Economics Letters 150 (2017) 27–33

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

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*

HIGHLIGHTS

- We investigate the informativeness of trade size in an electronic spot foreign exchange market.
- Large currency orders are likely placed by informed traders.
- Large trades are associated with increased exchange rate volatility.
- Small orders increase the likelihood of extreme events.
- Large orders from informed traders tend to be more concentrated.

ARTICLE INFO

Article history:
Received 12 September 2016
Received in revised form
1 November 2016
Accepted 9 November 2016
Available online 11 November 2016

JEL classification:
F3
F31
G0
G1

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

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 during increased price volatility episodes. 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.

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; 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

* Correspondence to: University of Guelph, College of Business and Economics, Department of Economics and Finance, 50 Stone Road East, Guelph, Ontario, N1G 2W1, Canada.

E-mail addresses: ngradoje@uoguelph.ca (N. Gradojevic), d.erdemlioglu@ieseg.fr (D. Erdemlioglu), rgencay@sfu.ca (R. Gençay).

[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.

http://dx.doi.org/10.1016/j.econlet.2016.11.010
0165-1765/

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 dataset, 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 on price volatility. This finding supports the *intraday trading invariance principle* proposed by Andersen et al. (2015). Finally, we assess the distributional characteristics of price increments 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.

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, 2002), 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.[3] 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.[4]

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

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

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

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

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

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

and in a good-event hour, this likelihood becomes

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

[2] Prior research on currency markets has focused on the link between expectations and shocks hitting currency markets (Evans, 2002; Evans and Lyons, 2002). While information-driven shocks change expectations and often increase market volatility (see, e.g., Jiang et al., 2011; Ederington and Lee, 1995; Ederington and Lee, 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).

[3] 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 or 10 min.

[4] 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.

Notably, 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. That is,

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

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

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

and the log likelihood function is

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

As in Easley et al. (2008), the log likelihood function, after dropping the constant and rearranging,[5] is given by

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

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 subsection, 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.[6] 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 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

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

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

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

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:

$$\begin{aligned} L((LB,LS,SB,SS)|\theta) &= (1-\alpha)L_n(.|\theta) + \alpha\delta L_b(.|\theta) \\ &\quad + \alpha(1-\delta)L_g(.|\theta). \end{aligned}$$

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,

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

and the log likelihood function is now

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

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.

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

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

3. Data

We obtain our dataset from the OANDA FXTrade internet trading platform, which 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 4983, and they mainly trade four major exchange rates.[7] As is standard in the literature, we eliminate weekends, 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. The final sample size is 3480 hourly data points covering 145 business days. In this range, 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.[8]

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. 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 5000, and it is subsequently shown that this does not affect the main results.[9]

Table 1 lists the average estimates of α_i, δ_i, ε_i, μ_i, ω_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.[10]However, the difference between the two sets of estimates of ω_i is statistically significant.[11] 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.

We now investigate whether the findings above are robust to the choice of the cutoff amount for a "large" trade. Table 2 reports the results for cutoff rates of 2000, 8000 and 12 000. The focus is on the difference column from Table 1 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 of cutoff values above 8000 (e.g., 12 000 in Table 2) 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 2) 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).[12] An interesting observation emerges from Table 2: 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 subsection 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 direction may also be considered. 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.[13]

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

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

---

7 By "active", the paper refers 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%.

8 This is one of the largest and most detailed tick-by-tick FX datasets ever to be used in an academic study. Please see (Gençay et al., 2015) for more information about the data.

9 Trade size is expressed in currency units of the base currency, i.e., the Euro.

10 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 $\bar{d}$.

11 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.

12 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.

13 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.

Table 1
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)***
$\overline{LLF}$	−15 111	−12 087.25	−12 087.46	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 ($\omega_i \neq \phi_i$; $i = 1, \ldots, 145$) 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.

* Indicate statistical significance at the 0.10 level.
** Indicate statistical significance at the 0.05 level.
*** Indicate statistical significance at the 0.01 level.

Table 2
The robustness of the estimates with respect to "large" trade size.

Parameter	2000 cutoff	8000 cutoff	12 000 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)	77.9 0.05 (0.19)
$\bar{\mu}$	74.54 0.42 (0.40)	83.11 1.34 (0.12)	82.6 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 12 000), 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 ($\omega_i \neq \phi_i$; $i = 1, \ldots, 145$). More precisely, each column represents the merged columns 3 and 4 from Table 1 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.

* Indicate statistical significance at the 0.10 level.
** Indicate statistical significance at the 0.05 level.
*** Indicate statistical significance at the 0.01 level.

Table 3
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.	190 195	232 661	10 659	10 307
$\chi^2(2)$	1498.30	715.00	170.21	55.54
	[0.00]	[0.00]	[0.00]	[0.00]
LL	−121 018	−148 503	−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.

* Indicate statistical significance at the 0.10 level.
** Indicate statistical significance at the 0.05 level.
*** Indicate statistical significance at the 0.01 level.

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 3 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).[14]

The third and fourth columns of Table 3 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 3). On

14 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.

Table 4
Distributional characteristics of price-contingent trades and transactional prices.

	ΔSLS	ΔP	ΔSLB	ΔP	ΔTPS	ΔP	ΔTPB	ΔP
Panel A. Large orders								
Obs.	20 028	20 028	17 728	17 728	14 032	14 032	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.	36 980	36 980	30 599	30 599	64 381	64 381	60 154	60 154
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. 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.

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 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).[15]

4.3. 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.

Table 4 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 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.

[15] The logit model results are robust to the choice of large trade size cutoff rates of 2000, 8000 and 12 000. These results can be available from the authors upon request.

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. Intuitively, informed retail FX traders attempt to camouflage their large trades during episodes of high volatility, where the potential impact of their (retail) 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).

References

Allen, F., Gale, D., 2000. Financial contagion. J. Polit. Econ. 108, 1–33.
Anand, A., Chakravarty, S., 2007. Stealth trading in options markets. J. Finan. Quant. Anal. 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.
Caballero, R.J., Krishnamurthy, A., 2008. Collective risk management in a flight to quality episode. J. Finance 63, 2195–2230.
Chakravarty, S., 2001. Stealth-trading: Which traders trades move stock prices? J. Financ. Econ. 61, 289–307.
Dungey, M., Milunovich, G., Thorp, S., 2010. Unobservable shocks as carriers of contagion. J. Bank. Finance 34, 1008–1021.

Easley, D., Engle, R., O'Hara, M., Wu, L., 2008. Time-varying arrival rates of informed and uninformed trades. J. Financ. Econom. 6, 171–207.
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. Financ. Stud. 10, 805–835.
Easley, D., O'Hara, M., 1987. Price, trade size, and information in securities markets. J. Financ. Econ. 19, 69–90.
Ederington, L.H., Lee, J.H., 1993. How markets process information: News releases and volatility. J. Finance 48, 1161–1191.
Ederington, L.H., Lee, J.H., 1995. The short-run dynamics of the price adjustment to new information. J. Finan. Quant. Anal. 30, 117–134.
Evans, M., 2002. FX trading and exchange rate dynamics. J. Finance 57, 2405–2447.
Evans, M., Lyons, R., 2002. Order flow and exchange rate dynamics. J. Polit. Econ. 110, 170–180.
Gençay, R., Gradojevic, N., 2013. Private information and its origins in an electronic foreign exchange market. Econ. Modell. 33, 86–93.
Gençay, R., Gradojevic, N., Olsen, R., Selçuk, F., 2015. Informed traders' arrival in foreign exchange markets: Does geography matter? Empir. Econom. 49, 1431–1462.
Gençay, R., Mahmoodzadeh, S., Rojcek, J., Tseng, M.C., 2016. Price impact of aggressive liquidity provision. Working Paper.
Jiang, G.J., Lo, I., Verdelhan, A., 2011. Information shocks, liquidity shocks, jumps, and price discovery: Evidence from the U.S. Treasury market. J. Finan. Quant. Anal. 46, 527–551.
Kalok, C., Fong, W.M., 2000. Trade size, order imbalance, and the volatility-volume relation. J. Financ. Econ. 57, 247–273.
Menkhoff, L., Schmeling, M., 2010. Whose trades convey information? Evidence from a cross-section of traders. J. Financ. Mark. 13, 101–128.
Osler, C., 2005. Stop-loss orders and price cascades in currency markets. J. Int. Money Finance 24, 219–241.
Osler, C., Savaser, T., 2011. Extreme returns: The case of currencies. J. Bank. Finance 35, 2868–2880.