Chaos and Nonlinear Dynamics: Application to Financial Markets

David A. Hsieh

The Journal of Finance, Vol. 46, No. 5. (Dec., 1991), pp. 1839-1877.

Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28199112%2946%3A5%3C1839%3ACANDAT%3E2.0.CO%3B2-P

The Journal of Finance is currently published by American Finance Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/afina.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Mon May 1 09:08:28 2006

THE JOURNAL OF FINANCE • VOL. XLVI, NO. 5 • DECEMBER 1991

Chaos and Nonlinear Dynamics: Application to Financial Markets

DAVID A. HSIEH*

ABSTRACT

After the stock market crash of October 19, 1987, interest in nonlinear dynamics, especially deterministic chaotic dynamics, has increased in both the financial press and the academic literature. This has come about because the frequency of large moves in stock markets is greater than would be expected under a normal distribution. There are a number of possible explanations. A popular one is that the stock market is governed by chaotic dynamics. What exactly is chaos and how is it related to nonlinear dynamics? How does one detect chaos? Is there chaos in financial markets? Are there other explanations of the movements of financial prices other than chaos? The purpose of this paper is to explore these issues.

CHAOS HAS CAPTURED THE fancy of many macroeconomists and financial economists. The attractiveness of chaotic dynamics is its ability to generate large movements which appear to be random, with greater frequency than linear models. As a result, there has been an explosion of papers searching for chaotic behavior in macroeconomic and financial time series. The purpose of this paper is to discuss some of the methodological issues in detecting chaotic and nonlinear behavior.

Section I provides a description of the key features of deterministic chaotic systems via a number of examples. Section II shows how deterministic chaos can, in principle, be detected using the method of correlation dimension proposed by Grassberger and Procaccia (1983). Section III deals with some limitations of this method. The Grassberger/Procaccia method requires a substantial number of data points, which is difficult to obtain in standard economic and financial time series. It also lacks a statistical theory for hypothesis testing. A different but related method has been proposed by Brock, Dechert, and Scheinkman (1987). Under the null hypothesis of independence and identical distribution (IID), the Brock, Dechert, and Scheinkman statistic has been shown to have good finite sample properties and good power against departures from IID. Some Monte Carlo evidence is

*Associate Professor, Fuqua School of Business, Duke University, Durham, NC 27706. The author is grateful to comments from workshop participants at Emory University, the Federal Reserve Bank of Atlanta, University of California at Berkeley, Harvard Business School, and The University of Michigan. The paper has also benefitted greatly from the comments of an anonymous referee.

provided. When applied to stock returns, this statistic rejects the null hypothesis of IID very strongly. The remainder of the paper investigates some of the causes of the rejection of IID: Section IV checks for nonstationarity; Section V, nonlinear conditional mean changes; and Section VI, conditional heteroskedasticity. Some concluding remarks are offered in Section VII.

I. What Is Chaos?

Chaos is a nonlinear deterministic process which "looks" random. There is a very good description of chaos and its origins in the popular book by James Gleick (1987), entitled *Chaos: Making a New Science*. Also, Baumol and Benhabib (1989) gives a good survey of economic models which produce chaotic behavior. Brock (1986) provides the exact mathematical definitions and formulations.

Chaos is interesting for several reasons. In the business cycle literature, there are two ways to generate output fluctuations. In the Box-Jenkins times-series models, the economy has a stable equilibrium, but is constantly being perturbed by external shocks (e.g., wars, weather). The dynamic behavior of the economy comes about as a result of these external shocks. In the chaotic growth models, the economy follows nonlinear dynamics, which are self-generating and never die down. The fact that economic fluctuations can be internally generated has a certain intuitive appeal.

It so happens that chaotic dynamics is necessarily nonlinear, which gives it a second appeal. It is well known that linear models can only generate four types of behavior: oscillatory and stable, oscillatory and explosive, nonoscillatory and stable, and nonoscillatory and explosive. On the other hand, nonlinear models can generate much richer types of behavior. For example, the system can have sudden bursts of volatility and occasional large movements. This has caught the attention of the financial press. Stock market analysts are always looking for explanations of large movements in asset prices, such as the October 19, 1987 stock market crash.

To get some ideas about the behavior of chaotic processes, we can consider several examples.

A. Tent Map

The simplest chaotic process is the tent map. Pick a number x_0 between 0 and 1. Then generate the sequence of numbers x_t using the following rule:

$$
\begin{aligned}
x_t &= 2x_{t-1}, \quad \text{if } x_{t-1} < 0.5, \\
x_t &= 2(1 - x_{t-1}), \quad \text{if } x_{t-1} \geq 0.5.
\end{aligned} \tag{1}
$$

The tent map is so named because the graph of x_t versus x_{t-1} is shaped like a "tent", as shown in Figure 1. Note that x_t is a *nonlinear* function of x_{t-1}.

Intuitively, the tent map takes the interval [0, 1], stretches it to twice the length, and folds it in half, as illustrated in Figure 2. Repeated application of

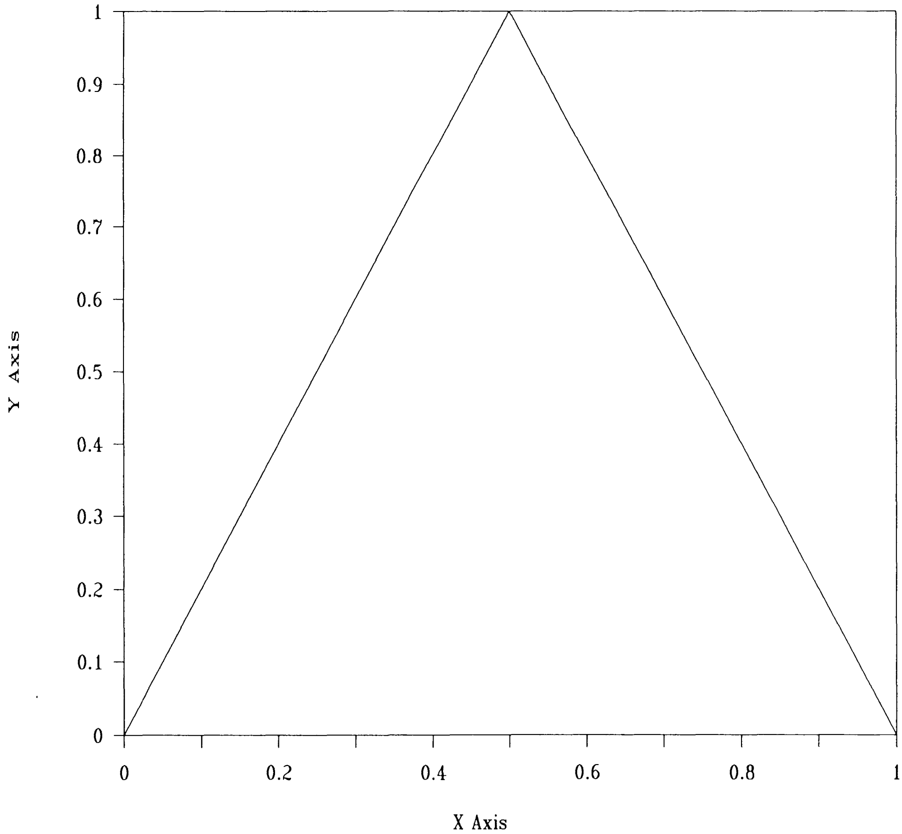

Figure 1. The tent map. The tent map is the graph of the function $y = 2x$ if $x < 0.5$ and $y = 2(1 - x)$ if $x \geq 0.5$, where x is the horizontal axis and y is the vertical axis.

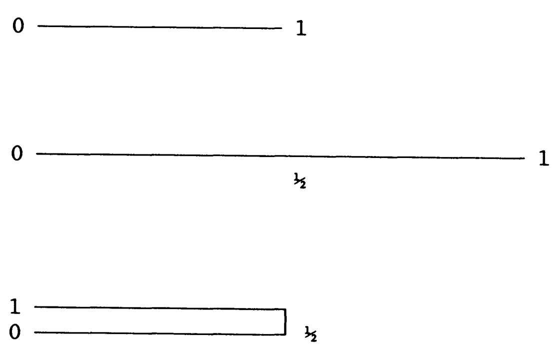

Figure 2. Stretch and fold action of the tent map. The tent map takes the unit interval [0,1], stretches it twice as long, and folds it back onto itself.

stretching and folding pulls apart points close to each other. This type of stretching and folding is characteristic of chaotic maps. It makes prediction difficult, thus creating the illusion of randomness.

There are four important properties of the tent map. One, $\{x_t\}$ fills up the unit interval [0, 1] uniformly as $t \to \infty$. Technically, this means that the fraction of points in $\{x_t\}$ falling into an interval [a, b] is (b − a) for any $0 < a < b < 1$. Two, any small error in measuring the initial x_0 will be compounded in forecasts of x_t exponentially fast. Three, x_t appears stochastic even though it is a deterministic process, in the sense that the empirical autocovariance function $\rho_{xx}(k) = E[x_t x_{t-k}] = \lim_{T\to\infty} \sum_{t=0}^{\infty} x_t x_{t-k}/T = 0$, which is the same as that of white noise. Four, x_t can have a series of small increases, and then it suddenly declines ("crashes?") sharply.

B. Pseudo Random Number Generators

A more "random" chaotic system can be obtained using the ideas of the tent map. Here is an example of a pseudo random number generator, which is very frequently used in computer programs. Take a number A (say 7^5) and a large prime number P (say $2^{32} - 1$). Pick any integer z_0, called a "seed," between 0 and P. Generate new seeds using the following rule:

$$z_t = Az_{t-1} \pmod{P}, \tag{2}$$

where the notation "x (mod y)" means "the remainder of x when divided by y." Generate the sequence:

$$x_t = z_t / P. \tag{3}$$

Then x_t is "uniformly distributed" on the interval [0, 1], in the same way as is the tent map.

It turns out that this method creates pseudo random numbers which are much more "random-looking" than the tent map. This pseudo random number generator can be related to the tent map as follows. First, we modify the "tent" pattern in Figure 1 to the "diadic map" in Figure 3. This changes the "stretch and fold" action of the tent map to "stretch, cut, and stack," as illustrated in Figure 4. Second, we increase the number of teeth from two to 7^5. By this time, the graph of this map appears to "fill up" the space in the unit square, and is the reason why it appears to be much more random.

C. Logistic Map

Other chaotic maps are frequently mentioned. The logistic map is slightly more complex than the tent map. Again, select x_0 between 0 and 1, and generate the sequence of x_t according to:

$$x_t = Ax_{t-1}(1 - x_{t-1}), \tag{4}$$

where A is between 0 and 4. For small values of A, the system is stable and well behaved. But as the value of A approaches 4, the system becomes

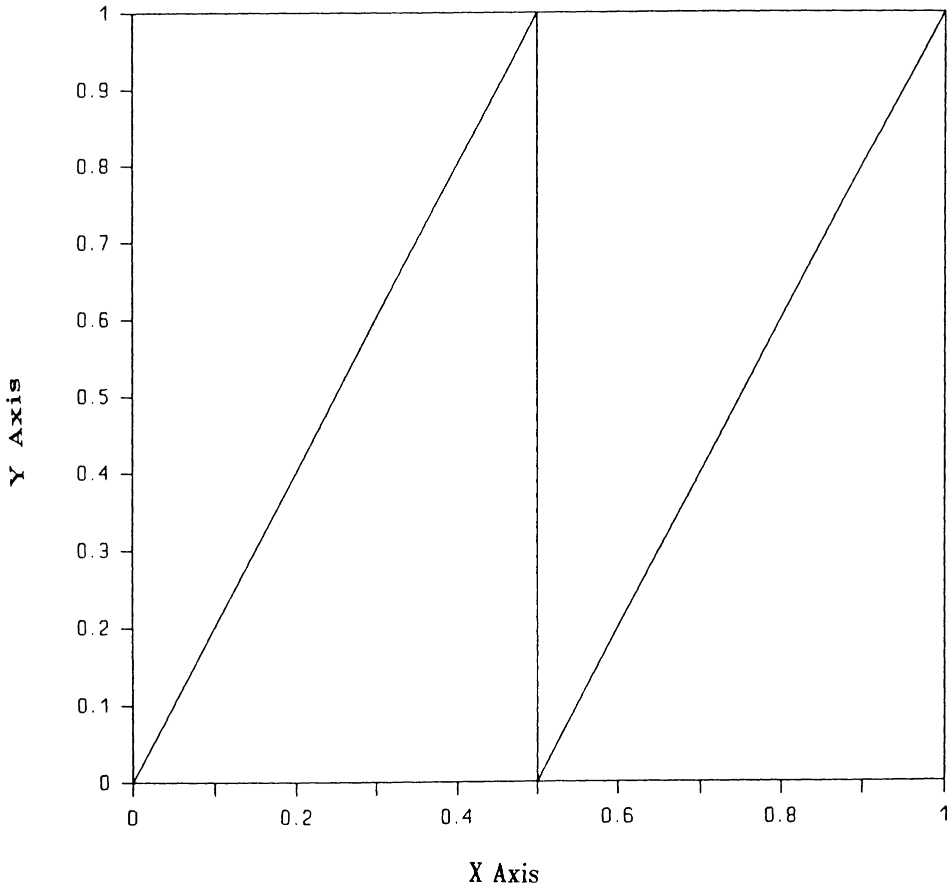

Figure 3. The diadic map. The diadic map is the graph of the function $y = 2x$ if $x < 0.5$, and $y = 2(x - 0.5)$ if $x \geq 0.5$, where x is the horizontal axis and y is the vertical axis.

chaotic. The logistic map adds a fifth property to chaotic behavior, that the dynamics of a system depends on a parameter (A in this case). For some values of the parameter, the dynamics may be simple, while for other values, the dynamics may be chaotic.

D. Hénon Map

Both the tent map and the logistic map are univariate chaotic systems. The Hénon (1976) map is a bivariate chaotic system, described by a pair of difference equations:

$$\begin{aligned} x_t &= y_{t-1} + 1 - Ax_{t-1}^2, \quad A = 1.4 \\ y_t &= Bx_{t-1}, \quad B = 0.3. \end{aligned} \tag{5}$$

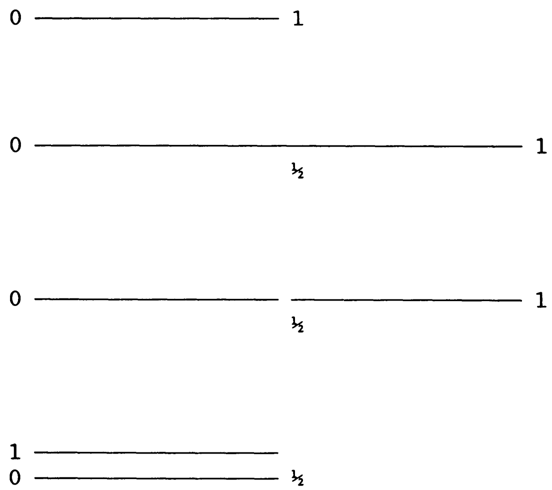

Figure 4. Stretch, cut, and stack action of the diadice map. The diadic map takes the unit interval [0,1], stretches it twice as long, cuts it in the middle, and stacks one piece on top of the other.

E. Lorenz Map

The Lorenz (1963) map is a trivariate chaotic system. Notice that it is a system of differential equations, rather than difference equations.

$$\begin{aligned} \dot{x} &= a(y - x), & a &= 10, \\ \dot{y} &= -y - xz - bx, & b &= 28, \\ \dot{z} &= xy - cz, & c &= 8/3. \end{aligned} \tag{6}$$

F. Mackey-Glass Equation

The above chaotic maps generate "low dimensional" chaos, which means that the nonlinear structure is easily detected, as we shall show later. There are, however, "high dimensional" chaotic systems which are much harder to detect, such as that in Mackey and Glass (1977). The Mackey-Glass equation is a delayed differential equation, given by:

$$\dot{x}(t) = \frac{ax(t-c)}{1 + x(t-c)^{10}} - bx(t) \quad a = 0.2, b = 0.1, c = 100. \tag{7}$$

G. General Chaotic Maps

There are many more examples of chaotic maps. In general, chaotic maps are obtained by a deterministic rule:

$$x_t = f(x_{t-1}, x_{t-2} \cdots). \tag{8}$$

Here, x_t can either be a scalar or a vector. In order to generate chaotic behavior, $f()$ must be a nonlinear function. Note, however, that nonlinearity alone is not sufficient to generate chaotic behavior. For example, $f(x) = x^3$ is a nonlinear map, but it is not chaotic.

II. Detecting Chaos

An important reason for the interest in chaotic behavior is that it can potentially explain fluctuations in the economy and financial markets which appear to be random. So there is need to test for the presence of chaos. We should, however, state clearly at the outset that we are interested only in low complexity chaotic behavior. If the world is truly governed by a highly complex chaotic process (e.g., an extremely good pseudo random number generator), we may never detect it using finite amounts of data. In this case, there is no practical difference between deterministic chaos and randomness. But if the world is governed by a not-too-complex chaotic process, it should have short-term predictability. However, traditional linear forecasting methods would not work; nonlinear models must be used.

How then can one test for low complexity chaos? Suppose we have a string of data, $x_1, x_2, \cdots, x_t, \cdots, x_T$. One method is to observe that chaotic maps do not "fill up" enough space in high dimension. To make this concrete, consider two sets of data: a_t is generated by the tent map, and b_t is a random variable which is uniform on the interval [0, 1]. If we plot a_t in one dimension, it is uniform over [0, 1], and so it fills up as much space as does b_t. However, consider the 2-vectors (a_{t-1}, a_t) and (b_{t-1}, b_t). If we plot them in two dimensions, the data from the tent map will fall on the tent, while the data from the uniform random variable will fall uniformly in the unit square $[0, 1] \times [0, 1]$. In other words, data from the tent map leave large "holes" in two dimensional space, while the random data do not.

When the chaotic process becomes more complex, we need to look at the data in higher dimensions. A chaotic process can fill up the first n dimensions, but leave large "holes" in the $(n + 1)$st dimension. Clearly it is not practical to do this type of graphical exercise in higher dimensions. Grassberger and Procaccia (1983) therefore developed the notion of correlation dimension. This is done in four steps.

Step 1: Remove autocorrelation, if present. Autocorrelation can affect some tests for chaos, so that we must remove it from the data. This is typically done by filtering the raw data using an autoregression, where the lag length is selected based on either the Akaike (1974) or Schwarz (1978) information criterion.

Step 2: From n-histories of the filtered data. These are denoted as follows:

$$\text{1-history: } x_t^1 = x_t.$$
$$\text{2-history: } x_t^2 = (x_{t-1}, x_t).$$
$$\vdots$$
$$n\text{-history: } x_t^n = (x_{t-n+1}, \cdots, x_t). \tag{9}$$

An n-history is a point in n-dimensional space; n is called the "imbedding dimension."

Step 3: Calculate the correlation integral:

$$C_n(\epsilon) = \lim_{T \to \infty} \# \{(t, s), 0 < t, s, < T : \| x_t^n - x_s^n \| < \epsilon\} / T^2, \quad (10)$$

where $\| \quad \|$ is the sup- or max- norm. In words, the correlation integral, $C_n(\epsilon)$, is defined as the fraction of pairs, (x_s^n, x_t^n), which are "close" to each other, in the sense that:

$$\max_{i=0,\cdots,n-1}\{ | x_{s-i} - x_{t-i} | \} < \epsilon.$$

Step 4: Calculate the slope of the graph of $\log C_n(\epsilon)$ versus $\log \epsilon$ for small values of ϵ. More precisely, we want to calculate the following quantity:

$$v_n = \lim_{\epsilon \to 0} \log C_n(\epsilon) / \log \epsilon. \quad (11)$$

If v_n does not increase with n, the data are consistent with chaotic behavior. In fact, the Grassberger-Procaccia correlation dimension is defined as:

$$v = \lim_{n \to \infty} v_n. \quad (12)$$

The meaning of the correlation dimension becomes clear when we consider the tent map. Since the tent map is uniformly distributed on the interval [0, 1], $C_1(\epsilon)$ doubles if ϵ doubles. Thus, for small values of ϵ,

$$v_1 = \log C_1(\epsilon) / \log \epsilon = 1. \quad (13)$$

But the 2-histories do not fill up the unit square [0, 1] × [0, 1]. In fact, all the points fall on the tent. For small values of ϵ, $C_2(\epsilon)$ doubles if ϵ doubles, and so

$$v_2 = \log C_2(\epsilon) / \log \epsilon = 1. \quad (14)$$

This continues to be true for any n, i.e.,

$$v_n = \log C_n(\epsilon) / \log \epsilon = 1. \quad (15)$$

So, for the tent map, the correlation dimension, v, is 1.

Next, apply this to data generated from the random variable uniformly distributed on the interval [0, 1]. Again, we would find that $C_1(\epsilon)$ doubles if ϵ doubles, so

$$v_1 = \log C_1(\epsilon) / \log \epsilon = 1. \quad (16)$$

However, since the 2-histories will uniformly fill up the unit square [0, 1] × [0, 1], $C_2(\epsilon)$ quadruples if ϵ doubles, and so

$$v_2 = \log C_2(\epsilon) / \log \epsilon = 2. \quad (17)$$

In fact,

$$v_n = \log C_n(\epsilon) / \log \epsilon = n. \quad (18)$$

For the random process, the correlation dimension, v, is ∞.

Using this method, Grassberger and Procaccia (1983) determine the correlation dimensions for the following chaotic systems: the logistic map, 1.00 ± 0.02, the Hénon map, 1.22 ± 0.01, the Lorenz map, 2.05 ± 0.01, and the Mackey-Glass equation, 7.50 ± 0.15. This shows that the chaotic maps do not fill up enough space at a sufficiently high imbedding dimension, which is a generic property of chaotic processes.[1]

It is important to remember that the correlation dimension is a measure of how much space is "filled up" by a string of data. That is why correlation dimensions need not be whole integers. We also need to point out that the correlation dimension is in no way related to the number of "independent factors" driving a system. To appreciate this point, note that the Mackey-Glass is a univariate process which has a correlation dimension around 7.

In principle, the four-step procedure to estimate correlation dimension sounds straightforward, and has been applied by scientists in many problems. In practice, however, a number of issues surface when dealing with economic and financial data. We shall discuss them in the context of the stock market.

III. What Do We Find in the Stock Market?

Scheinkman and LeBaron (1989) used the Grassberger-Procaccia plots and calculated the correlation dimension of weekly stock returns. They found that the slope of $\log C_n(\epsilon)$ versus $\log \epsilon$ appears to be around 6, even for dimensions as high as 25. They, however, noted that this is not sufficient evidence of chaos in stock returns, because there are a number of problems with this graphical procedure.

First, Scheinkman and LeBaron (1989) pointed out that some nonlinear stochastic model, such as Engle's (1982) autoregressive conditional heteroskedasticity (ARCH) model, exhibit "dependence" similar to that of chaotic maps. Using data from the ARCH model, they showed that the slopes of the graphs of $\log_n C(\epsilon)$ versus $\log \epsilon$ increase at a rate slower than n.

Second, there is no way to verify that a process has an infinite correlation dimension using a finite amount of data. Scientists typically use 100,000 or more data points to detect low dimensional chaotic system. Financial economists have substantially fewer points. The largest data sets generally have 2,000 observations. If we use the imbedding dimension of 10, we have only 200 nonoverlapping 10-histories. It is very hard to say whether 200 10-histories "fill up" a 10-dimensional space.

Third, we have to worry about biases in small data sets. Ramsey and Yuan (1989) show that the slope of the graph of $\log C_n(\epsilon)$ versus $\log \epsilon$ is biased downward in small data sets (even with as many as 2,000 observations). This biases the results in favor of finding chaos, even if there is none.

Fourth, the graphical procedure is not a statistical test. Ideally, we want a way to quantify the accuracy of the correlation dimension. This is not readily

[1] See the discussion in Brock (1986) and the proof in Takens (1980).

available for the correlation dimension plots, so we adopt a different method proposed by Brock, Dechert, and Scheinkman (1987).[2]

Before proceeding further, we digress here to deal with the naive view that there is no data limitation in finance since data are available at the tick by tick frequency. But this is merely an illusion. Tick by tick data capture bid-ask bounces and other dependencies which are caused by the micromarket structure, such as the sequential execution of limit orders on the books of the specialist as the market moves through those limit prices. These "artificial" dependencies will be picked up by any good test of nonlinear dynamics. The financial economist must increase the sampling interval in order to average out these "artificial" dependencies. Now, in order to obtain more observations, the researcher must look at longer histories, which runs into an entirely different problem. As one extends a data-set further and further back in time, nonstationarity (e.g., an unpredictable regime change) becomes increasingly more likely. As we shall see below, tests of nonlinear dynamics will detect nonstationarity. Thus, the requirements of long sampling intervals (to avoid micromarket structure dependencies) and short histories (to avoid nonstationarity) impose severe data limitations in finance. We now return from this digression to the Brock, Dechert, and Scheinkman (1987) statistic.

A. Statistical Test: The BDS Statistic

To deal with the problems of using the Grassberger-Procaccia plots, Brock, Dechert, and Scheinkman (1987) devised a statistical test. If $\{x_t: t = 1, \cdots, T\}$ is a random sample of independent and identically distributed (IID) observations, then:

$$C_n(\epsilon) = C_1(\epsilon)^n. \tag{19}$$

One can estimate $C_1(\epsilon)$ and $C_n(\epsilon)$ by the usual sample versions $C_{1,T}(\epsilon)$ and $C_{n,T}(\epsilon)$, and show that:

$$W_{n,T}(\epsilon) = \sqrt{T}\left[C_{n,T}(\epsilon) - C_{1,T}(\epsilon)^n\right]/\sigma_{n,T}(\epsilon) \tag{20}$$

has a limiting standard normal distribution. Here, $\sigma_{n,T}(\epsilon)$ is an estimate of the asymptotic standard error of $[C_{n,T}(\epsilon) - C_{1,T}(\epsilon)^n]$. We shall refer to $W_{n,T}(\epsilon)$ as the BDS statistic.

Note that the statement $C_n(\epsilon) = C_1(\epsilon)^n$ does not imply IID. Dechert (1988) has several counter examples.

Since the BDS statistic is a relatively new procedure, it is useful to study its finite sample distribution using Monte Carlo simulations. Some of them

[2] Denker and Keller (1986) and Brock and Baek (1991) provide ways to do this.

were reported in Hsieh and LeBaron (1988). The first set of results measure how well the asymptotic distribution approximates the finite sample distribution of the BDS statistic. We generate 1,000 IID observations (using a good pseudo random number generator), apply the BDS test, and repeat this 2,000 times. If we use a 5% significance level, we should reject 5% of the replications. Table I shows that the asymptotic distribution of the BDS test at dimension two is a reasonable approximation for IID data from four distributions (standard normal, Student t with 3 degrees of freedom, chi-square with 4 degrees of freedom, and Cauchy), when ϵ is set between one half to two standard deviations of the data. These distributions were selected for the following reasons: the standard normal is the base case; the Student t and the Cauchy have very fat tails; and the chi-square is strongly skewed. We also added two unusual distributions: the uniform and the bimodal, for which the asymptotic distribution of the BDS does not seem to fit the finite sample distribution. Fortunately, very little financial data look like these two distributions. Similar results are obtained for dimension five in Table II. We conclude that the BDS test avoids the biases of the correlation dimension estimates.

In the second set of simulations, we measure the ability of the BDS statistic to detect departures from IID. Given that there are uncountable ways to generate non-IID data, we select models which are interesting alternatives, and report the results in Table III.

The first two models represent time-series data with linear dependence. The AR1 is the first order autoregressive model, given by:

$$x_t = \rho x_{t-1} + u_t. \tag{21}$$

The MA1 is the first order moving average model, given by:

$$x_t = \theta u_{t-1} + u_t. \tag{22}$$

In the simulations, u_t is IID standard normal, $\rho = 0.5$, and $\theta = 0.5$. The point we wish to make here is that the BDS test can detect linear dependence easily. To employ BDS as a test for nonlinearity (whether chaotic or stochastic), we must remove any linear dependence in the data.[3]

The next two models represent data which violate the assumptions of strict stationarity and ergodicity. In the "2-mean" model, the data are independent and normally distributed, where the first 500 observations have mean -1 and variance 1, and the second 500 observations have mean $+1$ and variance

[3] It is not surprising to find that the BDS test has low power against the AR1 when ρ is small, say less than 0.2. If one is concerned about detecting linear dependence, it is best to use a test optimized for that alternative, such as the Durbin-Watson test, along with the BDS test. Since we are interested primarily in nonlinear dependence, we will just use a linear filter to remove any serial correlation.

Table I

Simulated Size of the BDS Statistic for Dimension 2

This table provides the percentage of BDS statistics (at dimension 2, ϵ equaling 0.25, 0.5, 1, 1.5, and 2 times the standard deviation of the data) rejecting the IID null hypothesis when it is true. The Monte Carlo simulation uses 2000 replications, each having 1000 observations, for six distributions: the standard normal, the Student t with 3 degrees of freedom, the chi-square with 4 degrees of freedom, the Cauchy, the uniform, and the bimodal distribution. $N(0, 1)$ denotes the percentage for a standard normal distribution.

	ϵ/σ					
	0.25	0.50	1.00	1.50	2.00	$N(0, 1)$
Standard normal						
% < −2.33	4.65	1.40	1.05	0.90	0.80	1.00
% < −1.96	8.95	3.25	2.90	2.45	2.65	2.50
% > 1.96	6.30	3.70	2.25	2.40	2.50	2.50
% > 2.33	3.60	1.55	0.90	0.70	0.90	1.00
$t(3)$						
% < −2.33	1.25	0.65	0.85	0.50	0.40	1.00
% < −1.96	3.25	2.50	2.20	2.05	1.20	2.50
% > 1.96	4.15	3.10	2.80	3.20	3.55	2.50
% > 2.33	1.90	1.50	1.10	1.45	1.80	1.00
$\chi^2(4)$						
% < −2.33	1.65	0.90	1.10	1.20	1.10	1.00
% < −1.96	5.00	3.05	3.00	3.35	2.45	2.50
% > 1.96	5.05	3.80	3.85	3.90	3.70	2.50
% > 2.33	3.25	2.10	1.90	1.65	2.15	1.00
Cauchy						
% < −2.33	0.40	0.35	0.45	0.70	0.75	1.00
% < −1.96	1.70	0.80	1.00	1.15	0.90	2.50
% > 1.96	4.30	4.75	4.75	4.55	4.90	2.50
% > 2.33	2.30	2.85	3.30	3.35	3.55	1.00
Uniform						
% < −2.33	44.95	21.75	1.45	1.40	1.40	1.00
% < −1.96	46.05	26.60	3.60	3.00	3.15	2.50
% > 1.96	42.45	24.30	5.05	2.85	2.85	2.50
% > 2.33	41.40	21.80	3.10	1.30	1.25	1.00
Bimodal						
% < −2.33	2.30	2.55	52.70	1.40	1.10	1.00
% < −1.96	5.45	5.00	54.70	3.85	3.10	2.50
% > 1.96	6.45	5.75	29.20	3.85	3.05	2.50
% > 2.33	3.45	3.05	28.10	2.05	1.40	1.00

Note: Approximate standard error is 1.12 for these probabilities.

1. In the "2-variance" model, the data are also independent and normally distributed, where the first 500 observations have mean 0 and variance 1, and the second 500 observations have mean 0 and variance 2. These models are examples of "structural changes" or "regime changes." Table III shows that BDS also has no trouble in detecting them.

Table II

Simulated Size of the BDS Statistic for Dimension 5

This table provides the percentage of BDS statistics (at dimension 5, ϵ equaling 0.25, 0.5, 1, 1.5, and 2 times the standard deviation of the data) rejecting the IID null hypothesis when it is true. The Monte Carlo simulation uses 2000 replications, each having 1000 observations, for six distributions: the standard normal, the Student t with 3 degrees of freedom, the chi-square with 4 degrees of freedom, the Cauchy, the uniform, and the bimodal distribution. $N(0, 1)$ denotes the percentage for a standard normal distribution.

	ϵ/σ					
	0.25	0.50	1.00	1.50	2.00	$N(0, 1)$
Standard normal						
% < −2.33	29.85	3.60	0.55	0.80	0.75	1.00
% < −1.96	32.75	7.40	2.35	2.35	2.55	2.50
% > 1.96	29.65	8.15	2.90	2.40	2.50	2.50
% > 2.33	26.95	5.30	1.40	1.10	1.30	1.00
$t(3)$						
% < −2.33	6.05	0.70	0.70	0.85	0.60	1.00
% < −1.96	9.55	2.25	2.30	2.55	2.50	2.50
% > 1.96	11.00	4.20	3.10	3.50	3.55	2.50
% > 2.33	7.55	2.25	1.95	1.70	1.60	1.00
$\chi^2(4)$						
% < −2.33	16.15	0.95	0.85	0.85	1.00	1.00
% < −1.96	20.65	3.45	2.30	2.25	2.30	2.50
% > 1.96	19.20	5.30	3.45	3.30	3.00	2.50
% > 2.33	15.40	2.95	1.75	1.25	1.50	1.00
Cauchy						
% < −2.33	0.70	0.55	0.90	0.85	1.20	1.00
% < −1.96	2.00	1.70	1.90	1.45	1.40	2.50
% > 1.96	3.50	3.80	4.40	4.75	4.50	2.50
% > 2.33	1.35	1.90	3.10	3.30	3.00	1.00
Uniform						
% < −2.33	49.20	35.50	4.05	1.50	1.30	1.00
% < −1.96	49.60	37.40	7.55	3.00	2.95	2.50
% > 1.96	48.05	38.50	6.85	3.75	3.40	2.50
% > 2.33	47.90	36.85	4.30	1.55	1.25	1.00
Bimodal						
% < −2.33	15.45	7.25	46.00	2.50	1.45	1.00
% < −1.96	20.00	10.80	47.20	5.70	3.30	2.50
% > 1.96	17.85	10.25	41.05	5.30	2.70	2.50
% > 2.33	13.65	6.85	39.95	2.80	1.40	1.00

Note: Approximate standard error is 1.12 for these probabilities.

We consider two nonlinear time-series models which have no autocorrelation but non-zero conditional means. Robinson (1977) proposed the nonlinear moving average (NMA) model:

$$x_t = u_t + \alpha u_{t-1} u_{t-2}, \tag{23}$$

Table III

Simulated Power of the BDS Statistic

This table provides the percentage of BDS statistics (at dimensions 2 through 5, ϵ equaling 0.5, 1, 1.5, and 2 times the standard deviation of the data) rejecting the IID null hypothesis when it is false. The Monte Carlo simulation uses 2000 replications, each having 1000 observations, for 11 non-IID alternatives: the first order autoregression (AR1), the first order moving average (MA1), the '2-mean' model (the first 500 observations have mean −1 and variance 1, the second 500 observations have mean 1 and variance 1), the '2-variance' model (the first 500 observations have mean 0 and variance 1, the second 500 observations have mean 0 and variance 2), the nonlinear moving average (NMA), the threshold autoregression (TAR), the autoregressive conditional heteroskedasticity (ARCH) model, the generalized autoregressive conditional heteroskedasticity (GARCH) model, the exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model, the Mackey-Glass data filtered by an autoregression of order 3, and the "Sine" model.

		ϵ/σ			
	m	0.50	1.00	1.50	2.00
AR1	2	100.00	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00
	4	100.00	100.00	100.00	100.00
	5	100.00	100.00	100.00	100.00
MA1	2	100.00	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00
	4	99.95	100.00	100.00	100.00
	5	99.95	100.00	100.00	100.00
2-mean	2	100.00	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00
	4	100.00	100.00	100.00	100.00
	5	100.00	100.00	100.00	100.00
2-variance	2	98.75	98.30	98.10	92.65
	3	99.00	100.00	99.95	99.80
	4	100.00	100.00	99.95	100.00
	5	100.00	100.00	99.95	100.00
NMA	2	99.95	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00
	4	100.00	100.00	100.00	100.00
	5	100.00	100.00	100.00	100.00
TAR	2	99.95	99.40	91.20	62.85
	3	99.90	98.80	89.20	61.75
	4	99.30	96.95	84.25	56.65
	5	97.30	94.50	78.00	49.10
ARCH	2	100.00	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00
	4	100.00	100.00	100.00	100.00
	5	100.00	100.00	100.00	100.00
GARCH	2	74.60	78.05	81.80	82.10
	3	88.10	91.80	93.85	94.35
	4	92.25	95.65	96.95	97.30
	5	93.50	97.05	97.75	98.25

Table III—*Continued*

	m	ϵ/σ 0.50	1.00	1.50	2.00
EGARCH	2	22.50	21.20	20.90	20.15
	3	32.05	31.10	30.15	28.65
	4	38.50	36.40	34.30	33.20
	5	43.10	39.85	37.70	36.65
Mackey-Glass	2	100.00	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00
	4	100.00	100.00	100.00	100.00
	5	100.00	100.00	100.00	100.00
Sine	2	100.00	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00
	4	100.00	100.00	100.00	100.00
	5	100.00	100.00	100.00	100.00

where u_t is IID standard normal.[4] In the simulations, $\alpha = 0.5$. The other nonlinear time-series model is the threshold autoregressive (TAR) model in Tong and Lim (1980):

$$x_t = \alpha x_{t-1} + u_t, \quad \text{if } x_{t-1} < 1,$$
$$x_t = \beta x_{t-1} + u_t, \quad \text{if } x_{t-1} \geq 1, \tag{24}$$

where u_t is IID standard normal. In the simulations, $\alpha = -0.4$ and $\beta = 0.5$. Table III shows that BDS can detect the nonlinearity in both the NMA and the TAR.

Next, we examine nonlinear time-series models, with no autocorrelation and zero conditional means, that exhibit conditional heteroskedasticity. As discussed earlier, Engle (1982) presented the autoregressive conditional heteroskedasticity (ARCH) model:

$$x_t = \sigma_t u_t,$$
$$\sigma_t^2 = \phi_0 + \phi x_{t-1}^2. \tag{25}$$

In our simulations, $\phi_0 = 1$ and $\phi_1 = 0.5$. Bollerslev (1986) turned ARCH into Generalized ARCH (GARCH) by making σ_t a function of its own past:

$$\sigma_t^2 = \phi_0 + \phi x_{t-1}^2 + \psi \sigma_{t-1}^2. \tag{26}$$

In our simulations, $\phi_0 = 1$ and $\phi = 0.1$, and $\psi = 0.8$. Nelson (1991) changed GARCH into exponential GARCH (EGARCH) by using $\log \sigma_t^2$ instead of σ_t^2.

$$\log \sigma_t^2 = \phi_0 + \phi \,|\, x_{t-1}/\sigma_{t-1} \,|\, + \psi \log \sigma_{t-1}^2 + \gamma x_{t-1}/\sigma_{t-1}. \tag{27}$$

[4] The nonlinear moving average is very similar to the bilinear model in Granger and Andersen (1978).

In our simulations, $\phi_0 = 1$ and $\phi = .1$, $\psi = .8$, and $\gamma = 0.1$. Unlike simple ARCH and GARCH, EGARCH is able to capture asymmetric response of the variance to the direction of x_t, e.g., a higher variance when x_t is negative, and a lower variance when x_t is positive, a phenomenon noted by Black (1976). We refer to all three as "ARCH-type" models. These models have enjoyed a great deal of attention in the econometric literature, particularly in applications to financial time series.[5] Table III shows that BDS can easily detect the simple ARCH and GARCH models, but has trouble detecting EGARCH.

For a chaotic (i.e., nonlinear deterministic) process, we use the Mackey-Glass equation. The results for the tent map, logistic map, and Hénon maps are similar, and available upon request. The Mackey-Glass is chosen, because it has the highest correlation dimension (7.5) among this group of chaotic processes, making it the most difficult to detect. In addition, its correlation dimension is similar to that of weekly stock returns as measured by Scheinkman and LeBaron (1989). To remove any evidence of linear dependence, we filter the data using an autoregression with three lags. Table III shows that BDS has no trouble in picking up the nonlinear dependence in the (filtered) Mackey-Glass data. (We will discuss the "sine" model later.)

The third set of simulation addresses the issue of "nuisance" parameters. We have already pointed out that we must remove any linear dependence from our data before applying the BDS test for nonlinearity. The question is: will linear filtering change either the asymptotic or the finite sample distribution of the test statistic? Brock (1987) proves that the asymptotic distribution of the BDS test is not altered by using residuals instead of raw data in linear models. In fact, Brock's theorem can be extended to residuals of some nonlinear models (such as the nonlinear moving average), but not to ARCH models. This is confirmed by the simulations in Table IV. The results show that the asymptotic distribution still approximates the finite sample distribution with the same degree of accuracy even when replacing raw data with residuals of the AR1, the MA1, and the NMA. The results also show that the BDS test may reject too infrequently in the case of standardized residuals from GARCH and EGARCH models.

B. Application to Stock Returns

We now apply the BDS test of IID to stock market returns. Our data are weekly stock returns provided by Peter Rossi using the data from the Center for Research in Securities Prices (CRSP) at the University of Chicago, beginning in 1963 and ending in 1987. These data have been carefully constructed to include dividends as well as capital gains, and they have also been made into different portfolios. We examine a value-weighted index (VW) and an equally weighted index (EW).[6] In addition, we use ten value-

[5] See the survey article by Bollerslev et al. (1990).

[6] We also examined the value and equally weighted indices in excess of a Treasury bill return. The results did not differ from the raw indices, and so were not reported.

Table IV

Simulated Size of the BDS Statistics for Residuals

This table provides the percentage of BDS statistics (at dimensions 2, ϵ equaling 0.5, 1.0, 1.5, and 2 times the standard deviation of the data) rejecting the IID null hypothesis when applied to residuals. The Monte Carlo simulation uses 2000 replications, each having 1000 observations, for data generated by 5 non-IID models: the first order autoregression (AR1), the first order moving average (MA1), the nonlinear moving average (NMA), the generalized autoregressive conditional heteroskedasticity (GARCH) model, and the exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model. $N(0, 1)$ denotes the percentage for a standard normal distribution.

	ϵ/σ				
	0.50	1.00	1.50	2.00	$N(0, 1)$
AR1 residuals ($\rho = 0.5$)					
% < −2.33	1.20	1.10	1.15	1.20	1.00
% < −1.96	3.25	2.90	2.60	2.65	2.50
% > 1.96	4.50	3.25	3.25	3.70	2.50
% > 2.33	1.90	1.65	1.20	1.50	1.00
MA1 residuals ($\theta = 0.5$)					
% < −2.33	1.30	1.00	1.25	1.20	1.00
% < −1.96	3.10	2.65	2.60	3.05	2.50
% > 1.96	4.40	3.20	3.30	3.75	2.50
% > 2.33	1.90	1.70	1.30	1.50	1.00
NMA residuals ($\alpha = 0.05$)					
% < −2.33	1.60	1.10	1.25	1.00	1.00
% < −1.96	3.75	2.70	2.90	3.00	2.50
% > 1.96	4.40	3.25	3.50	3.75	2.50
% > 2.33	2.05	1.70	1.65	1.95	1.00
GARCH standardized residuals ($\phi = 0.1$, $\psi = 0.8$)					
% < −2.33	0.40	0.30	0.20	0.25	1.00
% < −1.96	1.55	0.95	0.80	0.95	2.50
% > 1.96	1.80	1.15	0.80	0.45	2.50
% > 2.33	0.90	0.20	0.10	0.05	1.00
EGARCH standardized residuals ($\phi = 0.1$, $\psi = 0.8$)					
% < −2.33	0.20	0.00	0.00	0.05	1.00
% < −1.96	0.80	0.40	0.25	0.40	2.50
% > 1.96	3.35	2.50	1.85	1.95	2.50
% > 2.33	1.75	0.90	0.85	0.55	1.00

weighted decile portfolios in which firms are ranked by size every quarter. Results are reported for the first (smallest), fifth, and tenth (largest) decile portfolios, called DEC1, DEC5, and DEC10.[7] All data were first filtered by an autoregression whose lag length was determined by the Akaike (1974) information criterion.[8]

[7] The results are the same using equally weighted decile portfolios.

[8] The Schwarz (1978) information criterion was also used. The lags identified by the Akaike (Schwarz) information criterion are: VW 1 (1), EW 2 (1), DEC1 7 (1), DEC5 2 (1), DEC10 1 (0). Since there are large numbers of degress of freedom in our data, we used the longer lags identified by the Akaike information criterion.

Table V contains some descriptive statistics of these filtered series. The filtering procedure removes any nonzero mean from the data. (The means in the raw data are small to begin with.) The main point to note in this table is that all series are leptokurtic—with the coefficients of kurtosis much larger than 3—a fact which is well known.

Table VI gives the results of the BDS tests. They strongly reject the hypothesis that stock returns are IID. This is true for the market as a whole, as well as the decile portfolios.

What are the implications of the finding that stock returns are not IID? First, it does not contradict market efficiency. Market efficiency implies that forecast errors of returns are not predictable. The fact that returns themselves are not IID (and therefore *potentially* predictable) says nothing about the predictability of forecast errors.

Second, when returns are not IID, it is difficult to interpret unconditional density estimation. A number of studies have fit leptokurtic distributions to stock returns. For example, Blattberg and Gonedes (1974) found that the Student t distribution provides a better fit to stock returns than the symmetric stable paretian distribution of Mandelbrot (1963). Since both the stable paretian and the Student t are leptokurtic, the probability of observing large returns (in absolute values) is much higher than that from the normal distribution. One may therefore be tempted to "explain" crashes, such as that on October 19, 1987, as small but nonzero probability events.[9] The fact that returns are not IID,[10] however, makes this explanation for stock market crashes less plausible, because unconditional distributions will always have fatter tails than conditional distributions when the data have some type of conditional dependence.

Third, the rejection of IID does not provide direct evidence of chaos in the stock market. Our simulations in Table III show that BDS has good power to detect at least four types of non-IID behavior: linear dependence, nonstationarity, chaos, and nonlinear stochastic processes. We can rule out linear dependence, since there is little of it in the raw returns, and since we have removed whatever correlation there is by filtering the return series. We therefore concentrate on the remaining three causes.

The rejection of IID is consistent with the view that stock returns are nonstationary. Over a long time period, it is difficult to make a case that the behavior of stock returns remains unchanged. Changes in economic fundamentals, e.g., wars, can shift the mean return (represented by the "2-mean" model). Changes in the operating procedure of the Federal Reserve, e.g.,

[9] Table 1 in Fama and Roll (1968) shows that the probability of observing an outcome in excess of 6 standardized units is 5.36% for the Cauchy distribution, compared to almost 0% for the normal distribution. In fact, the probability of an outcome in excess of 20 standardized units is 1.59% for the Cauchy distribution!

[10] Note that the Cauchy distribution is a member of the stable paretian family. The simulations in Table I show that the asymptotic distribution of the BDS statistic can still approximate the finite distribution well, even though the Cauchy distribution has no moments.

Table V

Selective Statistics for Filtered Stock Returns

This table presents the mean, standard deviation, skewness, kurtosis, maximum, minimum, and the number of observations for 11 stock returns: the weekly value weighted portfolio (VW), the weekly equally weighted portfolio (EW), the weekly smallest decile portfolio (DEC1), the weekly fifth decile portfolio (DEC5), the weekly largest decile portfolio (DEC10), the weekly S&P500 index (SPW), the daily S&P500 index (SPD), and the 15-minute S&P500 indices for the first, second, third, and fourth quarter of 1988 (SPM1, SPM2, SPM3, and SPM4, respectively).

	Mean	Std Dev	Skewness	Kurtosis	Maximum	Minimum	No. of Observations
VW	0.0000	0.0202	−0.292	6.69	0.0875	−0.1538	1303
EW	0.0000	0.0221	0.049	9.01	0.1459	−1.1680	1302
DEC1	0.0000	0.0257	0.697	10.32	0.1797	−0.1704	1297
DEC5	0.0000	0.0239	0.001	7.13	0.1360	−0.1618	1302
DEC10	0.0000	0.0202	−0.240	6.09	0.0800	−0.1452	1303
SPW	0.0000	0.0203	−0.309	6.47	0.1243	−0.1321	1402
SPD	0.0000	0.0113	−4.275	88.99	0.0709	−0.2279	2017
SPM1	0.0000	0.00218	0.402	23.15	0.0258	−0.0158	1706
SPM2	0.0000	0.00170	−0.709	23.74	0.0150	−0.0195	1706
SPM3	0.0000	0.00138	0.139	13.60	0.0137	−0.0110	1706
SPM4	0.0000	0.00123	−0.501	11.20	0.0064	−0.0103	1707

switching from an interest rate to a money supply target during 1979–1982, can shift the volatility of financial markets (represented by the "2-variance" model).

The rejection of IID is also consistent with the view that returns are generated by nonlinear stochastic systems, e.g., NMA, TAR, and ARCH-type models. While there are few models in economics and finance which lead to nonlinear stochastic systems of these specific types, this observation does not imply that nonlinear stochastic models are not useful. The nonlinear moving average model can be regarded as a second order approximation of the Volterra representation, which all stationary (linear or nonlinear) time series possess. The threshold autoregressive process can result from an endogenous regime switching model.[11] The ARCH-type model can be thought of as approximating conditional variance changes.[12]

Finally the rejection of IID is also consistent with the presence of low complexity chaotic behavior in stock returns. Regardless of whether determinism is aesthetically appealing or not, there are many ways to generate economic models with chaotic dynamics, summarized by Baumol and Benhabib (1989). If a system is both chaotic and stochastic, we shall classify it (arbitrarily) as a stochastic system. What remains for us to do is to try to eliminate two of the three explanations for non-IID behavior of stock returns.

[11] See Hsieh (1990).
[12] See Nelson (1990) for a discussion.

Table VI

BDS Statistics for Filtered Stock Returns

This table presents the BDS statistics (at dimensions 2 through 5 and ϵ equaling 0.5, 1, 1.5, and 2 standard deviations of the data) for 11 stock returns: the weekly value-weighted portfolio (VW), the weekly equally weighted portfolio (EW), the weekly smallest decile portfolio (DEC1), the weekly fifth decile portfolio (DEC5), the weekly largest decile portfolio (DEC10), the weekly S&P500 index (SPW), the daily S&P500 index (SPD), and the 15-minute S&P500 indices for the first, second, third, and fourth quarter of 1988 (SPM1, SPM2, SPM3, and SPM4 respectively).

		ϵ / σ			
	m	0.50	1.00	1.50	2.00
VW	2	8.73	7.33	7.13	8.05
	3	13.32	10.31	9.21	9.53
	4	17.26	12.14	10.34	10.13
	5	22.32	14.25	11.49	10.71
EW	2	9.48	9.03	8.42	8.37
	3	14.15	11.95	10.52	9.69
	4	17.44	13.57	11.52	10.22
	5	21.88	15.43	12.55	10.84
DEC1	2	10.84	11.21	11.26	11.29
	3	13.68	13.53	12.56	11.44
	4	15.84	15.02	13.24	11.41
	5	18.72	16.45	13.82	11.49
DEC5	2	9.19	8.92	8.24	8.02
	3	13.13	11.78	10.43	9.78
	4	15.98	13.39	11.29	10.11
	5	19.38	15.14	12.24	10.75
DEC10	2	8.69	7.37	7.29	8.40
	3	13.06	10.12	9.18	9.62
	4	17.00	12.09	10.42	10.38
	5	21.87	14.23	11.61	11.02
SPW	2	8.78	9.04	8.92	8.67
	3	11.33	10.69	10.19	9.78
	4	14.53	12.74	11.57	10.73
	5	18.24	14.59	12.71	11.49
SPD	2	2.16	4.21	6.97	9.14
	3	2.28	4.86	8.15	10.37
	4	2.60	5.39	8.98	11.29
	5	2.93	5.96	9.69	12.04
SPM1	2	7.56	6.56	5.86	4.86
	3	10.75	8.51	7.46	6.59
	4	14.14	10.71	8.90	7.79
	5	17.25	11.69	9.21	7.99
SPM2	2	8.74	7.56	5.96	3.96
	3	11.81	9.22	7.17	4.75
	4	14.82	10.11	7.30	4.82
	5	18.00	11.12	7.69	5.07
SPM3	2	9.23	6.47	4.21	2.16
	3	12.10	8.04	5.50	3.28
	4	15.22	9.01	5.88	3.71
	5	18.46	10.09	6.68	4.40

Table VI—*Continued*

	m	ϵ/σ 0.50	1.00	1.50	2.00
SPM4	2	10.10	8.57	5.57	3.09
	3	14.30	11.91	8.34	5.47
	4	18.67	14.38	9.66	6.27
	5	23.49	16.19	10.45	6.74

IV. Is Nonstationarity Responsible for the Rejection of IID?

For financial economists, nonstationarity is synonymous with structural change. There may be many reasons for structural changes: technological and financial innovations, policy changes, etc. It would be difficult to argue that the structure of the economic and financial system has remained constant from 1963 to 1987. We must allow for the possibility that structural changes caused BDS to reject IID during this period.

In order to check this explanation, we look at the returns of the Standard & Poors 500 stock index (without dividends) for the following time periods: weekly returns from 1962 to 1989 (SPW), daily returns from 1983 to 1989 (SPD), and 15-minute returns during 1988 divided into 4 approximately equal subsamples (SPM1, SPM2, SPM3, SPM4).[13] Implicitly, we are assuming that structural changes occur infrequently. By going to higher and higher frequency data in shorter and shorter time periods, we should remove the effects of structural changes. But we stop well short of using tick by tick data to avoid picking up micromarket structure dependencies discussed in Section III.

Table VI indicates that the weekly S&P returns is not IID, the same as the value-weighted index over the same period. What is more interesting, however, is that the daily returns in 1982–1989 and the 15-minute returns in 1988 are also not IID.[14] This makes it unlikely that infrequent structural changes are causing the rejection of IID in weekly returns. It is, of course, possible that structural changes happen so frequently that they cause BDS to reject IID in the 15-minute returns over the course of 3 months. If this is the true, then econometric work on economic and financial data is virtually impossible.

[13] These are logarithmic differences of price changes. They are filtered by an autoregression whose lags are chosen by the Akaike (Schwarz) criterion to be: weekly returns, 6 (0), daily returns, 5(0), and 15-minute returns, 4 (1). Since we have a large number of degrees of freedom, we use the longer lag lengths.

[14] It is possible that the 15-minute return is capturing some nonlinear dynamics from the micromarket structure. This will have to be studied in the future. We checked that day-of-the-week and time-of-day effects are not responsible for the rejection of IID in the daily and 15-minute data.

V. Is Chaotic Dynamics Responsible for the Rejection of IID?

The rejection of IID is certainly consistent with the hypothesis that the stock market is governed by low complexity chaotic dynamics. The issue we raise here is—is there any direct evidence of chaotic behavior? In this section, we take two approaches to answer this question.

The first approach examines the unconditional third order moments of stock returns, following the method in Hsieh (1989). The motivation is as follows. If x_t is a chaotic process, it can be written as:

$$x_t = f(x_{t-1}, \ldots). \tag{28}$$

This is a special case of a more general category of nonlinear processes:

$$x_t = f(x_{t-1}, \ldots) + \epsilon_t, \tag{29}$$

where ϵ_t satisfies the condition that $E[\epsilon_t \mid x_{t-1}, \ldots] = 0$. For both models, we can consider $f()$ as the mean of x_t conditional on its own past. Since $f()$ is nonlinear, these models are "nonlinear-in-mean" (as opposed to "nonlinear-in-variance," which will be discussed later).

We can test for the null hypothesis that $f() = 0$ against the alternative that $f() \neq 0$. Under the null, the unconditional third order moments, $E[x_t x_{t-i} x_{t-j}] = 0$, for $i, j > 0$. Hsieh (1989) proposes the following test:

a) Define $\rho(i, j) = E[x_t x_{t-i} x_{t-j}]/\sigma^3$, where $\sigma^2 = V[x_t]$. Estimate $\rho(i, j)$ with the appropriate sample moments: $r(i, j) = [\sum x_t x_{t-i} x_{t-j} / T] / [\sum x_t^2 / T]^{1.5}$.

b) $\sqrt{[r(i, j) - \rho(i, j)]}$ has a limiting distribution $N(O, V(i, j))$, where $V(i, j)$ can be estimated by the method of moments:

$$\left[\sum\{x_t x_{t-i} x_{t-j}/T - r(i,j)\}^2\right] / \left[\sum x_t^2/T\right]^3.$$

While Hsieh (1989) tests $\rho(i, j) = 0$ individually using a t-statistic, we test the composite null hypothesis that $\rho(i, j) = 0$ for $0 < i \leq j \leq m$, for a given m, making use of the fact that the asymptotic covariance between $r(i, j)$ and $r(i', j')$ can be estimated using the obvious sample cross-moments:

$$\left[\sum\{x_t x_{t-1} x_{t-j}/T - r(i,j)\}\{x_t x_{t-i'} x_{t-j'}/T - r(i',j')\}\right] / \left[\sum x_t^2/T\right]^3.$$

The composite test can be conducted using the usual χ^2 statistic.[15]

This test statistic is designed so that it will not reject models which are "nonlinear-in-variance":

$$x_t = g(x_{t-1}, \ldots)\epsilon_t, \tag{30}$$

[15] The proof of this statement follows easily from Hsieh (1989), which can be viewed as a modification of Tsay (1986). We should note that the third order moment test can fail to detect a chaotic process whose odd product moments are zero. This can happen if the function f() is antisymmetric. This is not true for any of the chaotic examples in this paper.

where $g()$ is a nonlinear function. Since x_t and ϵ_t take on both positive and negative values, we cannot take logarithms of both sides and transform this model to one being nonlinear-in-mean. However, the third order moment test should detect hybrid models, those which are "nonlinear-in-mean" as well as "nonlinear-in-variance":

$$x_t = f(x_{t-1}, \dots) + g(x_{t-1}, \dots)\epsilon_t. \tag{31}$$

(The GARCH-in-mean model, where the conditional variance appears in the conditional mean, is such an example.)

As in the case of the BDS test, we perform simulations to evaluate the finite sample distribution of the third order moment test as well as its ability to detect nonlinearity-in-mean. The results are reported in Table VII.

The first 4 models use IID data generated by the standard normal, Student t with 3 degrees of freedom, Cauchy, and the chi-square with 4 degrees of freedom. They show that the asymptotic distribution of the third order moment test approximates the finite sample distribution for 1000 observations tolerably well for IID data generated by the standard normal and the $\chi^2(4)$, but rather poorly for the $t(3)$ and the Cauchy. The latter two distributions do not have fourth or higher moments, which are assumed to exist in the derivation of the asymptotic distribution of the third order moment test statistic. Thus care must be used when applying the third order moment test to very fat tailed data.[16]

The next 5 models have non-IID data, but do not have nonlinearity-in-mean. There is a slight tendency for the test to reject too infrequently. This is more so for the AR1, MA1, and MA1, and 2-mean, and less so for the 2-variance and EGARCH.

The next 4 models (NMA, TAR, filtered Mackey-Glass, and GARCH-in-mean), generate non-IID data which have nonlinearity-in-mean and nonzero third order moments. The third order moment test can detect the first 3 models nearly 100% of the time, but the power against the GARCH-in-mean model is low, probably because the high order moments of the GARCH-in-mean model do not exist.[17]

The last simulation uses the "sine" model which is nonlinear-in-mean but has zero third order moments:

$$x_t = \sin[x_{t-1}] + \epsilon_t, \tag{32}$$

where ϵ_t is IID standard normal. The simulation shows that the third order moment test, as expected, cannot pick up the nonlinearity in this model. The reason is quite simple. If the conditional means are zero, then the third order moments are zero. However, the converse is not true: if the third order

[16] The failure of existence of fourth moments can also affect the distribution of the Tsay (1986) nonlinearity test and the Hinich and Patterson (1985) nonlinearity test. It is difficult to test whether fourth moments exist in a finite data set. This points to one of the advantages of the BDS test, whose limiting distribution does not require the existence of any moments.

[17] We also reject 100% of the replications using the tent map, the Hénon map, and the logistic map (when $A = 4$).

Table VII

Simulated Size and Power of the Third Order Moment Test

This table presents the percentage of the third order moment test (using 5 lags). The Monte Carlo uses 2000 replications, each having 1000 observations, for data generated by: (a) nine models for which the conditional mean is zero or linear: the standard normal, the Student t with 3 degrees of freedom, the chi-square with 4 degrees of freedom, the Cauchy distribution, the first order moving average (MA1), the first order autoregression (AR1), the '2-mean' model (the first 500 observations have mean −1 and variance 1, the second 500 observations have mean 1 and variance 1), the '2-variance' model (the first 500 observations have mean 0 and variance 1, the second 500 observations have mean 0 and variance 2), and the exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model; and (b) five models for which the conditional mean is nonlinear: the threshold autoregression (TAR), the nonlinear moving average (NMA), the Mackey-Glass filtered by a third order autoregression, the generalized autoregressive conditional heteroskedasticity in mean model, and the "Sine" model. The "true size" denotes the percentage of rejections under the null hypothesis.

Test Statistic	Model					True Size
$\chi^2(15)$	$N(0, 1)$	$t(3)$	$\chi^2(4)$	Cauchy		
% > 22.31	9.05	6.75	12.05	18.25		10.00
% > 25.00	3.70	2.85	6.15	12.70		5.00
% > 27.49	1.70	1.15	3.05	9.60		2.50
% > 30.58	0.65	0.25	1.15	8.00		1.00
% > 32.80	0.35	0.05	0.70	7.25		0.50
	Model					
$\chi^2(15)$	MA1	AR1	2-mean	2-var	EGARCH	True Size
% > 22.31	5.60	5.40	4.90	7.95	8.55	10.00
% > 25.00	3.10	2.60	2.65	3.15	3.30	5.00
% > 27.49	1.45	1.95	1.60	1.25	1.55	2.50
% > 30.58	0.45	0.70	0.90	0.30	0.60	1.00
% > 32.80	0.15	0.40	0.55	0.20	0.40	0.50
	Model					
$\chi^2(15)$	TAR	NMA	Mackey-Glass	GARCH-M	Sine	True Size
% > 22.31	100.00	100.00	100.00	74.00	7.20	10.00
% > 25.00	100.00	100.00	100.00	61.85	4.00	5.00
% > 27.49	100.00	100.00	100.00	49.65	2.20	2.50
% > 30.58	100.00	99.85	100.00	36.35	0.80	1.00
% > 32.80	100.00	99.75	100.00	28.45	0.35	0.50

moments are zero, it does not imply that the conditional means are zero. This was first pointed out by Pemberton and Tong (1981). Note, however, that BDS has no trouble in detecting this type of nonlinearity. (See the results in Table III.)

We now apply the third order moment test to stock returns. Table VIII shows that there is no evidence to reject the null hypothesis that stock

Table VIII

Third Order Moments Test Statistics for Filtered Stock Returns

This table presents the third order moment test statistics (using 5 lags) for 11 stock returns: the weekly value-weighted portfolio (VW), the weekly equally weighted portfolio (EW), the weekly smallest decile portfolio (DEC1), the weekly fifth decile portfolio (DEC5), the weekly largest decile portfolio (DEC10), the weekly S&P500 index (SPW), the daily S&P500 index (SPD), and the 15-minute S&P500 indices for the first, second, third, and fourth quarter of 1988 (SPM1, SPM2, SPM3, and SPM4, respectively). These statistics are asymptotically $\chi^2(15)$, whose critical values (tail probabilities) are: 22.31 (10%), 25.00 (5%), 27.49 (2.5%), 30.58 (1%), and 32.80 (0.5%).

Stock Returns	Third Order Moment Statistic
VW	14.25
EW	15.43
DEC1	16.45
DEC5	15.14
DEC10	14.23
SPW	14.59
SPD	15.58
SPM1	12.05
SPM2	20.16
SPM3	10.21
SPM4	11.35

returns have zero third order moments. What does this mean? Had we rejected the null hypothesis of zero third order moments, we would have found evidence consistent with nonlinearity-in-mean (possibly chaotic dynamics). The failure to reject the null, however, does not allow us to rule out the presence of chaotic dynamics. We therefore turn to a second approach using nonparametric regressions to capture the conditional mean directly.

Suppose returns are generated by the following model:

$$x_t = f(x_{t-1}, \dots) + \epsilon_t, \tag{33}$$

where $f()$ is nonlinear and ϵ_t is IID. This includes chaotic models as special cases, if we set $\epsilon_t = 0$. When $f()$ is a smooth function, Stone (1977) showed that a large class of nonparametric regressions can be used to fit $f()$ consistently as the sample size increases. There are many ways to implement nonparametric regression; for example, kernel estimation, series expansion, neural network, and nearest neighbor. We select Cleveland's (1979) method of locally weighted regression (LWR), which is a generalization of nearest neighbor. LWR has been used to test for nonlinearity-in-mean by Diebold and Nason (1990) in weekly exchange rate changes, and LeBaron (1988) in weekly stock returns.

Diebold and Nason (1990) gave a very good description of locally weighted regression. Briefly, the idea is this. Suppose the data are generated according to:

$$x_{t+1} = f(x_t). \tag{34}$$

We have observed x_t, $x_{t-1}, \cdots$, and would like to forecast x_{t+1}. The BDS statistic indicates that, whenever x_{t-1} was close to x_{s-1}, x_t was also close to x_s. We can look at the history of returns, find those instances when x_s was close to x_t, run a nonparametric regression of x_{s+1} on x_s to estimate the function $f()$, and use $\hat{f}(x_t)$ to predict x_{t+1}, where $\hat{f}()$ denotes the nonparametric estimate of $f()$. The extension to the case where $f()$ contains more than one lag of x_t is straightforward. Locally weighted regressions uses the k nearest neighbors of x_t, and a scheme which gives more weight to closer observations and less weight to farther observations. There are a number of parameters to be selected: (a) The number of nearest neighbors to use: We try 10% of all observations, up to 90%, increasing in steps of 10%. (b) The number of lags of x_t to include as arguments of the unknown function $f()$: We use lags 1 through 5. (c) The weighting scheme: We use the "tricubic" scheme proposed by Cleveland and Devlin (1988).[18] (d) A period for out-of-sample forecasting: for the weekly returns (VW, EW, DEC1, DEC5, DEC10, SPW), we arbitrarily start the forecast at the 1001st observation and continued through the end. For the daily returns (SPD) we begin the forecast at the 1601st observation. For the 15-minute returns (SPM1, SPM2, SPM3, SPM4), we begin the forecast at the 1401st observation. This way, each series has at least 1000 observations for the locally weighted regression, and at least 300 observations for out-of-sample forecasting.

If stock returns are governed by low complexity chaos, we should be able to use locally weighted regression to forecast returns much better than simple methods, such as the random walk (RW) model of prices. In addition, our forecasts should improve as the forecast horizon becomes shorter and shorter. Neither implication is born out by the data. Table IX measures forecastability in terms of root mean squared errors. In most cases, the random walk model achieves the lowest root mean squared error. In a few instances, e.g., VW, EW, DEC1, and SPM4, the locally weighted regression has smaller forecast errors than the random walk, but the reduction in root mean squared error is less than 5%.[19] This can, however, happen by chance, given the wide range of parameter values in choosing the locally weighted regression.

One possible explanation of the inability of LWR to outperform random walk forecasts is that LWR is unable to capture conditional mean changes. We therefore perform a simulation using "2-mean," "2-variance," NMA, TAR, Sine, EGARCH, and Mackey-Glass. We generate 500 observations of each series, and begin out-of-sample forecasting at the 451st observations. The tricubic weighting function is used. Since the simulations are computationally intensive, we use only one choice of k—50 nearest neighbors (about 10% of the entire sample). We compare the root mean squared error of the LWR forecasts with that of the "random walk" forecast for 2000 replications. Table X shows that LWR beats the random walk 100% of the time in the

[18] We have experimented briefly with nearest neighbor, which is a rectangular weighting scheme. The results are similar to those using the tricubic weighting function.

[19] These results are consistent with the findings in LeBaron (1988).

Mackey-Glass equation, and 95% of the time for the "Sine" model, which was not detectable by the third order moment test. In addition, LWR outperforms random walk in the TAR model, even though the function f( ) is not smooth. This indicates that LWR has the ability to pick up conditional mean changes.

While we did not experiment with alternative methods of nonparametric regressions, other authors have had no more success. White (1988) found that forecasts of IBM stock returns using neural network did not outperform the random walk model. Prescott and Stengos (1988) found that forecasts of kernel estimators on gold and silver also could not outperform the random walk model.

The preponderance of the failure to outperform the random walk model in asset markets forces us to conclude that there is no strong evidence that the movements in stock market are primarily due to conditional mean changes, when conditioning on past returns.[20] In particular, there is no evidence of low complexity chaotic behavior in stock returns.[21]

VI. Is Conditional Heteroskedasticity Responsible for the Rejection of IID?

Next we proceed to consider whether stock returns are nonlinear-in-variance:

$$x_t = g(x_{t-1}, \ldots)\epsilon_t, \tag{35}$$

where $E[\epsilon_t \mid x_{t-1}, \ldots] = 0$ and $V[\epsilon_t \mid x_{t-1}, \ldots] = 1$ (without loss of generality). This is a general model of conditional heteroskedasticity, which includes ARCH-type models as special cases.

There is now growing evidence that stock market volatility is not only time-varying (e.g., French, Schwert, and Stambaugh (1987)) but is predictable (e.g., Schwert and Seguin (1990)). A number of papers have used ARCH-type models to describe conditional heteroskedasticity (e.g., Bollerslev (1987) and Nelson (1991)). We pose two questions in this section: (a) What is the evidence of conditional heteroskedasticity? (b) Does the conditional heteroskedasticity captured by ARCH-type models account for all the nonlinearity in stock returns?

To answer the first question, observe that if we take the absolute values of equation (35), we obtain:

$$|x_t| = |g(x_{t-1}, \ldots)| \; |\epsilon_t|. \tag{36}$$

If $g()$ is differentiable, a Taylor series expansion would yield the result that $|x_t|$ depends on $|x_{t-i}|$. Thus correlation of $|x_t|$ with $|x_{t-i}|$ is evidence of

[20] These results could change if we increase the information set to include variables other than past returns. For example, Gallant, Rossi, and Tauchen (1990) use returns and volume in a bivariate system.

[21] Even if we had found evidence of chaotic behavior, estimating the unknown parameters of a chaotic map is next to impossible. See Geweke (1989) for a discussion.

Table IX
Root Mean Squared Forecast Errors

The table presents the root mean squared forecast errors using the locally weighted regression with tricubic weights, using lags 1 through 5, and the number of nearest neighbors equaling the fraction (f), 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9 of the data. The root mean squared forecast errors of the random walk model (RW) are in parentheses. The smallest root mean squared forecast error for each series is underlined. All errors have been multiplied by 100. (See Table VIII for definitions of abbreviations.)

	Lags				
f	1	2	3	4	5
VW: (Random walk 0.05436081)					
0.10	<u>0.05198337</u>	0.05204231	0.05208826	0.05204822	0.05204699
0.20	0.05758597	0.05765421	0.05769833	0.05766710	0.05766342
0.30	0.05970985	0.05977849	0.05981333	0.05977608	0.05978498
0.40	0.06074411	0.06080367	0.06083238	0.06078065	0.06080701
0.50	0.06104386	0.06109431	0.06111920	0.06105552	0.06109534
0.60	0.06100911	0.06105054	0.06107035	0.06100228	0.06104746
0.70	0.06082516	0.06085788	0.06087175	0.06080501	0.06085057
0.80	0.06054519	0.06056918	0.06057733	0.06051347	0.06055829
0.90	0.06016496	0.06018105	0.06018456	0.06012428	0.06016784
EW: (Random walk 0.05320546)					
0.10	0.05317910	0.05313575	<u>0.05297957</u>	0.05310239	0.05310214
0.20	0.05399810	0.05394319	0.05383378	0.05387236	0.05386917
0.30	0.05475939	0.05469907	0.05462234	0.05461992	0.05461392
0.40	0.05531716	0.05526434	0.05521420	0.05518580	0.05517885
0.50	0.05571191	0.05566488	0.05563107	0.05558950	0.05558233
0.60	0.05600711	0.05596782	0.05594459	0.05589973	0.05589285
0.70	0.05620623	0.05617554	0.05616085	0.05611508	0.05610873
0.80	0.05630840	0.05628667	0.05627885	0.05623417	0.05622840
0.90	0.05630185	0.05628798	0.05628544	0.05624305	0.05623784
DEC1: (Random walk 0.06055333)					
0.10	0.05806902	0.05796735	0.05793429	0.05793747	<u>0.05788714</u>
0.20	0.05914330	0.05904963	0.05901102	0.05901133	0.05900607
0.30	0.05967503	0.05957489	0.05954111	0.05953993	0.05953798
0.40	0.06018445	0.06009454	0.06008802	0.06008652	0.06008414
0.50	0.06061733	0.06053545	0.06054605	0.06054455	0.06054260
0.60	0.06090465	0.06083168	0.06085730	0.06085586	0.06085410
0.70	0.06111063	0.06103978	0.06107295	0.06107149	0.06106924
0.80	0.06125867	0.06118942	0.06122721	0.06122567	0.06122274
0.90	0.06137222	0.06130418	0.06134270	0.06134127	0.06133782
DEC5: (Random walk <u>0.05593892</u>)					
0.10	0.05895377	0.05907995	0.05910458	0.05904954	0.05902467
0.20	0.05939389	0.05950339	0.05953901	0.05959645	0.05960005
0.30	0.05960399	0.05969049	0.05972788	0.05980120	0.05980946
0.40	0.05971944	0.05979293	0.05982902	0.05988964	0.05989699
0.50	0.05984676	0.05991038	0.05994459	0.05998930	0.05999401
0.60	0.05986240	0.05992130	0.05995344	0.05998826	0.05999137
0.70	0.05981183	0.05986855	0.05989852	0.05992758	0.05992973
0.80	0.05969451	0.05974899	0.05977675	0.05980114	0.05980247
0.90	0.05953412	0.05958599	0.05961180	0.05963290	0.05963368

Table IX—*Continued*

f	Lags 1	2	3	4	5
DEC10: (Random walk 0.05488350)					
0.10	0.05305787	0.05313033	0.05313959	0.05324359	0.05315436
0.20	0.05668828	0.05675475	0.05677510	0.05680402	0.05677795
0.30	0.05872669	0.05879617	0.05882264	0.05881426	0.05881659
0.40	0.05963001	0.05969522	0.05971931	0.05969445	0.05971353
0.50	0.06005248	0.06010943	0.06012974	0.06009160	0.06012147
0.60	0.06018747	0.06023472	0.06025017	0.06020321	0.06024141
0.70	0.06013509	0.06017344	0.06018397	0.06013311	0.06017672
0.80	0.05997842	0.06000830	0.06001420	0.05996200	0.06000904
0.90	0.05975363	0.05977644	0.05977930	0.05972548	0.05977443
SPW: (Random walk 0.05083948)					
0.10	0.05955174	0.05957110	0.05956666	0.05956406	0.05955766
0.20	0.05623793	0.05623091	0.05621985	0.05621961	0.05621936
0.30	0.05506007	0.05505484	0.05504434	0.05504418	0.05504591
0.40	0.05432155	0.05431458	0.05430424	0.05430363	0.05430552
0.50	0.05375797	0.05374680	0.05373616	0.05373535	0.05373691
0.60	0.05335459	0.05334193	0.05333248	0.05333188	0.05333303
0.70	0.05306897	0.05305650	0.05304932	0.05304907	0.05304987
0.80	0.05284070	0.05282971	0.05282462	0.05282451	0.05282518
0.90	0.05259664	0.05258771	0.05258438	0.05258441	0.05258503
SPD: (Random walk 0.00658354)					
0.10	0.00721163	0.00721740	0.00721715	0.00721743	0.00721916
0.20	0.00719038	0.00719276	0.00719432	0.00719232	0.00719306
0.30	0.00721627	0.00721747	0.00721892	0.00721708	0.00721742
0.40	0.00722019	0.00722111	0.00722225	0.00722079	0.00722102
0.50	0.00722375	0.00722433	0.00722536	0.00722404	0.00722417
0.60	0.00722701	0.00722731	0.00722822	0.00722706	0.00722712
0.70	0.00722815	0.00722836	0.00722908	0.00722816	0.00722821
0.80	0.00722651	0.00722674	0.00722722	0.00722661	0.00722666
0.90	0.00721953	0.00721972	0.00721988	0.00721968	0.00721972
SPM1: (Random walk 0.00029380)					
0.10	0.00029994	0.00030069	0.00030052	0.00030053	0.00030053
0.20	0.00029738	0.00029791	0.00029779	0.00029780	0.00029781
0.30	0.00029629	0.00029667	0.00029658	0.00029660	0.00029656
0.40	0.00029540	0.00029573	0.00029565	0.00029567	0.00029562
0.50	0.00029451	0.00029487	0.00029478	0.00029480	0.00029476
0.60	0.00029395	0.00029432	0.00029423	0.00029425	0.00029421
0.70	0.00029344	0.00029380	0.00029371	0.00029372	0.00029368
0.80	0.00029279	0.00029313	0.00029305	0.00029306	0.00029301
0.90	0.00029224	0.00029255	0.00029247	0.00029248	0.00029244
SPM2: (Random walk 0.00025644)					
0.10	0.00025903	0.00025900	0.00025898	0.00025899	0.00025906
0.20	0.00025804	0.00025805	0.00025799	0.00025799	0.00025807
0.30	0.00025800	0.00025803	0.00025796	0.00025797	0.00025805
0.40	0.00025804	0.00025807	0.00025800	0.00025801	0.00025809
0.50	0.00025803	0.00025807	0.00025799	0.00025799	0.00025808
0.60	0.00025798	0.00025802	0.00025794	0.00025794	0.00025803
0.70	0.00025788	0.00025792	0.00025784	0.00025784	0.00025793
0.80	0.00025786	0.00025789	0.00025783	0.00025782	0.00025792
0.90	0.00025797	0.00025800	0.00025794	0.00025794	0.00025805

Table IX—*Continued*

f	Lags 1	2	3	4	5
SPM: (Random walk 0.00009105)					
0.10	0.00009783	0.00009788	0.00009752	0.00009738	0.00009736
0.20	0.00009543	0.00009549	0.00009507	0.00009493	0.00009486
0.30	0.00009409	0.00009414	0.00009379	0.00009365	0.00009356
0.40	0.00009314	0.00009319	0.00009289	0.00009277	0.00009267
0.50	0.00009268	0.00009273	0.00009246	0.00009234	0.00009224
0.60	0.00009245	0.00009249	0.00009224	0.00009213	0.00009202
0.70	0.00009225	0.00009229	0.00009205	0.00009195	0.00009183
0.80	0.00009202	0.00009206	0.00009183	0.00009173	0.00009161
0.90	0.00009166	0.00009170	0.00009147	0.00009137	0.00009125
SPM4: (Random walk 0.00007927)					
0.10	0.00008360	0.00008355	0.00008367	0.00008365	0.00008365
0.20	0.00008165	0.00008164	0.00008171	0.00008169	0.00008169
0.30	0.00008064	0.00008063	0.00008068	0.00008066	0.00008066
0.40	0.00008023	0.00008024	0.00008027	0.00008026	0.00008026
0.50	0.00008004	0.00008007	0.00008009	0.00008008	0.00008008
0.60	0.00007993	0.00007996	0.00007998	0.00007996	0.00007996
0.70	0.00007985	0.00007989	0.00007991	0.00007990	0.00007989
0.80	0.00007980	0.00007985	0.00007986	0.00007985	0.00007984
0.90	0.00007974	0.00007980	0.00007981	0.00007980	0.00007979

Table X

Forecasting Simulated Data

This table reports the percentage when the root mean squared forecast errors of the locally weighted regression is smaller than that of the random walk model. The locally weighted regression forecasts for observation 451 through 500 were generated using 5 lags, 50 nearest neighbors, and the tricubic weighting function. The Monte Carlo simulations use 2000 replications, each having 500 observations, for data generated by seven models: the '2-mean' model (the first 500 observations have mean −1 and variance 1, the second 500 observations have mean 1 and variance 1), the '2-variance' model (the first 500 observations have mean 0 and variance 1, the second 500 observations have mean 0 and variance 2), the nonlinear moving average (NMA), the threshold autoregression (TAR), the "Sine" model, the exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model, and the Mackey-Glass filtered by a third order autoregression.

Model	Percentage
2-mean	96.80
2-variance	9.25
NMA	47.00
TAR	75.60
Sine	95.40
EGARCH	12.15
Mackey-Glass	100.00

Table XI

Testing for Conditional Heteroskedasticity

This table presents the autocorrelation coefficients of the absolute values of the data for 11 stock returns: the weekly value-weighted portfolio (VW), the weekly equally weighted portfolio (EW), the weekly smallest decile portfolio (DEC1), the weekly fifth decile portfolio (DEC5), the weekly largest decile portfolio (DEC10), the weekly S&P500 index (SPW), the daily S&P500 index (SPD), and the 15-minute S&P500 indices for the first, second, third, and fourth quarter of 1988 (SPM1, SPM2, SPM3, and SPM4, respectively).

	Lag					
	1	2	3	4	5	6
VW	0.242*	0.214*	0.144*	0.142*	0.133*	0.195*
EW	0.276*	0.232*	0.137*	0.112*	0.116*	0.148*
D1	0.339*	0.193*	0.140*	0.108*	0.103*	0.082*
D5	0.254*	0.224*	0.131*	0.126*	0.132*	0.152*
D10	0.244*	0.204*	0.157*	0.154*	0.138*	0.209*
SPW	0.245*	0.167*	0.168*	0.165*	0.123*	0.179*
SPD	0.228*	0.202*	0.226*	0.136*	0.217*	0.168*
SPM1	0.178*	0.149*	0.205*	0.127*	0.126*	0.042
SPM2	0.127*	0.114*	0.068*	0.094*	0.060	0.017
SPM3	0.126*	0.092*	0.070*	0.061	0.025	0.006
SPM4	0.148*	0.174*	0.092*	0.086*	0.066*	0.060

*Statistically significant at the 1% level (two-tailed test).

conditional heteroskedasticity (particularly when x_t is not correlated with x_{t-1}).[22] Table XI presents the autocorrelations of the absolute valued data. There is strong evidence of conditional heteroskedasticity in weekly and daily returns, and somewhat weaker evidence in 15-minute returns.[23]

ARCH-type models have been used to capture conditional heteroskedasticity in stock returns, and the typical diagnostic tests (e.g., autocorrelation of absolute values and squares of standardized residuals) show that they do. We are, however, interested in a deeper issue: does ARCH-type models capture all the nonlinear dependence in stock returns? To answer this question, we fit an EGARCH model to the data:

$$x_t \sim N(0, \sigma_t^2),$$

$$\log \sigma_t^2 = \phi_0 + \phi \mid x_{t-1}/\sigma_{t-1} \mid + \psi \log \sigma_{t-1}^2 + \gamma x_{t-1}/\sigma_{t-1}. \qquad (37)$$

EGARCH is chosen over the simpler ARCH or GARCH model for two reasons: (a) unlike the simple ARCH or GARCH model, EGARCH does not impose any restrictions on the signs of the parameters to guarantee that estimated variances are positive, thus avoiding numerical problems associated with constrained optimization, and (b) EGARCH can accommodate

[22] The same argument shows that x_t^2 would be correlated with $x_{t-\iota}^2$ under conditional heteroskedasticity. See Engle (1982) and McLeod and Li (1983).

[23] We point out here that the evidence is consistent with conditional heteroskedasticity. But it does not rule out even higher order dependence (e.g., conditional skewness, conditional kurtosis).

conditional skewness discussed in Black (1976) which is not allowed in the less flexible ARCH and GARCH models. We use the Berndt, Hall, Hall, and Hausman (1974) procedure with analytic first derivatives to estimate this model.

If the EGARCH model is correctly specified, the standardized residuals:

$$z_t = x_t / \hat{\sigma}_t, \tag{38}$$

should be IID in large samples. Here, $\hat{\sigma}_t$ is the fitted value of the standard deviation from the variance equation. Thus BDS can be used as on the standardized residuals to test if EGARCH captures all nonlinear dependence in stock returns.

Table XII shows that the BDS statistics on the standardized residuals are much smaller than those of the raw data. Only a few statistics are significant, if we use the asymptotic distribution. The trouble is that the asymptotic distribution of the BDS statistic cannot be used when dealing with ARCH, GARCH, and EGARCH standardized residuals, a point made in Table V. Therefore, we use the simulated critical values of the BDS statistic. The 2.5% and 97.5% critical values are given in Table XIII. Based on these critical values, the only series to pass the BDS diagnostic is the smallest decile portfolio, DEC1. All the other series contain several BDS statistics which are outside the 5% critical range. In particular, the daily S&P returns have the worst fit, failing the BDS diagnostic every time. There is sufficient evidence here to indicate that the EGARCH model cannot completely account for all nonlinearity in stock returns.

One problem with ARCH-type models is that the variance equation does not contain an innovation. To obtain a more general model, we add a stochastic term in the variance equation, leading to the following specification for stock returns:

$$x_t = \sigma_t z_t, \tag{39}$$

where z_t is an IID random variable, and σ_t evolves according to:

$$\log \sigma_t = \beta_0 + \Sigma_i \beta_i \log \sigma_{t-i} + v_t, \tag{40}$$

where v_t is IID, independent of z_t.

It is appropriate here to contrast this model with the mixture models in the earlier stock market literature. Blattberg and Gonedes (1974) pointed out that the symmetric stable distribution is obtained from a normal distribution whose variance is drawn from a strictly positive stable distribution, that the Student t is obtained from a normal distribution whose variance is drawn from an inverted gamma distribution, and that Clark's (1973) model is a normal distribution whose variance is drawn from a log normal distribution. Thus, all three mixture models can be written in the form: $x_t = \sigma_t z_t$, where z_t, is IID standard normal, and σ_t is another IID random variable. In these cases, x_t exhibits neither conditional heteroskedasticity or nonlinear dependence. Our more general specification allows for nonlinear dependence in the form of conditional heteroskedasticity.

To test the variance specification, we construct daily standard deviations of returns from April 21, 1982 to September 30, 1989, using the 15-minute data, after removing the serial correlation. Figure 5 is a plot of the natural logarithms of the daily standard deviations. Note that, while the volatility on October 19 and 20, 1987, were considered to be "huge" at the time, they did not show up as "outliers" in the logarithms. In fact, the volatility leading up to those days had been on the rise. This is consistent with the diagnostics on the least squares residuals. Using the Schwarz criterion, we determine the lag length to be 5.[24] The least squares fit is

$$\log \sigma_t = \underset{(0.1064)}{-0.8577} + \underset{(0.0229)}{0.2385} \log \sigma_{t-1} + \underset{(0.0236)}{0.1298} \log \sigma_{t-2}$$

$$+ \underset{(0.0236)}{0.1129} \log \sigma_{t-3} + \underset{(0.0236)}{0.1515} \log \sigma_{t-4} + \underset{(0.0229)}{0.1386} \log \sigma_{t-5}$$

$$R^2 = 0.4127. \tag{41}$$

The parentheses contain the standard errors of the estimated coefficients. Clearly, there is mean reversion in volatility. But the sum of the coefficients of this autoregressive process is 0.7713, which contains much less persistence than that of the GARCH model in Bollerslev (1987) and the EGARCH model in Nelson (1991).

We ran the BDS test on the residuals to test for the appropriateness of the linear model. Panel A in Table XIV shows that the BDS statistics are very small, giving no evidence of nonlinearity. Furthermore, it is interesting to note that the coefficient of kurtosis of the residuals is 3.49, not much higher than 3. There does not appear to be extreme points.

In the last step we check whether this model of conditional heteroskedasticity can capture the nonlinear dependence in stock returns. We standardize daily returns with the fitted values $\hat{\sigma}_t$ from the variance equation: $z_t = x_t / \hat{\sigma}_t$. (Note that x_t here is the raw data, not the linearly filtered data.) We then remove linear dependence in z_t (possibly due to asynchronous trading) using a first order autoregression. This lag length was identified by both the Akaike and the Schwarz criterion. Panel B in Table XIV contains the final diagnostics of this model. It shows that the BDS statistics are substantially lower than those in Table XII (for SPD). If we use the asymptotic distribution of the BDS test, we do not reject the model.[25] This lead us to conclude that the more flexible variance specification provides a much better description of the nonlinear dependence in daily stock returns. In addition, note that the kurtosis of 88.99 for SPD in Table V has been reduced to 8.816 in Panel B of Table XIV, implying that most but not all of the leptokurtosis of daily stock returns is due to variance changes.

[24] The Akaike criterion led to very long lags.

[25] Hsieh (1991) shows that the asymptotic distribution applies to the BDS statistic on the residuals of the generalized heteroskedasticity model in equations (39) and (40). Even if we apply a more stringent rejection criterion by using the critical values in Table XIII, we could reject only 1 BDS statistic, at dimension 2, when $\epsilon/\sigma = 2$.

Table XII

BDS Statistics for EGARCH Standardized Residuals

This table presents the BDS statistics (at dimensions 2 through 5 and ϵ equaling 0.5, 1, 1.5, and 2 standard deviations of the data) for EGARCH standardized residuals of 11 stock returns: the weekly value-weighted portfolio (VW), the weekly equally weighted portfolio (EW), the weekly smallest decile portfolio (DEC1), the weekly fifth decile portfolio (DEC5), the weekly largest decile portfolio (DEC10), the weekly S&P500 index (SPW), the daily S&P500 index (SPD), and the 15-minute S&P500 indices for the first, second, third, and fourth quarter of 1988 (SPM1, SPM2, SPM3, and SPM4, respectively).

		ϵ/σ			
	m	0.50	1.00	1.50	2.00
VW	2	0.49	−1.47	−2.16*	−2.03*
	3	0.92	−0.96	−2.08*	−2.11*
	4	1.16	−1.18	−2.38*	−2.44*
	5	0.88	−1.02	−2.41*	−2.45*
EW	2	0.14	−0.58	−1.20	−1.23
	3	1.30	−0.09	−1.05	−1.40*
	4	1.50	−0.34	−1.42*	−1.75*
	5	2.20	−0.34	−1.52*	−1.80*
DEC1	2	0.83	1.01	1.11	1.16
	3	0.78	0.83	0.69	0.42
	4	0.17	0.60	0.35	−0.06
	5	−0.25	0.43	0.09	−0.35
DEC5	2	0.29	−0.43	−0.94	−1.04
	3	0.67	−0.31	−1.00	−1.37*
	4	0.75	−0.64	−1.56*	−2.00*
	5	0.88	−0.71	−1.74*	−2.13*
DEC10	2	0.08	−1.23	−1.82*	−1.64*
	3	0.46	−1.07	−1.88*	−1.84*
	4	0.68	−1.26*	−2.16*	−2.09*
	5	0.44	−1.14*	−2.19*	−2.14*
SPW	2	−0.26	−0.46	−0.82	−0.83
	3	−0.37	−0.89	−1.59*	−1.69*
	4	0.18	−0.76	−1.50*	−1.75*
	5	1.22	−0.39	−1.27*	−1.64*
SPD	2	−3.46*	−3.44*	−2.88*	−1.78*
	3	−4.44*	−4.39*	−3.95*	−2.74*
	4	−4.36*	−4.48*	−4.11*	−2.84*
	5	−4.05*	−4.30*	−4.06*	−2.77*
SPM1	2	0.42	−0.99	−1.31	−0.98
	3	1.72	0.11	−0.54	−0.58
	4	3.16*	1.50	0.68	0.23
	5	4.04*	1.97	0.95	0.26
SPM2	2	2.08	0.95	0.02	−0.21
	3	3.10*	1.13	−0.11	−0.63
	4	3.82*	1.32	−0.35	−0.92
	5	4.90*	1.84	0.03	−0.60

Table XII—*Continued*

	m	ϵ/σ 0.50	1.00	1.50	2.00
SPM3	2	5.41*	2.36*	0.32	−0.58
	3	7.32*	3.59*	1.26	0.10
	4	9.34*	4.36*	1.66	0.47
	5	11.56*	5.34*	2.51*	1.19
SPM4	2	−0.32	−0.99	−1.59*	−2.12*
	3	1.36	0.54	−0.06	−0.64
	4	2.65*	1.59	0.78	−0.09
	5	3.68*	2.20	1.22	0.10

*Significant at the 5% (two-tailed) test.

Table XIII

Simulated BDS Critical Values for EGARCH Standardized Residuals

This table presents the simulated 2.5% and 97.5% critical values of the BDS statistic (at dimensions 2 through 5 and ϵ equaling 0.25, 0.5, 1, 1.5, and 2 standard deviations of the data) when applied to EGARCH standardized residuals. The Monte Carlo simulation uses 2000 replications, each with 1000 observations. $N(0, 1)$ denotes the critical values of a standard normal distribution.

n	ϵ/σ 0.25	0.50	1.00	1.50	2.00	$N(0, 1)$
2.5% crticial values						
2	−2.21	−1.61	−1.52	−1.47	−1.49	−1.96
3	−2.86	−1.65	−1.29	−1.29	−1.29	−1.96
4	−4.46	−1.63	−1.17	−1.13	−1.12	−1.96
5	−7.77	−1.94	−1.11	−1.00	−0.99	−1.96
97.5% critical values						
2	2.88	2.11	1.96	1.85	1.88	1.96
3	3.56	2.34	2.14	2.01	2.00	1.96
4	5.42	2.49	2.25	2.17	2.14	1.96
5	9.34	2.90	2.40	2.28	2.22	1.96

This model gives rise to some interesting possibilities. The mean reversion in volatility implies that one can forecast future volatility based on past volatility. In addition, the standardized data (after dividing by expected volatility) are IID, so we can obtain a nonparametric estimate of their density, which can then be used to make probability statements that are useful in, say, setting margin requirements for stocks.

VIII. Concluding Remarks

We have found strong evidence to reject the hypothesis that stock returns are IID. The cause does not appear to be either regime changes or chaotic

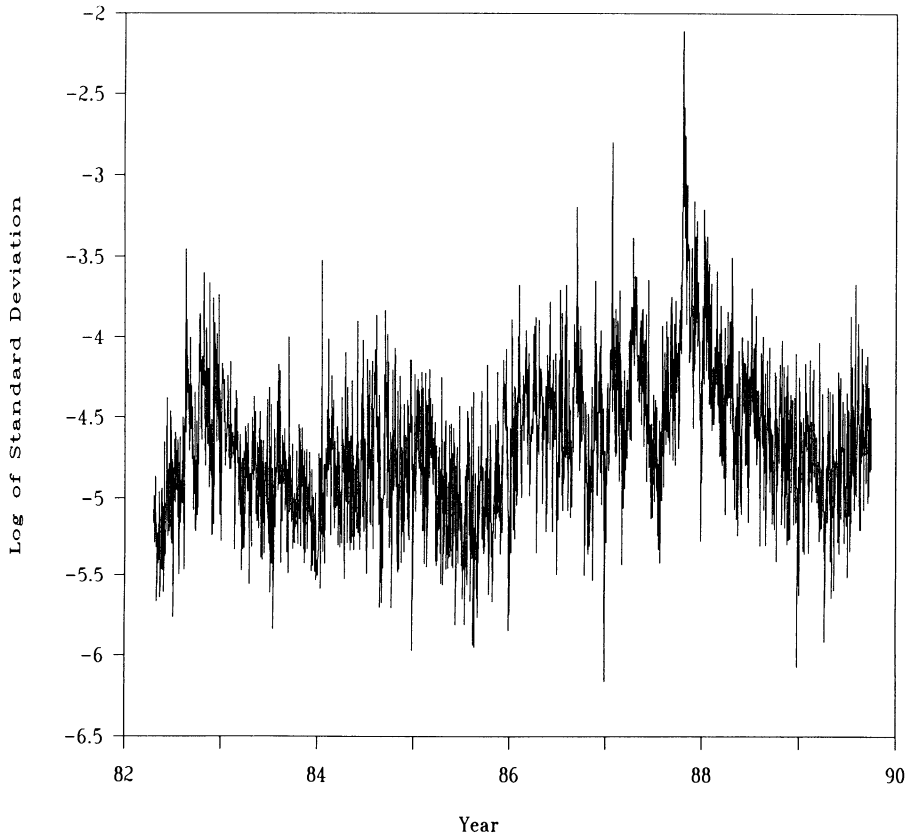

Figure 5. Logarithm of daily standard deviations of the S&P500 index. This plots the natural logarithm of the daily standard deviations of the S&P500 index, computed from fifteen minute returns during each trading day.

dynamics. Rather, the cause appears to be conditional heteroskedasticity (e.g., predictable variance changes). While we find that ARCH-type models do not fully capture the nonlinearity in stock returns, a more flexible model of conditional heteroskedasticity can. These findings have many interesting implications. One, if we want to fit conditional density functions on stock returns, we must account for their nonlinear dependence. Two, if we are interested to model the nonlinearity in stock returns, we should direct our efforts at conditional heteroskedasticity rather than conditional mean changes (which include chaotic dynamics). Three, if the flexible conditional heteroskedasticity model holds up under future analysis, it can provide conditional volatility forecasts. Those, together with a nonparametric estimate of the density of the standardized residuals, can deliver a conditional probability distribution which would be useful in many applications. Lastly, it would be interesting to see if this model can capture nonlinearity found in other

Table XIV

Diagnostics for a Generalized Heteroskedasticity Model

This table presents diagnostics for a generalized heteroskedasticity model for daily S&P500 returns. Panel A contains the diagnostics for residuals from a fifth order autoregression for the natural logarithm of daily standard deviations of the S&P500 index, computed from 15-minute returns during each trading day. Panel B contains the diagnostics from a first order autoregression for the daily S&P500 returns standardized by the fitted values of standard deviations based on the regression in Panel A.

Panel A: Residuals from: $\log \sigma_t = \beta_0 + \sum_{i=1}^{5} \beta_i \log \sigma_{t-i} + v_t$					
Mean	0.0000				
Std dev	0.3666				
Skewness	0.214				
Kurtosis	3.49				
Maximum	1.481				
Minimum	−1.309				
BDS test:		ϵ/σ			
m		0.50	1.00	1.50	2.00
2		−0.21	0.03	0.18	0.39
3		−0.02	0.09	0.14	0.28
4		0.16	0.16	0.30	0.33
5		0.29	0.31	0.40	0.37
Panel B: Residuals from: $z_t = \alpha_0 + \alpha_1 z_{t-1} + e_t$, where $z_t = x_t/\hat{\sigma}_t$					
Mean	0.0000				
Std dev	1.0268				
Skewness	−0.213				
Kurtosis	8.816				
Maximum	6.410				
Minimum	−9.349				
BDS test:		ϵ/σ			
m		0.50	1.00	1.50	2.00
2		−0.20	0.30	0.94	1.94
3		−1.25	−0.59	0.14	1.19
4		−1.42	−0.66	0.06	1.14
5		−1.18	−0.57	0.12	1.28

financial data such as exchange rates and interest rates. This is left for future research.

REFERENCES

Akaike, H., 1974, A new look at the statistical model identification, *IEEE Transactions on Automatic Control* 19, 716–723.

Baumol, W. and J. Benhabib, 1989, Chaos: significance, mechanism, and economic applications, *Journal of Economic Perspectives* 3, 77–105.

Berndt, E. K., B. H. Hall, R. E. Hall, and J. A. Hausman, 1974, Estimation and inference in nonlinear structural models, *Annals of Economic and Social Measurement* 4, 653–665.

Black, F., 1976. Studies of stock market volatility changes, *Proceedings of the American Statistical Association, Business and Economic Statistics Section* 177-181.

Blattberg, R. C. and N. Gonedes, 1974, A comparison of the stable paretian and student distribution as statistical model for prices, *Journal of Business* 47, 244-280.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307-327.

——, 1987, A conditional heteroskedastic time series model for speculative prices and rates of return, *Review of Economics and Statistics* 69, 542-547.

——, R. Chow, and K. Kroner, 1990, ARCH modeling in finance: a selective review of the theory and empirical evidence, with suggestions for future research, working paper, Northwestern University, Georgia Tech, and University of Arizona.

Brock, W., 1986, Distinguishing random and deterministic systems: abridged version, *Journal of Economic Theory* 40, 168-195.

——, 1987, Notes on nuisance parameter problems in BDS type tests for IID, working paper, University of Wisconsin at Madison.

—— and E. Baek, 1991, Some theory of statistical inference for nonlinear science, *Review of Economic Studies,* Forthcoming.

——, W. Dechert, and J. Scheinkman, 1987, A test for independence based on the correlation dimension, Working Paper, University of Wisconsin at Madison, University of Houston, and University of Chicago.

—— and C. Sayers, 1988, Is the business cycle characterized by deterministic chaos? *Journal of Monetary Economics* 22, 71-90.

Clark, P. K., 1973, A subordinate stochastic process model with finite variance for speculative prices, *Econometrica* 41, 135-155.

Cleveland, W. S., 1979, Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association* 74, 829-836.

Cleveland, W. S. and S. J. Devlin, 1988, Locally weighted regression: an approach to regression analysis by local fitting, *Journal of the American Statistical Association* 83, 596-610.

Dechert, W., 1988, A characterization of independence for a gaussian process in terms of the correlation dimension, SSRI working paper 8812, University of Wisconsin at Madison.

Denker, G. and G. Keller, 1986, Rigorous statistical procedures for data from dynamical systems, *Journal of Statistical Physics* 44, 67-93.

Diebold, F. X. and J. A. Nason, 1990, Nonparametric exchange rate prediction? *Journal of International Economics* 28, 315-332.

Engle, R., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of U. K. inflations, *Econometrica* 50, 987-1007.

Fama, E. and R. Roll, 1968, Some properties of symmetric stable distributions, *Journal of American Statistical Association* 63, 817-837.

French, K., G. W. Schwert, and R. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3-29.

Gallant, R., P. Rossi, and G. Tauchen, 1990, Stock prices and volume, Working paper, North Carolina State University, University of Chicago, and Duke University.

Geweke, J., 1989, Inference and forecasting for chaotic nonlinear time series, Working paper, Duke University.

Gleick, J., 1987, *Chaos: Making a New Science* (Viking Press, New York, NY).

Granger, C. and A. Andersen, 1978, *An Introduction to Bilinear Time Series Models* (Vandenhoeck & Ruprecht, Göttingen).

Grassberger, P. and I. Procaccia, 1983, Measuring the strangeness of strange attractors, *Physica* 9D, 189-208.

Hénon, M., 1976, A two-dimensional mapping with a strange attractor, *Communications in Mathematical Physics* 50, 69-77.

Hinich, M. and D. Patterson, 1985, Evidence of nonlinearity in stock returns, *Journal of Business and Economic Statistics* 3, 69-77.

Hsieh, D., 1989, Testing for nonlinearity in daily foreign exchange rate changes, *Journal of Business* 62, 339-368.

——, 1990, A nonlinear stochastic rational expectations model of exchange rates, Working paper, Duke University.

——, 1991, Implications of nonlinear dynamics for financial risk management, Working paper, Duke University.

——, and B. LeBaron, 1988, Finite sample properties of the BDS statistic, Working paper, University of Chicago and University of Wisconsin at Madison.

LeBaron, B., 1988, The changing structure of stock returns, Working paper, University of Wisconsin.

Lorenz, N., 1963, Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences* 20, 130-141.

Mackey, M, and L. Glass, 1977, Oscillation and chaos in physiological control systems, *Science* 50, 287-289.

Mandelbrot, B., 1963, The variation of certain speculative prices, *Journal of Business* 36, 394-419.

McLeod, A. J. and W. K. Li, 1983, Diagnostic checking ARM time series models using squared-residual autocorrelations, *Journal of Time Series Analysis* 4, 269-273.

Nelson, D., 1990, ARCH models as diffusion approximations, *Journal of Econometrics* 45, 7-38.

——, 1991, Conditional heteroskedasticity in asset returns: a new approach, *Econometrica* 59, 347-370.

Pemberton, J. and H. Tong, 1981, A note on the distribution of non-linear autoregressive stochastic models, *Journal of Time Series Analysis* 2, 49-52.

Prescott, D. M. and T. Stengos, 1988, Do asset markets overlook exploitable nonlinearities? The case of gold, Working paper, University of Guelph.

Priestley, M., 1980, State-dependent models: A general approach to non-linear time series analysis, *Journal of Time Series Analysis* 1, 47-71.

Ramsey, J. and H. Yuan, 1989, Bias and error bias in dimension calculation and their evaluation in some simple models, *Physical Letters A* 134, 287-297.

Robinson, P., 1977, The estimation of a non-linear moving average model, *Stochastic Processes and Their Applications* 5, 81-90.

Scheinkman, J. and B. LeBaron, 1989, Nonlinear dynamics and stock returns, *Journal of Business* 62, 311-337.

Schwarz, G., 1978, Estimating the dimension of a model, *Annals of Statistics* 6, 461-464.

Schwert, G. W. and P. J. Seguin, 1990, Heteroskedasticity in stock returns, *Journal of Finance* 45, 1129-1155.

Stone, C. J., 1977, Consistent nonparametric regressions, *Annals of Statistics* 5, 595-620.

Takens, F., 1980, Detecting strange attractors in turbulence, in D. Rand and L. Young, eds.: *Dynamical Systems and Turbulence* (Springer-Verlag, Berlin).

Tong, H. and K. Lim, 1980, Threshold autoregression, limit cycles, and cyclical data, *Journal of the Royal Statistical Society* Series B, 42, 245-292.

Tsay, R., 1986, Nonlinearity tests for time series, *Biometrika* 73, 461-466.

White, H., 1988, Economic prediction using neural networks: The case of IBM daily stock returns, Working paper, University of California, San Diego.