Bid-ask spread, order size and volatility in the foreign exchange market: an empirical investigation

Saida Gtifa¹ and Naoufel LIOUANE²

¹ University of Science Economic and Management of Tunisia
²High Institute of Finance and Taxation, University of Sousse, Tunisia.

Accepted 18 November, 2013

This paper provides empirical evidence on the relationship between order size, volatility and spread in the foreign exchange market based on a FX dealer’s quotes. It uses a new data set that includes intra-daily data on trading volumes. The results are broadly consistent with the findings of the literature. It is found that spreads are independent of order size in both markets for the two currencies.

Keywords: bid-ask spread; trading volume; foreign exchange market.

JEL Classification: F31; G14

INTRODUCTION

This paper looks at the relationship between bid-ask spreads, order size and volatility in foreign exchange markets. Theoretical and empirical researches on the microstructure of foreign exchange markets have analysed the importance of this issue. From a policy perspective, the issue is important because of its affect for the analysis of risk liquidity. Indeed, a market can be considered to be liquid when large transactions can be executed with a small impact on prices (BIS (1999a)).

A number of studies on the microstructure of foreign exchange markets have looked at this issue from both a theoretical and an empirical point of view. Cost models show that order size and spread should be negatively related, while inventory holding risk and information cost models suggest that spread and order size are positively related. Lyons (1995) concluded that spread and order size are positively related, while Bjonnes and Rime (2005) showed that currency spread bears little relation to order size. Ding (2007) found that spreads are independent of order sizes in the inter-dealer market, but they are negatively correlated in the customer market.

This paper uses a new data set that matches intra-daily data on spread, trading volumes and volatility. The data were collected from a foreign exchange dealer who displayed both customer and inter-dealer bid-ask quotes. The data set covers the EUR/TND and USD/TND parities. The data cover the period from December 15, 2009 to January 15, 2010. The data were then tested by an econometric model to determine the relationship between spread, order size and volatility, revealing that spread is negatively related to order size in the inter-dealers and customer markets. I provide evidence of a positive correlation between spreads and volatility for inter-dealers and customer market for EUR/TND. However, the results do not show a significant impact of volatility on spreads for USD/TND.

The remainder of the paper is organized as follows: section 2 represent the literature review; section 3 summarizes the data base; section 4 examines the features of exchange rate and displays results; and section 5 concludes.

LITERATURE REVIEW

The literature on market microstructure is well developed as it pertains to the centralized market structure of NYSE specialists. In the earliest literature on spread determination, Demsetz (1968) presented the first formal
model for the stock market bid-ask spread. Finding that buy and sell orders generally do not reach the market at the same time, Demsetz assumed a separate class of market participants who provide immediacy by standing ready to buy and sell. There are two principal approaches to modeling market making behavior. First, inventory control models O'Hara and Oldfield (1986) consider the pricing problem faced by risk averse dealers to keep their inventories within bounds. Admeti and Pfleiderer (1988) focus on adverse selection problems when there are traders with heterogeneous information. Madhavan and Smidt (1991) develop theoretical model that incorporate both effects and test them using inventory and trade data. They found that price changes reflect significant information effects but weak inventory control effects.

Lyons (1995) is the first study that utilizes dealer intraday inventory and trade data in foreign exchange market. first, his dataset contain only dealers transactions during five trading days. while, Yao (1998) utilizes dealers and customer trades, he found that there is a positive relation between order flows and spread, he concluded that customer trades are the major source of private information in the FX.

Database
The data used in this paper were collected from a foreign exchange dealer. This dealer displayed both customer and inter-dealer quotes for several major currencies. I focused on the rates of the USD versus TND (USD/TND) and EUR versus TND (EUR/TND), currently the most frequently traded currencies pairs in Tunisia exchange market.

The market maker has a substantial customer order flow. He made 1820 and 1236 customer transactions respectively for the euro and dollar and 131 and 79 inter-dealers transactions. These observations were collected during the period of December 15, 2009 through January 15, 2010, with weekend days excluded

Features of exchange rates and estimation results
Given the mobility of the capital and the electronic trading system, competition among market makers becomes important. Indeed, quotes tend to be equal to market rate. We start with the first assumption of the model; we study the relationship between customer and inter-dealers average quotes. As preliminary tests, we examine the paired T-test for equal variances and its parametric test, the test of Wilcoxon Mann-Whitney (tables 1 and 2). This test used to verify if two independent samples result from the same population (homogeneity of the independent samples).

The results show that the values are not significant, the null assumption can not be rejected and the customer and inter-dealers average quotes are statistically equivalent.

To verify this hypothesis \((\alpha = 0 \text{ and } \beta = 1)\), we estimate the following regression:

\[
p_i^c = \alpha + \beta p_i^d + \epsilon_i
\]

(1)

Where \(p_i^c\), \(p_i^d\) are respectively the customer and inter-dealers average quotes. Then, we know that foreign exchange rate is first-order integrated. Indeed, the Dickey-Fuller and Philips-Peron tests confirm the same result (See Tables 3 and 4).

Compared to the data used in previous literature, the most important feature of this paper’s dataset that it contains orders size and matching spreads. Table 5 and 6 display the results of descriptive statistics. As shown in the tables the mean value of customer spread is significantly large than inter-dealers spread. In fact, the inter-dealers average spreads are approximately 0.0056 and 0.0029 respectively for EUR/TND and USD/TND, in the same way, the customer spreads are approximately 0.0249 and 0.0140 respectively for the two currencies. The explanation provided by microstructure theory is that bid-ask spreads are determined inter alia by inventory costs, which widen when exchange rate volatility increases. It is generally thought that such cost is significantly weak for dealers than customers due to electronic inter-dealers trading system.

The quotes have the advantage of announcing the heterogeneity of the operators and the direction in which each market maker wishes to operate. In other words, the market maker can anticipate competitive prices. Several authors show that the frequency of posted quotes constitutes a good indicator of the market activity, Danielsson and Payne (2002), see tables 5 and 6.

This paper estimates the following equation in first difference:

\[
p_{t+1}^c - p_t^c = \alpha + \beta (p_{t+1}^d - p_t^d) + \epsilon_{t+1}
\]

(2)

If customer quotes and inter-dealers quotes are equivalent, the null assumption: \(H_0: \alpha = 0 \text{ et } \beta = 1\) should not be rejected. According to tables 7 and 8, the coefficient \(\alpha\) is close to zero and the coefficient \(\beta\) is close to 1 for two parities. Indeed, both average quotes are equivalent; however, the spread of customer market is broader than the inter-dealers spread. Yao (1998), advance that the customer transactions are less transparent than inter-dealers transactions. In this respect, the market maker determines a large spread for the customers. Galati (2000), advance that the spread can measure the frequency of arrival of intra-day information.
Table 1. Test Wilcoxon Mann-Whitney for EUR/TND

<table>
<thead>
<tr>
<th>Comparison of average quotes</th>
<th>Two-sample Wilcoxon (Mann-Whitney) test</th>
</tr>
</thead>
<tbody>
<tr>
<td>A- comparison of average quotes</td>
<td>T=0.38, P-value=0.69, z=-0.78, P-value=0.53</td>
</tr>
<tr>
<td>B- Comparison of spreads</td>
<td>t=86.265, P-value=0.000, z=-46.566, P-value=0.000</td>
</tr>
</tbody>
</table>

A- customer average quote - inter-dealers average quotes: $H_0: \text{diff} = 0$, $H_A: \text{diff} \neq 0$

B- customer average spread - inter-dealers average spread: $H_0: \text{diff} = 0$, $H_A: \text{diff} > 0$

Table 2. Test Wilcoxon Mann-Whitney for USD/TND

<table>
<thead>
<tr>
<th>Comparison of average quotes</th>
<th>Two-sample Wilcoxon (Mann-Whitney) test</th>
</tr>
</thead>
<tbody>
<tr>
<td>A- comparison of average quotes</td>
<td>t=0.43, P-value=0.57, z=-0.73</td>
</tr>
<tr>
<td>B- Comparison of spreads</td>
<td>t=91.29, P-value=0.000, z=-42.12</td>
</tr>
</tbody>
</table>

A- customer average quote - inter-dealers average quotes: $H_0: \text{diff} = 0$, $H_A: \text{diff} \neq 0$

B- customer average spread - inter-dealers average spread: $H_0: \text{diff} = 0$, $H_A: \text{diff} > 0$

Table 3. Test ADF and PP for EUR/TND : I(1)

<table>
<thead>
<tr>
<th>Test ADF</th>
<th>Test Phillips-Perron</th>
</tr>
</thead>
</table>

Table 4. Test ADF and PP for USD/TND : I(1)

<table>
<thead>
<tr>
<th>Test ADF</th>
<th>Test Phillips-Perron</th>
</tr>
</thead>
</table>

Table 5. Descriptive statistics EUR/TND

<table>
<thead>
<tr>
<th>Variables</th>
<th>mean</th>
<th>Std.div</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>customer Average quotes</td>
<td>1.812954</td>
<td>0.020161</td>
<td>1.752083</td>
<td>1.854750</td>
</tr>
<tr>
<td>customer spread</td>
<td>0.024926</td>
<td>0.012449</td>
<td>-0.02912</td>
<td>0.086500</td>
</tr>
<tr>
<td>inter-dealers Average quotes</td>
<td>1.814268</td>
<td>0.022307</td>
<td>1.767900</td>
<td>1.864500</td>
</tr>
<tr>
<td>Inter-dealers spread</td>
<td>0.005653</td>
<td>0.006178</td>
<td>0.000500</td>
<td>0.032000</td>
</tr>
<tr>
<td>Trading volume</td>
<td>173650.0</td>
<td>670736.4</td>
<td>9565.240</td>
<td>16004514</td>
</tr>
</tbody>
</table>
Table 6. Descriptive statistics USD/TND

<table>
<thead>
<tr>
<th>Variables</th>
<th>Moyenne</th>
<th>Std. div</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>customer Average quotes</td>
<td>1.356771</td>
<td>0.036593</td>
<td>1.239630</td>
<td>1.412333</td>
</tr>
<tr>
<td>customer spread</td>
<td>0.026685</td>
<td>0.061453</td>
<td>-0.131259</td>
<td>1.34582</td>
</tr>
<tr>
<td>inter-dealers Average quotes</td>
<td>1.339323</td>
<td>0.039352</td>
<td>1.264100</td>
<td>1.415800</td>
</tr>
<tr>
<td>Inter-dealers spread</td>
<td>0.002905</td>
<td>0.000844</td>
<td>-0.004500</td>
<td>0.003000</td>
</tr>
<tr>
<td>Trading volume</td>
<td>353292.8</td>
<td>1118697</td>
<td>10760.540</td>
<td>13055623</td>
</tr>
</tbody>
</table>

Table 7. Relationship between customer and inter-dealers average quotes

\[ p_{t+1}^c - p_t^c = \alpha + \beta(p_{t+1}^d - p_t^d) + \epsilon_{t+1} \]

<table>
<thead>
<tr>
<th>EUR/TND</th>
<th>USD/TND</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>0.43 \times 10^{-06} (0.0324)</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.978 (1.079268)</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.0904</td>
</tr>
<tr>
<td>( F - statistique )</td>
<td>1.1648</td>
</tr>
<tr>
<td>( P \sim F )</td>
<td>0.18067</td>
</tr>
</tbody>
</table>

To test this assumption, we use two tests to compare the spreads: the paired T-test with unequal variances and the test Wilcoxon Mann Whitney. The preliminary results are reported in tables 1 and 2. Results show that the difference between customer and inter-dealers spreads is significantly positive.

According to the model, the spread can be written as follows:

\[ S_t^c - S_t^d = 2\sigma^c + 2(\beta^c - \beta^d)Q \tag{3} \]

Since the last equation describes a simple linear relationship, I just use OLS model to estimate the following equation:

\[ S_t^c - S_t^d = \alpha_0 + \alpha_1 Q_t + \epsilon_t \quad ; \quad \epsilon_t = N(0, \sigma^2) \tag{4} \]

where \( T \) is the sample size. \( Q \) is the logarithm of the order size, and the spreads in both markets are represented by the relative spreads, which are computed from customer and inter-dealer bid-ask prices as follows:

\[ S = \log(a) - \log(b) \tag{5} \]

where \( a \) and \( b \) are ask and bid prices respectively. Based on the analysis of previous section the customer spread is generally greater than the inter-dealer spread, but their differential decreases with a rise in order size, Ding (2009).

The estimation of the difference between two spreads show that the constant \( \alpha_0 \) is positive and significant for both currencies. Indeed, the customer spread is more important than inter-dealers spread. The coefficient reflecting the order size is negative for EUR/TND; this reflects that there exists a negative correlation between the spread and the order size. On the other hand, this result is not validated for USD/TND. Galati (2000) explains the negative relation between the order size and the price differential by that the sample period may be too short to allow for changes in foreign exchange market that lead to more efficient trade processing and higher competition among market-makers. In other words, the order size should be negatively correlated with spreads to the extent that they reflect economies of scale and are associated with higher competition among market-makers, Cornell (1978). Bollerslev and Domowitz (1993) show that the spread of the dollar-mark tend to increase during the lunch break in Tokyo exchange market. In European markets the spread remain very weak at midday, however the activity of the markets culminates. But, it increases until the close of North American markets.

Generally, we cannot release a systematic relationship between the market activity and spread. The conclusions are divergent; all depends on the sample period, the market and the importance of the market maker (See Table 8).

Other variables can affect the price differential more than the market activity. Ding (2009) show that the volatility of exchange rate and order size can affect the spread. This paper estimates the spread as dependent...
variable, while the volatility and order size are the independent variables (Bollerslev et al., 1994). Since many studies have demonstrated that the volatility of spot exchange rate can be modelled as a GARCH process, this paper estimates the volatility through a MA(1)-GARCH(1,1) specification:

\[
10,000 \Delta M_t = \mu + \theta \epsilon_{M,t-1} + \epsilon_{M,t},
\]
\[
\sigma^2_{M,t} = \sigma^2 + \alpha \epsilon^2_{M,t-1} + \beta \sigma^2_{M,t-1},
\]
\[
\eta_{M,t}/I_{t-1} = N(0, \sigma^2_{M,t}).
\]

Where \( M \) stands for the spot exchange rate and \( \Delta M \) is the change of the rate. \( I \) represents the information set, and \( \mu, \theta, \sigma, \alpha \) et are the parameters to be estimated. The time \( t \) subscript refers to the place in the order of the series of quotes, so that \( \hat{\sigma}^2_{M,t} \) provides an estimate of the exchange rate volatility. Since the magnitude of mid-quote fluctuations is very small, the exchange rate change \( \Delta M \) is multiplied by 10,000 to enlarge the effects of dependent variables so that estimated parameters will not be too small. In this estimation model, the observation of the exchange rate is measured as the logarithm of the mid-quote of bid-ask prices. Given that \( a_t \) and \( b_t \) denote ask and bid prices respectively, \( M \) is computed by the following formula:

\[
M_t = \log \left( \frac{a_t + b_t}{2} \right)
\]

Tables 9 and 10 report the results of the volatility estimation. The parameters are significantly different from zero and statistically significant for EUR/TND. This suggests that the volatility depend of all variables and that the volatility of spot exchange rates is fairly stable and does not possess the features of autocorrelation and heteroskedasticity. By contrast these results show that the volatility in time \( t \) is dependant for volatility in time \( t-1 \).

After obtaining the exchange rate volatility, a simple linear model was applied to estimate the impact of order size on spread:

\[
s_t = \gamma_0 + \gamma_1 \hat{\sigma}_t + \gamma_2 O_t + \epsilon_t; \quad \epsilon_t \sim N(0, \sigma^2_t), t = 1,2,...,T.
\]

Where \( s_t \) denotes the spread at time \( t \), \( O_t \) it's the order size and \( \hat{\sigma}_t \) denotes the exchange rate volatility obtained from GARCH estimation. \( \gamma_0, \gamma_1, \gamma_2 \) are constant parameters, and \( \epsilon_t \) is an error item following normal distribution with a zero mean. The variable \( O_t \) is taken as the logarithm of the original order sizes. The spreads of customer and inter-dealers markets are computed as the logarithm of the spread measured such that:

\[
s_t = \log [(a_t - b_t) \times 10000]
\]

The choice of estimation method is determined by the nature of the spread data. It is well known that high frequency financial data such as exchange rate spreads usually exhibit non-normality and high autocorrelation. Therefore, OLS or Maximum Likelihood Estimation might be inefficient to estimate the model. The statistical assumptions required by GMM for hypothesis testing are quite weak and neither autocorrelation of the data nor non-normality of the residuals jeopardizes its estimation (Ding, 2009).

\[
J(\hat{\gamma}) = M(\gamma)WM(\gamma)
\]
Table 9. Volatility estimation of EUR/TND

\[
10,000 \Delta M_t = \mu + \theta e_{M,t-1} + \epsilon_{M,t} \\
\sigma_{M,t}^2 = \varpi + \alpha \epsilon_{M,t-1}^2 + \beta \sigma_{M,t-1}^2, \\
\eta_{M,t} / I_{t-1} = N\left(0, \sigma_{M,t}^2\right)
\]

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>$\theta$</th>
<th>$\varpi$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.44E-05</td>
<td>0.6073</td>
<td>5.76E-07</td>
<td>0.9454</td>
<td>0.1073</td>
</tr>
<tr>
<td>(1.724)</td>
<td>(36.024)</td>
<td>(27.587)</td>
<td>(15.936)</td>
<td>(3.421)</td>
</tr>
</tbody>
</table>

Table 10. Volatility estimation of USD/TND

\[
10,000 \Delta M_t = \mu + \theta e_{M,t-1} + \epsilon_{M,t} \\
\sigma_{M,t}^2 = \varpi + \alpha \epsilon_{M,t-1}^2 + \beta \sigma_{M,t-1}^2, \\
\eta_{M,t} / I_{t-1} = N\left(0, \sigma_{M,t}^2\right)
\]

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>$\theta$</th>
<th>$\varpi$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0018</td>
<td>0.4423</td>
<td>-8.18E-09</td>
<td>2.341</td>
<td>0.345</td>
</tr>
<tr>
<td>(-129.92)</td>
<td>(10.138)</td>
<td>(-1.372)</td>
<td>(11.68)</td>
<td>(20.445)</td>
</tr>
</tbody>
</table>

Suppose that the conditions of moment are noted matrix $Z$, and it represents the vector of the residues, the conditions of moment of the linear model used can be written as follows:

\[
T^{-1}Z' \epsilon = 0 \tag{11}
\]

Substitute the equation of the conditions of moment in the equation (10), the objective function becomes:

\[
J(\hat{\phi}) = (T^{-1}Z' \epsilon)W(T^{-1}Z' \epsilon) \tag{12}
\]

The weighting matrix $W$ of GMM objective function determines the importance of the various conditions of moment. The weighting matrix $W$ can be estimated by several approaches that can account for various forms of heteroskedasticity and/or serial correlation. Meanwhile, instrument variables are chosen from the explanatory variables themselves. The first instrument matrix employs one period lag of explanatory variables (GMM1), while the other is a two period lag of explanatory variables (GMM2).

Tables 11 and 12 display the results of equation 8 respectively for tow parities and for both customer and inter-dealers spread. As shown in tables, the results suggest a significantly negative coefficient for order size for both markets. This means that the spreads quoted by the dealer are negatively related to the order sizes. In principle, this result is consistent with Ding (2009), who also found that larger order sizes lead to narrower spreads for both customer and inter-dealers markets. Osler et al. (2006) advance that larger order sizes lead to narrow spreads for both commercial and financial customers.

Jorion (1996) and Galati (2000) show that by split the data relative to trading volumes into expected and unexpected components. They found positive and significant coefficients on expected volumes. However, in contrast to the predictions of theory, they do not find positive and significant coefficients on unexpected volumes.

In regard to the impact of exchange rate volatility shown in tables 11 and 12, both GMM1 and GMM2 obtain positive results, the coefficient $\gamma_1$ is statistically different from zero in the inter-dealers market for the EUR/TND, and the values of t-statistics are around 2 it’s large enough to suggest that exchange rate volatility affects spreads significantly. For customer market the results are similar for GMM1, but it’s negative for GMM2 and the t-statistics are close to zero. In contrast, the results for USD/TND show that exchange rate volatility is negative related to spread in the customer market. For GMM1 the volatility affects positively the spread in the inter-dealers market, but the t-statistics are not significant. These results can be explained by the importance of the exchange with European market, since the European Union is the largest partner for Tunisia.
Table 11. Relationship between bid-ask spreads, order size and volatility of EUR/TND

<table>
<thead>
<tr>
<th></th>
<th>GMM1</th>
<th></th>
<th>GMM2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \gamma_0 )</td>
<td>( \gamma_1 )</td>
<td>( \gamma_2 )</td>
</tr>
<tr>
<td>Inter-dealers</td>
<td>0.0066</td>
<td>46.003</td>
<td>-2.9^{E-10}</td>
</tr>
<tr>
<td></td>
<td>(1.190)</td>
<td>(2.293)</td>
<td>(-0.530)</td>
</tr>
<tr>
<td>customer</td>
<td>0.0066</td>
<td>0.2052</td>
<td>-6.2^{E-9}</td>
</tr>
<tr>
<td></td>
<td>(12.08)</td>
<td>(0.049)</td>
<td>(-1.346)</td>
</tr>
</tbody>
</table>

Table 12. Relationship between bid-ask spreads, order size and volatility of USD/TND

<table>
<thead>
<tr>
<th></th>
<th>GMM1</th>
<th></th>
<th>GMM2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \gamma_0 )</td>
<td>( \gamma_1 )</td>
<td>( \gamma_2 )</td>
</tr>
<tr>
<td>Inter-dealers</td>
<td>0.0017</td>
<td>-1.221</td>
<td>-8.4^{E-10}</td>
</tr>
<tr>
<td></td>
<td>(1.717)</td>
<td>(-0.256)</td>
<td>(-0.860)</td>
</tr>
<tr>
<td>customer</td>
<td>0.014</td>
<td>-0.0023</td>
<td>-2.9^{E-8}</td>
</tr>
<tr>
<td></td>
<td>(0.198)</td>
<td>(-0.301)</td>
<td>(-0.093)</td>
</tr>
</tbody>
</table>

Conclusions

This article empirically examines the relationship between order size and spread in the foreign exchange market. Based on quotes from an individual FX dealer, spread is independent of order size in both inter-dealers and customer markets. I find that exchange rate volatility of EUR/TND affects spreads significantly. I also find evidence of a positive relation between spreads and volatility, as suggested by findings of the literature.

Although this paper focuses on the relationship between spread, order size and volatility in the foreign exchange market, it also reveals additional areas to be explored. First, this study focused on one individual dealer and two parities. Whether the conclusions apply to other dealers and other currencies deserves more research. Second, the dataset used in this paper is not large. More data need to be tested to verify whether this article’s finding is robust.

REFERENCES


BIS (1999a): “Market liquidity: research findings and selected policy implications”. Report for the CGFS.


Endnotes

I. See the econometric model of L.Ding (2009), Bid-ask spread and order size in the foreign exchange market: an empirical investigation.

II. Ding (2009) suggests that order sizes are much larger than spreads in terms of magnitude: estimated parameters would be very small if the model was estimated by the original data. To balance the magnitude of variables on both sides of the equation, spreads measured in pips are used in the estimation.
III.

1 See the econometric model of L. Ding (2009), Bid-ask spread and order size in the foreign exchange market: an empirical investigation.

2 Ding (2009) suggests that order sizes are much larger than spreads in terms of magnitude; estimated parameters would be very small if the model was estimated by the original data. To balance the magnitude of variables on both sides of the equation, spreads measured in pips are used in the estimation.

http://pointblanknews.com/pbn/articles-opinions/the-crimes-of-buhari/