HAL open science

Tick size reduction and price clustering in a FX order book.

Mehdi Lallouache, Frédéric Abergel

To cite this version:

Mehdi Lallouache, Frédéric Abergel. Tick size reduction and price clustering in a FX order book.. Physica A: Statistical Mechanics and its Applications, 2014, 416, pp.488-498. 10.1016/j.physa.2014.09.016 . hal-01006414

HAL Id: hal-01006414

https://hal.science/hal-01006414v1

Submitted on 16 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Empirical properties of the foreign exchange interdealer market

Mehdi Lallouache*[a] and Frédéric Abergel[a]

December 12, 2013

[a] *Chaire de Finance Quantitative, Laboratoire de Mathématiques Appliquées aux Systèmes Ecole Centrale Paris, Châtenay-Malabry, 92290, France*

Abstract

Using a new high frequency quality data set we provide a precise empirical study of the interdealer spot market. We check that the main stylized facts of financial time series are valid for the FX market: fat-tailed distribution of returns, aggregational normality and volatility clustering. We report two standard microstructure phenomena: microstructure noise effects in the signature plot and the Epps effect. We find an unusual shape for the average book, the spread distribution being bimodal. We construct the order flow and analyse its main characteristics: volume, placement, arrival intensity and sign. Many quantities have been dramatically affected by the decrease of the tick size in March 2011. We argue that the coexistence of manual traders and algorithmic traders, who react differently to the new tick size, leads to a strong price clustering in all types of orders and affects the price formation.

*Corresponding author. Email: mehdi.lallouache@ecp.fr

1 Introduction

Despite its macroeconomic role, the foreign exchange (FX) market is not well studied from an empirical point of view when compared to equity markets. This can be explained by the lack of quality data due to the over-the-counter[1] nature of this market. Fortunately, electronic trading has transformed the market over the past decades. This formerly opaque market is becoming fairly transparent and it is now possible to study its microstructure. Studies by **?**Ohnishi et al. (2008), Hashimoto et al. (2010) and Kozhan and Salmon (2012) are relevant examples of previous works using recently available high-frequency datasets . In this paper, we are only interested in the electronic interdealer market, where dealing bank traders (and also hedge funds via prime brokerage) conduct high volume transactions in order to liquidate unwanted accumulated inventory. For a detailed description of the foreign exchange market structure, see King et al. (2012). The interdealer market operates as a limit order market and it represents around one third of spot FX trades BIS (2010). Dealing banks quote a bid and ask price to their customers based on the interdealer market price. This market can be seen as the heart of the FX market; therefore, it is the most relevant for FX quantitative studies.

Using a better data set than any other in the existent literature, the aim of this paper is to do a precise empirical study of the interdealer spot market. First, we test the validity of the stylized facts concerning returns against our data. Then, we present order book and order flow statistics. When it is relevant, we highlight the changes triggered in those statistics by the tick decimalization of March 2011. Goldstein and A. Kavajecz (2000) analyzed similar tick size reduction in the equity context. Finally, we document a strong price clustering property due to this tick decimalization and we prove that this clustering stems from the fact that there are two types of traders: manual traders and automatic traders (computer algorithms), who react differently to decimalization.

2 Data

Most of spot interdealer trading occurs on two competing platforms: EBS Spot (Electronic Broking Service) and Reuters D-3000. Due to network externalities, liquidity naturally gravitated to just one platform for each currency. EBS has long dominated interdealer trading for the EUR, JPY, and CHF, while Reuters dominates the GBP, AUD, CAD, and the Scandinavian currencies. In this paper, we study two major currency pairs for which EBS is the leader: EUR/USD and USD/JPY. The following periods of historical data were bought from EBS: from August 1, 2010 to July 31, 2011 and from January 1, 2012 to March 31, 2012. The dataset contains a *Quote Record* and a *Deal Record* on a $0.1s$ time-slice basis. The Quote Record is a snapshot of the ten best levels of the book at the end of a time-slice (if a price or a volume in the book changed within the time-slice). The Deal Record lists the highest buying deal price and the lowest selling deal price (with the dealt volumes) during the time-slice. For the second time-period, we also know the total signed volume of trades in a time-slice. This is so far the best available data from EBS in terms of frequency (almost tick by tick) and depth (10 levels). In March 2011, EBS decided to set in a tick *decimalization*. The EUR/USD tick size changed from 10^{-4} to 10^{-5} and the USD/JPY one from 10^{-2} to 10^{-3}. In FX terminology, the market went from *pip-pricing* to *decimal pricing*. This structural change led to major consequences in many market variables (spread, order book shape, orders placement...) which are analysed throughout the paper.

3 Stylized facts

Statistical regularities in the price series have been noticed in financial data in a wide range of different markets. In this section, we examine if the EBS market follows these facts.

[1] Unregulated and decentralized market.

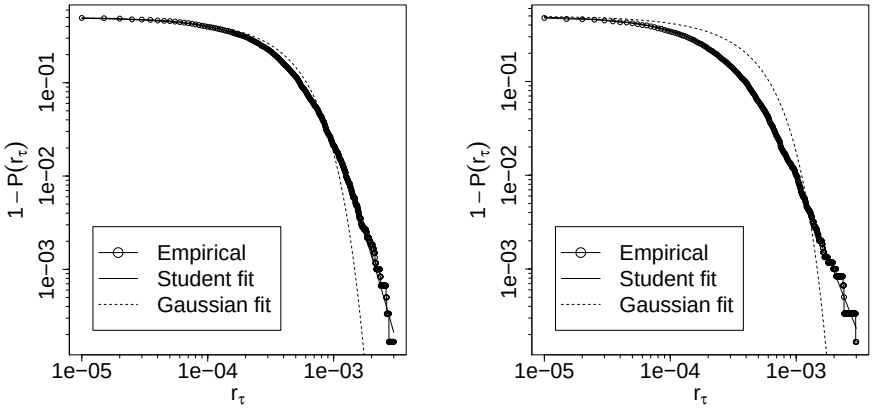

Figure 1: (Left) Empirical cumulative distribution of the EUR/USD 5-minute log-returns between January and March 2012. (Right) Same distribution for USD/JPY. We observe the same shape for negative returns.

3.1 Returns distribution

3.1.1 Fat-tails

Let p_t be the price of a financial asset at time t. We define its log-return over a period of time τ to be:

$$r_\tau(t) = \log(p(t+\tau)) - \log(p(t)) \tag{1}$$

Perhaps the most important stylized fact, first observed by Mandelbrot (1963), is the fact that the empirical distributions of financial returns and log-returns are fat-tailed. There is however no consensus on the exact form of the distribution, which varies with the timescale. The mechanisms behind FX rates tail events are studied by Osler and Savaser (2011), they have found that price-contingent trading may be a major source of extreme returns. Figure 1 shows the empirical cumulative distribution of the EUR/USD and USD/JPY positive log-returns. Log-returns are useful for FX rates because they can be inverted easily ($r_\tau(USD/JPY) = -r_\tau(JPY/USD)$) and one can then directly compare different pairs and different tick size periods. We compute the log-returns by sampling (every 5 mins) our high-frequency dataset from 8 a.m. to 6 p.m. from 1 January 2012 to 31 March 2012 (except weekends and holidays). The distributions are clearly not-gaussian and fat-tailed as expected. A Student fit is quite satisfactory. The tail exponent (obtained through Student distribution maximum likelihood estimation) is roughly equal to 5 for EUR/USD and to 3.5 for USD/JPY. The exponents are fairly stable through time.

3.1.2 Aggregational normality

It has been observed by Gopikrishnan et al. (1999) that as τ increases, the distribution of returns becomes better approximated by a normal distribution. This *aggregational normality* is easier to see using trade time (Huth and Abergel (2012)). Once again, we compute the log-returns distribution but this time we sample our data every 1000 trades. Jarque-Bera and Shapiro normality tests were not able to reject the normal hypothesis. Figure 2 shows QQ-plots for visual confirmation.

3.2 Auto-correlation of returns

Another widely known empirical fact is that there is no evidence of a linear correlation between successive returns (except on very short timescales, see Chakraborti et al. (2011)). In figure 3, we plot the autocorrelation of log-returns. The autocorrelation function is not statistically different from zero.

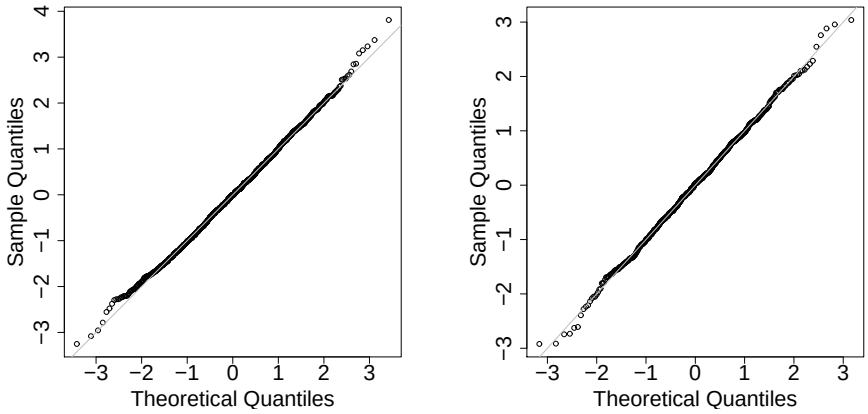

Figure 2: QQplots: Log-returns (sampled every 1000 trades) quantiles against normal quantiles. (Left) EUR/USD case. (Right) USD/JPY case.

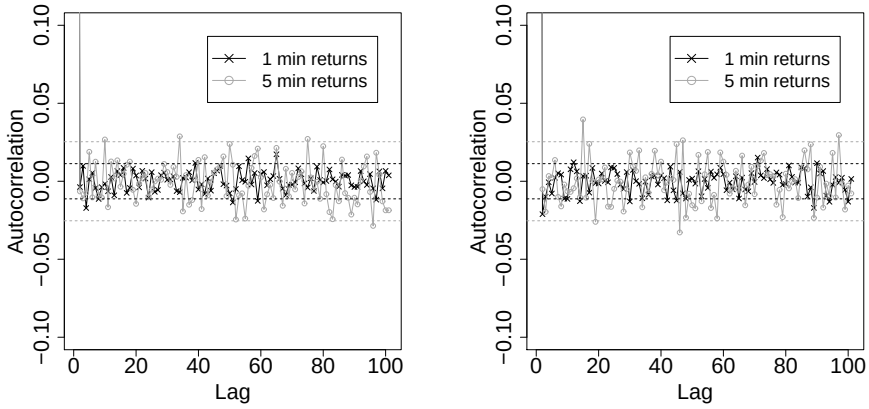

Figure 3: (Left) Autocorrelation function of the EUR/USD returns between January and March 2012. (Right) ACF of USD/JPY returns.

3.3 Volatility

In this section, we present some major empirical findings about volatility.

3.3.1 Volatility clustering

Time series of absolute (or square) mid-price returns has been found to display positive autocorrelation that is slowly decaying. This phenomenom is called *volatility clustering* . In the FX context, Berger et al. (2009) analyzed the factors driving the volatility persistence. Interestingly, they have shown that variations in market sensitivity to information play at least as large a role as do variations in the flow of information reaching the market through the trading process. Figure 4 shows the autocorrelation function of absolute returns. The function is statistically different from zero for a significant time horizon, demonstrating the existence of volatility clustering in our data. The effect is stronger for the USD/JPY case.

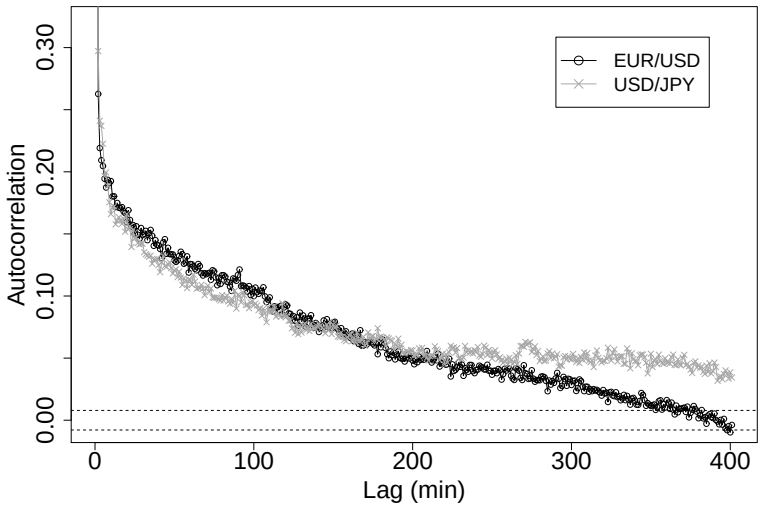

Figure 4: Autocorrelation function of 1-min absolute returns between January and February 2012.

3.3.2 Signature plot

If returns were i.i.d., the signature plot (realized volatility as a function of τ) would be a flat line, but Andersen et al. (2000) showed that financial returns realized volatility grows strongly as τ goes to zero due to microstructure noise. Figure 5 displays the EUR/USD signature plot with two different prices: last traded price and mid-price. We estimate the realized volatility using the two most active hours each day (14h-16h) and averaging over 65 trading days. Last traded price realized volatility is stronger than the mid-quotes one for small τ because of the bid-ask bounce effect. We notice two shape changes due to the tick size decimalization. First, the final value is reached faster before decimalization. Second, at high frequency before decimalization, mid-quote realized volatility explodes. This signals of the passage from a large tick asset to a small tick one. These two empirical facts are contradictory regarding the choice of an optimal tick size.

3.4 Correlation

The two currency pairs are correlated but the correlation vanishes at very high frequency (Epps effect). Figure 6 is a plot of the correlation between EUR/USD and JPY/USD as a function of τ. After the tick change, the correlation is lower and it takes longer to reach the final state.

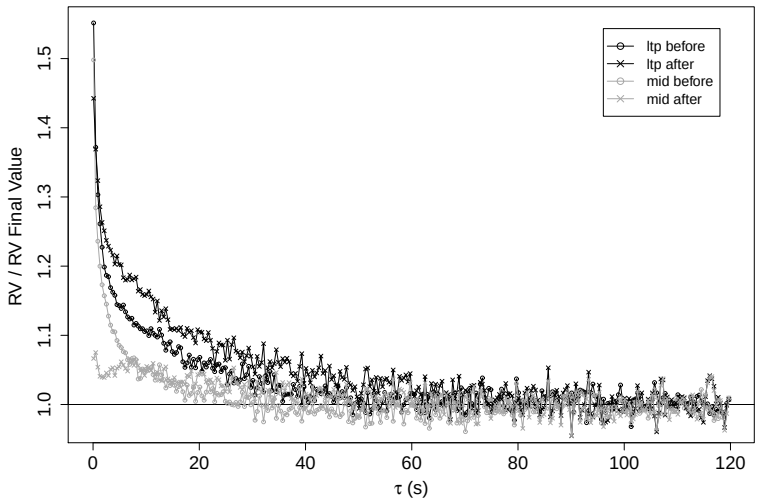

Figure 5: EUR/USD signature plot. Similar results for USD /JPY. Final value is reached faster before decimalization. The sampling period covers 65 trading days from January to March 2012.

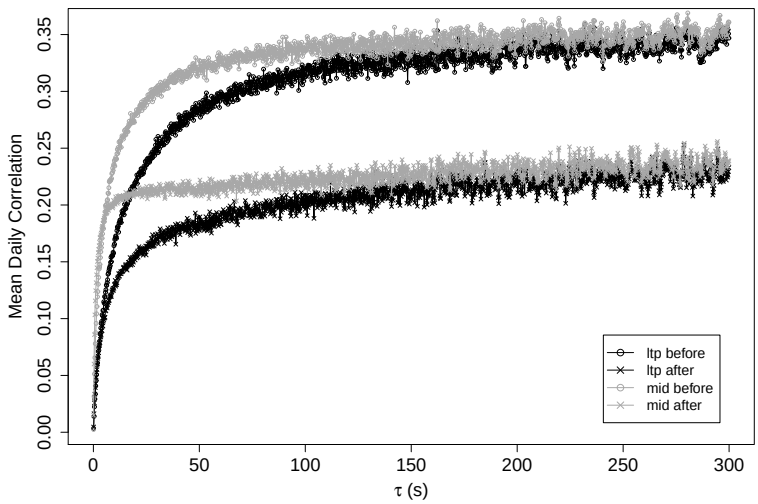

Figure 6: Correlation between EUR/USD and JPY/USD as a function of τ. Clear Epps effect. The sampling period covers 65 trading days from January to March 2012.

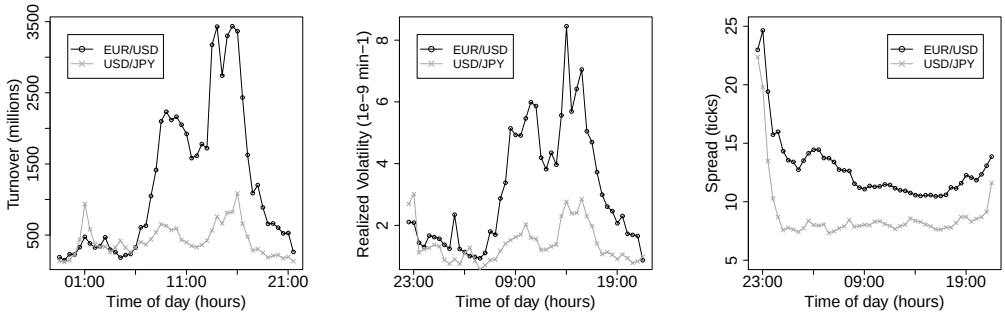

Figure 7: Intraday seasonalities. (Left) Average turnover in a 30-minutes interval. (Middle) Average realized volatility in a 30-minutes interval. (Right) Average spread in a 30-minutes interval.

4 Statistics of order books and order flow

In this section, we study various variables crucial to a good understanding of order book dynamics: placement of orders, size of orders, order book shape, etc. Thanks to our recent quality dataset, we provide a new empirical plot for each variable. Comparing these plots with previous studies on order books (Bouchaud et al., 2002, **?**),may reveal some distinctive features of the EBS market.

4.1 Intraday seasonality

EBS market operates 24/5 but the activity is of course not constant throughout the day. Figure 7 plots the average turnover and average spread in a 30-minute interval. We can clearly see a night and day pattern: from 8 a.m to 6 p.m. the volatility and turnover are much higher and the spread is lower than in the rest of the time; it corresponds to London working hours. The afternoon session shows a greater turnover (and volatility) than the morning session because it overlaps with New-York trading. Such global intra-day variation of market activity is a well-known fact on equities (see Chakraborti et al. (2011) and references therein) and it had been studied on the EBS market by Ito and Hashimoto (2006).

4.2 Order reconstruction

As explained in section 2, our data contains order book snapshots and deal records but not limit orders and cancellations. The first challenge is to infer the events from these two sources. Perfect reconstruction is impossible: the 10-levels limitation and the $0.1s$ time-slicing imply loss of information. Nevertheless, in most cases nothing happen within a time-slice. For example, between 2 p.m. and 4 p.m. (two most active hours of the day) nothing change in the book in half of the time-slices. Moreover market participants also face this 0.1 s time-slicing so they can't act at a much higher frequency. We therefore capture most of the order book changes. To build the order flow, we first split the data according to their side (bid or ask) and then we go sequentially through the quotes and compare two subsequent snapshots. When there is no transaction between two subsequent snapshots, all the changes between the snapshots are easily explained by limit orders and cancellations. When there is a transaction within a time-slice, we face two cases. Case 1 (around 75 per cent of the cases): the total traded volume is equal to the reported trade[2] volume. In this case, we know that there is a unique trade in that timeslice (for the considered side), then we match the deal volume with the corresponding volume decrease between the two corresponding subsequent snapshots. The rest of the liquidity changes are explained by limit orders ans cancellation. Case 2: the total traded volume is greater than the reported trade volume. We proceed like the previous case for the reported trade and we randomly distribute the remaining dealt volume among the available prices (from best price to the reported trade price) . Again the remaining liquidity

[2]By reported trade, we mean the lowest selling deal or highest buying deal depending on the considered side, see section 2

	Limit orders	Cancellations	Trades
EUR/USD	220	200	25
USD/JPY	75	70	7

Table 1: Average number of events per day (in thousands).

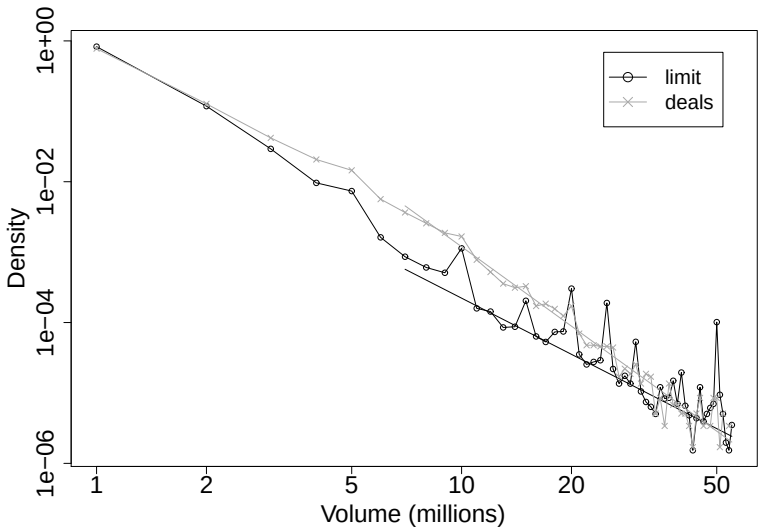

Figure 8: EUR/USD orders volume distribution in loglog scale. Similar results for USD/JPY. Sample period: march 2012.

changes are attributed to limit orders and cancellations. Some order of magnitudes obtained with this procedure are given in table 1 along with the number of trades for comparison.

4.3 Order characteristics

Thanks to our order reconstruction we can investigate the EBS order statistical properties. We have all the required data for 2012 only so we cannot look into the order characteristics before the decimalization.

4.3.1 Volume

In the EBS market, order size must be a multiple of 1 million (of the base currency). Figure 8 shows the distribution of EUR/USD order volumes (data from March 2012). The majority (around 80%) of the order sizes is at the minimal value (1 million euros). We observe a strong representation of limit orders and cancellations around 5, 10, 15, 20, 25, etc. Although weaker, the effect is also present for deals volumes. We observe roughly a power-law decay with exponent 2.6 for limit orders and cancellations and exponent 3.7 for deals. The exponents are stable through time and are a bit smaller for USD/JPY (2.3 and 3.5, respectively).

4.3.2 Placement

It is very interesting to know how traders place their orders. When a limit order arrives we can compute its distance to the best quote. We plot the distribution of the following quantity in our data set, i.e. using only the first ten limits of the order book: $\Delta p = b_0(t-) - b(t)$ (resp. $a(t) - a_0(t-)$) if a bid (resp. ask) order arrives at price $b(t)$ (resp. $a(t)$), where $b_0(t-)$ (resp.$a_0(t-)$) is the best bid (resp. ask) before the arrival of this order. Results are plotted in figure 9.

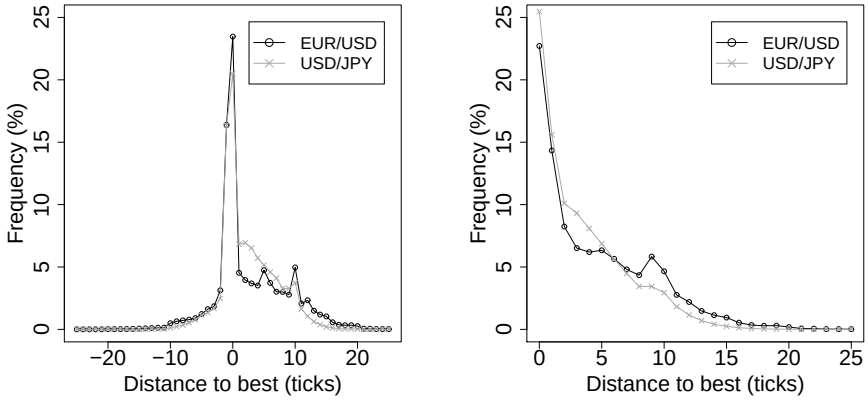

Figure 9: Placement of orders using the same best quote reference for February 2012. (Left) Limit orders. (Right) Cancellations. Ask side and bid side are similar.

	Limit Ask	Cancel Ask	Market Ask	Limit Bid	Cancel Bid	Market Bid
Limit Ask	1.000	0.837	0.561	0.563	0.750	0.716
Cancel Ask	0.837	1.000	0.602	0.751	0.716	0.571
Market Ask	0.561	0.602	1.000	0.726	0.585	0.644
Limit Bid	0.563	0.751	0.726	1.000	0.841	0.566
Cancel Bid	0.750	0.716	0.585	0.841	1.000	0.600
Market Bid	0.716	0.571	0.644	0.566	0.600	1.000

Table 2: Correlation matrix of the number of events time-series for EUR/USD. Similar values for USD/JPY.

These graphs being computed with incomplete data (ten best limits), we do not observe a placement as broad as in Bouchaud et al. (2002). We notice an asymmetry in the empirical distribution: the left side is less broad than the right side. Since the left side represent limit orders submitted *inside* the spread, this is expected. We observe the maximum at zero (best quote) and a high value at -1, which is the smallest price improvement. In the EUR/USD case, we see two clear peaks at 5 and 10, which corresponds to a half-pip and a pip distances. The origin of these peaks is discussed in section 5.

4.3.3 Arrival times

It is now clearly established that the Poisson hypothesis for the arrival times of orders is not empirically verified (see Chakraborti et al. (2011) and references therein). Our data are sampled, thus we can't study directly the inter-arrival times but we can compute the number of events in a 10 s window[3]. We consider 6 types of events : limit orders, market orders and cancellations on each side of the book. We suspect clustering and inter-dependence phenomena; to investigate this hypothesis, we compute autocorrelation and cross-correlations for time-series of the number of events. Figure 10 (left) shows that the autocorrelation is statistically different from 0 at several lags. The function seems to decrease as a power-law but we do not have enough data to see if this assumption holds on many decades. Table 2 is the correlation matrix for the six time-series. Significant correlations (as high as 0.84) are present, demonstrating the inter-dependence between order arrival processes. The independent Poisson processes hypothesis is clearly rejected. As suggested by Large (2007), Hawkes processes are better candidates for stream of orders modeling.

[3]We chose the window length as a trade-off between the number of windows (wanted) and the number of empty windows (unwanted).

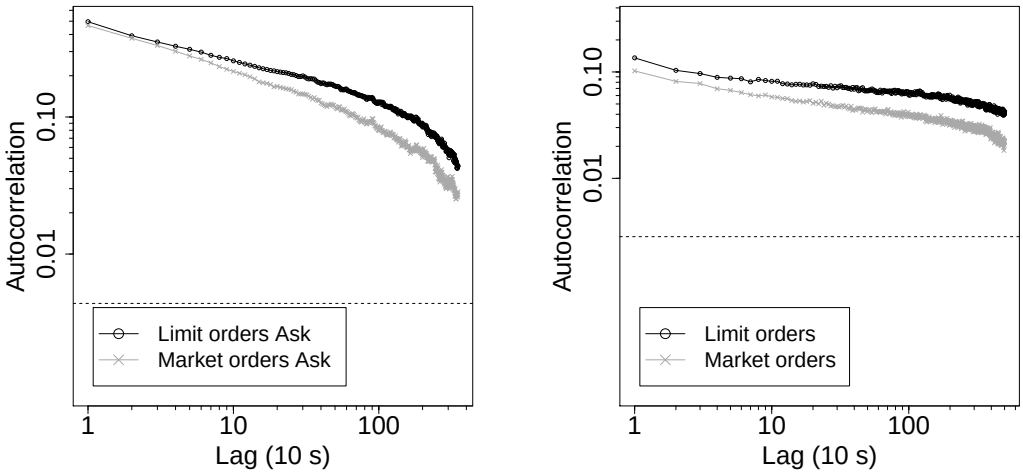

Figure 10: Autocorrelation function of the number of orders in a 10 s window. (Left) EUR/USD. (Right) Autocorrelation function of the mean orders sign in a 10 s window. Cancellations and limit orders are similar. The sampling period covers 65 trading days from January to March 2012.

4.3.4 Signs

Lillo and Farmer (2003) demonstrated that the signs of orders in the London Stock Exchange obey a long-memory process. In a similar fashion, we compute the mean sign (+1 for buy orders, −1 for sell orders) in a 10 s window for limit orders, cancellations and market orders and look at the autocorrelation function of these time-series. The results are plotted in figure 10 (right). We observe a slowly decaying function for limit orders and market orders. Again, we do not have enough data to observe the behavior on many decades.

After the review of the orders characteristics, we ask the question: How are these properties expressed in the shape of the order book ?

4.4 Average shape of the order book

A usual way to represent the shape of the book is to compute the time-averaged volume in the order book as a function of the distance from the current bid (or ask). Here, the shape is limited because we only have access to the ten first levels. In figure 11 we plot the shape before and after the decimalization. As already highlighted by other studies (see, e.g., Bouchaud et al. (2002)) the maximum is not located at the best quote. Before the decimalization, we can see the well-known hump-shaped curve with a maximum at two pips. After decimalization, the shape is unusual. The volume decreases after the best quotes and then increases to reach a maximum at 10 (which corresponds to one pip). For EUR/USD we can also see small peaks at 5,15 and 20 ticks. We investigate these peculiarities in section 5. The results are similar for the bid side and for other time period.

Another interesting quantity related to the shape of the book is the average gap, i.e. the price distance between two levels. We plot this quantity in figure 12 after decimalization. Before decimalization, the gap is almost always equal to one-tick. After decimalization, it decreases with the level. The results are similar for the ask side and do not change in time.

4.5 Spread

One of the most important quantity for traders is the *spread* (difference between best ask and best bid) because it measures the cost of making a transaction immediately (through a market order). Before decimalization, the spread was equal to one tick 65% of the time and to two ticks otherwise.

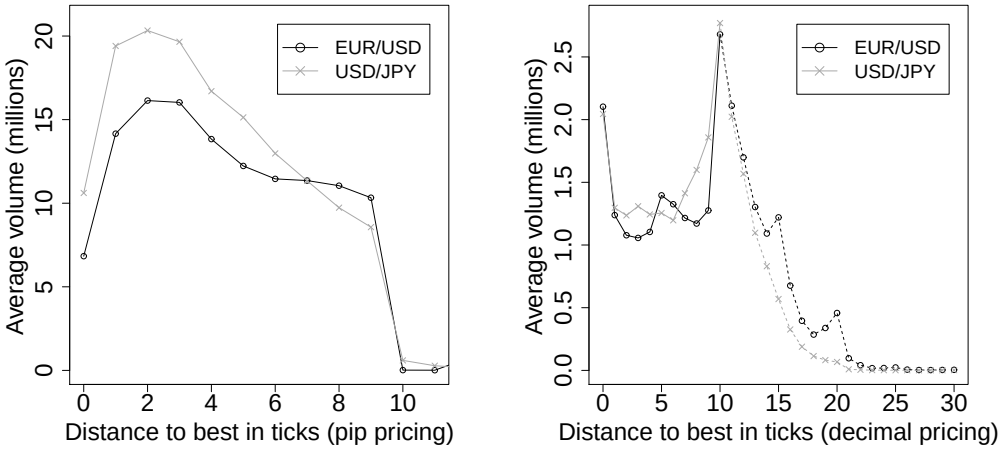

Figure 11: Average shape of the book ask-side. (Left) Before decimalization: February 2011. (Right) After decimalization: March 2012. The shape is exact up to 10 ticks

We want to know the extent of the impact of the decimalization on the spread. For this purpose, we compute the spread distribution using one-month data sampled every second from 8 a.m. to 6 p.m. The results are presented in figure 12. The spread can now take many values around the previous typical spread: 10. For the EUR/USD the distribution presents a bimodality. The first mode is at 9 and the second is at 13. The second mode position may vary with the time period considered, but we always have the first mode at 9 and the second one greater than 10. The results are similar for different sampling frequencies. The USD/JPY distribution seems more "natural" and the USD/JPY spread is smaller than the EUR/USD one. We explain the bimodality in section 5.5.

5 Price clustering

Using previously gathered empirical facts, especially the average shape of the book and the limit order placement in section 4, we can suspect a strong price clustering due to the tick decimalization. Price clustering is the tendency of prices to center around certain prices and to avoid others, it was recently studied by Cellier and Bourghelle (2009) on Euronext and Onnela et al. (2009) on the NYSE. For the FX spot market, the first clustering analysis was done by Goodhart and Curcio (1991) and, more recently, we can cite Sopranzetti and Datar (2002). However, these studies concerns time periods before the EMU and the electronic revolution, and use indicative quotes. For liquid currency pairs, if the tick size is appropriate, no clustering should occur. As we are interested in clustering due to the decimalization, we are going to focus on the price last digit.[4] A spurious bid/ask asymmetry may appear if one looks at the last digit directly for both sides. For example, lets suppose that the best quotes are on prices with a 0 last digit (they are called round or "integer" prices). In this situation, if a limit order improves the best quote by one tick at the bid, its last digit is 1, whereas if it is posted at the ask, its last digit is 9. In our view, it is the same situation and should give the same "decimal part". In the following, we will use the term last digit for both sides, but when it concerns a price on the ask side, it will actually designate the distance (number of ticks) to the smallest following "integer". In this section, we present results about trade price and order price clustering and, in the end, we argue that price clustering comes from the coexistence of two types of traders in the market with different abilities to make the most of the new tick size.

[4]We have checked that there is no clustering on the other digits. Of course the first digit seems clustered but it is just an intrinsic value of the exchange rate. It is very unlikely that EUR/USD rise above 2 for example.

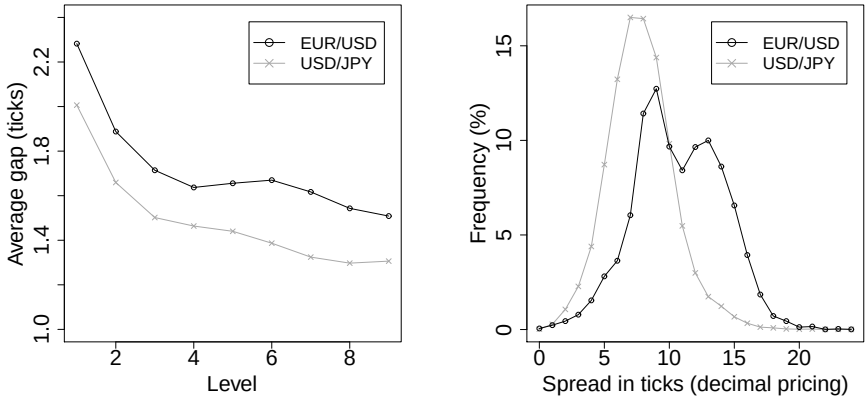

Figure 12: (Left) Average gap in the book bid-side. (Right) Spread distribution. Data: March 2012.

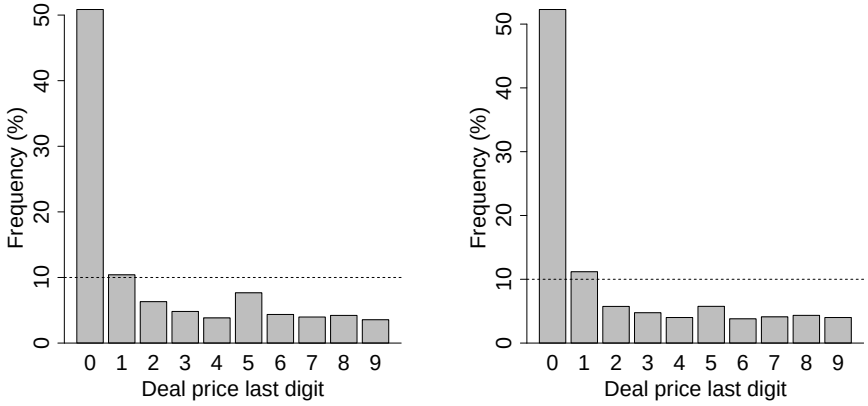

Figure 13: Trade prices (ask side) last digit distribution for March 2012. The dashed line represents the theoretical frequency under the uniform hypothesis. (Left) EUR/USD. (Right) USD/JPY. Around 50% of the trades occurs at integer prices. Same results for bid side.

5.1 Trade price

Under active trading it is natural to assume that realized trades should not cluster at certain prices and therefore, the distribution of the last digit of price should be uniform. In figure 13, we see the last digit distribution for deal prices, and we observe an uneven use of price fractions allowed by the tick size. The distribution is very far from uniformity[5]. Deal prices with a 0 as last digit (integer prices) are over-represented. In other words, the old tick is somehow still present. We also checked that the distribution was uniform before the decimalization.

5.2 Limit orders

We now look at price clustering for limit orders. The last digit distribution for limit order prices is plotted in figure 14. Again, the fractions are not equally frequent. The clustering is less pronounced than in trade prices but it is still strong. Around 20% of the orders are posted at integer prices and a half-integer peak is present in the EUR/USD case. The next prominent last digit is 1, this

[5]The clustering is so strong that we do not report statistical tests results. We performed χ^2 goodness of fit tests and they always reject the uniform distribution hypothesis. This is true for all the following subsections

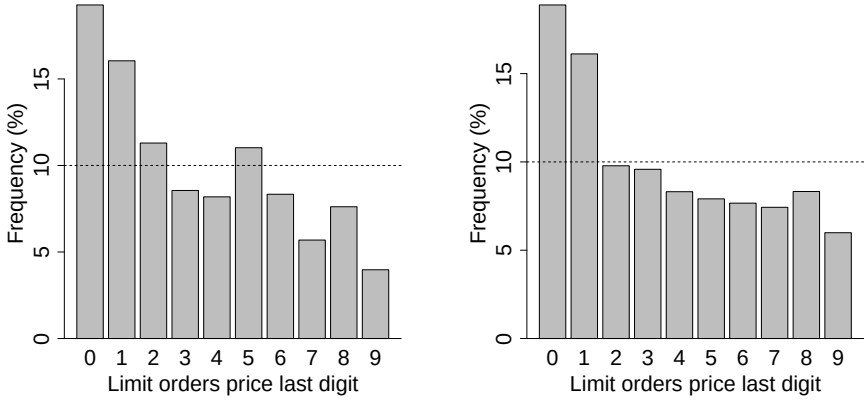

Figure 14: Distribution of last digits in limit order prices in March 2012. (Left) EUR/USD. (Right) USD/JPY. Around 20% of the orders are posted at integer prices.

is certainly related to the strategic behavior of some traders, who anticipate clustering tendencies and step-ahead round prices to obtain priority.

The limit order relative price distribution peaks or the average shape peaks (both at 5 and 10, see figure 9 and 11) might rise the question: is there also a round distance preference? The answer is negative, these peaks come from the price clustering. We verified this by computing the aforementioned distributions conditionally on the best quote last digit. The peaks positions change in a way that favors round prices and not round distances. The depth accumulation at 5 and 10 in the average shape comes directly from price clustering.

5.3 Best quote - Price Barriers

Limit orders have most of the time a size of 1 million (figure 8), then a clustering in terms of number of orders is also a clustering in terms of volume. Consequently, price clustering generates depth accumulation at round prices, affecting the best quote dynamics. Since more volume will be necessary to push the price through integers and halves, round best quotes may constitute "price barriers". The fact that there are more transactions at round prices may offset the depth accumulation. We show that this is not the case by recording the best quote last digit every 0.1 s and plotting its distribution in figure 15. A congestion effect is present, since more time is spent on round prices. This is an important point since a change in a microstructure parameter affected the price formation process in contradiction with classic financial theory.

5.4 Two types of traders

The price clustering can be explained by the presence of two types of traders: manual traders and automatic traders (computers algorithms), who reacted differently to the decimalization. According to King et al. (2012) there is now a 50/50 split (in orders volume) between algorithmic traders and manual traders with a keypad. Automatic traders adapted quickly to the new tick (just an algorithm adjustment) whereas manual traders refused to use the new tick [6]. We present here two arguments which corroborate this explanation.

Firstly, clustering effects are expected just after the decimalization, but they should decrease regularly as traders get used to the new tick. But for EBS the clustering is strong and stabilized in time. Therefore, part of the traders took into account the new situation whereas the others did not and will not. Secondly, it is interesting to look at the order volume depending on their last

[6] From discussions with people working in major banks, we know that manual traders do not care about price improvement to that fifth decimal point if they are trying to trade in large sizes. Beside, they have been used to the pip-pricing for many years and do not want to adapt.

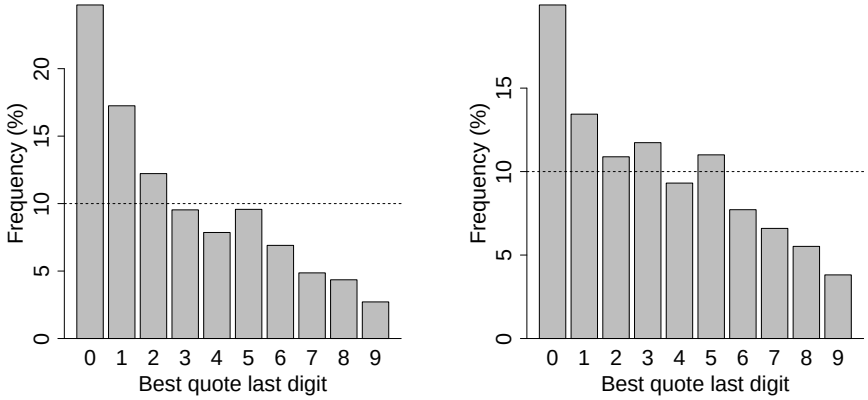

Figure 15: Best quote last digit distribution for January 2012 (1 s sampling). (Left) EUR/USD. (Right) USD/JPY. The best price spend most of the time on integers. Similar results with bid side and ask side.

digit value. If our analysis is true, round price orders should have a broader volume distribution because manual traders tend to post big orders compared to automatic traders. Figure 16 plots the limit order volume distribution but first we have separated the data into two samples: orders with integer price and orders with non-integer price. For integer orders the distribution is broad (power-law) and there are peaks on big round volumes (5,10,15,20 millions...) which is a trace of manual trading. On the other hand, for decimal orders the distribution is exponentially decreasing, typical of algorithmic trading.

5.5 EUR/USD post-decimalization spread and clustering

The EUR/USD spread bimodality (section 4.5) is intriguing and deserves more attention. We argue that the bimodality is a consequence of price clustering and notably the half-integer preference (which is not present for USD/JPY). To understand qualitatively the spread distribution shape, it is important to notice that the integer preference (or, equivalently, manual traders behavior) leads to a "natural" spread value of 10, when the best bid and the best ask are on integers (it corresponds to the pre-decimalization minimal spread). Traders who anticipate clustering tendencies can easily obtain priority by posting limit orders just above the best bid or just below the best ask[7]. More precisely, some traders try to genuinely take priority in order to deal at better price (the proportion of deals with last digit 1 is above 10%) while others are just using flash trading[8]. This is shown by the -1 strong value in the limit orders placement distribution (figure 9) and the prominence of 1 in the distribution of last digits in limit order prices (figure 14). The existence of congestion on integer values combined with agents "step-ahead" strategy explains the first peak at 9 (and the strong value at 8). To confirm the analysis, one can look at the quotes when the spread equals 8 or 9. From the different ways that the spread could be equal to 8 or 9 (in terms of ask/bid last digit) only one or two are really present (see figure 17). For example, when the spread is equal to 9, the configuration with one of the best quote lying on an integer value corresponds to 74% of the time.

Usual spread distributions are unimodal, but in figure 12 we can see a second peak with a local maximum at 13. This hump arise because of the combination of clustering at integer value (for

[7]We can also add that of course traders are reluctant to post orders just behind clusters, it would imply losing price priority over the large depth available at integer prices.

[8]Flash traders send an order at the top of the book (new best price) followed by a cancellation to lure the other book observers. The goal is to make people believe that there is a bid (during the 250 ms minimum quote life) at a certain level to trigger sales at this level. Those who give themselves a small margin to be sure of making the sale by showing a lower price to bid will automatically be executed with lower bids leaved in the book by the flash trader.

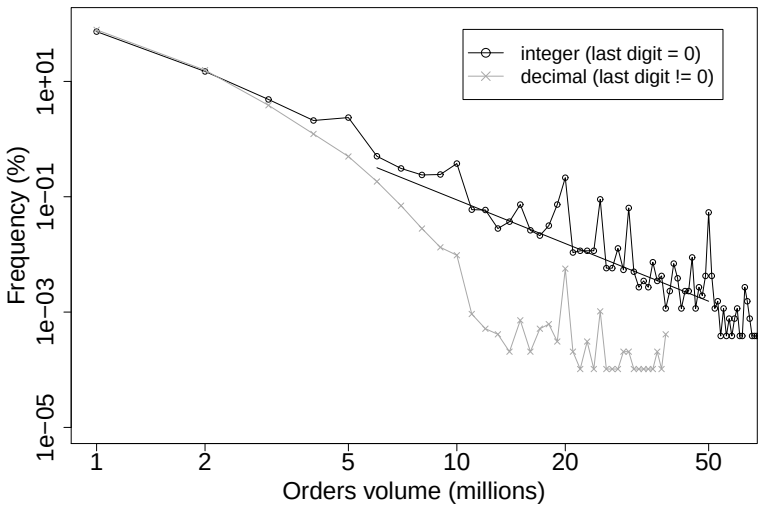

Figure 16: Limits orders volume distribution. The distribution depends on the order price last digit. Big volumes for integer orders with a 2.6 power-law exponent. Same results for both pairs.

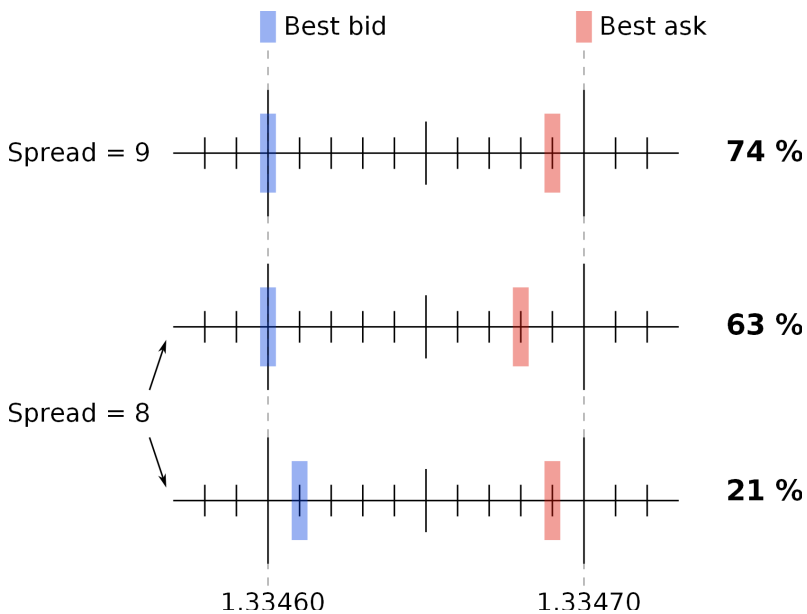

Figure 17: Small spreads typology. Price clustering leads to a drastic limitation of spread configurations in terms of bid and ask last digit. We do not draw symmetric configurations (ask on integer instead of bid) but we count them. If the configurations were equally likely one would expect 20% per configuration for 9 (odd spread : 5 possible configurations) and 16.67% per configuration for 8 (even spread : 6 possible configurations), instead of the observed higher percentage.

bid or ask) combined with the smaller clustering at half-integer values (for the other side). Again, some traders take price priority by posting limit orders above or below price barriers, which favors spreads slightly below 15. This second peak in spread distribution is not as high as the first one because half-integer clustering is weaker than integer clustering and there is a natural tendency for the spread to revert to smaller values near 10[9].

6 Conclusion

The main contribution of this paper is to provide a proper high-frequency empirical study of the FX market that was lacking in the literature. This study was feasible thanks to a new high quality data set.

We have confirmed that the main stylized facts were also present in the FX interdealer market: fat-tailed distribution of returns, aggregational normality and volatility clustering. We have also observed two standard microstructure phenomena: microstructure noise effects in the signature plot and the Epps effect. We have found an unusual shape of the book and spread distribution. We have inferred and analyzed the order flow characteristics (volume, placement, intensity...). The most striking fact is that many quantities depends dramatically on the tick size.

One of the most interesting features of the EBS market is the coexistence of two types of traders (manual traders and computers) with different abilities to make the most of the new tick size. This mixture of behavior leads to a strong price clustering property in all types of orders and affects the price formation (price barriers).

All the changes triggered by the change in tick size lead naturally to the question of the optimal tick size for the EBS market. The pip-pricing is praised by traditional banks because their traders were used to this tick size, whereas hedge-funds prefer the decimal pricing in order to use high-frequency strategies. Each tick size has advantages and drawbacks. With the old tick size the book was displaying high liquidity at each level, a stable price but the spread was almost always equal to one, suggesting that a tick size reduction could reduce the spread. With decimal pricing the spread is smaller, the mid-quote signature plot is flattened but at the cost of a flickering price and a strong price clustering phenomenon. The USD/JPY situation is slighly better than the EUR/USD one: smaller spread and less pronounced price barriers. One possible explanation is that the USD/JPY relative tick size[10] is bigger than the EUR/USD one. All these facts suggest that the optimum tick size lie somewhere between pip and decimal. Interestingly in December 2012, during the writing of this paper, EBS decided to change again the tick size and to go for half-pip pricing. It would be very relevant to see if it improved market quality. Notably, we can hope for a reduction in price clustering without a major increase in spread. We leave this question for future studies.

Acknowledgments

We would like to thank Joao Batista, Damien Challet, Guillaume Pons, Aymen Jedidi and Ilija Zovko for helpful discussions and suggestions.

[9] Due to the existence of market making effects (also known as order book resiliency).

[10] By relative tick size we mean $\delta = \frac{tick}{price}$. We find $\delta_{eurusd} \simeq 7 \times 10^{-6}$ and $\delta_{usdjpy} \simeq 1.2 \times 10^{-5}$, so $\frac{\delta_{usdjpy}}{\delta_{eurusd}} \simeq 1.7$.

References

T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. Realized Volatility and Correlation. *Manuscript, Northwestern University, Duke University and University of Pennsylvania. Published in revised form as "Great Realizations," Risk, March 2000, 105-108.*, (March), 2000.

D. Berger, A. Chaboud, and E. Hjalmarsson. What drives volatility persistence in the foreign exchange market? *Journal of Financial Economics*, 94(2):192–213, Nov. 2009. ISSN 0304-405X. doi: http://dx.doi.org/10.1016/j.jfineco.2008.10.006. URL `http://www.sciencedirect.com/science/article/pii/S0304405X09001196`.

BIS. Triennial-Central Bank Survey-Report on global foreign exchange market activity in 2010. Technical Report December, 2010.

J.-P. Bouchaud, M. Mézard, and M. Potters. Statistical properties of stock order books: empirical results and models. *Quantitative Finance*, 2(4):251–256, Aug. 2002. ISSN 1469-7688. doi: 10.1088/1469-7688/2/4/301. URL `http://dx.doi.org/10.1088/1469-7688/2/4/301`.

A. Cellier and D. Bourghelle. Limit Order Clustering and Price Barriers on Financial Markets: Empirical Evidence from Euronext. In *AFFI, Conférence internationale Brest*, Feb. 2009. doi: 10.2139/ssrn.966454. URL `http://papers.ssrn.com/abstract=966454`.

A. Chakraborti, I. M. Toke, M. Patriarca, and F. Abergel. Econophysics review: I. Empirical facts. *Quantitative Finance*, 11(7):991–1012, June 2011. ISSN 1469-7688. doi: 10.1080/14697688.2010.539248. URL `http://dx.doi.org/10.1080/14697688.2010.539248`.

M. A. Goldstein and K. A. Kavajecz. Eighths, sixteenths, and market depth: changes in tick size and liquidity provision on the NYSE. *Journal of Financial Economics*, 56(1):125–149, Apr. 2000. ISSN 0304-405X. doi: http://dx.doi.org/10.1016/S0304-405X(99)00061-6. URL `http://www.sciencedirect.com/science/article/pii/S0304405X99000616`.

C. Goodhart and R. Curcio. The clustering of bid-ask prices and the spread in the foreign exchange market. *LSE Financial Market Group Discussion Paper*, 110, 1991.

P. Gopikrishnan, V. Plerou, L. A. Nunes Amaral, M. Meyer, and H. E. Stanley. Scaling of the distribution of fluctuations of financial market indices. *Physical Review E*, 60(5):5305–5316, Nov. 1999. URL `http://link.aps.org/doi/10.1103/PhysRevE.60.5305`.

Y. Hashimoto, T. Ito, T. Ohnishi, M. Takayasu, H. Takayasu, and T. Watanabe. Random walk or a run. Market microstructure analysis of foreign exchange rate movements based on conditional probability. *Quantitative Finance*, 12(6):893–905, Dec. 2010. ISSN 1469-7688. doi: 10.1080/14697681003792237. URL `http://dx.doi.org/10.1080/14697681003792237`.

N. Huth and F. Abergel. The times change: multivariate subordination. Empirical facts. *Quantitative Finance*, 12(1):1–10, Jan. 2012. ISSN 1469-7688. URL `http://dx.doi.org/10.1080/14697688.2010.481635`.

T. Ito and Y. Hashimoto. Intraday seasonality in activities of the foreign exchange markets: Evidence from the electronic broking system. *Journal of the Japanese and International Economies*, 20(4):637–664, Dec. 2006. ISSN 0889-1583. URL `http://www.sciencedirect.com/science/article/pii/S0889158306000463`.

M. R. King, C. Osler, and D. Rime. Foreign Exchange Market Structure, Players, and Evolution. In *Handbook of Exchange Rates*, pages 1–44. John Wiley & Sons, Inc., 2012. ISBN 9781118445785. URL `http://dx.doi.org/10.1002/9781118445785.ch1`.

R. Kozhan and M. Salmon. The information content of a limit order book: The case of an FX market. *Journal of Financial Markets*, 15(1):1–28, Feb. 2012. ISSN 1386-4181. doi: http://dx.doi.org/10.1016/j.finmar.2011.07.002. URL `http://www.sciencedirect.com/science/article/pii/S1386418111000292`.

J. Large. Measuring the resiliency of an electronic limit order book. *Journal of Financial Markets*, 10(1):1–25, Feb. 2007. ISSN 1386-4181. URL http://www.sciencedirect.com/science/article/pii/S1386418106000528.

F. Lillo and J. D. Farmer. The long memory of the efficient market. *Studies in Nonlinear Dynamics Econometrics*, 8(3), 2003. URL http://www.mendeley.com/research/long-memory-efficient-market/.

B. Mandelbrot. The Variation of Certain Speculative Prices. *The Journal of Business*, 36(4): 394–419, Oct. 1963. ISSN 00219398. URL http://www.jstor.org/stable/2350970.

T. Ohnishi, H. Takayasu, T. Ito, Y. Hashimoto, T. Watanabe, and M. Takayasu. Dynamics of quote and deal prices in the foreign exchange market. *Journal of Economic Interaction and Coordination*, 3(1):99–106, 2008. ISSN 1860-711X. doi: 10.1007/s11403-008-0033-7. URL http://dx.doi.org/10.1007/s11403-008-0033-7.

J.-P. Onnela, J. Töyli, and K. Kaski. Tick size and stock returns. *Physica A: Statistical Mechanics and its Applications*, 388(4):441–454, Feb. 2009. ISSN 0378-4371. URL http://www.sciencedirect.com/science/article/pii/S0378437108008406.

C. Osler and T. Savaser. Extreme returns: The case of currencies. *Journal of Banking & Finance*, 35(11):2868–2880, Nov. 2011. ISSN 0378-4266. doi: http://dx.doi.org/10.1016/j.jbankfin.2011.03.016. URL http://www.sciencedirect.com/science/article/pii/S037842661100121X.

B. J. Sopranzetti and V. Datar. Price clustering in foreign exchange spot markets. *Journal of Financial Markets*, 5(4):411–417, Oct. 2002. ISSN 1386-4181. URL http://www.sciencedirect.com/science/article/pii/S1386418101000325.