Chapter 22

NOISE IN FOREIGN EXCHANGE MARKETS

George G. Szpiro
Jerusalem
Israel

"Noise makes financial markets possible, but also makes them imperfect."

—Fischer Black

This chapter employs a new technique to compare the level of noise in financial markets. The data that are analyzed consist of high-frequency time series of three foreign exchange markets. It is shown that noise in the Dollar–Deutschmark series is least intense, that the Yen–Dollar market has about 10 percent more noise, and that there is about 70 percent more noise in the Deutschmark–Yen time series. On average, the noise level is higher in the late summer and fall than in the winter. The differentials may be related to the amount of news and the timing of its arrival.

1 INTRODUCTION

Noise, omnipresent in realistic models of nature, is generally considered a nuisance because it keeps researchers from completely explaining natural phenomena and prevents practitioners from making exact predictions. In economics this is more true than in the exact sciences, where conditions can be controlled and experiments can be shielded from outside influences. On the other hand, in economic and in financial markets, noise may be one of the reasons that profits can be made. [2]; [21].

Usually the variance that remains after the effects of all variables of the model have been taken into account is defined as noise. This remainder is considered to be due to variables that are unknown, or that are disregarded by the model: external shocks, random fluctuations, or the arrival of new information. Market behavior is often described and studied by using autoregressive models with noise,[1]

$$x_t = f(x_{t-1}, x_{t-2}, ..., x_{t-n}) + \varepsilon_t. \quad (1)$$

When confronted with a time-series of observed financial or economic data, the usual procedure is to specify a model, $f(x_{t-1}, x_{t-2}, ..., x_{t-n})$, and then to define noise as the unexplained remainder. Hence, in a fully specified autoregressive time-series, noise is due to factors other than x_{t-j}. While there exist relatively simple methods to estimate and fit linear models to time series, the problem of model specification becomes especially difficult when the data-generating process may be nonlinear.

Recently a new method has been proposed that reverses the process [25] instead of first specifying a model and then defining noise as the remaining variance, the technique permits the measurement of noise even before a linear or nonlinear autoregressive model is specified. All that is required in this case is that the noise's distribution be known. By adapting a method borrowed from theoretical physics—specifically, a well-known algorithm for the determination of the dimension of a chaotic system [14]—the size of the noise can be computed. But we can go one step further. When not only the structure of the model is unknown but, also one has no knowledge about the noise's distribution, the levels of noise in different time series can still be compared, and their relative intensities can be inferred. The only prerequisite in this case is that the noises have the same distributions, and this is the only assumption made in this chapter.

One drawback is that massive amounts of data are needed. Therefore, this chapter applies the methodology to high-frequency time series from foreign exchange markets (containing between several hundred thousand and 1.5 million data points). The intensity of the noise levels in three markets is compared, and then possible differences between months of the year and days of the week are investigated. As will be shown, evidence exists that the noise level is higher in the late summer and fall than in the winter. No significant differences were found for the days of the week.

This chapter provides a first attempt at comparing the amount of noise in financial time series. It does not try to explain the reasons why the level of noise in one series is higher or lower than in another. Further research about the intensity of noise, its timing, and its causes (e.g., arrival of new information, external shocks) is warranted.

The next section explains the difference between two types of noise and discusses the related problem of the tick size in foreign exchange quotes. Section 3 gives a brief description of the method used to measure noise. This method is related to techniques used to search for nonlinear dynamics and "strange attractors" that may underlie financial and economic time-series.[2] Section 4 examines

[1]This can be justified by Takens' embedding theorem [27], which states that the lagged values of one variable suffice to characterize the dynamics of a multivariable system.

[2]Some studies that attempted to discover nonlinearities in the foreign exchange rates include Hsieh [16], Meese and Rose [18], Guillaume [15], Evertsz [11], Mizrach [19], Cecen and Erkal [6], and Brooks [5]. On the role of "chartists," who search for simple patterns in one

the database, and Section 5 presents the results. Section 6 concludes with a brief summary of findings and suggestions for further research.

2 MEASUREMENT NOISE VERSUS DYNAMIC NOISE, FINITE TICK SIZES

In experimental science, say, in physics or chemistry, one distinguishes between two types of noise: dynamic noise and measurement noise [20]. The first, dynamic noise δ, derives from outside influences that enter the model,

$$x_t = f(x_{t-1} + \delta_t), \tag{2}$$

while the latter, measurement noise μ, is due to the finite resolution of the measuring apparatus or to roundoff errors [23],

$$\begin{aligned} &x_t = y_t + \mu_t \\ &\text{where } y_t = f(y_{t-1}). \end{aligned} \tag{3}$$

Some attempts have been made to distinguish between the two (e.g., [24]). In economics and finance, it is generally believed that the latter noise presents few problems, since the resolution of the measurements is small relative to the numerical values involved. The cost of high-priced goods is generally given to the nearest dollar; for low-priced goods, the dollar is divided into cents. Stock prices are quoted in eighths of dollars; national accounts data involving trillion dollar figures are often given to the nearest million. Exchange-rate data are also quite accurate: the Dollar–Deutschmark exchange rate, for example, is given to four digits after the comma. Hence, with a DM/$ rate of, say, 1.4111, the precision is on the order of 0.007 percent.

Even though price quotes are generally quite precise, they are nevertheless finite. Indeed, the number of possible prices for a good or a commodity must be limited, because if there were an infinite number of pricing possibilities – say, all rational numbers between some upper and lower limits – buyers and sellers would have great difficulty matching their bids. Hence markets can only exist if there is a convention to meet at a finite number of prices. As pointed out above, the cost of a good is usually determined in dollars and cents. If this resolution is too coarse, market participants agree on ticks of, say, tenths or hundreths of a cent. For high-priced goods, on the other hand, dollars and cents may provide too many pricing possibilities for the market, and a convention evolves to trade only at prices rounded to the next five, ten, or hundred units of the currency. The shares of the Swiss newspaper "Neue Zürcher Zeitung," for example, priced at around $40,000, are traded at intervals of 250 Francs (about $180) by market makers in Switzerland.

The size of the tick, seemingly minute when compared to the price of the commodity, does present problems for research, however. In Figure 22.1, we create a

---

dimension, see Allen and Taylor [1]. Engle et al. [10] and Goodhart [13], for example, deal with news in the foreign exchange market. For a survey of the recent literature on exchange rate economics, see Taylor [28].

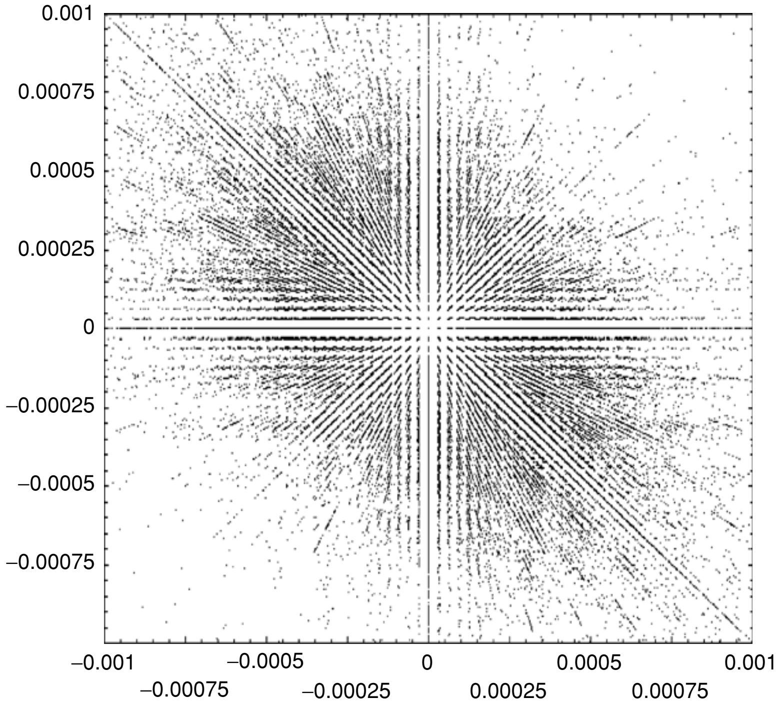


Figure 22.1. The Compass Rose (250,000 points of the $-DM data) ($x$-axis: x_t, y-axis: x_{t+1})

large sample of two-dimensional vectors from the $-DM data[3] and plot the x_{t+1}-values against their x_t-values. (In technical parlance, we embed the time series into two-dimensional space.)[4] The pattern that emerges, the so-called "compass rose," was discovered by Crack and Ledoit [7] and further analyzed by Szpiro [26]. As these authors showed, the pattern arises because of the discreteness of the quotes, i.e., the finite resolution of the prices. Finite tick-sizes may lead particular tests to indicate the existence of structure when, in fact, there is none. The problem is compounded by the fact that for such tests – and the method to measure noise uses a variant of such a test – one needs to "embed" the time series in high dimensions. (See Section 3.) As is well known, the lengths of objects generally expand when they are projected upwards into higher dimensions. In fact, the distance between two points grows on average with $\sqrt{d}$, where d is the dimension. Hence, for d=50, the effect of measurement errors increases about sevenfold. In conclusion, it may safely be assumed that noise in financial markets consists of both measurement and dynamic noise.

[3]Of the 1,472,241 observations, approximately every sixth vector was used, to give a sample-size of 250,000.

[4]The somewhat visible predominance of points in the North–West and South–East directions is due to the negative correlation between consecutive observations in high-frequency data. [13, 3].

3 CHAOS, DIMENSION, AND MEASUREMENT OF NOISE

The proposed technique [25] borrows from mathematical physics. Specifically, an algorithm is employed that was suggested by Grassberger and Procaccia [14] to measure the dimension of a chaotic system by embedding it in increasingly high dimensions.[5] To illustrate the technique, let us assume a very low-dimensional system and dynamical noise that is uniformly distributed in [–M,+M]. We embed the data into two-dimensional space, take a reference point, and count all neighbors that are contained within a circle of radius r. Let us call this number $C_2(r)$. Now we move to three-dimensional space. When "unfolding" the attractor, we realize that in three-dimensional space the nearest neighbors in the circle are actually downward projections of points that are contained in a cylinder above the circle. The as-yet unknown height M of this cylinder corresponds to the noise level. Let us count the neighbors that are contained in a ball around the reference point and call this number $C_3(r)$. Since the volume of the cylinder is $2M\pi r^2$, and the volume of the ball is $4/3\pi r^3$, the ratio of the number of points in two-dimensional space to those in three-dimensional space gives an indication as to the noise level. In this case, we have

$$C_2(r) = 2M\pi r^2 \quad \text{and} \qquad C_3(r) = \frac{4}{3}\pi r^3 \tag{4}$$

Computing the ratio, we get

$$\frac{C_3(r)}{C_2(r)} = \eta_{2,3}(r) = \frac{2}{3}\frac{r}{M} \tag{5}$$

We now run a homogenous regression between the ratio $\eta_{2,3}(r)$ and the radius r:

$$\eta_{2,3}(r) = \gamma r. \tag{6}$$

From the last two equations it follows that the level of noise, M, can be computed as

$$M = \frac{2/3}{\gamma}. \tag{7}$$

By adding and subtracting 1.96 standard errors of the γ-estimates, an approximate 95 percent confidence interval can be determined for the estimated noise level. For reasons that will not be discussed here,[6] the regressions are actually run with both r and r^2 as independent variables, and of course the technique must be applied in dimensions much higher than just two or three. As was pointed out above, if the noise distribution is unknown, the numerical results per se have no meaning, but nevertheless do provide an index of the amount of noise present.

---

[5]The algorithm is also used in a test that was devised to search for nonlinearities and chaos in economic time series [4]. There have been numerous attempts to determine whether a chaotic system underlies the data-generating process of financial or economic time series. [22, 12, 16]
[6]See Szpiro [25] for a more detailed description of the method.

Hence, the method can be used to compare relative intensities of noise in different time series, even if the noise's distribution is unknown.

When plotting the results of the computations against the embedding dimension, we will see high levels of noise at first that decrease as dimension increases, and that finally converge to a constant level. The reason for this phenomenon is that in a low-embedding dimension, several "strings" of the system are superimposed on each other, but the system "unfolds" when one moves to higher dimensions [29]. In the present study of the foreign exchange markets, we run the regressions with twenty r-values – for embedding dimensions 1 to 50 – starting with the smallest sphere that contains at least 10^5 data points.

4 DATA

The data for this study consist of the foreign exchange rates in three markets for the time period October 1st, 1992, to September 30th, 1993: the Dollar–German Mark ($–DM), the German Mark–Yen (DM–Y), and the Yen–Dollar (Y–$) markets. The data do not consist of actual trades, but of bid and ask quotes, the means of which will serve as proxies for actual trades. The quotes were collected by Olsen & Associates, who made it available to the academic community.[7] The log-differences of the mean between the bid and the ask price were calculated.[8] The series for the $–DM market consists of 1,472,241 entries (mean 0.00992, variance 835.2, and skewness 531.7). The Y–$ series has 570,814 entries (–0.02152, 2594.6, –1324.0), and the DM–Y series has 158,979 entries (–0.17354, 3615.1, –23174.0).

An argument could be made to use only data at specified points in time—say, at five-minute intervals—or to employ "business time" to filter the data [8]. The present study uses every observation in the series (i.e., "quote time"), the reason being that a higher frequency of quotes implies more hectic activity or, in other words, a speeding up of time. Furthermore, there seems to be no empirical evidence for intraquote dynamics of any relevance [11].

From the time-series $x_1, x_2, x_3, \ldots$ we derive k-dimensional vectors $\langle x_1, x_2, \ldots, x_k\rangle$, $\langle x_2, x_3, \ldots, x_{k+1}\rangle$, *etc., and compute the Euclidean distance between them, for dimensions 1 to 50. A total of* (n-k)(n-k-1)/2 *pairs could be calculated, but we make do with a sample, albeit a large one. For the $–DM series, the interpair distances of every vector with every 100th other vector were calculated; in the Y–$, every 15th vector was used; and in the DM–Y, distances were calculated with 4/5ths of the possible vectors. This sampling ensured that an equal amount of pairs (approximately 10^{10}) was used in each market to compute the amount of noise.[9] The pairs were grouped into spheres with radii between 1 and 150. For each embedding dimension, regressions (Eq. 6) were run for twenty consecutive spheres, starting with the smallest one that contained at least 10^5 pairs.*

[7]Olsen & Associates Ltd., Research Institute for Applied Economics, Zürich.
[8]The results are multiplied by 10^5 in order to receive numerical values that are more manageable.
[9]The method is very computer intensive. The calculation of the 10 billion interpair distances in embedding dimensions 1 to 50, and their grouping into spheres, took more than 200 hours on a Pentium-133 computer for each of the three series.

5 FINDINGS

In Figure 22.2, the results of the noise measurements for the three markets are depicted. Let us recall that since we do not know the distribution of the underlying noise, the numerical values only give an indication as to the relative levels of noise in the three markets, i.e., they represent indexes. As expected, the noise estimates are high in the low embedding dimension, but then gradually decrease as the embedding dimension increases. We see in the figure that in all three examples the estimates eventually converge.[10] From the evidence we conclude that the \$–DM market has a noise index of 12.2 (95 percent confidence interval: ±0.2), the Y–\$ market a level of 13.5 (±0.1), and the DM–Y market a level of 20.7 (±0.1). In other words, with the least noisy \$–DM market as a baseline, we may conclude that the Y–\$ market has about 10 percent more noise, and that there is about 70 percent more noise in the DM–Y time-series.

For comparison purposes, the identical operations are performed with time series whose entries have the same distribution as the original series but whose order is random. The results for the scrambled time series can serve as a benchmark. [4] The only difference in the scrambled series is the removal of any possible structure that existed in the order of the original entries. In Figure 22.3, the estimated noise levels for the scrambled series are presented. Even for low embedding dimensions, these are much higher than in the original series. For the \$–DM market, the level is about 35 (±0.3), and for the other two markets it lies above 45 (±0.5 and ±0.4). Hence, by scrambling the entries, two to three times as much noise was introduced into the series as there was in the original data.

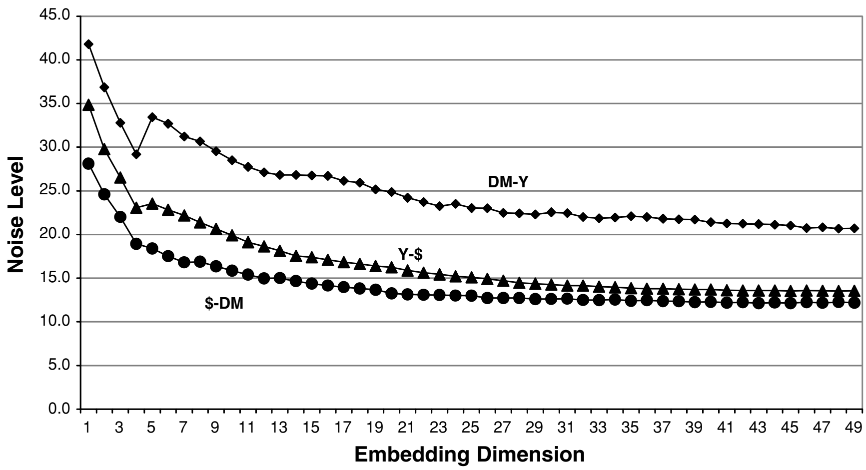


Figure 22.2. Levels of noise in three FX markets

---

[10]Convergence occurs above an embedding dimension of about 35, which would indicate that the time series have dimensions of not more than 17 [29]. However, such dimension estimates may be questioned. [9].

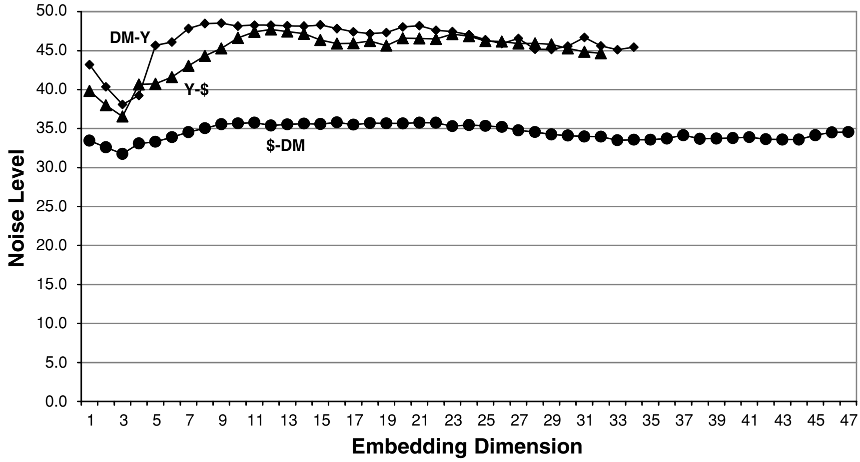
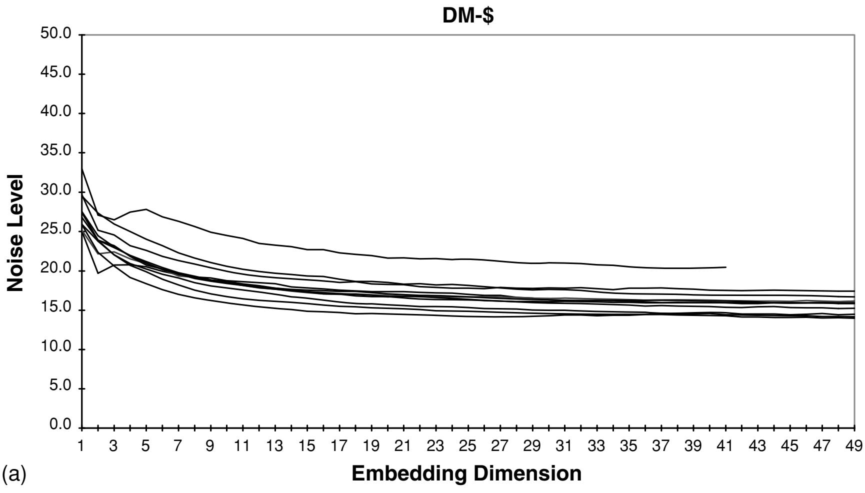


Figure 22.3. Levels of noise in scrambled time series

Let us now turn to more narrow data and compute the noise level within each month of the year, for each of the three markets. Figure 22.4 plots the development of the measured levels of noise against the embedding dimension. Again, the estimates converge,[11] and in Figure 22.5 a summary of the results is presented for embedding dimension 20. On average, the noise level for the months November to January is lower than for the period July to October. A tentative



(a)

Figure 22.4. Noise level in intramonth data

[11]For some months the calculations were not made for all embedding dimensions up to 50, since there was not a sufficient number of spheres that contain a minimum of 10^5 points. The confidence intervals are again very narrow and won't be given.

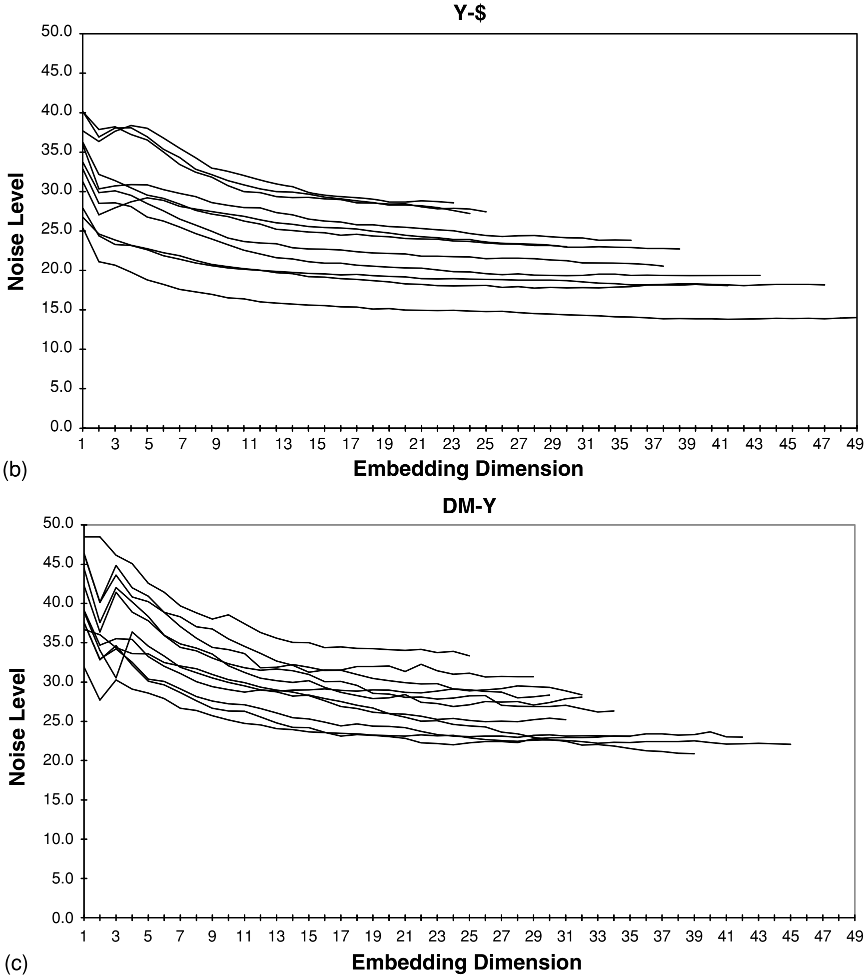


Figure 22.4. Noise level in intramonth data

explanation of this phenomenon could be that more information arrives during the late summer and fall than during the winter.

A comparison between Figures 22.2 and 22.5 shows that the estimates for the single months lie above the results for the whole years (for embedding dimension 20). It may seem surprising that the noise levels, as measured in the monthly data, do not average out to the noise level for the year as a whole. The answer to this puzzle goes to the heart of the nature of noise. The latter is generally defined as the data's variance that is not explained by a certain model. With a better model, unexplained variance decreases, and the data have less noise. In the method used in this chapter, vectors of the time series are compared with previous vectors, and similarities (i.e., closeness in the Euclidean norm) are sought. If one time series is a small subseries of the other, it is more difficult to find similar vectors. On the other hand, the longer the time series, the better known does the underlying structure become. Hence noise decreases. The time series of the data for single months

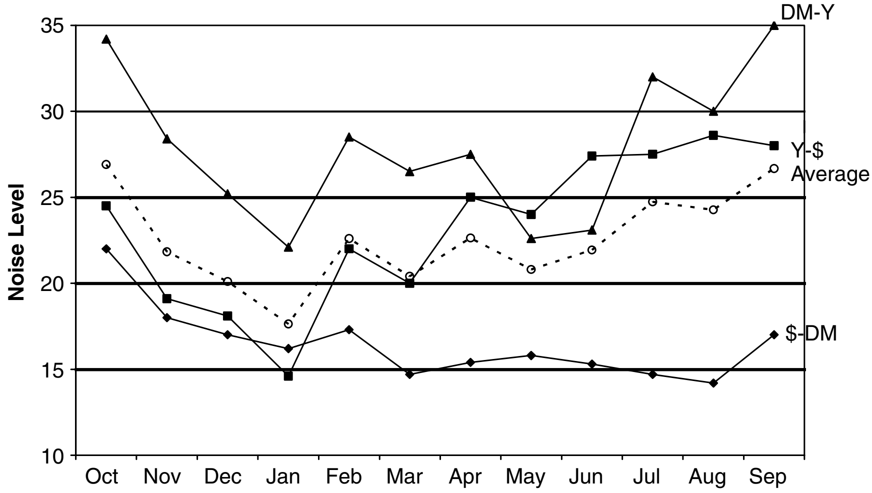


Figure 22.5. Noise level in intramonth data (Embedding dimension = 20)

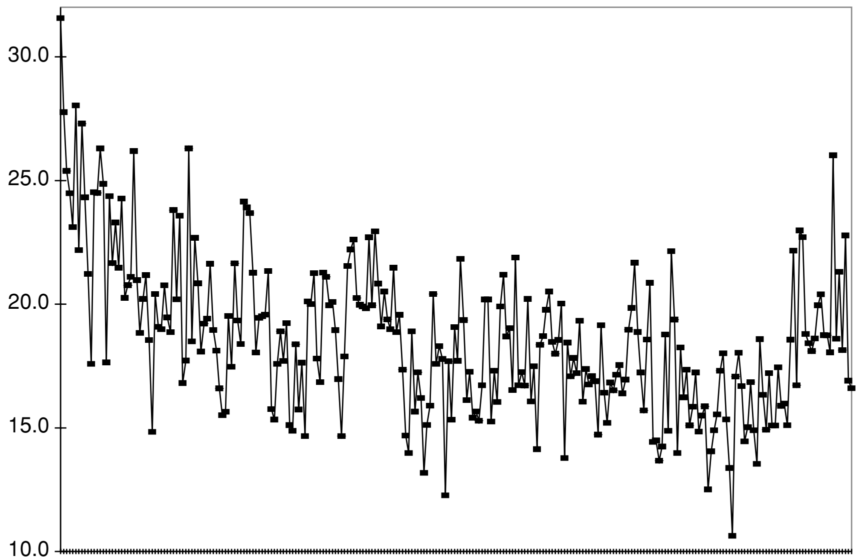


Figure 22.6. Noise level intraday data (Embedding dimension = 20; Mondays indicated by rectangles)

gives less opportunity to "learn" the structure, and a larger part of the data's variance is identified as noise. Incidentally, the reason why an approximately equal number of pairs is used for all markets in the estimations of the noise level for the full year was to counteract this effect.

Finally, we turn to intraday data. The noise level for 258 weekdays for the $–DM data is estimated and depicted in Figure 22.6.[12] Visual inspection does not

[12]There were no quotes for two Fridays during the year, 25th of December and 1st of January. For the following analysis, interpolated values were entered for these days.

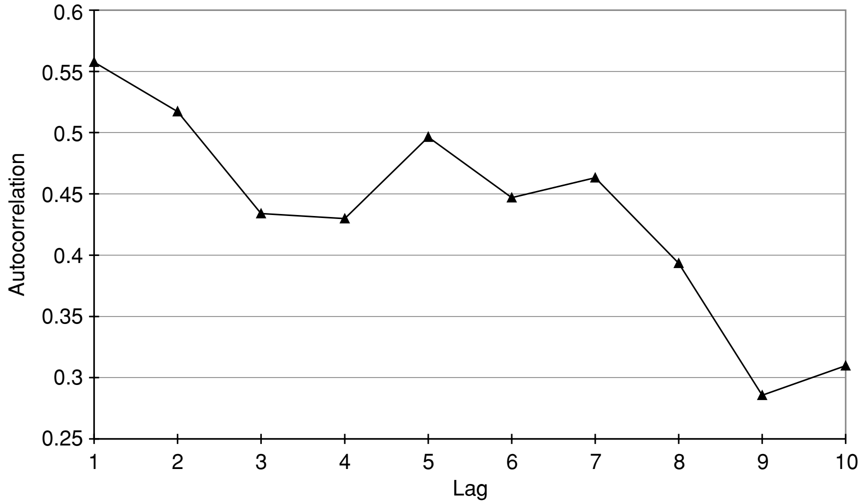


Figure 22.7. Autocorrelation for noise level in intraday data (Embedding dimension = 20)

reveal any significant pattern, and the average noise levels for the days of the week do not differ significantly. However, a common thread may nevertheless run through the days of the week. A plot of the first ten autocorrelations reveals strong "seasonality" for the fifth lag: $\rho = 0.496$ (see Figure 22.7). This suggests that the noise on a certain day of the week is correlated somehow to the same weekday's noise of the previous week. Autocorrelation is especially strong for Tuesdays and Wednesdays ($\rho = 0.519$ and 0.501, respectively). Possible reasons for this phenomenon may be that external shocks (of unknown origin) occur on certain weekdays, or that information arrives at specific times during the week. Further research is warranted.

6 CONCLUSIONS

This chapter employs a new technique to compare the levels of noise present in different time series. Since the method requires very long series, high-frequency data from foreign exchange markets are used. It is shown that of the three markets analyzed, the \$–DM market has the least amount of noise, the Y–\$ market has about 10 percent more noise, and there is about 70 percent more noise in the DM–Y time series. We also see that, on average, there is less noise during the winter than during the late summer and fall. Intraday data show some autocorrelation for the days of the week. The amount of noise may be related to the amount of news that arrives at certain times during the week, during the months, or in various markets.

The analysis presented in this chapter is meant to be an attempt at the comparative study of noise. It does not provide explanations as to why the level of noise in one series is higher or lower than in another. Further research about the intensity of noise and related issues, for example, the arrival and assimilation of new information, is warranted.

REFERENCES

[1] A.L. Helen and P.M. Taylor (1990): Charts, noise and fundamentals in the foreign exchange markets, *Economic J. 100,* 49–59.
[2] F. Black (1986): Noise, *J. Finance 41*, 529–543.
[3] T. Bollerslev and I. Domowitz (1993): Trading patterns and prices in the interbank foreign exchange market, *J. Finance 48*, 1421–1449.
[4] A. Brock, W.A.D. Hsieh, and B. LeBaron (1991): *Nonlinear Dynamics, Chaos, and Instability*, MIT Press, Cambridge.
[5] C. Brooks (1997): Linear and non-linear (non)-forecastability of high-frequency exchange rates, *J. Forecasting 16*, 125–145.
[6] A. Aydin Cecen and C. Erkal (1996): Distinguishing between stochastic and deterministic behavior in foreign exchange rate returns: further evidence, *Economics Letters 51*, 323–329.
[7] T.F. Crack and O. Ledoit (1996): Robust structure without predictability: the compass rose pattern of the stock market, *J. Finance 51*, 751–762.
[8] M. M. Dacarogna, U. A. Müller, R. J. Nagler, R. B. Olsen, and O. V. Pictet (1993): A geographical model for the daily and weekly seasonal volatility in the FX market, *J. International Money and Finance 12*, 413–438.
[9] J.-P. Eckmann and D. Ruelle (1992): Fundamental limitations for estimating dimensions and Lyapounov exponents in dynamical systems, *Physica D 56*, 185–187.
[10] R. F. Engle, T. Ito, and W. L. Lin (1990): Meteor showers or heat waves? Heteroskedastic intra-dayly volatility in the foreign exchange market, *Econometrica 58*, 525–542.
[11] J.G. Carl Evertsz (1996): Self-similarity of high-frequency USD-DEM exchange rates, in *Proceedings of the HFDF Conference* 1996, Olsen & Associates Ltd., Research Institute for Applied Economics, Zürich.
[12] M. Frank and T. Stengos (1989): Measuring the strangeness of gold and silver rates of return, *Review of Economic Studies 56*, 553–567.
[13] C. A. E. Goodhart and L. Figiuoli (1991): Every minute counts in financial markets, *J. International Money and Finance 10*, 24–52.
[14] P. Grassberger and I. Procaccia (1983): Characterization of strange attractors, *Physical Review Letters 50*, 346–349.
[15] M. Dominique Guillaume (1994): A low-dimensional fractal attractor in the foreign exchange markets?, working paper, Olsen & Associates Ltd., Research Institute for Applied Economics, Zürich.
[16] A. David Hsieh (1989): Testing for non-linear dependence in daily foreign exchange rates, *J. Business 62*, 339–368.
[17] _____, (1991): Chaos and nonlinear dynamics: application to financial markets, *J. Finance 46*, 1839–1877.
[18] A.R. Meese and K.A. Rose (1991): An empirical assessment of non-linearities in models of exchange rate determination, *Review of Economic Studies 58*, 603–619.
[19] B. Mizrach (1996): Determining delay times for phase space reconstruction with application to the FF/DM exchange rate, *J. Economic Behavior and Organization 30*, 369–381.

[20] M. Möller, W. Lange, F. Mitschke, N. B. Abraham, and U. Hübner (1989): Errors from digitizing and noise in estimating attractor dimension, *Physics Letters A 138*, 176–182.

[21] F. Palomino (1996): Noise trading in small markets, *J. Finance 51*, 1537–1550.

[22] J. Scheinkman and B. LeBaron (1989): Nonlinear dynamics and stock returns, *J. Business 62*, 311–337.

[23] G.G. Szpiro (1993a): Cycles and circles in roundoff errors, *Physical Review E 47*, 4560–4563.

[24] G.G. Szpiro (1993b): Measuring dynamical noise in dynamical systems, *Physica D 65*, 289–299.

[25] G.G. Szpiro (1997): Noise in unspecified, non-linear time series, *J. Econometrics 78*, 229–255.

[26] G.G. Szpiro (1998): Tick size, the compass rose and market nanostructure, *J. Banking and Finance 22*, 1559–1570.

[27] F. Takens (1981): Detecting strange attractors in turbulence, in D. Rand and L. Young (eds), *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, Springer Verlag *898*, 366–381.

[28] M.P. Taylor (1995): The economics of exchange rates, *J. Economic Literature 33*, 13–47.

[29] H. Whitney (1944): The self-intersection of a smooth *n*-manifold in 2*n*-space, *Ann. Mathematics 45*, 220–246.