Chaos and Nonlinear Dynamics: Application to Financial Markets

by

David A. Hsieh
Fuqua School of Business
Duke University
Durham, NC 27706

October 1990

1. Introduction

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.

2. 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.

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 facing external shocks (e.g. wars, weather) which perturb it from the equilibrium. The economy fluctuates because of these external shocks, in the absence of which the economy will settle into a steady state. In the chaotic growth models, the economy follows nonlinear dynamics, which are self-generating and never die down. External shocks are not needed to cause

economic fluctuations, which are part of the dynamics of the economy.

In 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. One explanation of the crash was that there was some (unanticipated) news which caused investors to drastically mark down the value of equities. Another explanation was that the stock market is a chaotic process which, as we shall see below, is characterized by occasional large movements.

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

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:

$$x_t \quad = \quad 2\, x_{t-1}, \quad\quad \text{if } x_{t-1} < 0.5,$$

$$x_t \quad = \quad 2\, (\, 1 - x_{t-1}\,), \text{ if } x_{t-1} \geq 0.5.$$

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 <u>nonlinear</u> 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 stretching and folding pulls apart points close to each other. This 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 $t \rightarrow \infty$. 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. Suppose we only know that x_0 is in [a-

δ,a+δ]. If we try to forecast x_t into the future, all we know is that x_t lies

in [0,1] as t$\rightarrow\infty$. Three, x_t appears stochastic even though it is a

deterministic process, in the sense that the empirical autocovariance function

$\rho_{xx}(k) = E[x_t x_{t-k}] = \lim_{r\rightarrow\infty} \Sigma_{t0} x_t x_{t-k}/T = 0$, which is the same as that of white

noise. Four, x_t will have a series of small increases, and then it suddenly

declines ("crashes?") sharply.

<u>Pseudo Random Number Generators</u>

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 number z_0, called a "seed", between

0 and P. Generate new seeds using the following rule:

z_t =   A z_{t-1} (mod P).

Generate the sequence:

x_t = z_t/P.

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.

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:

$$x_t = A x_{t-1} (1-x_{t-1}),$$

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 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.

Hénon Map

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

$$x_t = y_{t-1} + 1 - A x_{-1}, \qquad A = 1.4$$

$$y_t = B x_{t-1}, \qquad B = 0.3.$$

Lorenz Map

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

$$= a (y - x), \qquad a = 10,$$

$$= - y - x z - b x, \qquad b = 28,$$

$$= x y - c z, \qquad c = 8/3.$$

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. The Mackey-Glass equation is such an example. It is a "delayed"
differential equation, given by:

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

General Chaotic Maps

In general, chaotic maps are obtained by a deterministic rule:

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

In order to generate chaotic behavior, $f(\;)$ must be a nonlinear function.

If $f(\;)$ is linear, then either x_t will converge to a number (called a
fixed point), or x_t will cycle with a fixed period, or x_t will explode.
Therefore, if $f(\;)$ is a chaotic map, it must be nonlinear. 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.]

3. 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.
Suppose we have a string of data, x_1, x_2, ..., x_t, ..., x_T. [Say these are
stock returns.] How can we tell if the data are chaotic?

One way to detect chaos 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, the data is
uniform over [0,1], and so they fill 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

$[0,1]_X[0,1]$. In other words, data from the tent map leave large "holes" in

two dimensional space, while the random data do not.

Clearly it is not practical to do this exercise in higher dimensions.

In order to do detect chaotic behavior, Grassberger and Procaccia (1983a)

developed the notional of correlation dimension. This is done in four steps.

Step 1. Remove autocorrelation, if present. Autocorrelation can effect

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. Form n-histories of the filtered data. These are denoted as

follows:

1-history: x $= x_t$.

2-history: x $= (x_{t-1}, x_t)$.

.

n-history: x $= (x_{t-n+1}, ..., x_t)$.

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

dimension."

Step 3. Calculate the correlation integral:

$$C_n(\varepsilon) = \#\{ (t,s), 0<t,s,<T: \ |x - x| < \varepsilon \} / T^2,$$

where $|\ |$ is the sup- or max- norm. In words, the correlation integral,

$C_n(\varepsilon)$, is defined as the fraction of pairs, (x,x), which are "close" to each

other, in the sense that:

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

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

$$v_n = \lim_{\varepsilon \to 0} \log C_n(\varepsilon) / \log \varepsilon.$$

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

$$v = \lim_{n \to \infty} v_n.$$

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), $C_1(\varepsilon)$ doubles if ε doubles. Thus, for small values of ε,

$$v_1 = \log C_1(\varepsilon) / \log \varepsilon = 1.$$

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

$$v_2 = \log C_2(\varepsilon) / \log \varepsilon = 1.$$

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

$$v_n = \log C_n(\varepsilon) / \log \varepsilon = 1.$$

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 $C_1(\varepsilon)$ doubles if ε doubles, so

$$v_1 = \log C_1(\varepsilon) / \log \varepsilon = 1.$$

However, $C_2(\varepsilon)$ quadruples if ε doubles, and so

$$v_2 = \log C_2(\varepsilon) / \log \varepsilon = 2.$$

In fact,

$$v_n = \log C_n(\varepsilon) \,/\, \log \varepsilon = n.$$

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

The correlation dimension therefore is a measure of how much space is "filled up" by a string of data. Here are the Grassberger-Procaccia correlation dimensions for the other chaotic maps, given in Grassberger and Procaccia (1983a):

Logistic	1.00 ± 0.02
Hénon	1.22 ± 0.01
Lorenz	2.05 ± 0.01
Mackey-Glass	7.50 ± 0.15

From this, we conclude that the chaotic maps do not full up enough space. This is in fact a generic property of chaotic processes, whenever the imbedding dimension is more than twice the correlation dimension.[1]

Although this 4 step procedure sounds straight forward, and has been applied by scientists in many problems, a number of issues surface when dealing with economic and financial data.

4. 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 $\log C_n(\varepsilon)$ versus $\log \varepsilon$ 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.

Firstly, the characterization that chaotic processes have finite correlation dimension is not a sufficient and necessary condition. There are

degenerate random processes which have finite correlation dimensions. Also,

some nonlinear stochastic model, such as Engle's (1982) autoregressive

conditional heterskedasticity model, exhibit "dependence" in the sense that

the slope of the graph of $\log_n C(\varepsilon)$ versus $\log \varepsilon$ increases at a rate slower

than n, as pointed out by Scheinkman and LeBaron (1989).

Secondly, we have only finite amounts of data, which means that there is

no way to verify that a processes has an infinite correlation dimension.

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 non-overlapping 10-histories. It is very

hard to say whether 200 10-histories "fill up" a 10-dimensional space. In

other words, there is no practical way to distinguish between a chaotic

process with a high correlation dimension (say larger than 10) and a truly

random process.

Thirdly, we have to worry about biases in small data sets. Ramsey and

Yuan (1989) show that the slope of the graph of $\log_n C(\varepsilon)$ versus $\log \varepsilon$ is

biased downward in small data sets (2,000 or fewer observations). This biases

the results in favor of finding chaos, even if there is none.

Fourthly, 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 available.[2]

Statistical Test: the BDS Statistic

To deal with these problems, Brock, Dechert, and Scheinkman (1987) have

devised a statistical test. If $\{x_t: t=1,\ldots,T\}$ is a random sample of

independent and identically distributed (IID) observations, then:

$$C_n(\varepsilon) = C_1(\varepsilon)^n.$$

One can estimate $C_1(\varepsilon)$ and $C_n(\varepsilon)$ by $C_{1,T}(\varepsilon)$ and $C_{n,T}(\varepsilon)$, and show that:

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

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

Note that the statement $C_n(\varepsilon) = C_1(\varepsilon)^n$ 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. 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), apply the BDS test, and repeat this 2,000 times. If we use a 5% significance level, we should reject 5% of the replications. Most of these simulations were reported in Hsieh and LeBaron (1988). Table 1 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 with the following intentions: 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. The results for dimension five in Table 2 are similar with some slight deterioration. We conclude that the BDS test avoids the biases of the correlation dimension estimates.

In a 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 3.

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

$$x_t \quad = \quad \rho \; x_{t-1} \quad + \quad u_t.$$

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

$$x_t \quad = \quad \theta \; u_{t-1} \quad + \quad u_t.$$

In the simulations, u_t is IID standard normal, $\rho = .5$, and $\theta = .5$. 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.

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 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 3 shows that BDS also has no trouble in detecting them.

We consider two nonlinear time series models which have no

-12-

autocorrelation but non-zero conditional means. Robinson (1977) proposed the

nonlinear moving average (NMA) model:

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

where u_t is IID standard normal.[3] In the simulations, $\alpha=.5$. The other

nonlinear time series model is the threshold autoregressive (TAR) model in

Tong and Lim (1980):

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

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

where u_t is IID standard normal. In the simulations, $\alpha=-.4$ and $\beta=.5$. Table 3

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, but exhibits conditional heteroskedasticity. As

discussed earlier, Engle (1982) presented the autoregressive conditional

heteroskedasticity (ARCH) model:

$$x_t = \sigma_t \, u_t,$$

$$\sigma = \varphi_0 + \varphi \, x_{-1}.$$

In our simulations, $\varphi_0=1$ and $\varphi_1=.5$. Bollerslev (1986) turned ARCH into

Generalized ARCH (GARCH) by making σ_t a function of its own past:

$$\sigma = \varphi_0 + \varphi \, x_{-1} + \psi \, \sigma_{-1}.$$

In our simulations, $\varphi_0=1$ and $\varphi=.1$, and $\psi=.8$. Nelson (1988) changed GARCH into

exponential GARCH (EGARCH) by using $\log s_t$ instead of σ:

$$\log \sigma = \varphi_0 + \varphi \, |x_{t-1}/\sigma_{t-1}| + \psi \, \log \sigma_{-1} + \gamma \, x_{t-1}/\sigma_{t-1}.$$

In our simulations, $\varphi_0=1$ and $\varphi=.1$, $\psi=.8$, and $\gamma=0$. 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.[4] Table 3 shows that BDS can easily detect the simple ARCH and the 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 3 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 address 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 4. The results show that the asymptotic distribution still approximates the finite sample distribution with the same degree of accuracy even when replacing raw data

-14-

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.

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 dividend 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).[5] In addition, we use ten valued-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.[6] All data were first filtered by an autoregression whose lag length was determined by the Akaike (1974) information criterion.[7]

Table 5 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 6 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 is the implication of the finding that returns are not IID? Firstly, 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 forecast errors.

Secondly, when returns are not IID, it is not appropriate to estimate (unconditional) density. 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 using the normal distribution. One may therefore be tempted to "explain" crashes, such as that on October 19, 1987, as small but nonzero probability events.[8] The fact that returns are not IID,[9] however, makes this explanation for stock market crashes less plausible, because the parameters of these distributions were estimated assuming that returns are IID.

Thirdly, the rejection of IID does not provide direct evidence of chaos in the stock market. Our simulations in Table 3 show that there are 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. Shift in economic fundamentals, e.g. wars, can result in a different mean return (represented by the "2-mean" model). Changes in the operating procedure of the Federal

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

The rejection of IID is also consistent with returns being 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.[10] The ARCH-type model can be thought of as approximating conditional variance changes.[11]

Finally the rejection of IID is also consistent with 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.

5. 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 1962 to 1987. So 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).[12] 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.

Table 6 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.[13] 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 occur in

minute-to-minute intervals over the course of 3 months. If this is the true,

then econometric work on economic and financial data is almost impossible.

6. 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 chaotic dynamics. The issue we raise here is

--- is there any direct evidence of chaotic behavior? Before we proceed, we

should point out that there is no way to distinguish complex chaotic dynamics

(e.g. sophisticated pseudo random number generators) from truly random

behavior. But our task is a little bit simpler. If stock returns are

-18-

governed by chaotic dynamics, the process generating these dynamics must have low complexity, as indicated by the strong rejection of IID by the BDS statistics, even at dimension 2.

The initial attempt we make is to see if the third order moments of stock returns are nonzero, following the suggestion in Hsieh (1989). If x_t is a chaotic process, then we can write:

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

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

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

where ε_t is satisfies the condition that $E[\varepsilon_t | x_{t-1}, \ldots] = 0$. 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). The nonlinearity of $f(\,)$ can generate much richer dynamics than linear models, and can give rise to "large" moves, such as crashes, in stock markets.

To test for nonlinearity-in-mean, we pose the null hypothesis that $f()$ is identically zero. This implies that 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: $r(i,j) = [\Sigma\, x_t x_{t-i} x_{t-j}/T]/[\Sigma x/T]^{1.5}$.

b) $\sqrt{[r(i,j) - \rho(i,j)]}$ has a limiting distribution $N(0, V(i,j))$, where $V(i,j)$ can be estimated by the method of moments:

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

Hsieh (1989) then tests the null hypothesis that $\rho(i,j) = 0$ individually using a t-statistic. The contribution of this paper is to recognize that one can test the composite null hypothesis that $\rho(i,j) = 0$ for $0 < i \leq j \leq m$, for a given m, making

-19-

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:

$$[\Sigma \{x_t x_{t-i} x_{t-j}/T - r(i,j)\}\{x_t x_{t-i'} x_{t-j'}/T - r(i',j')\}] / [\Sigma x/T]^3.$$

Then the composite test can be conducted using the usual χ^2 statistic.[14]

We point out that the test statistic is designed so that it will not reject models which are "nonlinear-in-variance":

$$x_t = g(x_{t-1},\ldots)\varepsilon_t.$$

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 having nonlinear-in-mean. However, the test should detect hybrid models, those which are "nonlinear-in-mean" as well as "nonlinear-in-variance":

$$x_t = f(x_{t-1},\ldots) + g(x_{t-1},\ldots)\varepsilon_t.$$

[The GARCH-in-mean model, where the conditional variance appears also 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 7.

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 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

-20-

to very fat tailed data.

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 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,[15] but the power against the GARCH-in-mean model is low, probably because the high order moments of the GARCH-in-mean model does not exist.

The last simulation uses a model which has nonlinearity-in-mean but zero third order moments:

$$x_t = sin[x_{t-1}] + \varepsilon_t,$$

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 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 3.)

We now apply the third order moment test to stock returns. If we had found large test statistics, we would have concluded that stock returns have nonlinearity-in-mean. But Table 8 shows that there is no evidence to reject the null hypothesis that stock returns have zero third order moments. As

discussed above, this is not sufficient to rule out that stock returns have nonlinearity-in-mean.

We therefore turn to nonparametric regressions to try to capture the conditional mean directly. Suppose returns are generated by the following model:

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

where f() is nonlinear and ε_t is IID. This includes chaotic models as special cases, if we set $\varepsilon_t=0$. 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 the method of locally weighted regression (LWR), which is a generalization of nearest neighbor. LWR has been employed by Diebold and Nason (1990) in testing for nonlinearity-in-mean in exchange rates, and LeBaron (1988) in stock returns.

Diebold and Nason (1990) gave a very good description of locally weighted regression. Briefly, the idea is this. Suppose we want to forecast $x_{t+1}=f(x_t,\ldots)$. We have already observed x_t, x_{t-1}, etc. Since the BDS statistic indicates that, whenever x_{t-i} was close to x_{s-i}, x_t was also close to x_s. We can look at the history of returns, find those instances which are close to $(x_t,\ldots)$, and run a regression to predict x_{t+1}. Locally weighted regressions uses the k nearest neighbors, 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) We must pick the number of nearest neighbors to use. We have tried 10% of all observations, up to 90%, increasing in steps of 10%. (b) We have to select the number of lags of x_t to use in

-22-

approximating $f(x_t, \ldots)$. We use 1 through 5. (c) We have to choose a

weighting scheme. We use the "tricubic" scheme proposed by Cleveland and

Devlin (1988).[16] (d) We need to choose a period for out-of-sample forecasting.

 For the weekly returns (VW, EW, DEC1, DEC5, DEC10, and SPW), we arbitrarily

started the forecast at the 1001-st observation and continued through the end.

 For the daily returns (SPD) we begin the forecast at the 1601-st observation.

 For the 15-minute returns (SPM1, SPM2, SPM3, SPM4), we begin the forecast at

the 1401-st 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.

    Forecastability is measured in terms of root mean squared errors, which

are reported in Table 9. The locally weighted regression forecasts are

compared to the bench mark forecast using a random walk model in prices. In

most cases, the random walk model achieves the lowest root mean squared error.

 In certain instances, e.g., VW, EW, DEC1, and SPM4, some forecasts using

locally weighted regressions were able to beat the random walk, but the

reduction in root mean squared error is less than 5%.[17] Given the wide range

of parameter values in choosing the locally weighted regression, we would

interpret this only as "data-mining" rather than "forecastability."

    If stock returns are chaotic with low complexity, we should be able to

use locally weighted regression to forecast returns much better than the

random walk model of prices. In addition, our forecasts should improve as the

forecast horizon becomes shorter and shorter. Neither of these implications

seem to be born out by the data. We do not observe consistently better

forecasting performance by the locally weighted regressions, and we certainly

do not improve our forecasting performance when we shorten the forecasting

horizon.

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 451-st 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 10 shows that LWR beat the random walk forecast 100% of the time in the 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 f( ) function is not smooth. This indicates that LWR has the ability to pick up conditional mean changes.

While we did not experiment with different methods of nonparametric regressions, we note that 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 and random walk model.

The preponderance of the failure to outperform the random walk model in asset markets force us to conclude that there is no strong evidence that the movements in stock market is primarily due to conditional mean changes, when conditioning on past returns.[18] In particular, there is no evidence of low complexity chaotic behavior in stock returns.[19] Furthermore, this casts doubt

on models trying to explain stock price movements as conditional mean changes, e.g. time varying risk premium in expected stock returns.

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

Next we proceed to consider whether stock returns have nonlinearity-in-variance:

$$x_t = g(x_{t-1}, \ldots)\, \varepsilon_t,$$

where $E[\varepsilon_t | x_{t-1}, \ldots] = 0$ and $V[\varepsilon_t | x_{t-1}, \ldots] = 1$ (without loss of generality). This is a general model of conditional heteroskedasticity, which contain 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 Sequin (1990)]. A number of papers have used ARCH-type models to describe conditional heteroskedasticity [e.g. Bollerslev (1987) and Nelson (1988).] 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 the previous equation, we obtain:

$$|x_t| = |g(x_{t-1}, \ldots)|\, |\varepsilon_t|.$$

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 conditional heteroskedasticity (particularly when x_t is not correlated with x_{t-i}).[20] Table 11 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.[21]

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:

$x_t \sim N(0, \sigma),$

$\log \sigma = \varphi_0 + \varphi \left| x_{t-1} / \sigma_{t-1} \right| + \psi \log \sigma_{-1} + \gamma \, x_{t-1} / \sigma_{t-1}.$

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

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

$z_t = x_t /_t,$

should be IID in large samples. Here, $_t$ 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 12 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

-26-

ARCH, GARCH, and EGARCH standardized residuals, a point made in Table 5. Therefore, we use the bootstrap critical values of the BDS statistic from the simulations. The 2.5% and 97.5% critical values are given in Table 13. 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:

$$x_t = \sigma_t \, z_t,$$

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

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

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 there 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 conditional heteroskedasticity (which is a form of nonlinear dependence).

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 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 [the Akaike criterion led to very long lags.] The least squares fit is as follows:

$$\log \sigma_t = -.8577 + .2385 \log \sigma_{t-1} + .1298 \log \sigma_{t-2} + .1129 \log \sigma_{t-3}$$
$$\quad (.1064)\ (.0229) \qquad\qquad (.0236) \qquad\qquad (.0236)$$
$$\quad + .1515 \log \sigma_{t-4} + .1386 \log \sigma_{t-5}$$
$$\quad (.0236) \qquad\qquad (.0229)$$
$$\quad R^2 = .4127.$$

There is clearly mean reversion in volatility.

We ran the BDS test on the residuals to test for the appropriateness of the linear model. Table 14 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 $_t$ from the variance equation: $z_t = x_t/_t$. [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 an AR1. This lag length was identified by both the Akaike and the Schwarz criterion. Table 14 contains the final diagnostics of this model. It shows that the BDS statistics are substantially lower than those in Table 12 (for SPD). If we use the asymptotic distribution of the BDS test, we do not reject the model. Even if we apply a more stringent rejection criterion and use the critical values in Table 13, we could reject only 1 BDS statistic --- at dimension 2, when $\varepsilon/\sigma=2$. In either case, we believe that the more flexible variance specification provides a much better description of the nonlinear dependence in daily stock returns.

This model gives rise to some interesting possibilities. First, the more flexible model of conditional heteroskedasticity fits the data much better, in the sense that it captures all of the nonlinear dependence in daily stock returns. Second, the mean reversion in volatility means that one can forecast future volatility based on past volatility. Third, the standardized data (after dividing by expected volatility) still has fatter tails than the normal distribution, as indicated in Table 14. Since these standardized data are IID, 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 stock returns.

8. Concluding Remarks

We have found that stock returns are not IID. The evidence points to conditional heteroskedasticity as the cause of the rejection of IID. While we

-29-

find that ARCH-type models do not fully describe the nonlinearity in stock returns, a more flexible model of conditional heteroskedasticity can. These findings have many interesting implications. One, we should not fit unconditional density functions on stock returns when these procedures assume that returns are IID. We must first remove the nonlinear dependence. Two, if we are interested to model nonlinearity in stock returns, we should direct our efforts not at modeling conditional mean changes (which include chaotic dynamics), but at modeling conditional heteroskedasticity. 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 give a conditional probability distribution, which would be useful in various contexts.

Notes

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

2. Denker and Keller (1986) and Brock and Baek (1990) provide ways to do this.

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

4. See the survey article by Bollerslev, et. al. (1990).

5. 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.

6. The results are the same using equally weighted decile portfolios.

7. 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 degrees of freedom in our data, we used the longer lags identified by the Akaike information criterion.

8. 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!

9. Note that the Cauchy distribution is a member of the stable paretian family. The simulations in Table 1 show that the asymptotic distribution of the BDS statistic can still approximate the finite distribution well, even though the Cauchy distribution has no moments.

10. See Hsieh (1990).

11. See Nelson (1989) for a discussion.

12. 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.

13. 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.

14. We note that this test can fail to detect a chaotic process whose odd product moments are zero. This can happen if the function f() is anti-symmetric. This is not true for any of the chaotic models in this paper.

15. We also reject 100% of the replications using the tent map, the Hénon map, and the logistic map (when $\alpha=4$).

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

17. These results are consistent with the findings in LeBaron (1988).

18. These results could change if we increase the information set to include variables other than past returns.

19. 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.

20. The same argument shows that x would be correlated with x_{-i} under conditional heteroskedasticity. See Engle (1982) and McLeod and Li (1983).

21. 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).

References

Akaike, H., (1974), "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control* 19: 716-723.

Baumol, W., and Benhabib, J., (1989), "Chaos: Significance, Mechanism, and Economic Applications," *Journal of Economic Perspectives* 3: 77-105.

Berndt, E.K., Hall, B.H., Hall, R.E., and Hausman, J.A., (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 Gonedes, N., (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.

Bollerslev, T., Chow, R., Jayaraman, N., and Kroner, K., (1990), "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence, with Suggestions for Future Research," unpublished manuscript, Northwestern University, Georgia Tech, Georgia Tech, and University of Arizona.

Brock, W., (1986), "Distinguishing Random and Deterministic Systems: Abridged Version," *Journal of Economic Theory* 40: 168-195.

Brock, W., (1987), "Notes on Nuisance Parameter Problems in BDS Type Tests for IID," unpublished manuscript, University of Wisconsin at Madison.

Brock, W., and Baek, E., (1990), "Some Theory of Statistical Inference For Nonlinear Science," unpublished manuscript, University of Wisconsin at Madison.

Brock, W., Dechert, W., and Scheinkman, J., (1987), "A Test for Independence Based On the Correlation Dimension," University of Wisconsin at Madison, University of Houston, and University of Chicago.

Brock, W., and Sayers, C., (1988), "Is The Business Cycle Characterized by Deterministic Chaos?" *Journal of Monetary Economics* 22: 71-90.

Clark, P.K., (1973), "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica* 41: 135-155.

Cleveland, W.W., (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association* 74: 829-836.

Cleveland, W.S., and Devlin, S.J., (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.

Diebold, F.X., and Nason, J.A., (1990), "Nonparametric Exchange Rate Prediction?" *Journal of International Economics* 28: 315-332.

Denker, G., and Keller, G., (1986), "Rigorous Statistical Procedures for Data from Dynamical Systems," *Journal of Statistical Physics* 44: 67-93.

Engle, R., (1982), "Autoregressive Conditional Heteroscedasticity With Estimates of The Variance of U. K. Inflations," *Econometrica* 50: 987-1007.

Fama, E., and Roll, R., (1968), "Some Properties of Symmetric Stable Distributions," *Journal of American Statistical Association* 63: 817-837.

French, K., Schwert, G.W., and Stambaugh, R., (1987), "Expected Stock Returns and Volatility," *Journal of Financial Economics* 19: 3-29.

Geweke, J., (1989), "Inference and Forecasting for Chaotic Nonlinear Time Series," unpublished manuscript, Duke University.

Gleick, J., (1987), *Chaos*, New York: Viking.

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

Grassberger, P., and Procaccia, I., (1983a), "Measuring the Strangeness of Strange Attractors," *Physica* 9D, 189-208.

Grassberger, P., and Procaccia, I., (1983b), "Estimation of the Kolmogorov Entropy From a Chaotic Signal," *Physical Review A*, 28: 2591-2593.

Hinich, M., and Patterson, D., (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.

Hsieh, D., (1990), "A Nonlinear Stochastic Rational Expectations Model of Exchange Rates," unpublished manuscript, Duke University.

Hsieh, D., and LeBaron, B., (1988), "Finite Sample Properties of the BDS Statistic," University of Chicago and University of Wisconsin at Madison.

LeBaron, B., (1988), "The Changing Structure of Stock Returns," unpublished manuscript, University of Wisconsin.

Mandelbrot, B., (1963), "The Variation of Certain Speculative Prices," *Journal of Business* 36: 394-419.

McLeod, A.J., and Li, W.K., (1983), "Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations," *Journal of Time Series Analysis* 4: 269-273.

Nelson, D., (1988), "Conditional Heteroskedasticity in Asset Returns: A New Approach," unpublished manuscript, Massachusetts Institute of Technology.

Nelson, D., (1989), "ARCH Models as Diffusion Approximations," unpublished manuscript, University of Chicago.

Pemberton, J., and Tong, H., (1981), "A Note on the Distribution of Non-linear Autoregressive Stochastic Models," *Journal of Time Series Analysis* 2: 49-52.

Prescott, D.M., and Stengos, T., (1988), "Do Asset Markets Overlook Exploitable Nonlinearities? The Case of Gold," unpublished manuscript, 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 Yuan, H., (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 LeBaron, B., (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 Sequin, P.J., (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 Rand, D., and Young, L., ed., *Dynamical Systems and Turbulence* (Berlin: Springer-Verlag), p. 366-382.

Tong, H., and Lim, K., (1980), "Threshold Autoregression, Limit Cycles, and Cyclical Data," *Journal of the Royal Statistical Society*, series B, 42: 245-292.

White, H., (1988), "Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns," unpublished manuscript, University of California, San
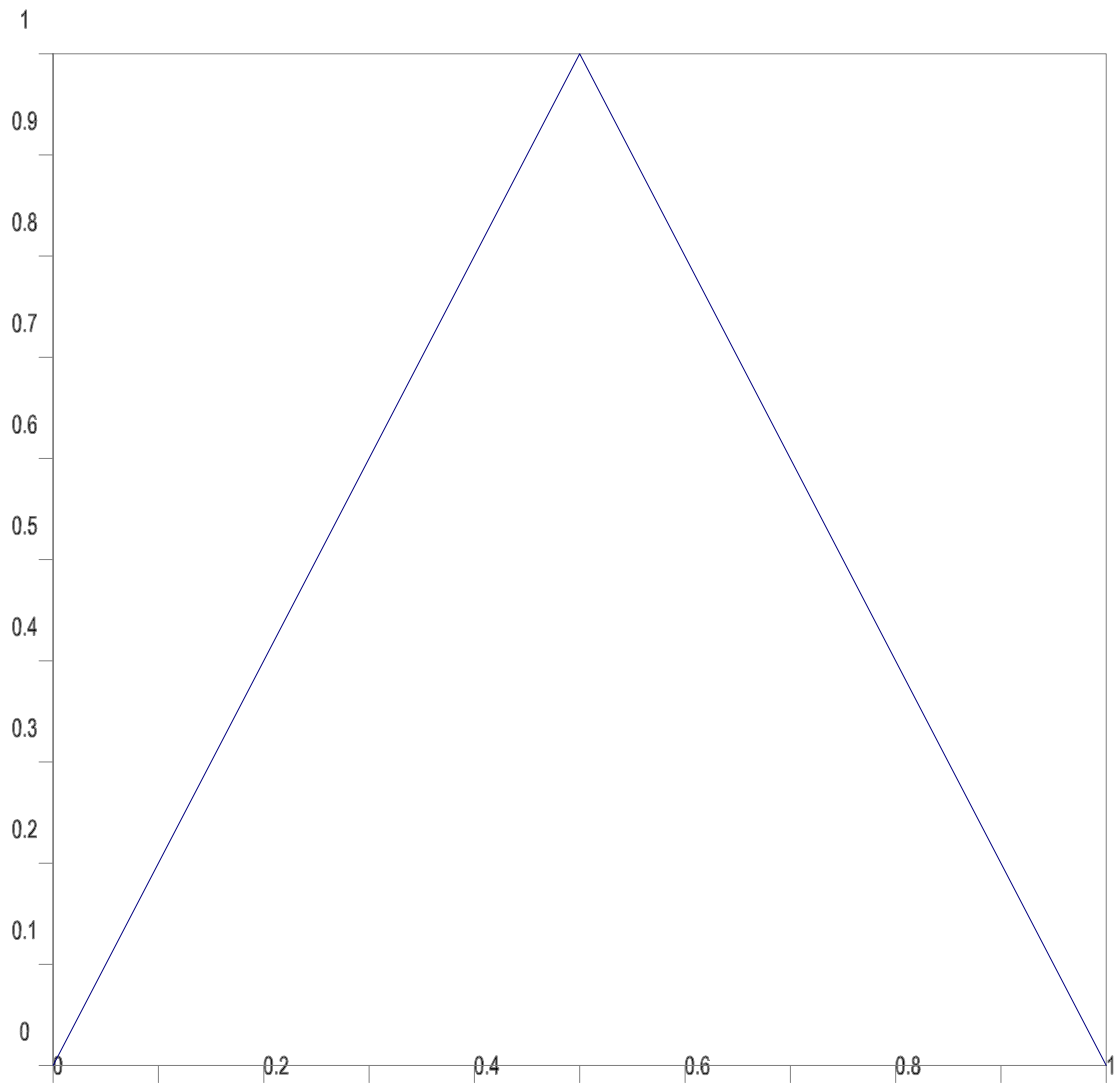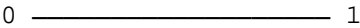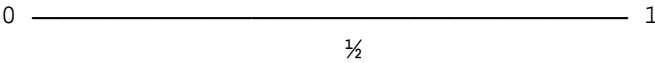
Diego.

Figure 1. Tent Map

Start with the unit interval

0 ——————————— 1

Stretch to 2 times

0 ————————————————— 1
                    ½
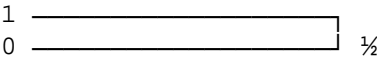
Fold

1 ———————————
0 ——————————— ½

Figure 2. Stretch and Fold
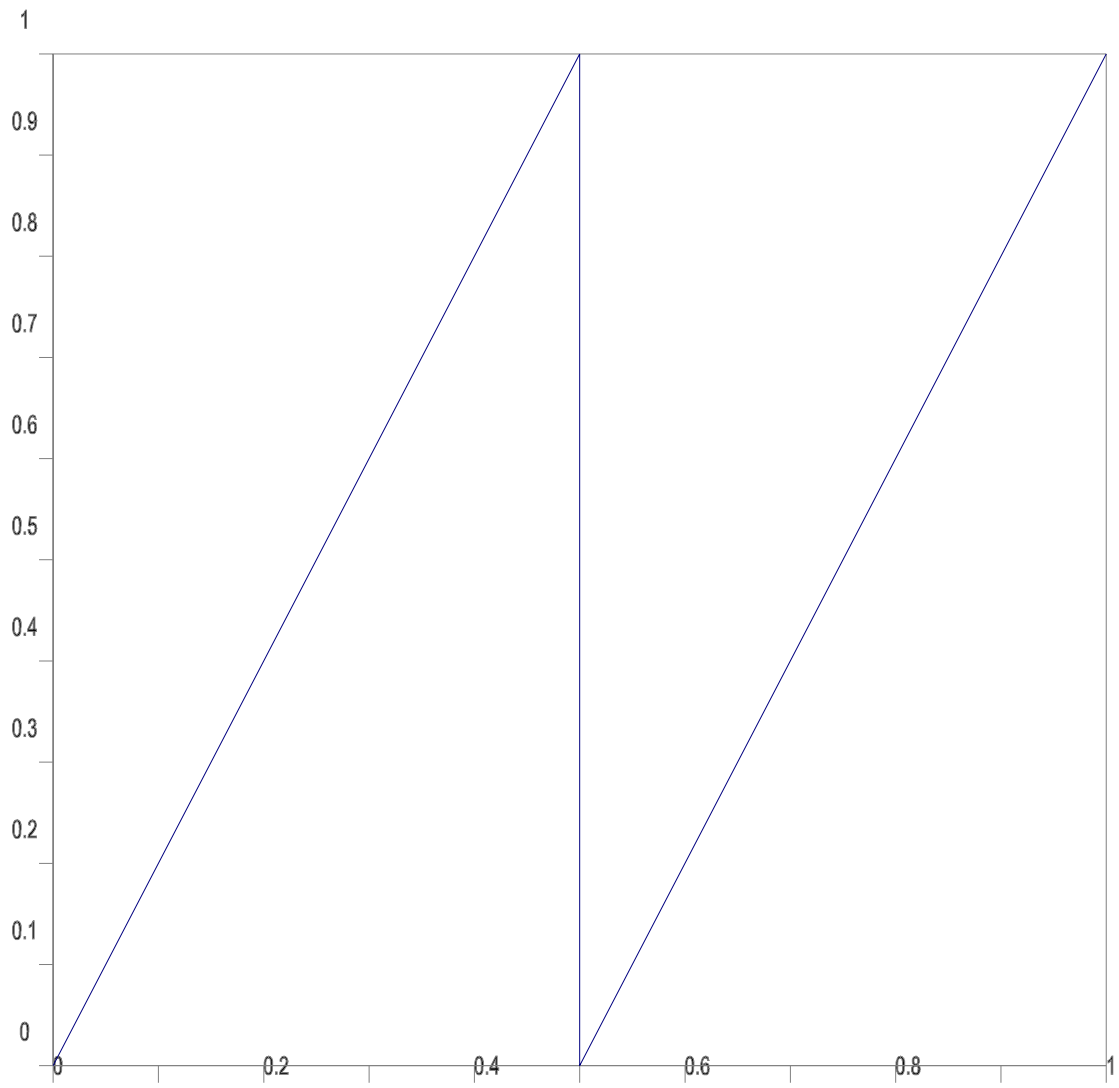
Figure 3. Diadic Map

Start with the unit interval

0 ————————————— 1

Stretch to 2 times

0 ————————————————————— 1
                    ½

Cut into 2 pieces

0 ———————————   ——————————— 1
                    ½

Stack

1 ———————————
0 ——————————— ½
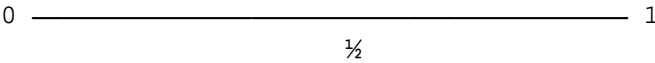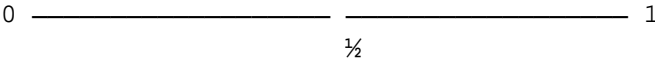
Figure 4. Stretch, Cut, and Stack

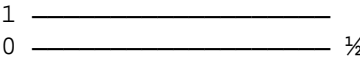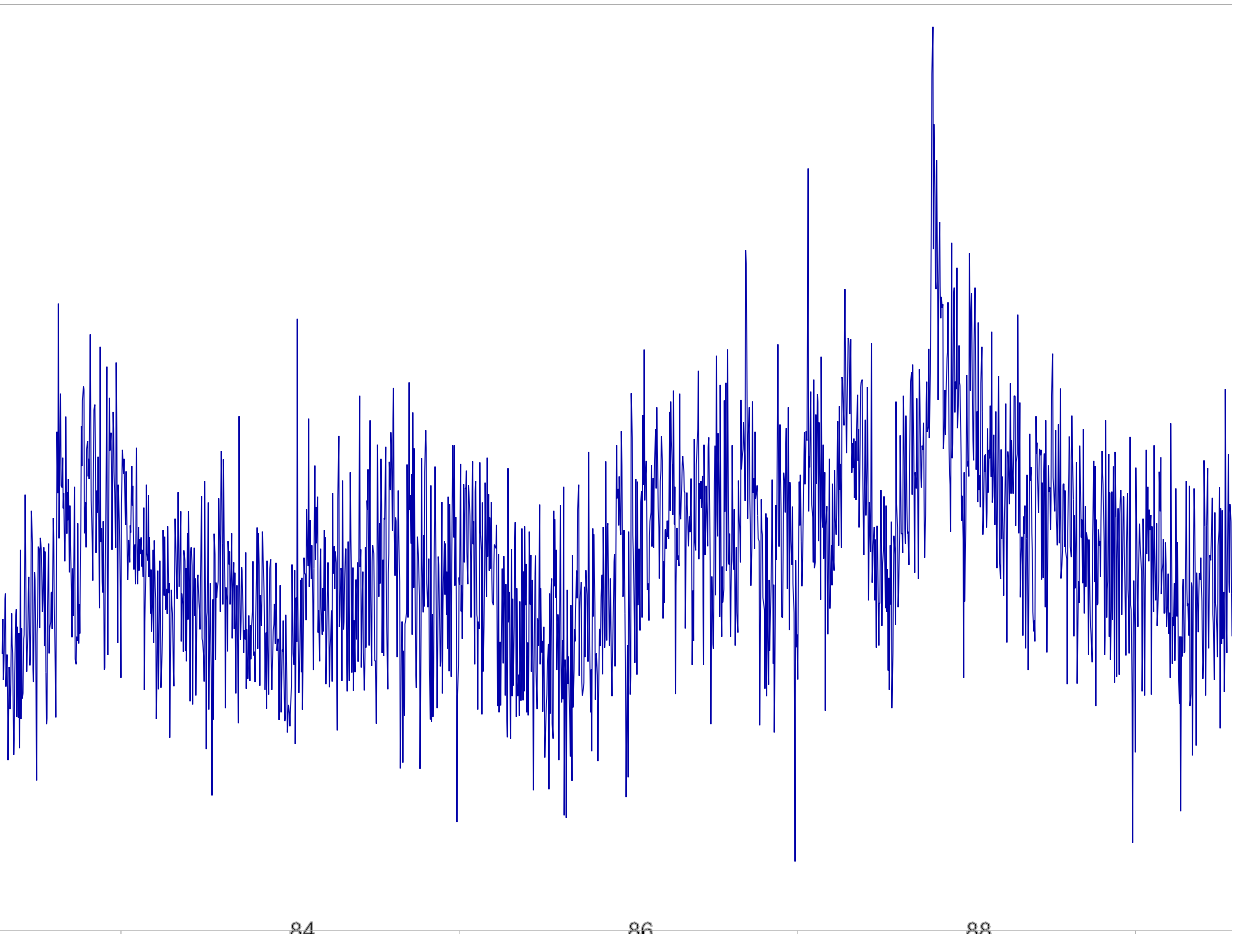84   86   88

Figure 5. Logarithm of Daily Standard Deviation of S&P500 Index