

Bid–ask spreads and volatility in the foreign exchange market

An empirical analysis

Tim Bollerslev

Department of Finance, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208, USA

Michael Melvin*

Department of Economics, Arizona State University, Tempe, AZ 85287, USA

Received December 1992, revised version received June 1993

Consistent with the implications from a simple asymmetric information model for the bid–ask spread, we present empirical evidence that the size of the bid–ask spread in the foreign exchange market is positively related to the underlying exchange rate uncertainty. The estimation results are based on an ordered probit analysis that captures the discreteness in the spread distribution, with the uncertainty of the spot exchange rate being quantified through a GARCH type model. The data sets consist of more than 300,000 continuously recorded Deutschmark/dollar quotes over the period from April 1989 to June 1989.

Key words: Exchange rates; Market micro-structure

JEL classification: F3

1. Introduction

The bid–ask spread component of transactions costs in the foreign exchange market has received little attention in the literature. Early studies on the subject, such as Glassman (1987) and Boothe (1988), concentrated on the own statistical properties of the spread. Only recently have researchers attempted to offer empirical as well as theoretical analyses of the determinants of foreign exchange market spreads. These studies include Goodhart

Correspondence to: Michael Melvin, Department of Economics, Arizona State University, Tempe, AZ 85287, USA.

*The first author gratefully acknowledges the generous hospitality during a visit to Arizona State University, where some of the work was completed. Excellent research assistance was provided by Bettina Peiers. Peter Luan aided in adapting our computer programs to handle the more than 300,000 time-series observations utilized in this study. We thank Charles Goodhart for providing the data. Torben Anderson, Richard Brecher (the editor), Bettina Peiers, and two anonymous referees all made valuable comments on an earlier draft of the paper.

(1990), Bossaerts and Hillion (1991), Black (1991), Melvin and Tan (1992) and Bollerslev and Domowitz (1993). Yet no explicit analysis of the relationship between the magnitude of foreign exchange market spreads and the underlying exchange rate volatility is currently available. Intuitively, unambiguous 'good' or 'bad' news regarding the fundamentals of the exchange rate should have no systematic effect on the spread. Both the bid and the ask prices should adjust in the same direction in response to the traders receiving buy or sell orders that reflect the particular news event. Greater uncertainty regarding the future spot rate, as associated with greater volatility of the spot rate, is likely to result in a widening of the spread, however.

The next section provides a simple theoretical framework that illustrates this role of uncertainty, or volatility, in determining the spread. In section 3 we empirically test the implications of the theory using ordered probit analysis and GARCH estimates of the exchange rate volatility. The tests are based on a high-frequency data set that consists of every Deutschemark/dollar quote posted on the Reuters screen for the interbank foreign exchange market during a 3-month period in 1989. The total number of such quotes exceeds 300,000.¹ Section 4 summarizes and concludes the study.

2. A simple model of volatility and the bid-ask spread

The formal setup for our stylized market microstructure model is motivated by the analysis in Glosten and Milgrom (1985), Admati and Pfleiderer (1989), and Andersen (1993). We assume that the foreign exchange market is comprised of two kinds of traders: liquidity traders and information-based traders. Liquidity traders' transactions are driven by the needs for buying and selling goods and services and financial assets internationally. These traders do not speculate, but buy or sell currencies due to the financing needs of their normal business activity. Informed traders profit by intermediating the demands and supplies of foreign exchange for the liquidity traders. This group of traders also take positions in the foreign exchange market based on information advantages received through their dealings with the liquidity traders or, more generally, information asymmetries regarding fundamentals underlying the determination of the spot exchange rate.

The liquidity traders constitute the proportion $(1 - \lambda)$ of the total market participants. The liquidity traders receive a signal to either buy or sell foreign

¹The foreign exchange market is by far the largest financial market in the world. According to recent estimates, on a busy day more than \$600 billion worth of currencies are traded. The Deutschemark/dollar is the most actively traded spot rate, accounting for more than one-third of the total volume. Although the posted quotes on the Reuters screen are only indicative, with all the transactions done over the telephone, reputation effects preclude the posting of prices at which a bank would subsequently refuse to deal. For a more thorough description of the important institutional features characteristic to the foreign exchange market, see Goodhart (1988), Tygier (1988), and Lyons (1991).

currency regardless of the actual value of the currency in comparison with the bid or the ask prices prevailing at the time. Informed traders constitute the remaining λ proportion of the market.² This group of traders receives some information about the true underlying fundamental value of the exchange rate, say s_t . This fundamental value is assumed to evolve over time according to a martingale model,

$$s_t = s_{t-1} + \varepsilon_t, \quad (1)$$

where $E_{t-1}(\varepsilon_t) = 0$, $E_{t-1}(\varepsilon_t^2) \equiv \sigma_t^2$, and $E_{t-1}(\cdot)$ denotes the conditional expectation based on the information set generated by the past values of s_t . To simplify the analysis below we furthermore assume that the standardized innovations, $\varepsilon_t \sigma_t^{-1}$, are independent and symmetrically, but not necessarily identically, distributed through time.

At time $t-1$, one of the many market-making traders will set bid and ask quotes, B_t and A_t , good for trading at time t .³ The bid-ask spread is assumed to be set symmetrically around the known fundamental price prevailing at the time of quote formation, i.e. $A_t = s_{t-1} + k_{t,t-1}$ and $B_t = s_{t-1} - k_{t,t-1}$. Thus, the quoted spread for trades at time t , $K_t \equiv A_t - B_t = 2k_{t,t-1}$, depends on time $t-1$ information only.

Trades at existing quotes will generate losses, on average, for the market-maker when the party acting in response to the quote is an informed trader. Informed traders, who received the signal ε_t , buy currency if $A_t < s_t$ and sell currency if $s_t < B_t$. For values of $B_t \leq s_t \leq A_t$, the informed traders cannot profit from knowing the true fundamental value revealed by ε_t .⁴ The uninformed traders only know s_{t-1} , and expect ε_t to equal zero.

Trader positions are limited by the convention that existing quotes are only good for up to some maximum quantity of currency. Assuming that the market-makers limit trading to one unit of currency at existing quotes, the loss for the quoting trader relative to the true value s_t arising from informed trading is therefore

$$\pi_t^i = \min [s_t - B_t, 0, A_t - s_t] = \min [\varepsilon_t + k_{t,t-1}, 0, k_{t,t-1} - \varepsilon_t]. \quad (2)$$

Let $P_{t-1}(\cdot)$ denote the probability conditional on the time $t-1$ information.

²For simplicity we assume that the proportion of information-based traders is time invariant. The model could be generalized, at the expense of notational complexity, to accommodate endogenous information acquisition following the analysis in Admati and Pfleiderer (1989). The positive relationship between the spread and the uncertainty of the fundamental value remains intact under fairly general assumptions, however.

³In practice, the time between the quotes may depend on the information arrivals. We abstract from this complexity here. We also note that the model of trading activity in the foreign exchange market developed by Lyons (1991) ignores any effect of informational heterogeneity, or exchange rate volatility, in determining the size of the bid-ask spread.

⁴Assuming only partial information about ε_t , an expected positive profit condition could be invoked for small marginal risk-neutral trades, resulting in the same qualitative conclusions.

Since the standardized innovations, $Z_t \equiv \varepsilon_t \sigma_t^{-1}$, are assumed to be independent and symmetrically distributed through time, the expected loss from informed trading may be expressed as

$$\begin{aligned} E_{t-1}(\pi_t^i) &= E_{t-1}(\varepsilon_t + k_{t,t-1} \mid \varepsilon_t + k_{t,t-1} < 0) P_{t-1}(\varepsilon_t + k_{t,t-1} < 0) \\ &\quad + E_{t-1}(k_{t,t-1} - \varepsilon_t \mid k_{t,t-1} - \varepsilon_t < 0) P_{t-1}(k_{t,t-1} - \varepsilon_t < 0) \\ &= 2[k_{t,t-1} - E_{t-1}(\varepsilon_t \mid k_{t,t-1} < \varepsilon_t)] P_{t-1}(k_{t,t-1} < \varepsilon_t) \\ &= 2[k_{t,t-1} - \sigma_t E(Z_t \mid k_{t,t-1} \sigma_t^{-1} < Z_t)] [1 - P(Z_t < k_{t,t-1} \sigma_t^{-1})] < 0. \end{aligned} \quad (3)$$

Assuming an equal probability of a buy or a sell order from the liquidity traders, it follows that the expected profit for the quoting trader conditional on an uninformed trade equals:

$$E_{t-1}(\pi_t^u) = E_{t-1}(\frac{1}{2}(A_t - s_t) + \frac{1}{2}(s_t - B_t)) = k_{t,t-1} > 0. \quad (4)$$

Combining the expected trading loss in Eq. (3) with the gain in Eq. (4) yields the expected profit for the market-maker conditional on time $t-1$ information:

$$\begin{aligned} E_{t-1}(\pi_t) &= E_{t-1}(\pi_t^i + \pi_t^u) \\ &= 2\lambda[k_{t,t-1} - \sigma_t E(Z_t \mid Z_t > k_{t,t-1} \sigma_t^{-1})] [1 - P(Z_t < k_{t,t-1} \sigma_t^{-1})] \\ &\quad + (1 - \lambda)k_{t,t-1}. \end{aligned} \quad (5)$$

In equilibrium, competition from other banks or market-makers will drive this expected profit to zero.⁵ Expressing this zero profit condition in terms of the total spread, $K_t \equiv 2k_{t,t-1}$, yields

$$\begin{aligned} K_t &= \sigma_t 4\lambda E(Z_t \mid Z_t > k_{t,t-1} \sigma_t^{-1}) [1 - P(Z_t < k_{t,t-1} \sigma_t^{-1})] \\ &\quad \times [1 + \lambda - 2\lambda P(Z_t < k_{t,t-1} \sigma_t^{-1})]^{-1}. \end{aligned} \quad (6)$$

Since the conditional expectation and probabilities on the right-hand side of

⁵In practice the benefits from posting quotes and being perceived as an active market participant and a better provider of liquidity by third-party customers is likely to outweigh the direct costs of quote formation. Banks might therefore be willing to post quotes even though the direct expected profit from providing liquidity is zero in the absence of any superior information. Also, as noted by Lyons (1991), if a bank continuously refuses to quote bid and ask prices, or quote uncompetitive spreads, other traders might reduce their contact, thereby effectively excluding the bank from the network that makes up the market.

Eq. (6) only depend on the time $t-1$ information set through $\sigma_t^{-1}k_{t,t-1}$, it follows that in equilibrium the spread must move proportional to the conditional standard deviation of the true fundamental value of the exchange rate.⁶ Of course, in a more general model with endogenous information acquisition, this simple proportionality condition no longer holds true. The result that an increase in σ_t^2 leads to an increase in K_t remains generally valid, however. This relationship between exchange rate volatility and the spread is the focus of the empirical analysis to which we now turn.

3. Empirical evidence in volatility and the spread

Section 2 derives the theoretical proposition that greater exchange rate volatility is associated with a greater spread. Before describing the econometric methodology employed in testing this hypothesis, we briefly discuss our data.

3.1. Continuous Deutschemark/dollar spot rate quotations

The data set consists of every Deutschemark/dollar quote that appeared on the Reuters screen over the period from 9 April to 30 June 1989.⁷ During these 3 months a total of 305,604 quotes were posted on the screen by some 125 different banks. Each observation on a quote lists the time of day, the Reuters code for the name of the bank making the quote, the city where the bank is located, together with the bid and ask prices. To illustrate, consider the following four consecutive quotes for Monday 27 June 1989:

| | | | |
|--------|----------|------|-----------|
| 100248 | PRIVAT | COP | 1.9515/25 |
| 100300 | COMMERZB | H.K. | 1.9530/37 |
| 100306 | CHASE | FFT | 1.9530/40 |
| 100312 | ROY SCOT | LDN | 1.9529/34 |

The time of day is GMT, or Greenwich Mean Time; for the first observation it is 10:02:48. That is, 10:02 a.m. and 48 seconds. The second observation is at 10:03:00 a.m., or just 12 seconds later. The second field refers to the bank making the quote. The third field is the location of the quoting bank. The major banks participating in the interbank foreign exchange market have

⁶Given a particular distributional assumption for Z_t , it would be possible to solve for K_t as an implicit function of σ_t and λ .

⁷The data set was provided by Charles Goodhart. The same set of data has previously been analyzed by Goodhart (1990), Goodhart and Demos (1990), Goodhart and Figliuoli (1991) and Bollerslev and Domowitz (1993) in exploring other high-frequency intra-day aspects of the foreign exchange market. We note again that the quotes on the Reuters screen are only indicative of current prices and not actual transactions prices. The actual buying and selling is done through telephone conversations.

branches all over the world. The fourth field gives the bid and ask prices. Thus, at 10:02:48 Privat Banken (PRIVAT) in Copenhagen (COP) was willing to buy marks at 1.9515 per dollar, and sell marks at 1.9525 per dollar. For the four observations listed here the spreads are 0.0010, 0.0007, 0.0010, and 0.0005, respectively.

As documented by Goodhart and Demos (1990) and Bollerslev and Domowitz (1993), the number of quote arrivals have a very distinct time pattern across the hours of the day and the different days of the week. For the present analysis, we note the fact that Saturday and Sunday quote volume is sharply lower than weekdays. The number of quotes by the days of the week are: Monday, 60,095; Tuesday, 66,109; Wednesday, 63,812; Thursday, 61,521; Friday, 53,082; Saturday, 98; and Sunday 887. With the exception of lunchtime in the Far Eastern markets, on weekdays the interval between quotes is typically measured in seconds. On weekends, hours may elapse without any quote activity. Because of this lack of continuity in the weekend quotes, the subsequent estimation is done excluding all Saturday and Sunday observations.

The model in section 2 related changes in the spread to changes in exchange rate volatility. The bid and ask may move together, so that there is no change in the spread when trading is simply driven by 'good' or 'bad' news. For instance, if market-makers are receiving one-sided buy (sell) orders, they may increase (decrease) the bid and the ask by the same amount in response to the market pressure.⁸ If, however, there is more uncertainty regarding the future spot rate, our stylized theoretical model suggests that the spread should widen. The data indicate that for 8.0 percent of the observations, the bid and the ask move *up* together with no change in the spread. In 8.3 percent of the observations, the bid and the ask move *down* together resulting in no change in the spread. Thus, the majority, or 83.7 percent, of the observations actually involve a change in the spread compared with the previous pair of quotes.

In addition to examining the data across all quoting banks, the data set allows us to analyze the major market-makers individually. Table 1 lists the five banks with the most quotes posted. Not surprisingly, Deutsche Bank is the leading market-maker, with more than 30,000 quotes during the time under study. The next four banks are Morgan Guaranty, Societe Generale, Citibank, and Credit Suisse. These five banks account for 30.4 percent of the overall quote activity. The data on the frequency of change in the spread across across the large banks are very similar to the findings for the full data set. Only between 14 and 21 percent of the observations result in no change

⁸We focus here on variations in the quoted spreads as a function of volatility. Of course, imbalanced buy or sell orders could also affect the quoted spreads as market-makers try to encourage a more balanced order flow.

Table 1
Number of quotes, and frequency of no spread change.

| | Number | No change in spread | |
|-------------------|---------|---------------------|--------------|
| | | Bid-ask rise | Bid-ask fall |
| All quotes | 304,619 | 8.0 | 8.3 |
| Deutsche Bank | 30,901 | 7.3 | 7.2 |
| Morgan Guaranty | 23,210 | 9.7 | 10.3 |
| Societee Generale | 13,874 | 7.9 | 7.1 |
| Citibank | 13,665 | 6.5 | 7.5 |
| Credit Suisse | 11,267 | 9.8 | 11.2 |

in the spread from the same banks' most recent quote pair. The distribution of no change in the spread is roughly symmetric. In summary, the overwhelming majority of the observations do involve a shift in the spread.

3.2. GARCH estimation of volatility

The empirical investigation of the relationship between the spread and the associated exchange rate uncertainty requires an explicit proxy for the time-varying volatility of the spot rate. Numerous studies have demonstrated that exchange rate volatility may be conveniently modeled as a GARCH process.⁹ Following this literature we here employ a two-stage estimation procedure in which the conditional variance for the spot exchange rate is first estimated as a GARCH process. These estimates for the conditional variance are then used as our proxy for exchange rate volatility in the second-stage model for the temporal behavior of the spread. The implicit assumption is that the market-makers use GARCH models to forecast volatility.

We use the ask price for estimation purposes; the bid and ask prices have virtually identical higher order moments and differ only very slightly in their conditional means. The particular MA(1)-GARCH(1,1) model found to fit the data takes the form

$$\begin{aligned}
 10,000\Delta \log A_t &= \mu + \theta \varepsilon_{A,t-1} + \varepsilon_{A,t}, \\
 \sigma_{A,t}^2 &= \omega + \alpha \varepsilon_{A,t-1}^2 + \beta \sigma_{A,t-1}^2, \\
 \varepsilon_{A,t} | I_{t-1} &\sim N(0, \sigma_{A,t}^2),
 \end{aligned}
 \tag{7}$$

where I_{t-1} denotes the time $t-1$ information set, and μ , θ , ω , α , and β are the parameters to be estimated.¹⁰ The time t subscript refers to the place in

⁹See Bollerslev et al. (1992) for a recent survey of this extensive literature.

¹⁰Since the second to second movements in the ask prices are typically very small, we scaled the logarithmic first differences by 10,000 for numerical reasons. For a more complete description of the quasi-maximum likelihood estimation procedure employed see Bollerslev et al. (1992).

the order of the series of quotes, so that $\hat{\sigma}_{A,t}^2$ provides an estimate of the price volatility between quotes.

The particular specification for the conditional variance in Eq. (7) may be justified by the theoretical arguments in Nelson (1990, 1992). Intuitively, if the sample path for the true unobservable volatility process is continuous, it follows that on interpreting the GARCH(1,1) model as a non-parametric estimator, or a one-sided filter, the resulting estimates for the conditional variance will generally be consistent as the length of the sampling interval goes to zero.¹¹

Panel A of Table 2 reports the maximum likelihood estimates from the model in Eq. (7) for all of the quotes. In order to avoid any discontinuities associated with less reliable quotes originating during thin weekend trading, or potential big price jumps from Friday to Monday, the estimation was done for each of the 12 weeks in turn. Also, from a computational perspective the estimation of a GARCH model with more than 300,000 observations is not feasible in practice. Examining the table, it is obvious that the results are quite robust across the different weeks, although there appears to be a slight decline in the absolute magnitude of the estimated moving average coefficients from week 1 to week 7. The GARCH effect is highly significant, and the coefficient estimates are very similar across all of the weeks. Standard diagnostic test statistics also indicate that this relatively simple time-series model does a reasonably good job of tracking the own temporal dependencies in the conditional mean and variance of the continuously recorded spot exchange rate.¹²

In addition to the estimates for each of the weeks, panel B of Table 2 reports the estimates for the same MA(1)-GARCH(1,1) model for all the ask prices from the two banks with the most quotes over the sample period. Comparing these estimates for Deutsche Bank and Morgan Guaranty with the estimates for the different weeks, the GARCH coefficients, α and β , are seen to be quite similar. The most marked difference between the results in panels A and B concerns the magnitude of the negative moving average coefficient, θ . As illustrated by Bollerslev and Domowitz (1993), the time interval between quotes differs markedly across the trading day. The negative estimates for θ may therefore be partly attributed to a non-synchronous quoting phenomenon; see Lo and MacKinlay (1990) for a formal analysis. The more pronounced negative serial correlation for all the quotes reported in panel A when compared with the estimates in panel B may be due to

¹¹For all the data the average time interval between new quotes is only about 17 seconds. Zhou (1992) develops an alternative estimator for hourly volatility that exploits the 'continuous' nature of foreign exchange market quotation data.

¹²Even though some of the portmanteau tests for remaining residual autocorrelation in $\hat{\epsilon}_{A,t}/\hat{\sigma}_{A,t}$ and $\hat{\epsilon}_{A,t}^2/\hat{\sigma}_{A,t}^2$ are significant at the conventional 5 percent critical level, with 20,000 or more observations a careful balance between type I and II errors would dictate the use of much lower significance levels.

Table 2
GARCH estimates.

$$10,000\Delta \log A_t = \mu + \theta \varepsilon_{A,t-1} + \varepsilon_{A,t},$$

$$\sigma_{A,t}^2 = \omega + \alpha \varepsilon_{A,t-1}^2 + \beta \sigma_{A,t-1}^2, \quad t = 1, 2, \dots, T.$$

$$\varepsilon_{A,t} | I_{t-1} \sim N(0, \sigma_{A,t}^2),$$

Panel A: Weekly estimates

| Week | μ | θ | ω | α | β | T |
|------|-------------------|-------------------|------------------|------------------|------------------|--------|
| 1 | 0.006 (0.003) | -0.680 (0.005) | 0.049 (0.004) | 0.050 (0.002) | 0.933 (0.003) | 20,408 |
| 2 | 0.007 (0.003) | -0.686 (0.005) | 0.063 (0.003) | 0.050 (0.001) | 0.928 (0.002) | 25,018 |
| 3 | 0.016 (0.003) | -0.661 (0.005) | 0.187 (0.009) | 0.076 (0.002) | 0.858 (0.005) | 22,717 |
| 4 | 0.009 (0.004) | -0.638 (0.006) | 0.157 (0.008) | 0.095 (0.003) | 0.845 (0.005) | 18,382 |
| 5 | 0.010 (0.003) | -0.620 (0.005) | 0.058 (0.003) | 0.060 (0.002) | 0.919 (0.002) | 25,421 |
| 6 | 0.015 (0.004) | -0.572 (0.005) | 0.111 (0.004) | 0.078 (0.001) | 0.893 (0.002) | 26,283 |
| 7 | 0.002 (0.005) | -0.542 (0.005) | 0.131 (0.007) | 0.079 (0.002) | 0.897 (0.003) | 29,279 |
| 8 | -0.006 (0.005) | -0.547 (0.005) | 0.082 (0.005) | 0.062 (0.002) | 0.921 (0.003) | 27,580 |
| 9 | 0.010 (0.005) | -0.543 (0.005) | 0.128 (0.007) | 0.069 (0.002) | 0.904 (0.003) | 27,804 |
| 10 | 0.004 (0.005) | -0.550 (0.005) | 0.053 (0.003) | 0.049 (0.001) | 0.942 (0.003) | 29,350 |
| 11 | 0.001 (0.005) | -0.548 (0.005) | 0.058 (0.004) | 0.047 (0.001) | 0.941 (0.002) | 29,546 |
| 12 | 0.004 (0.005) | -0.555 (0.006) | 0.132 (0.006) | 0.067 (0.002) | 0.904 (0.003) | 22,820 |

Panel B: Individual banks

| Bank | μ | θ | ω | α | β | T |
|-----------------|------------------|-------------------|------------------|------------------|------------------|--------|
| Deutsche Bank | 0.023 (0.010) | -0.241 (0.005) | 0.129 (0.003) | 0.111 (0.001) | 0.897 (0.001) | 30,900 |
| Morgan Guaranty | 0.046 (0.013) | -0.090 (0.006) | 0.180 (0.004) | 0.119 (0.001) | 0.892 (0.001) | 23,209 |

Notes: The estimates in panel A for the different weeks are based on all the quotes for the particular week. The estimates in panel B are based on all the quotes for the two banks over the full 12-week period. Asymptotic standard errors are reported in parentheses.

differences in information, or the different interpretation of the same information, across traders. In the process of arriving at a common consensus regarding the true fundamental value, the quotes are likely to oscillate between different traders. This would lead to significantly higher negative serial correlation in the consecutive ask quotes from all the banks than for the individual bank's quotes. Another related explanation for the negative serial correlation may arise from the individual banks adjusting to order imbalances. For instance, when a bank has recently sold Deutsche marks, resulting in a long dollar position, it may try to offset that position by skewing the spread towards a more favorable bid price. Finally, we note that even in the absence of any differences in beliefs regarding the true fundamental value or any adjustments for order imbalances, if the average bid-ask spreads are systematically higher for some of the banks, the more pronounced negative serial correlation for all of the quotes may simply be the result of the ask quotes bouncing back and forth between the different banks.¹³

The primary purpose of the GARCH estimation was to create proxies for the conditional variance of the exchange rate to be used in the investigation of the determinants of the spread. To this end, we saved all the estimates for the conditional variances from the models in Table 2, with the individual weekly results combined into a single time series of volatility estimates for the full set of weekday quotes.

3.3. Frequency distribution of the spread

The choice of estimation strategy used in examining the statistical relationship between the spreads and the volatility estimates from the previous section is predicated on the nature of the spread data. The distribution of the foreign exchange market spreads clearly does not have continuous support. The difference between the bid-ask quote pairs tend to take on only a few distinct values. This feature of the quotes is obvious from a cursory glance at the frequency distribution given in Table 3, which lists the frequencies for the complete set of quotes and for Deutsche Bank and Morgan Guaranty separately. The raw spreads have been converted into basis points by multiplication with 10,000.

The 'normal' spread in the Deutschmark/dollar market is 10 basis points. In the whole data set, 170,892 observations, accounting for 56.1 percent of

¹³A highly significant first-order serial correlation coefficient has also been reported by Zhou (1992) for continuously recorded Deutschmark/dollar quotes for all of 1990. On temporally aggregating the quotes, Zhou (1992) finds that the magnitude of the serial correlation decreases with the sampling frequency, consistent with the continuous sample path generated by a Brownian motion contaminated by 'noise'. This also could explain the lower serial correlation for the individual banks, for which the average time interval between the quotes is longer.

Table 3
Frequency distribution of spreads.

| Spread | All quotes | Deutsche Bank | Morgan Guaranty |
|-------------|-----------------|----------------|-----------------|
| 0 < · < 5 | 2,304 (0.8%) | 17 (0.1%) | 21 (0.1%) |
| 5 | 77,856 (25.6%) | 5,887 (19.1%) | 7,423 (32.0%) |
| 5 < · < 7 | 607 (0.2%) | 28 (0.1%) | 6 (0.0%) |
| 7 | 34,878 (11.5%) | 1,633 (5.3%) | 1,713 (7.4%) |
| 7 < · < 10 | 2,977 (1.0%) | 65 (0.2%) | 16 (0.1%) |
| 10 | 170,892 (56.1%) | 22,174 (71.8%) | 13,091 (56.4%) |
| 10 < · < 15 | 329 (0.1%) | 206 (0.7%) | 11 (0.1%) |
| 15 | 11,534 (3.8%) | 702 (2.3%) | 777 (3.4%) |
| 15 < · < 20 | 39 (0.0%) | 2 (0.0%) | 2 (0.0%) |
| 20 | 2,616 (0.9%) | 152 (0.5%) | 126 (0.5%) |
| 20 < · | 572 (0.2%) | 29 (0.1%) | 23 (0.1%) |

Note: The spreads are converted into basis points.

the cases, took on a value of 10. This is almost exactly the same proportion as for the Morgan Guaranty quotes. For Deutsche Bank, 71.8 percent of their quotes took on a value of 10 basis points. The next most frequent spread was 5 basis points, followed by 7, and 15. These four spreads account for 97.0 percent of the totals in the entire data set, and 98.5 and 99.2 percent for Deutsche Bank and Morgan Guaranty, respectively. This lack of continuity suggests that pursuing stochastic processes with continuous state spaces will not represent the spread data very well. Addressing this issue of discreteness in U.S. stock market data, Hausman et al. (1992) used an ordered probit model in their analysis of continuously recorded transactions price changes.¹⁴ We adopt the same statistical procedure here.

3.4. Ordered probit analysis

Before turning to the results, we will briefly outline the ordered probit methodology employed in the estimation.¹⁵ Consistent with the discussion in the previous section, the observed spread, K_t , is assumed to take on only a fixed number of discrete values, a_1, a_2, \dots, a_J . The unobservable continuous random variable, K_t^* , is defined by

$$K_t^* = \delta_0 + \delta' X_t + \varepsilon_{K,t}. \quad (8)$$

The vector X_t denotes a set of predetermined variables that affect the

¹⁴As noted by a referee, for simultaneity reasons Hausman et al. (1992) do not include contemporaneous trading volume in their model of tick-by-tick stock price movements, thus effectively eliminating important information from the analysis. In the foreign exchange market, however, only the quotes as analyzed here are directly available.

¹⁵Our discussion is intuitive. For a more thorough review of the ordered probit methodology, see, for example, Maddala (1983) and Greene (1990).

conditional mean of K_t^* , and $\varepsilon_{K,t}$ is conditionally normally distributed with mean zero and variance, $\sigma_{K,t}^2$,

$$\varepsilon_{K,t} | I_{t-1} \sim N(0, \sigma_{K,t}^2). \quad (9)$$

To allow for conditional heteroskedasticity in the spread, we parameterize the logarithm of $\sigma_{K,t}^2$ as a linear function of the same explanatory variables that enter the conditional mean of K_t^* , i.e.

$$\sigma_{K,t}^2 = [\exp(\gamma' X_t)]^2. \quad (10)$$

As for the GARCH model in subsection 3.2, the time t subscript refers to the place in the order of the series of posted quotes, as opposed to clock time.

The ordered probit model relates the observed spreads to K_t^* via

$$K_t = a_j, \quad \text{iff } K_t^* \in A_j, \quad j = 1, 2, \dots, J, \quad (11)$$

where the A_j 's form an ordered partition of the real line into J disjoint intervals. The probability that the spread takes on the value a_j is equal to the probability that K_t^* falls into the appropriate partition, A_j . For tractability reasons, we base the empirical analysis on a classification of the spread into only four different categories. From the discussion in subsection 3.3 the four most commonly observed spreads account for 97.0 percent of the total quotes. In the categorization, the group a_1 is for small spreads less than or equal to 5 basis points; a_2 is for spreads greater than 5 but less than 10; a_3 denotes the 'normal' spread of 10 basis points; and a_4 is for large spreads greater than 10 basis points. The corresponding intervals for the unobservable latent variable K_t^* are defined by

$$\begin{aligned} A_1 &\equiv]\infty, \mu_0], \\ A_2 &\equiv]\mu_0, \mu_1], \\ A_3 &\equiv]\mu_1, \mu_2], \\ A_4 &\equiv]\mu_2, \infty[. \end{aligned} \quad (12)$$

The partition parameters, μ_i , are estimated jointly with the other parameters of the model.

The ordered probit model defined by Eqs. (8)–(12) allows us to estimate the probability of a particular spread being observed as a function of the predetermined variables, X_t . In order to test the hypothesis that the spread is partly determined by the volatility of the spot rate, the GARCH estimate of the conditional variance for the ask prices is included as one of the elements in X_t . Given the partition boundaries determined by the data, if a higher

conditional mean $\delta'X_t$, is caused by a larger conditional variance of the spot rate, and this raises the probability of observing a higher spread, we will infer that the hypothesized theoretical link is supported by the empirical analysis. Of course, to the extent that the estimated GARCH conditional variances only serve as a proxy for the true unobservable spot exchange rate volatilities used by agents, the results should be interpreted carefully. However, as long as $\hat{\sigma}_{A,t}^2$ is a consistent estimator, the second-stage ordered probit model generally yields consistent parameter estimates, although the associated standard errors may be downward biased compared with the results using the true unobservable $\sigma_{A,t}^2$ process. Pagan (1984) provides a formal econometric analysis of issues related to the use of generated regressors.

The evidence presented in Bollerslev and Domowitz (1993) indicates a distinct intra-day pattern in the spread distribution. It is possible that any significant effect of the conditional variance in isolation may merely reflect this dependence rather than provide an independent influence on the spread process. In order to take account of this own temporal dependence, we also include K_{t-1} as an element of the X_t vector in the joint estimation of the conditional mean and variance functions for K_t^* :

$$K_t^* = \delta_0 + \delta_1 \hat{\sigma}_{A,t}^2 + \delta_2 K_{t-1} + \varepsilon_{K,t},$$

$$\sigma_{K,t} = \exp(\gamma_1 \hat{\sigma}_{A,t}^2 + \gamma_2 K_{t-1}). \quad (13)$$

All of the variables are expressed in logarithmic form. Note that the conditional variance for the spot exchange rate, $\hat{\sigma}_{A,t}^2$, is predetermined and by definition only depends on information up through time $t-1$. In addition to the parameters in (13), we also estimated the threshold parameters, μ_1 and μ_2 . Following standard practice, and without loss of generality, the value for μ_0 was normalized to zero.¹⁶

3.4.1. Estimates for all quotes

Maximum likelihood estimates for the ordered probit model for the full data set are presented in the first column of Table 4. The boundaries for partitioning the data, μ_1 and μ_2 , are estimated with a high degree of precision as seen by the large t -statistics. The δ_1 coefficient suggests that there is a significantly positive effect of exchange rate volatility on the spread. The conditional mean of K_t^* is an increasing function of $\hat{\sigma}_{A,t}^2$. This is consistent with the implications drawn from the model of section 1, and in support of the main thesis of our study. The estimate for δ_2 is indicative of the strong intra-day persistence in the spread process; if the current quoted

¹⁶Maximum likelihood estimates were obtained using the LIMDEP computer program for an ordered probit model with multiplicative heteroskedasticity. A description of the particular optimization method is provided in Greene (1992, ch. 2).

Table 4
Ordered probit estimates.

$$K_t^* = \delta_0 + \delta_1 \hat{\sigma}_{A,t}^2 + \delta_2 K_{t-1} + \varepsilon_{K,t}$$

$$\sigma_{K,t} = \exp(\gamma_1 \hat{\sigma}_{A,t}^2 + \gamma_2 K_{t-1})$$

| | All quotes | Deutsche Bank | Morgan Guaranty |
|------------|---------------------|---------------------|---------------------|
| δ_0 | -1.826 (-66.431) | -6.978 (-21.647) | -6.479 (-28.527) |
| δ_1 | 1.011 (70.629) | 0.664 (15.074) | 0.214 (19.333) |
| δ_2 | 0.713 (60.231) | 3.730 (22.718) | 3.208 (27.811) |
| γ_1 | 0.099 (25.855) | 0.143 (15.399) | -0.171 (-29.743) |
| γ_2 | 0.050 (8.050) | 0.294 (18.104) | 0.118 (6.276) |
| μ_1 | 0.485 (77.862) | 0.518 (22.259) | 0.386 (24.891) |
| μ_2 | 3.381 (75.555) | 8.596 (21.765) | 3.010 (26.577) |
| LR | 58,789 | 5,444 | 20,169 |

Notes: Maximum Likelihood estimates with individual t -statistics reported in parentheses. The estimates for All quotes, Deutsche Bank and Morgan Guaranty are based on 304,607, 30,900 and 23,209 quotes, respectively. The μ_1 and μ_2 parameters define the partition intervals. LR denotes the likelihood ratio statistics for the joint significance of all the explanatory variables i.e. $\delta_1 = \delta_2 = \gamma_1 = \gamma_2 = 0$.

spread is large, the following spread will also tend to be large. The γ_1 and γ_2 coefficients highlight the importance of heteroskedasticity in the spread equation. Both the current conditional variance of the exchange rate and the lagged spread have a positive influence on $\sigma_{K,t}^2$.

An overall measure of the joint significance of the explanatory variables in the mean and variance part of the model is provided by the likelihood ratio statistic for $\delta_1 = \delta_2 = \gamma_1 = \gamma_2 = 0$ reported at the bottom of the table. The test statistic takes the value 58,789, which is overwhelmingly significant in the corresponding asymptotic chi-squared distribution with four degrees of freedom. We note that with more than 300,000 observations, considerable care should be exercised in interpreting this LR test, as well as the t -statistics for the individual coefficient estimates. For conventional significance levels the probability of a type II error is virtually zero. Along these lines, we also caution that the estimated effects are likely to have been overstated if important unobservable variables have been omitted from the analysis.

The actual magnitude of the ordered probit coefficients in Table 4 are not

easily interpreted. For instance, in a dynamic linear model for the spread, a first-order autoregressive coefficient of 3, corresponding to $\hat{\delta}_2$ for the individual banks, would clearly be problematic. Thus, in order to provide an indication of the economic significance of the model estimates for the ordered probit model, we also carried out a simulation exercise. The model, assuming homoskedasticity, was evaluated at the mean values of the explanatory variables through simulation methods. This gave us a baseline for the probabilities of the spread falling into one of the four categories, i.e. a_1 , a_2 , a_3 , or a_4 . The simulations were then repeated with each of the explanatory variables increased by one standard deviation, holding everything else constant. The resulting changes in the model-generated probabilities are listed in panel A of Table 5. If the contemporaneous conditional variance of the exchange rate increases one standard deviation, then the probability of the spread falling into the lowest a_1 category falls by 10.8 percent. The probability of the spread being in the next highest a_2 category falls by 3.6 percent. The probability of the spread being in the 'normal' spread category of a_3 increases by 10.3 percent, whereas the probability of the spread being in the largest a_4 category increases by 4.1 percent. Clearly, these are economically meaningful shifts in the spread implied by the model.¹⁷

3.4.2. Individual bank estimates

The results for the whole data set are interesting and in support of our theoretical model, relating volatility to the size of the bid-ask spread. However, it is useful to examine how a single bank's posted quotes respond to volatility. Of course, the individual bank's quotes are not updated as frequently as those for the entire data set. For comparison purposes, we therefore estimate the same ordered probit model for the spreads quoted by the two most active banks using the conditional variance estimates reported in panel B of Table 2.

These estimates for Deutsche Bank and Morgan Guaranty are given in the last two columns of Table 4. The qualitative results are the same as for the entire set of quotations. Most importantly, the δ_1 coefficients indicate that exchange rate volatility has a positive and significant effect on the spread for the individual banks also. The estimates for δ_2 are again positive and somewhat larger than for all the quotes. The significance of the γ_1 and γ_2 coefficients is evidence of time-varying heteroskedasticity. Both likelihood ratio statistics for the joint significance of all the explanatory variables in the mean and variance part of the model are highly significant at any levels.

¹⁷To account for any additional temporal dependencies in the data, we also estimated an augmented version of the basic model in Eq. (13) that included two lags of both $\hat{\sigma}_{A,t}^2$ and K_t . The overall conclusion from this more general specification is broadly consistent with the findings for the simpler model reported on above. Details are available on request.

Table 5
Sensitivity analysis.

| | a_1 | a_2 | a_3 | a_4 |
|--------------------------|--------|--------|-------|-------|
| Panel A: All quotes | | | | |
| $\hat{\sigma}_{A,t}^2$ | -0.108 | -0.036 | 0.103 | 0.041 |
| K_{t-1} | -0.058 | -0.016 | 0.058 | 0.016 |
| Panel B: Deutsche Bank | | | | |
| $\hat{\sigma}_{A,t}^2$ | -0.046 | -0.010 | 0.044 | 0.012 |
| K_{t-1} | -0.086 | -0.022 | 0.075 | 0.033 |
| Panel C: Morgan Guaranty | | | | |
| $\hat{\sigma}_{A,t}^2$ | -0.061 | -0.023 | 0.081 | 0.003 |
| K_{t-1} | -0.188 | -0.114 | 0.238 | 0.064 |

Notes: The table reports the change in the probabilities of the spread falling into each of the four different categories for the ordered probit models in Table 4 as a result of a one standard deviation increase in the mean of the various explanatory variables.

The economic significance of the coefficient estimates is illustrated through the simulations reported in panels B and C of Table 5. As the volatility of the exchange rate increases by one standard deviation, the probability of the spread falling into the lowest category drops by 4.6 and 6.1 percent for Deutsche Bank and Morgan, respectively. The other categories reflect changes in the probabilities that are consistent with greater volatility causing economically meaningful increases in the spread. The effects are similar for the lagged spread. Overall, the results are quite encouraging in terms of their support for the model sketched in section 1.

4. Conclusion

We have utilized every Deutschmark/dollar quote that crossed the Reuters screen during 3 months in the spring of 1989 to explore the link between exchange rate volatility and the bid-ask spread quoted by banks. Our results show that the foreign exchange market bid-ask spreads are characterized by strong temporal dependencies.

The quoted spreads tend to cluster about a few values. The four spread values of 5, 7, 10, and 15 basis points account for 97 percent of the quotes observed over the 3-month period. Due to this limited number of spread values quoted by banks, we used an ordered probit analysis to estimate the relationship between volatility and spreads. Separate estimation was carried out using all of the more than 300,000 weekday quotes, as well as using quotes from the two major quoting banks only.

Measuring exchange rate volatility as the conditional variance of the ask

price estimated by an MA(1)-GARCH(1,1) model, we find that there is a strong positive relationship between volatility and spreads. The coefficient estimates of the effect of the conditional variance on the spread are highly statistically significant. Simulating the effects of a one standard deviation increase in volatility also indicates a strong economic effect on the spread. For instance, using all the observations, if the volatility increases by one standard deviation, the probability that the next quote will fall in the lowest spread category declines by almost 11 percent points.

Since the quotes average 17 seconds apart for the entire sample, it is, perhaps not surprising that there exists a strong temporal persistence in the quoted spreads. In addition to the estimated statistical significance of the autocorrelation in the quotes, simulating the model yields inference regarding the economic significance. The results for the entire data set indicate that if the current quoted spread increases by one standard deviation, then the probability that the next quote will fall in the lowest spread range falls by almost 6 percent.

The findings of this paper add a new layer to our knowledge of exchange rate dynamics and their determinants at the microstructure level. Motivated by a stylized model of quote formation, the empirical investigations were conducted using a unique data set of continuously recorded bid-ask quotations. As such, the study provides a new look at the foreign exchange market, and it is our hope that it will stimulate further research on the distinctive microstructure characteristics of this global financial market.

References

- Admati, A.R. and P. Pfleiderer, 1989, Divide and conquer: A theory of intra-day and day of the week effects, *Review of Financial Studies* 2, 189-223.
- Anderson, T.G., 1993, Return volatility and trading volume: An information flow interpretation of stochastic volatility, Manuscript, Department of Finance, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL.
- Black, S.W., 1991, Transactions costs and vehicle currencies, *Journal of International Money and Finance* 10, 512-526.
- Bollerslev, T. and I. Domowitz, 1993, Trading patterns and prices in the interbank foreign exchange market, *Journal of Finance*, 48, 1421-1443.
- Bollerslev, T., R.Y. Chou and K.F. Kroner, 1992, ARCH modeling in finance: A review of the theory and empirical evidence, *Journal of Econometrics* 52, 5-59.
- Boothe, P., 1988, Exchange rate risk and the bid-ask spread: A seven country comparison, *Economic Inquiry* 26, 485-492.
- Bossaerts, P. and P. Hillion, 1991, Market microstructure effects of government intervention in the foreign exchange market, *Review of Financial Studies* 4, 513-541.
- Glassman, D., 1987, Exchange rate risk and transactions costs: Evidence from bid-ask spreads, *Journal of International Money and Finance* 6, 479-490.
- Glosten, L. and P. Milgrom, 1985, Bid ask, and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 13, 71-100.
- Goodhart, C.A.E., 1988, The foreign exchange market: A random walk with a dragging anchor, *Economica* 55, 437-460.
- Goodhart, C.A.E., 1990, News and the foreign exchange market, Manuscript, London School of Economics, London.

- Goodhart, C.A.E. and A. Demos, 1990, Reuter screen images of the foreign exchange market: The deutschemark/dollar spot rate, *Journal of International Securities Markets*, 333-348.
- Goodhart, C.A.E. and L. Figliuoli, 1991, Every minute counts in financial markets, *Journal of International Money and Finance* 10, 23-52.
- Greene, W.H., 1990, *Econometric analysis* (Macmillan, New York, NY).
- Greene, W.H., 1992, *LIMDEP user's manual and reference guide* (Econometric Software Inc., Bellport, NY).
- Hausman, J.A., A.W. Lo and A.C. MacKinlay, 1992, An ordered probit analysis of transaction stock prices, *Journal of Financial Economics* 31, 319-379.
- Lo, A.W. and A.C. MacKinlay, 1990, An econometric analysis of nonsynchronous trading, *Journal of Econometrics* 45, 181-210.
- Lyons, R.K., 1991, Private beliefs and information externalities in the foreign exchange market, Manuscript, Department of Finance, Columbia University, New York, NY.
- Maddala, G.S., 1983, *Limited dependent and qualitative variables in econometrics* (Cambridge University Press, Cambridge, UK).
- Melvin, M. and K.H. Tan, 1992, Foreign exchange market bid-ask spreads and the market price of social unrest, Manuscript, Department of Economics, Arizona State University, Tempe, AZ.
- Nelson, D.B., 1990, ARCH models as diffusion approximations, *Journal of Econometrics* 45, 7-38.
- Nelson, D.B., 1992, Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model, *Journal of Econometrics* 52, 61-90.
- Pagan, A.R., 1984, Econometric issues in the analysis of regressions with generated regressors, *International Economic Review* 25, 221-247.
- Tygiel, C., 1988, *Basic handbook of foreign exchange* (Euromoney Publications Limited, London, UK).
- Zhou, B., 1992, High frequency data and volatility in foreign exchange rates, Manuscript, Department of Finance, Sloan School of Management, MIT, MA.