convergence in the presence of a large number of buys and sells.

[8] For simplicity, the no-trade outcome considered in Easley et al. (1997b) for a much smaller dataset of stock prices is ignored.

probability of a small uninformed trade, and $(1-\omega)$ is the probability of a small informed trade. All other parameters (α, μ, δ and ε) follow the previous notation. The likelihood of observing a sequence of orders with LB large buys, SB small buys, LS large sells and SS small sells in a bad news hour is

$$\begin{aligned} & L_b((LB,LS,SB,SS)|\theta) = L_b(LB|\theta)L_b(LS|\theta)L_b(SB|\theta)L_b(SS|\theta) \\ &= e^{-(\mu+2\varepsilon)T}\frac{(\varepsilon\phi)^{LB}[\varepsilon(1-\phi)]^{SB}(\varepsilon\phi+\mu\omega)^{LS}[\varepsilon(1-\phi)+\mu(1-\omega)]^{SS}T^{LB+LS+SB+SS}}{LB!LS!SB!SS!}, \end{aligned}$$

where $\theta=(\alpha,\delta,\varepsilon,\mu,\omega,\phi)$. On a no-event day, the likelihood of observing a sequence of LB large buys, SB small buys, LS large sells and SS small sells is

$$\begin{aligned} L_n((LB,LS,SB,SS)|\theta) &= L_n(LB|\theta)L_n(LS|\theta)L_n(SB|\theta)L_n(SS|\theta) \\ &= e^{-2\varepsilon T}\frac{\phi^{LB+LS}(1-\phi)^{SB+SS}(\varepsilon T)^{LB+LS+SB+SS}}{LB!LS!SB!SS!}. \end{aligned}$$

On a good-event day, the likelihood is

$$\begin{aligned} & L_g((LB,LS,SB,SS)|\theta) = L_g(LB|\theta)L_g(LS|\theta)L_g(SB|\theta)L_g(SS|\theta) \\ &= e^{-(\mu+2\varepsilon)T}\frac{(\varepsilon\phi)^{LS}[\varepsilon(1-\phi)]^{SS}(\varepsilon\phi+\mu\omega)^{LB}[\varepsilon(1-\phi)+\mu(1-\omega)]^{SB}T^{LB+LS+SB+SS}}{LB!LS!SB!SS!}. \end{aligned}$$

As before, the likelihood of observing LB large buys, SB small buys, LS large sells and SS small sells is the weighted average of the above equations:

$$L((LB,LS,SB,SS)|\theta) = (1-\alpha)L_n(.|\theta) + \alpha\delta L_b(.|\theta) + \alpha(1-\delta)L_g(.|\theta)$$

Because this work uses hourly data, the likelihood of observing the data $D=(LB_i,LS_i,SB_i,SS_i)_{i=1}^{I}$ over twenty-four hours ($I=24$) is the product of the hourly likelihoods. That is,

$$L(D|\theta) \quad = \quad \prod_{i=1}^{I} L(\theta|LB_i,LS_i,SB_i,SS_i),$$

and the log likelihood function is now

$$\begin{aligned}
\ell(D|\theta) &= \sum_{i=1}^{I} \ell(\theta|LB_i, LS_i, SB_i, SS_i) \\
&= \sum_{i=1}^{I} [-2\varepsilon + M_i \ln x + N_i \ln y] \\
&\quad + \sum_{i=1}^{I} [(LB_i + LS_i) \ln(\varepsilon\phi + \mu\omega) + (SB_i + SS_i) \ln(\varepsilon(1-\phi) + \mu(1-\omega))] \\
&\quad + \sum_{i=1}^{I} \ln[(1-\alpha) x^{LB_i + LS_i - M_i} y^{SB_i + SS_i - N_i} + \alpha\delta e^{-\mu} x^{LB_i - M_i} y^{SB_i - N_i} \\
&\quad + \alpha(1-\delta) e^{-\mu} x^{LS_i - M_i} y^{SS_i - N_i}],
\end{aligned}$$

where $M_i \equiv \min(LB_i, LS_i) + \max(LB_i, LS_i)/2$, $N_i \equiv \min(SB_i, SS_i) + \max(SB_i, SS_i)/2$, $y = \frac{\varepsilon(1-\phi)}{\varepsilon(1-\phi)+\mu(1-\omega)} \in [0,1]$ and $x = \frac{\varepsilon\phi}{\varepsilon\phi+\mu\omega} \in [0,1]$. Here, to obtain the final expression, the terms $\ln[x^{M_i}(\mu\omega + \varepsilon\phi)^{LB_i+LS_i}]$ and $\ln[y^{N_i}(\mu(1-\omega) + \varepsilon(1-\phi))^{SB_i+SS_i}]$ are added to and subtracted from the right-hand side of the likelihood equation. Before proceeding with the estimations, we first describe our trading database.

3. Data

In the early 1990s, the practice of switching from voice brokers to electronic trading systems rendered the FX market more transparent. However, early over-the-counter (OTC) FX market participants had no means of observing the market-wide order flow. The introduction of centralized electronic broking systems such as Reuters and EBS thus provided a new platform for research on the FX market microstructure. Although Reuters and EBS are dominant in electronic FX markets, they do not publicly report high-frequency volume data or trader identities.

The dataset from the OANDA FXTrade internet trading platform consists of tick-by-tick foreign exchange transaction prices and the corresponding volumes for several exchange rates from October 1, 2003, to May 14, 2004. The number of active traders during this period is 4,983, and they mainly trade four major exchange rates.[9] The data show that the overall trading frequency increases from

[9] By "active", the paper refers to traders that did not simply receive interest on their positions but placed orders

Monday to Wednesday (the peak) and falls from Thursday to Saturday. Using the trader's identity (trader ID), we find that approximately 33% of the investors specialize in exactly one currency pair, approximately 11% specialize in two currency pairs, and approximately 9% specialize in three currency pairs. This decreasing trend leads to only 2-4% of active traders who deal in 10 to 13 currency pairs. Hence, traders appear to specialize in a small number of currency pairs, in line with OANDA FXTrade's intention to attract small investors.

[Insert Figure 1 here]

Because the bulk of all transactions (approximately 40 percent[10]) involve only U.S. Dollar-Euro (USD-EUR) trading, the focus is on transactions involving only USD-EUR. In particular, the research analyzes all USD-EUR buy and sell transactions (market, limit order executed, margin call executed, stop-loss, and take-profit transactions). In addition to price and volume, the trader ID for each transaction is known, and it ranges from 123 to 5904. The average number of USD-EUR transactions per trader is 512. One can observe that a few traders transact very frequently in this currency pair (between 10,000 and 25,000 times) over the time period of the data (Figure 1). The day-of-the-week trading patterns for the USD-EUR transactions (trading frequency and volume) are similar to those of other currency pairs.

[Insert Table 1 here]

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 or by the analysis of prices from the Interbank market. The OANDA FXTrade policy is to offer the tightest possible bid/ask spread (e.g., 0.0009% spread on the USD-EUR, regardless of transaction size). These traders, similar to most market makers, profit from the spreads. Some of the other market features include continuous, second-by-second interest

during this period. The market share of these traders is approximately 86.4%.

[10]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).

rate payments, no limit on the transaction size, no requirement for 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 findings in Gençay et al. (2015), it is reasonable to assume that informed traders are also present in this market.

The theoretical model by Easley et al. (2008) 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 hours and is decentralized. Further, unlike the NYSE, the FX market does not involve a so-called specialist responsible for maintaining fairness and order, with 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 (from below five minutes to half an hour). This process is sometimes referred to as "hot-potato trading" (Evans and Lyons, 2002a; Bjønnes and Rime, 2005). On the NYSE, however, inventory has an average half-life of more than a week (Madhavan and Smidt, 1993). Thus, inventory management is an important component of intraday FX trader activity.

Given its features, the OANDA FXTrade can be viewed as a "special case" of the FX market that can be approached using the model by Easley et al. (2008). 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. With regard to trader behavior, as the only focus is 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.

The preliminary analysis (unreported for the sake of brevity) indicates that overall market activity was extremely low on certain days or during certain weeks. Therefore, the weekends are eliminated, 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 Independence Day (April 5-9). This leaves 145 24-hour periods. To avoid extremely high-frequency noise and no-activity periods in small time windows, we aggregate the data over one-hour intervals. Aggregating over trading intervals smaller than one hour is not feasible, as this would not cover a sufficient number of buy and sell transactions for the model to be empirically applicable. By contrast, 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). The flow of information is unlikely to 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 EST, 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-EUR contracts. According to the BIS Triennial Survey for 2004, the daily average turnover in the USD-EUR currency pair was 501 billion USD. Hence, on average, the data represent approximately 0.045% of the global daily USD-EUR trading volume. Nevertheless, it is one of the largest tick-by-tick FX datasets to be used in an academic study.

Because the trader's identity is known in each transaction, one is able to identify the number of unique traders in each one-hour 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.

4. Empirical results

To present the empirical results, we proceed as follows. First, we compare the estimated ω and ϕ over 145 days in our sample. If $\omega > \phi$, then trade size conveys additional information to market participants. If, however, $\omega < \phi$, then one can conclude that traders (either informed or uninformed) do not significantly benefit from trade size. Following that, our second objective is to examine how changes in the cutoff trade size impact the estimates. Third, we investigate whether trade size responds to price volatility or vice versa. Finally, we link trade size to the empirical properties of exchange rate data and particularly assess the implications for excess kurtosis.

4.1. Testing for the impact of trade size

The procedure of testing for trade size effects involves comparing estimates of the restricted ($\omega = \phi$) and unrestricted ($\omega \neq \phi$) models. The cutoff amount that differentiates large from small trades is initially set to 5,000, and it is subsequently shown that this does not affect the main results.[11]

[Insert Table 2 here]

Table 2 lists the average estimates of $\alpha_i, \delta_i, \varepsilon_i, \mu_i, \omega_i$, and ϕ_i ($i = 1, \ldots, 145$). The paired t-test of the equality of the means of the constrained and unconstrained models shows no significant difference for the first four parameters.[12] However, the difference between the two sets of estimates of ω_i is statistically significant.[13] Furthermore, including trade size effects (unconstrained model) significantly increases the absolute value of the log likelihood function, thus indicating that the constraint is binding. The informativeness of trade size is also confirmed by the unpaired t-test of the equality of $\bar{\omega}$ and $\bar{\phi}$ for the unconstrained model. The model concludes that $\bar{\omega}$ is significantly greater than $\bar{\phi}$. On approximately 68% of the days in the sample, $\omega_i > \phi_i$, and the difference in the probabilities ($\omega_i - \phi_i$) ranges from -0.12 to 0.14 ($i = 1, \ldots, 145$). Although there are 47 days when the probability of uninformed large trading exceeds the probability of informed large trading, this is not the typical case.

[Insert Table 3 here]

We now investigate whether the findings above are robust to the choice of the cutoff amount for a "large" trade. Table 3 reports the results for cutoff rates of 2000, 8000 and 12000. The focus is on the difference column from Table 2 and the mean values of the unconstrained estimates.

The results indicate that in the cutoff range between 3000 and 8000, all estimates are stable, and the informed large trade size is more informative than the uninformed large trade size. The choice

[11] Trade size is expressed in currency units of the base currency, i.e., the Euro.

[12] The null hypothesis for this test is that the mean difference ($\bar{d}$) between the paired observations (constrained and unconstrained) of estimated parameters is zero. The test statistic is calculated as $t = \frac{\bar{d}}{\sqrt{s_{\bar{d}}/145}}$, where $s_{\bar{d}}$ is the sample standard deviation for $\bar{d}$.

[13] Additionally, the standard errors of $\hat{\omega}_i$ and $\hat{\phi}_i$ for the unconstrained model are consistently on the order of 10^{-4} and 10^{-5}, respectively, thus indicating statistically significant differences in the probabilities.

of cutoff values above 8000 (e.g., 12000 in Table 3) distorts the results due to the low frequency of such large trades. Similarly, it is unreasonable to consider trades above small cutoff values (e.g., 2000 in Table 3) to be "large," in which case the observed effects diminish. These findings confirm the work of Chakravarty (2001) and Anand and Chakravarty (2007), who find that medium-sized trades are the most informative. This finding can also be interpreted as a "separating equilibrium" outcome in which informed traders submit mainly large orders (Easley and O'Hara, 1987).[14] An interesting observation emerges from Table 3: the probability of both informed and uninformed large trading declines with the cutoff value. This result can be explained by the fact that increasing the cutoff value eliminates the majority of the transactions that qualify as "large" trades.

4.2. Does trade size respond to price volatility?

The previous section provides evidence that large trades are likely to be placed by informed traders in the marketplace. Given this finding, it is natural to investigate whether trade size reacts to market conditions or uncertainty. To assess the relationship between trade size and volatility, Kalok and Fong (2000), for instance, consider trade size to be a *control* variable contributing rather than responding to volatility. In the empirical framework of Kalok and Fong (2000), treating size as a predictor is undoubtedly plausible because their objective is to explain how trade size affects the volatility-volume relationship in the equity market (NASDAQ and NYSE).

Nevertheless, our approach differs from Kalok and Fong (2000) in two important respects. First, we focus on the FX market microstructure, and thus, the volatility-trade size relationship could be different here. Second, we are primarily interested in characterizing the properties of (large) trade size and analyzing its reaction to price volatility, although the reverse causality will be tested in the following subsection. Our (a priori) expectation is that if informed traders appear to place large orders, consistent with the evidence, then investors may strategically adjust their trade sizes during market stress or excess volatility.[15]

[Insert Table 4 here]

[14]Suppose that the constrained model is found to be more appropriate. This would indicate a "pooling equilibrium," where informed traders submit both large and small orders roughly equally.

[15]In our view, this intuition is also in line with the *intraday trade invariance* argument of Andersen et al. (2015), who empirically show that trade size increases with price volatility in the E-mini S&P 500 futures market.

To test how volatility impacts trade size, we estimate a standard logit model that links large sizes (binary dependent variable) to intraday absolute log-price changes (predictor) as a proxy for spot volatility. For our binary variable, we choose two categories such that

$$y_i = \begin{cases} 1 & \text{if trade size} > 5000 \text{ (cutoff)}, \\ 0 & \text{otherwise.} \end{cases}$$

Based on these choices, our interest lies in modeling the probabilities of observing a large trade (exceeding 5000), that is, $p_i = \Pr\{y_i = 1\}$. To accomplish this objective for both buy and sell sides, we consider four transaction types for closing prices: [1] market buy side, [2] market sell side, [3] limit order buy side (executed), and [4] limit order sell side (executed).

Table 4 reports the estimation results of the baseline logit model with the trade size cutoff 5000. The table indicates that price volatility significantly increases the probability of observing *large* trades in the market transaction data (models [1] and [2]). The estimated parameters (i.e., log-odds) for variable VOL is statistically significant, and its impact is the largest (15.98) for market buy-side transactions (model [1]) compared with other parameter estimates, e.g., in [2] and [4] (7.96 and 2.90, respectively).[16]

The third and fourth columns of Table 4 further show that the reaction of trade size to volatility is *asymmetric* across the buy and sell sides of executed limit orders (i.e., models [3] and [4], respectively). For instance, while volatility tends to lower the likelihood of observing large trades on the *buy side* (with log-odds of -4.89), trade size increases with price volatility in *sell-side* limit orders (log-odds of 2.90). The estimates are economically significant, suggesting potential differences in traders' risk aversion (upside or downside) based on market conditions.

Moreover, we provide the estimates of quadratic price volatility ($SQVOL$) for all four specifications (the third row of Table 4). On both the market buy and sell sides (models [1] and [2]), the sign of the volatility impact becomes significantly negative (at the 1% level), and the likelihood of

[16]Notably, the interpretation of parameter estimates (log-odds) in binary choice models differs from that in standard linear regression modeling. Nevertheless, one can calculate the exponential of the estimates as $(e^x/(1+e^x))$, which transfers the estimate to the corresponding empirical (success) likelihood. For the sake of brevity, we report only the parameter estimates.

large trade arrivals decreases with $SQVOL$. One potential reason for this evidence could be related to *order-splitting* activity: while volatility (VOL) results in large trade execution, a high degree of market uncertainty ($SQVOL$) might cause informed traders to split their large orders into small units. These results, however, do not hold when transactions are (executed) buy-side limit orders (model [3]) for which the $SQVOL$ estimate has a positive sign (4.05).

[Insert Table 5 here]

We now turn to examine the sensitivity of our logit results to the choice of trade size cutoffs. Table 5 reports the estimates for cutoff rates of 2000, 8000 and 12000. Panels A to C in the table suggest that our main results remain qualitatively similar irrespective of the trade size thresholds. For instance, as in our baseline findings (reported in Table 4), the largest (log-odds) impact comes from the market buy side (model [1]) across all three cutoff choices. The probability of large trades increases with volatility, yet the impact diminishes with high volatility levels ($SQVOL$). Another noticeable feature is that the parameter estimate of VOL ($SQVOL$) has a positive (negative) sign when the cutoff level is 12000 (model [3] in Panel C). This result may suggest that the limit order is more sensitive to market conditions (and hence volatility) relative to the results for market transactions (i.e., models [1] and [2]).

4.3. Reverse direction: does price volatility respond to trade size?

Given the evidence that large trades vary with volatility, one can also argue whether the volatility-size relationship is bidirectional. This assessment can be motivated by the fact that large trades are informative and that price volatility may thus reflect information embedded in trades. To test for this reverse hypothesis, we now consider price volatility as a response variable and trade size as a predictor in our logit regressions. We choose various *high volatility* thresholds (k) that rely on a certain number of standard deviations (e.g., $k = 1, 3, 6, 9$) from the mean volatility estimate over the sample.[17]

[17]Notably, this truncation approach is relatively *ad hoc*. In our analysis, we are solely interested in identifying excessive (or abnormal) volatility compared to normal (mean/median) levels. Despite its apparent simplicity, this approach is still broadly in line with the techniques implemented in the jump literature (see, e.g., Aït-Sahalia and Jacod, 2012, 2010 and Aït-Sahalia, 2004). In these studies, α typically represents the size of the log return observed in terms of the number of standard deviations from the mean (see Aït-Sahalia, 2004) or from the Brownian (diffusion) part of prices—linked to long-term volatility (see, e.g., Aït-Sahalia and Jacod, 2012, 2010).

[Insert Table 6 here]

Table 6 reports the estimation results for five different choices of truncation levels (from [1] to [5] in the top row). The estimates indicate a positive relationship between trade size and the volatility of market transaction prices (third row). The degree of impact appears to be significant for episodes of high volatility (i.e., between [1] and [3]), but the evidence seems to be relatively weak when volatility reaches extreme levels ([4] and [5]). For example, if we measure abnormal volatility as nine standard deviations away from the mean (i.e., by setting $k = 9$), then order size fails to influence volatility patterns (model [5]).[18] These reverse-causality findings remain robust to various subsample selections.[19]

4.4. Trade size, price-contingent orders and excess kurtosis

We complete our analysis by examining the link between trade size and empirical characteristics of exchange rate data. This is motivated by Osler and Savaser (2011) and Osler (2005), whose evidence indicates that price-contingent trading helps explain the excess kurtosis in currency returns. In the same vein, we thus assess how trade size is associated with excess kurtosis. For this objective, we first split our sample into large and small orders with the cutoff level of 5000. As in Osler and Savaser (2011) and Osler (2005), we then consider the log-price changes (i.e., exchange rate returns) corresponding to stop-loss and take-profit trades in each subsample.

[Insert Table 7 here]

Table 7 reports the sample moments computed on the log-price differences of price-contingent trades. The table indicates that excess kurtosis in large trade samples (Panel A) is significantly smaller than that in small trade samples (Panel B). This pattern is particularly noticeable with take-profit transactions (Panel B). For example, the excess kurtosis of ΔTPS (ΔTPB) is around 228 (332) in small orders, whereas the estimate decreases to 50-60 in large order samples (Panel A). Relying on our main empirical results, we can explain these features in two ways. First, as we

[18] One explanation for this finding could be linked to the nature of the trading environment. Our database (OANDA) is a retail trading platform, and hence, while informed traders adjust their orders depending on market volatility, the reverse may not necessarily hold. The impact of retail trading per se on the entire FX trade is limited in practice.

[19] These additional robustness tests are unreported for the sake of brevity, but they are available upon request.

show that small orders are typically placed by uninformed traders, they might take a wide range of aggressive positions that increase the likelihood of extreme events. Second, large orders from informed traders tend to be more concentrated and less intrusive because these traders use their private information to place consistent large orders that are rarely extreme.

5. Conclusions

This paper provides evidence that the transactions of informed FX traders are related to larger trade sizes. These findings are robust with regard to reasonable choices for cutoff points that define a “large” trade, although some trade sizes that are found to be informative can also be interpreted as medium-sized. Extending and complementing the extant literature, our analysis provides evidence on the link between informed trading and larger trade sizes (e.g., Easley et al., 1997a, Menkhoff and Schmeling, 2010, Chakravarty, 2001 and Anand and Chakravarty, 2007). The observed behavior can be described as a strong strategic component in the activity of informed traders that is not observed for uninformed traders (Gençay and Gradojevic, 2013). In contrast, uninformed traders submit smaller currency orders while acting in a “dispersed manner” that increases the likelihood of extreme events in the FX market.

Our empirical analysis further shows that large trades are associated with local price volatility representing market uncertainty at high frequency. In the electronic foreign exchange market, the probability of large trade arrivals *increases* with EUR/USD volatility but *decreases* when volatility is severe; this finding suggests a nonlinear interaction between these two factors. While this evidence indirectly supports the *trading invariance principle* recently outlined by Andersen et al. (2015) for stock futures, our results contrast with Kalok and Fong (2000) regarding the properties of the volatility-trade size relationship. Specifically, FX data suggest that trade size depends on price volatility, and hence, from an econometric viewpoint, size may not be considered an *exogenous* microstructure parameter. Moreover, reversing the direction of causality shows that large trade sizes positively affect FX market volatility, albeit less strongly. Intuitively, informed retail FX traders attempt to camouflage their large trades during episodes of high volatility, where the impact of their trading on FX volatility is relatively small. These conclusions remain valid regardless of changes in large trade sizes (e.g., from relatively medium to large trades and from large to extra-large trades).

Although our results generally suggest that information leads to large order execution, there could be other potential factors affecting the trade size−information relationship. For instance, informed currency traders may prefer to split orders and front-run large institutional traders. In the presence of such order-splitting actions, the orders of informed traders might become smaller, and hence, traders with large orders could be uninformed. This possibility may arise particularly in data with jump-type price movements (Tseng et al., 2015). Building on the predictions of our model, future research may thus shed more light on the role of order-splitting and excess volatility in the informativeness of trade size.

References

Admati, A., Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. Review of Financial Studies 1, 3–40.

Aït-Sahalia, Y., 2004. Disentangling diffusion from jumps. Journal of Financial Economics 74, 487–528.

Aït-Sahalia, Y., Jacod, J., 2010. Is Brownian motion necessary to model high-frequency data? Annals of Statistics 38, 3093–3128.

Aït-Sahalia, Y., Jacod, J., 2012. Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. Journal of Economic Literature 50, 1007–1050.

Allen, F., Gale, D., 2000. Financial contagion. Journal of Political Economy 108, 1–33.

Anand, A., Chakravarty, S., 2007. Stealth trading in options markets. Journal of Financial and Quantitative Analysis 42, 167–188.

Andersen, T.G., Bondarenko, O., Kyle, A.S., Obizhaeva, A.A., 2015. Intraday trading invariance in the E-mini S&P 500 futures market. Working Paper.

Banerjee, S., Green, B., 2015. Signal or noise? Uncertainty and learning about whether other traders are informed. Journal of Financial Economics 117, 398–423.

Baruch, S., Glosten, L.R., 2013. Fleeting orders. Working Paper.

Bjønnes, G., Rime, D., 2005. Dealer behavior and trading systems in foreign exchange markets. Journal of Financial Economics 75, 571–605.

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A., Sokolov, K., 2015a. High frequency trading and extreme price movements. Working Paper.

Brogaard, J., Hendershott, T.J., Riordan, R., 2014. High-frequency trading and price discovery. Review of Financial Studies 27, 2267–2306.

Brogaard, J., Hendershott, T.J., Riordan, R., 2015b. Price discovery without trading: Evidence from limit orders. Working Paper.

Caballero, R.J., Krishnamurthy, A., 2008. Collective risk management in a flight to quality episode. Journal of Finance 63, 2195–2230.

Chakravarty, S., 2001. Stealth-trading: Which traders trades move stock prices? Journal of Financial Economics 61, 289–307.

Duffie, D., Zhu, H., 2015. Size discovery. Working Paper.

Dufour, A., Engle, R.F., 2000. Time and the price impact of a trade. Journal of Finance 55, 2467–2498.

Dungey, M., Milunovich, G., Thorp, S., 2010. Unobservable shocks as carriers of contagion. Journal of Banking and Finance 34, 1008–1021.

Easley, D., Engle, R., O'Hara, M., Wu, L., 2008. Time-varying arrival rates of informed and uninformed trades. Journal of Financial Econometrics 6, 171–207.

Easley, D., Kiefer, N., O'Hara, M., 1997a. The information content of the trading process. Journal of Empirical Finance 4, 159–185.

Easley, D., Kiefer, N., O'Hara, M., 1997b. One day in the life of a very common stock. Review of Financial Studies 10, 805–835.

Easley, D., O'Hara, M., 1987. Price, trade size, and information in securities markets. Journal of Financial Economics 19, 69–90.

Ederington, L.H., Lee, J.H., 1993. How markets process information: News releases and volatility. Journal of Finance 48, 1161–1191.

Ederington, L.H., Lee, J.H., 1995. The short-run dynamics of the price adjustment to new information. Journal of Financial and Quantitative Analysis 30, 117–134.

Evans, M., 2002. FX trading and exchange rate dynamics. Journal of Finance 57, 2405–2447.

Evans, M., Lyons, R., 2002a. Order flow and exchange rate dynamics. Journal of Political Economy 110, 170–180.

Evans, M.D.D., Lyons, R.K., 2002b. Informational integration and FX trading. Journal of International Money and Finance 21, 807–831.

Gençay, R., Gradojevic, N., 2013. Private information and its origins in an electronic foreign exchange market. Economic Modelling 33, 86–93.

Gençay, R., Gradojevic, N., Olsen, R., Selçuk, F., 2015. Informed traders' arrival in foreign exchange markets: Does geography matter? Empirical Economics 49, 1431–1462.

Gençay, R., Mahmoodzadeh, S., Rojcek, J., Tseng, M.C., 2016. Price impact of aggressive liquidity provision. Working Paper.

Glosten, L.R., 1989. Insider trading, liquidity, and the role of the monopolist specialist. Journal of Business 62, 211–235.

Glosten, L.R., 1994. Is the electronic open limit order book inevitable? Journal of Finance 49, 1127–1161.

Glosten, L.R., Milgrom, P.R., 1985. Bid, ask and transaction prices in a market-maker market with heterogeneously informed traders. Journal of Financial Economics 14, 71–100.

Goodhart, C.A.E., Payne, R., 1996. Multistructural dynamics in foreign exchange electronic broking system. Journal of International Money and Finance 15, 829–852.

Hautsch, N., Huang, R., 2012. The market impact of a limit order. Journal of Economic Dynamics and Control 36, 501–522.

Jiang, G.J., Lo, I., Verdelhan, A., 2011. Information shocks, liquidity shocks, jumps, and price discovery: Evidence from the U.S. Treasury market. Journal of Financial and Quantitative Analysis 46, 527–551.

Kalok, C., Fong, W.M., 2000. Trade size, order imbalance, and the volatility-volume relation. Journal of Financial Economics 57, 247–273.

King, M., Osler, C.L., Rime, D., 2013. The market microstructure approach to foreign exchange: Looking back and looking forward. Journal of International Money and Finance 38, 95–119.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Madhavan, A., Smidt, S., 1991. A Bayesian model of intraday specialist pricing. Journal of Financial Economics 30, 99–134.

Madhavan, A., Smidt, S., 1993. An analysis of changes in specialist inventories and quotations. Journal of Finance 48, 1595–1628.

Menkhoff, L., Schmeling, M., 2008. Local information in foreign exchange markets. Journal of International Money and Finance 27, 1383–1406.

Menkhoff, L., Schmeling, M., 2010. Whose trades convey information? Evidence from a cross-section of traders. Journal of Financial Markets 13, 101–128.

Osler, C., 2005. Stop-loss orders and price cascades in currency markets. Journal of International Money and Finance 24, 219–241.

Osler, C., Savaser, T., 2011. Extreme returns: The case of currencies. Journal of Banking and Finance 35, 2868–2880.

Osler, C.L., Mende, A., Menkhoff, L., 2011. Price discovery in currency markets. Journal of International Money and Finance 30, 16961718.

Tseng, M.C., Mahmoodzadeh, S., Gençay, R., 2015. White noise jumps and adverse selection in FX market. Working Paper.

Zhang, M.Y., Russell, J.R., Tsay, R.S., 2001. A nonlinear autoregressive conditional duration model with applications to financial transaction data. Journal of Econometrics 104, 179–207.

Table 1: OANDA FXTrade institutional characteristics

Variable	Sample
Hours of operation	24 hours/7 days per week
Number of currency pairs	30
Number of active traders	4,983
Number of trades (USD-EUR)	667,030 sell transactions
	666,133 buy transactions
Average number of trades per day (USD-EUR)	192 sell transactions
	191 buy transactions
Total volume (USD-EUR)	32.6 billion USD
Average volume per day (USD-EUR)	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

Notes: The table presents the institutional characteristics of the OANDA FXTrade trading platform. OANDA FX-Trade is an electronic market making system that executes orders using bid/ask prices that are realistic and prevalent in the marketplace. Our sample consists of tick-by-tick foreign exchange transaction prices and the corresponding volumes for several exchange rates from October 1, 2003 to May 14, 2004.

Table 2: The information role of trade size

Parameter	Benchmark model	Constrained	Unconstrained	Difference (p-value)
$\bar{\alpha}$	0.30	0.30	0.30	0.00 (0.17)
$\bar{\delta}$	0.47	0.47	0.47	0.00 (0.42)
$\bar{\varepsilon}$	77.8	77.16	77.12	0.04 (0.19)
$\bar{\mu}$	83.5	82.5	82.5	0.00 (0.99)
$\bar{\omega}$	-	0.43	0.45	-0.02 (0.00)***
$\bar{\phi}$	-	0.43	0.42	0.01 (0.00)***
$\overline{LLF}$	-15111	-12087.25	-12087.46	0.21 (0.08)*

Notes: The first column lists the average estimates for the model, which do not account for the trade size. The second and third columns represent the average estimates of the parameters in the constrained ($\omega_i = \phi_i$) and unconstrained ($\omega_i \neq \phi_i$; $i = 1, \dots, 145$) versions of the model, respectively. The last column contains the differences in mean value between the 145 parameters estimated from the constrained and unconstrained models. The p-value comes from the paired t-test for the null hypothesis of the difference being equal to zero. LLF denotes the value of the log likelihood function. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table 3: The robustness of the estimates with respect to "large" trade size

Parameter	2000 cutoff		8000 cutoff		12000 cutoff	
$\bar{\alpha}$	0.31	0.00 (0.43)	0.30	0.00 (0.14)	0.30	0.00 (0.33)
$\bar{\delta}$	0.48	0.00 (0.53)	0.47	0.00 (0.66)	0.47	0.02 (0.38)
$\bar{\varepsilon}$	74.43	0.02 (0.59)	78	0.00 (0.15)	77.93	0.05 (0.19)
$\bar{\mu}$	74.54	0.42 (0.40)	83.11	1.34 (0.12)	82.65	0.82 (0.19)
$\bar{\omega}$	0.73	-0.02 (0.00)***	0.36	-0.01 (0.00)***	0.24	-0.00 (0.02)**
$\bar{\phi}$	0.70	0.01 (0.00)***	0.34	0.01 (0.00)***	0.23	0.00 (0.00)***
$\overline{LLF}$	0.12	(0.15)	1.66	(0.09)*	0.97	(0.17)

Notes: For each cutoff amount for a "large" trade (2000, 8000 and 12000), this table presents the average parameter estimates from an unconstrained model along with the average difference between the estimates from the two versions of the model: constrained ($\omega_i = \phi_i$) and unconstrained ($\omega_i \neq \phi_i$; $i = 1, \dots, 145$). More precisely, each column represents the merged columns 3 and 4 from Table 2 for different cutoff amounts. LLF denotes the average value of difference between the log likelihood function for the two models. The p-value reported in the brackets comes from the paired t-test for the null hypothesis of the difference being equal to zero. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table 4: Logit model estimation results for trade size: impact of price volatility

	[1]	[2]	[3]	[4]
c	-0.823***	-0.754***	1.210***	-1.022***
	(0.006)	(0.005)	(0.028)	(0.027)
VOL	15.976***	7.962***	-4.885***	2.898***
	(0.508)	(0.315)	(0.516)	(0.503)
$SQVOL$	-79.689***	-7.009***	4.046***	-2.588**
	(5.911)	(1.381)	(1.333)	(1.272)
Obs.	190195	232661	10659	10307
$\chi^2(2)$	1498.30	715.00	170.21	55.54
	[0.00]	[0.00]	[0.00]	[0.00]
LL	-121018	-148503	-6090	-6153
R^2	0.006	0.002	0.014	0.004

Notes: The table reports the logit model estimation results for a fixed trade size cutoff. The response variable trade size (y_i) is $y_i = 1$ if the trade size exceeds 5000, and $y_i = 0$ otherwise. Estimation is based on Newton's method. The predictors are absolute log-price changes (VOL), as proxy for intraday price volatility, and its squared version to capture potential nonlinear effects ($SQVOL$). c denotes the constant. The table presents the estimated coefficients (i.e., the log-odds) and standard errors in parenthesis. For closing prices, we consider four models with different transaction types. [1]: Market buy-side, [2]: Market sell-side, [3]: Buy-side limit order executed, [4]: Sell-side limit order executed. The bottom three rows in the table report the number of observations (Obs), $\chi^2(2)$ with its rejection probability (in square brackets), log-likelihood value (LL) and pseudo (McFadden's) R^2 values calculated as 1-[$LL(full)/LL(baseline)$]. The sample covers the periods from October 1, 2003 to May 14, 2004. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table 5: Logit model estimation results for different large trade size cutoffs

	Panel A. cutoff 2000				Panel B. cutoff 8000				Panel C. cutoff 12000			
	[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]
c	-0.209***	-0.117***	0.568***	-0.321***	-0.982***	-0.905***	1.289***	-1.120***	-1.460***	-1.435***	-1.976***	-1.710***
	(0.006)	(0.005)	(0.025)	(0.024)	(0.007)	(0.006)	(0.029)	(0.028)	(0.008)	(0.006)	(0.034)	(0.033)
VOL	15.410***	8.323***	-4.297***	2.010***	16.217***	7.796***	-4.605***	2.431***	14.511***	6.960***	4.084***	2.648***
	(0.464)	(0.311)	(0.478)	(0.434)	(0.528)	(0.320)	(0.521)	(0.459)	(0.583)	(0.346)	(0.427)	(0.587)
$SQVOL$	-63.520***	-8.342***	3.210**	-1.782***	-84.297***	-6.975***	3.696***	-1.757**	-72.655***	-4.487***	-1.816***	-2.267
	(5.081)	(1.396)	(1.265)	(1.031)	(6.229)	(1.365)	(1.333)	(1.033)	(6.788)	(1.243)	(0.639)	(1.441)
Obs.	190195	232661	10659	10307	190195	232661	10659	10307	190195	232661	10659	10307
$\chi^2(2)$	1675.90	801.26	155.69	31.11	1431.10	657.43	149.10	41.93	952.65	441.31	110.98	33.76
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
LL	-130911	-160834	-7097	-7051	-116208	-142748	-5910	-5936	-97662	-117498	-4321	-4641
R^2	0.006	0.002	0.011	0.002	0.006	0.002	0.012	0.004	0.005	0.002	0.013	0.004

Notes: The table reports the logit model estimation results for different large trade size cutoffs (2000 in Panel A., 8000 in Panel B., 12000 in Panel C.). The response variable trade size (y_i) is $y_i = 1$ if the trade size exceeds the cutoff level (reported in each panel), and $y_i = 0$ otherwise. Estimation is based on Newton's method. The predictors are absolute log-price changes (VOL), as proxy for intraday price volatility, and its squared version to capture potential nonlinear effects ($SQVOL$). c denotes the constant in all models. The table presents the estimated coefficients (i.e., the log-odds) and standard errors in parenthesis. For closing prices, we consider four models with different transaction types. [1]: Market buy-side, [2]: Market sell-side, [3]: Buy-side limit order executed, [4]: Sell-side limit order executed. The bottom three rows in the table report the number of observations (Obs), $\chi^2(2)$ with its rejection probability (in square brackets), log-likelihood value (LL) and pseudo (McFadden's) R^2 values calculated as 1-[$LL(full)/LL(baseline)$]. The sample covers the periods from October 1, 2003 to May 14, 2004. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table 6: Logit model estimation results for price volatility: impact of trade size

	[1]	[2]	[3]	[4]	[5]
c	0.011**	-1.454***	-3.904***	-6.456***	-7.893***
	(0.005)	(0.006)	(0.017)	(0.058)	(0.119)
$SIZE$	0.004***	0.009***	0.025***	0.045**	0.041
	(0.002)	(0.002)	(0.005)	(0.019)	(0.039)
Obs.	190195	190195	190195	190195	190195
$\chi^2(1)$	7.46	21.13	20.82	5.83	1.08
	[0.01]	[0.00]	[0.00]	[0.02]	[0.30]
LL	-131826	-92303	-18492	-2239	-630
R^2	0.000	0.000	0.001	0.001	0.001

Notes: The table reports the logit model estimation results for volatility impact of trade size. The response variable high volatility (y_i) is $y_i = 1$ if local volatility exceeds the threshold (k), and $y_i = 0$ otherwise. Estimation is based on Newton's method and only for market buy-side. The predictor is the log-differences of trade size ($SIZE$) and c denotes the constant. In model [1], high volatility threshold k is the median of sample volatility. In other specifications (from [2] to [5]), the threshold is the k standard deviation above the mean value (i.e., $k = 1, 3, 6, 9$), respectively. The bottom four rows in the table report the number of observations (Obs), $\chi^2(2)$ with its rejection probability (in square brackets), log-likelihood value (LL) and pseudo (McFadden's) R^2 values calculated as 1-$[LL(full)/LL(baseline)]$. The sample covers the periods from October 1, 2003 to May 14, 2004. *, **, and *** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Table 7: Distributional characteristics of price-contingent trades and transactional prices

	ΔSLS	ΔP	ΔSLB	ΔP	ΔTPS	ΔP	ΔTPB	ΔP
Panel A. Large orders								
Obs.	20028	20028	17728	17728	14032	14032	9597	9597
Mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Std.Dev.	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Skewness	0.848	0.915	-2.953	-4.036	-4.958	-5.683	2.404	2.799
Excess Kurtosis	**48.231**	**53.973**	**180.920**	**59.250**	**62.554**	**72.046**	**52.566**	**55.701**
Minimum	-0.018	-0.019	-0.024	-0.019	-0.019	-0.019	-0.019	-0.019
Maximum	0.009	0.010	0.024	0.005	0.007	0.004	0.016	0.016
Median	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Panel B. Small orders								
Obs.	36980	36980	30599	30599	64381	64381	60154	60154
Mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Std.Dev.	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000
Skewness	0.983	1.259	-4.120	-4.796	-6.980	-10.498	1.385	2.758
Excess Kurtosis	**78.165**	**87.766**	**68.496**	**78.866**	**228.900**	**288.090**	**332.050**	**171.620**
Minimum	-0.018	-0.018	-0.018	-0.018	-0.019	-0.019	-0.021	-0.018
Maximum	0.007	0.010	0.005	0.004	0.015	0.004	0.020	0.009
Median	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: The table reports the sample moments computed on the log-differences of stop-loss and take-profit prices. Panels A and B report the moments for the sample of large orders and small orders, respectively. The order size cutoff is 5000. The sample covers the periods from October 1, 2003 to May 14, 2004. ΔSLS (ΔSLB): log-differences in stop-loss sell (buy) prices, ΔTPS (ΔTPB): log-differences in take-profit sell (buy) prices. The sampling frequency is tick-by-tick. "ΔP" further denotes the log-differences in transaction market closing prices corresponding to each transaction type (i.e., stop-loss or take-profit). All trades are closing values. Significant excess kurtosis values are reported in bold.

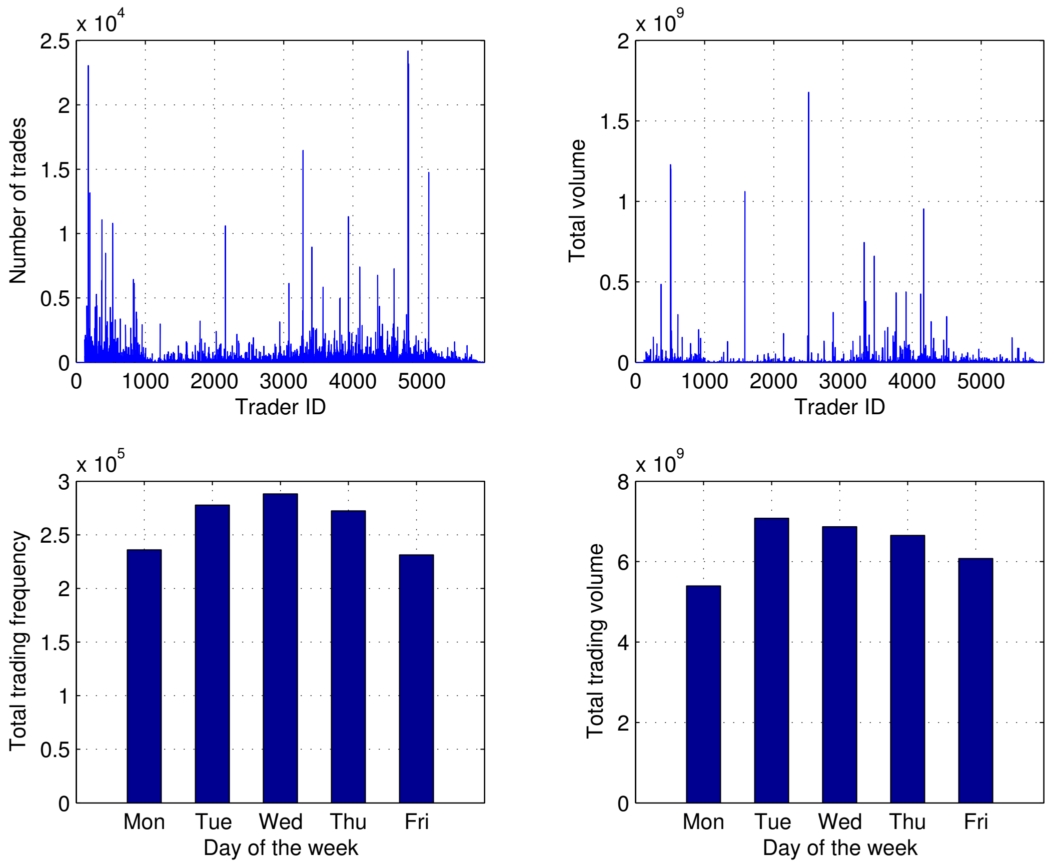

Figure 1: OANDA FXTrade trading. Notes: Top left: total trading frequency (number of trades) per trader (trader ID). Top right: total trading volume (in the units of base currency - EUR) per trader (trader ID). Bottom left: total trading frequency (number of trades) for each day of the week. Bottom right: total trading volume (in the units of base currency - EUR) for each day of the week.