APPEARS IN THE PROCEEDINGS OF COMPUTATIONAL INTELLIGENCE IN FINANCIAL ENGINEERING
IEEE PRESS (PISCATAWAY, NJ), PP. 26-30, 1995.

PRICE BEHAVIOR AND HURST EXPONENTS OF TICK-BY-TICK INTERBANK FOREIGN EXCHANGE RATES

John Moody and Lizhong Wu

Oregon Graduate Institute, Computer Science Dept., Portland, OR 97291-1000
Email: moody@cse.ogi.edu and lwu@cse.ogi.edu

ABSTRACT

Our previous analysis of tick-by-tick interbank Foreign Exchange (FX) rates has suggested that the market is not efficient on short time scales. We find that the price changes show mean-reverting rather than random-walk behavior [4]. The results of rescaled range and Hurst exponent analysis presented in the first part of this paper further confirms the mean-reverting attribute in the FX data. In the second part of this paper, we report the highly significant correlations between Bid/Ask spreads, volatility and forecastability that we have found in the FX data. These interactions show that higher volatility results in higher forecast error and increased risk for market makers, and that to compensate for this increase in risk, market makers increase their Bid/Ask spreads.

1. Introduction: Tick-by-Tick Interbank FX Rates

Tick-by-tick interbank FX data consists of a sequence of Bid/Ask prices quoted by various firms that function as market makers. While Bid/Ask price quotes from many market makers are displayed simultaneously by wire services such as Reuters and Telerate, a single price series can be constructed from the sequence of newly updated quotes.

We are analyzing a full year of such tick-by-tick Interbank FX price quotes for three exchange rates: the Deutschmark / US Dollar rate (DEM/USD), the Japanese Yen / US Dollar rate (JPY/USD), and the Deutschmark / Yen (DEM/JPY) cross-rate. The data were obtained from Olsen & Associates of Zürich. The data sample includes every tick from October 1992 through September 1993. For the DEM/USD, the year has 1,466,946 ticks.

For some purposes, it is useful to first "reduce" the dual Bid and Ask series to a single average price series. While we have taken this approach in some of our work, we have also studied the Bid/Ask series separately and other quantities such as the spread and volatility.

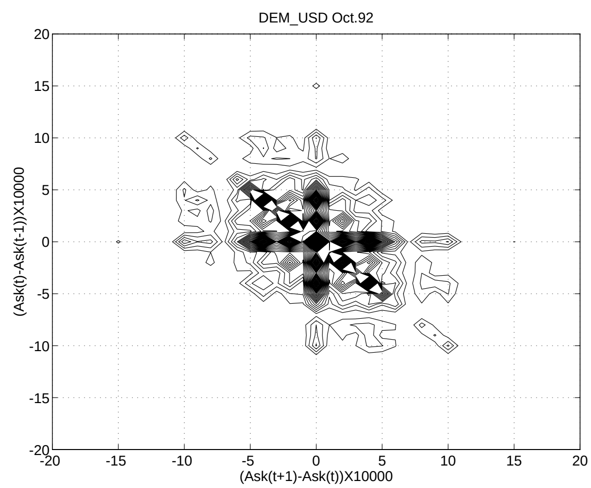

Figure 1: Contour plot of two dimensional probability distributions for successive Ask returns for the DEM/USD exchange rate during October 1992, which consists of $134,813$ ticks. The black thick line with a slope of about -1 shows the anticorrelation of successive Ask returns. Similar results obtain for the Bid series and the data in other time periods.

In this paper, we present some of our results on the price behavior of tick-by-tick interbank FX data. We report the temporal correlation structure of the data via rescaled range analysis in section 2. In section 3. we demonstrate the interactions of Bid/Ask spreads, volatility and forecastability by observing their cross-correlation coefficients. While our studies have been extended to other exchange rates, all the examples given in this paper are with DEM/USD.

2. Temporal Correlation Structure: Rescaled Range Analysis and Hurst Exponents

In our previous studies [4], we have investigated the short term price behavior of tick-by-tick interbank

FX data. We examined the one and two tick probability distributions, the autocovariance functions, the distribution of zero-crossing times and the short term price patterns. The results of this analysis show the presence of short term mean reversion in the tick-by-tick data. As an example to demonstrate this mean-reverting attribute, Figure 1 plots contours of a two dimensional histograms of frequencies of returns for DEM/USD.

To study the behavior of returns on time horizons longer than a few ticks, we perform rescaled range analysis and compute Hurst exponents (see [1] and [2]). This analysis is able to quantify long-term correlations and deviations from Gaussian behavior in a stochastic process.[1]

The rescaled range R/S on time scale N ticks is computed as follows. Define the detrended cumulative return from time t_0 on time scale N as

$$X(N, t_0, \tau) = \sum_{t=t_0+1}^{t_0+\tau} (r_t - m(N, t_0)) \text{ for } \tau \in \{0, N\} \ , \quad (1)$$

where r_t is the one tick return at time t and the mean return is $m(N, t_0) = \sum_{t=t_0+1}^{t_0+N} r_t/N$. Define $X(N, t_0, 0) \equiv 0$, and note that $X(N, t_0, N) = 0$. The range is then

$$R(N, t_0) = \max_{\tau} X(N, t_0, \tau) - \min_{\tau} X(N, t_0, \tau) \ , \quad (2)$$

and the rescaled range $R(N, t_0)/S(N, t_0)$ is obtained by dividing by the standard deviation of the returns computed for the same interval of length N beginning at time t_0.

The average rescaled range denoted $[R/S](N)$ is obtained by averaging $R(N, t_0)/S(N, t_0)$ over t_0 for a long sample of ticks $M \gg N$. If $[R/S](N)$ follows an approximate scaling law, then the Hurst exponent H is defined by

$$[R/S](N) \approx (CN)^H \ , \quad (3)$$

where C is a constant. Note that for a stochastic process, $H \in (0, 1)$. The case of $H = 0.5$ corresponds to a Gaussian and statistically independent process. Other values of H indicate the presence of statistical correlations, with the cases $H \in (0, 0.5)$ and $H \in (0.5, 1)$ corresponding to mean-reverting and mean-averting (trending) processes respectively.

Figure 2 shows the rescaled range analysis for October 1992 DEM/USD Bid returns. Both the original data and scrambled data were analyzed. The upward shift in the curve for the scrambled data is evidence for mean reversion of the original series on all time scales measured. Figure 3 shows Hurst exponents measured on data samples of varying lengths up to 10,000 ticks.

[1] A related analysis involving the computation of drift exponents has been conducted by researchers at Olsen & Associates, as described for example in [5] and references therein.

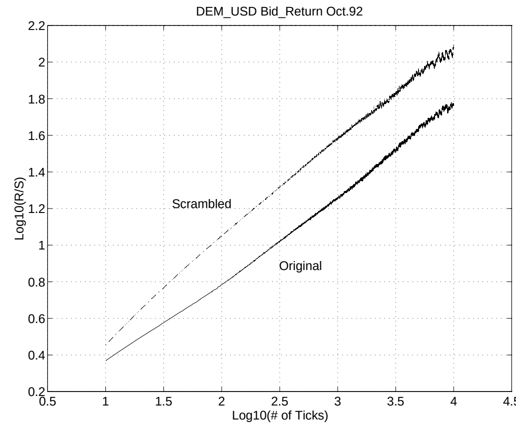

Figure 2: Rescaled range analysis for DEM/USD during October 1992. The lower slope for the original series for time scales less than 1000 ticks suggests that mean reversion is present within these time scales. For reference, a typical business day has about $6,000$ ticks in DEM/USD market.

To explain the above results, we have performed rescaled range analysis and computed Hurst exponents for two AR(1) returns processes, one which is mean averting and a second which is mean reverting. (Here, we use the terms "mean averting" and "mean reverting" in the very short term sense as determined by the sign of the coefficient of the AR(1) returns process.) Figure 4 shows the results of this analysis for both processes. It is clear from these results that the DEM/USD series behaves like the mean reverting process analyzed on the right hand graphs.

It is interesting to note that the observed Hurst exponents for FX data and for synthetic data depend on both the time scale window N and the length of the data M considered. Both N and M determine the estimation accuracy. In general, two types of estimation errors exist in rescaled range analysis and computation of Hurst exponents. One is referred to as low frequency distortion, which happens when the time scale N is large and the assumption $M \gg N$ is hardly satisfied. As shown in Figure 2, the average rescaled range $[R/S](N)$ becomes less smooth when the time scale N increases. Therefore, the estimation accuracy of Hurst exponents will also fall down. The low frequency distortion can be overcome by us-

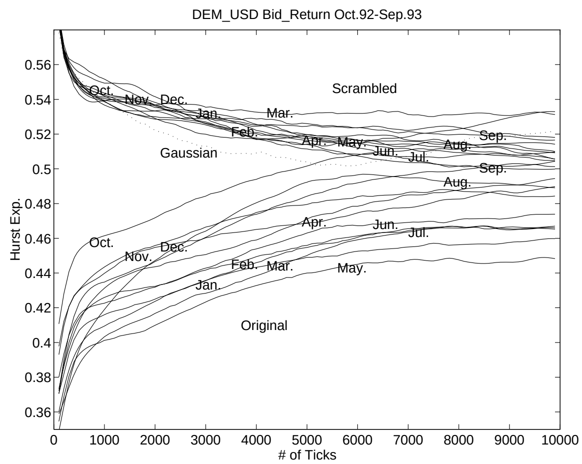

Figure 3: Hurst exponents for DEM/USD estimated for each month for different time scales ranging from 100 to 10,000 ticks. Curves for both original and scrambled returns are shown. The dotted line shows Hurst exponents estimated for a simulated Gaussian process. The differences between the original and [scrambled / Gaussian] exponents are due to mean reversion in the DEM/USD series.

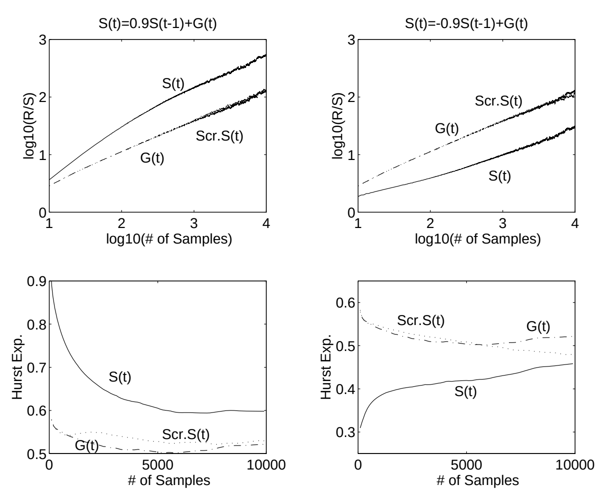

Figure 4: Rescaled range analysis (upper) and Hurst exponents (lower) of simulated returns processes: Gauss-Markov process ($S(t)$, solid), Gaussian process ($G(t)$, dashed), and Scrambled Gauss-Markov ($Scr.\ S(t)$, dotted) process. The left (right) plots are for positively (negatively) correlated processes $S(t)$. Samples of 60,000 points were used in these experiments. The minimum time scale on which Hurst exponents were calculated is 100 ticks.

Table 1: Hurst exponents computed in three different time-scale windows: $10-100$, $100-1000$ and $1000-10000$. The results given are for DEM/USD data from Oct. 1992 to Sep. 1993. We computed the Hurst exponents month by month and then averaged over 12 months.

Time-Scale	10-100	100-1000	1000-10000
	Original		
Mean	0.366	0.438	0.499
STD	0.021	0.018	0.023
Max	0.410	0.472	0.539
Min	0.340	0.409	0.464
Median	0.363	0.440	0.497
	Scrambled		
Mean	0.588	0.532	0.505
STD	0.003	0.008	0.028
Max	0.595	0.553	0.538
Min	0.583	0.524	0.438
Median	0.588	0.531	0.506

ing more data. Another estimation error is referred to as high frequency distortion, which happens when the time scale N is very small. In theory, the Hurst exponents of Gaussian series and the scrambled data should always be equal to 0.5. However, from Figure 3 and 4, we find that their Hurst exponents are larger than 0.5 when the time scale is small. In [3], Mandelbrot and Wallis analyzed this high frequency distortion. They found that the error occurs due to a quantization effect in the discrete data, which results in the "grid population range" being smaller than the "true population range". As the time scale increases, the "grid population range" approaches the "true population range".

Due to the existence of estimation errors, one should be very careful when interpreting the results obtained from rescaled range analysis and Hurst exponent computations. In our analysis, we always compare the results obtained from the original data to those obtained from the scrambled data and some synthetic series. Furthermore, as listed in Table 1, we computed the Hurst exponents in three different time-scale windows: $10-100$, $100-1000$ and $1000-10000$ and again compared to their corresponding figures obtained from the scrambled data. For the time scale from $10-1000$, we have more than 99% confidence to accept the hypothesis that the Hurst exponents of the original data are different from those of scrambled data.

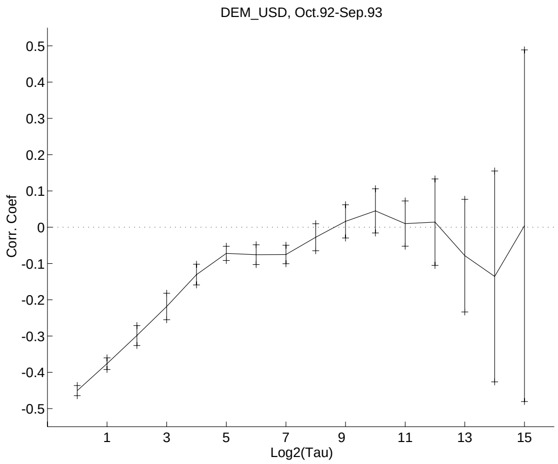

Figure 5: Long term autocorrelation analysis for DEM/USD data from Oct. 1992 to Sep. 1993. The autocorrelation coefficients are computed month by month and then averaged over 12 months. The error bars are ± one standard deviation.

The results in Table 1 show that the mean-reverting attribute of tick-by-tick interbank FX data exists even when the time scales are up to 1000 ticks. This can further be confirmed by long-term correlation analysis. We computed the autocorrelation coefficients between $[p(t+\tau)-p(t)]$ and $[p(t)-p(t-\tau)]$ (p stands for FX quotes) with different τ. Figure 5 shows the results. We see that significant negative correlations exists for $\tau \leq 256$.

To observe whether the above rescaled range analysis and Hurst exponent computations are affected by possible outliers in the data, we used robust estimation and replaced means by medians. Now, the $[R/S](N)$ is the median, not average, rescaled range, and the Hurst exponent is obtained with least median of squares regression instead of ordinary least squares regression. Table 2 compares the results obtained with robust approach to those with ordinary, non-robust approach. We see that both approaches obtain very similar results for small time scales, but there is a difference for the time scale $1000-10000$.

In summary, the behavior of the Hurst exponents in the returns of DEM/USD exchange rates is qualitatively different from the Hurst exponents of Gaussian series and scrambled series of the returns of DEM/USD exchange rates. It is similar to that of an AR(1) process with negative coefficient. Further investigation is required to estimate the two cutoff time scales within which the estimation error is small and both low and high frequency distortions can be ignored.

Table 2: Comparison of Hurst exponents on different time scales estimated with robust and non-robust regression methods. The results are for the Oct. 1992 DEM/USD data.

Time-Scale	10-100	100-1000	1000-10000
Robust	0.396	0.485	0.560
Non-robust	0.410	0.472	0.524

Table 3: Interactions of monthly and daily average Bid/Ask Spreads, Volatility, and Forecastability. The 95% confidence limits for the cross-correlation coefficients are {-.49,+.66} for monthly average and {-.12,+.13} for daily average. Thus, all of the cross-correlations are significantly different from zero.

Monthly	Forecastability	Volatility	Spreads
Forecastability	1.000		
Volatility	-0.797	1.000	
Spreads	-0.912	0.856	1.000
Daily	Forecastability	Volatility	Spreads
Forecastability	1.000		
Volatility	-0.402	1.000	
Spreads	-0.566	0.796	1.000

Our present results, however, support the conclusion that the DEM/USD series is mean-reverting for the period studied.

3. Bid/Ask Spreads, Volatility, and Forecastability

We constructed recursive AR predictors of the Bid and Ask returns series. The predictors were univariate predictors and were re-estimated for each tick based on a moving window of the 1024 previous ticks. The detailed designs and performances for predictors on various time scales ahead can be seen in [4].

One interesting phenomenon that we have found is a very high degree of correlation between average Bid/Ask spreads, volatility, and forecastability. Figure 6 shows our one-tick ahead prediction accuracy (as measured by the percentage of correctly predicted ups, downs, or no-changes), the volatility (defined as the standard deviation of price changes) and the Bid/Ask spreads. Table 3 gives their correlation coefficients. Note the highly significant correlations for both monthly and daily averages.

These results support the notion that higher volatility results in higher forecasting error and increased risk for market makers. To compensate for this increase in risk, market makers increase their Bid/Ask spreads.

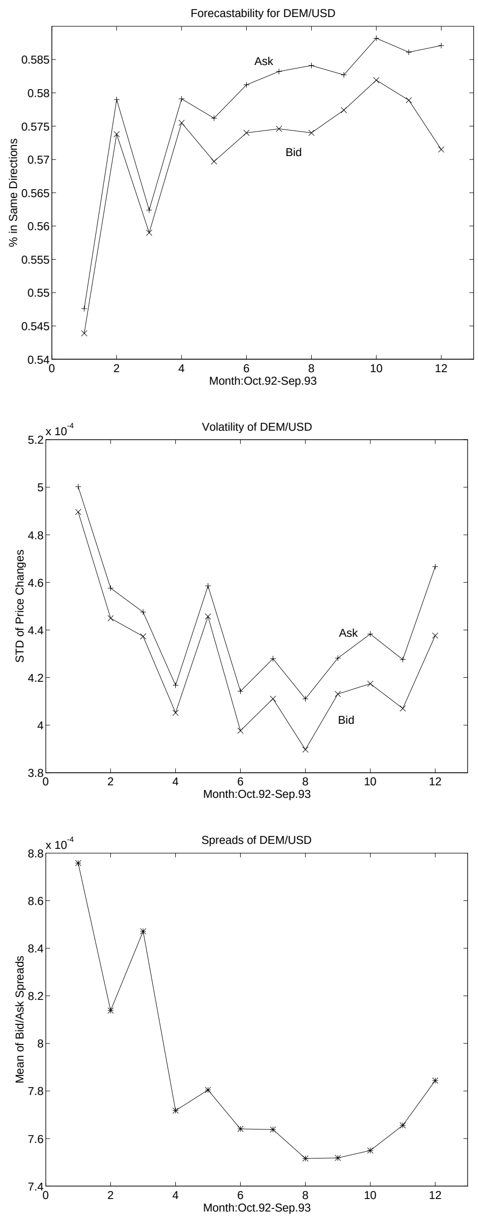

Figure 6: Forecastability, Volatility and Bid/Ask Spreads computed monthly for the DEM/USD market during Oct. 1992 to Sep. 1993. Significant correlations between the curves are apparent. See Table 3 for the cross-correlation coefficients.

4. Concluding Remarks

We have demonstrated convincing evidence for statistically-significant mean reversion on short time scales in tick-by-tick foreign exchange spot rates. Our results include correlation analysis, rescaled range analysis, and computation of Hurst exponents.

Moreover, we have demonstrated the presence of significant statistical relationships in the tick-by-tick data between forecastability, volatility, and Bid/Ask spreads. Spreads are positively correlated with volatility, while forecastability is negatively correlated with both spreads and volatility. These relationships agree with economic intuition about the rational behavior of FX dealers.

5. References

[1] H.E. Hurst. Long-term storage of reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 1951.

[2] B. Mandelbrot and J. Van Ness. Fractional Brownian motion, fractional noise, and applications. *SIAM Review*, 10, 1968.

[3] B. Mandelbrot and J. Wallis. Computer experiments with fractional Gaussian noises. Part 3, mathematical appendix. *Water Resources Research*, 3(1):260–267, Feb. 1969.

[4] John Moody and Lizhong Wu. Statistical analysis and forecasting of high frequency foreign exchange rates. Presented at *the Second International Workshop on Neural Networkd in the Capital Markets*, Nov. 1994.

[5] U.A. Muller, M.M. Dacorogna, R.D. Dave, O.V. Pictet, R.B. Olsen, and J.R. Ward. Fractals and intrinsic time - a challenge to econometricians. Technical Report UAM 1993-08-16, Olsen & Associates, Zurich, Switzerland, 1993.

Acknowledgements

We thank Olsen & Associates, Research Institute for Applied Economics of Zürich for providing us with the tick-by-tick FX spot rate data.