JOURNAL OF ECONOMIC THEORY **40**, 168–195 (1986)

Distinguishing Random and Deterministic Systems: Abridged Version

W. A. BROCK*

Department of Economics,
University of Wisconsin, Madison, Wisconsin 53706

Received September 4, 1985; revised January 8, 1986

This paper attempts to give a mathematically precise version of recent tests on an observed time series $\{a_t\}$ for the presence of low dimensional deterministic chaos. Three practical tests for chaos are discussed: (a) the correlation dimension of $\{a_t\}$ must be low, (b) the estimated largest Lyapunov exponent must be positive, (c) the residuals $\{a_t\}$ of an estimated linear time (or nonlinear) series model (for a large class of such models) must have the same dimension and largest Lyapunov exponent as $\{a_t\}$. Based on (a)–(c) evidence for chaos in post war II, U.S. quarterly real GNP is weak. *Journal of Economic Literature* Classification Numbers: 023, 211, 213.

1. INTRODUCTION

Recently there has been a lot of interest in nonlinear deterministic economic models that generate highly irregular trajectories, for example, [4, 15, 16, 38, 17, 19, 8]. Intense exploration of low dimensional deterministic dynamical systems models has been going on in physics and chemistry, [40], ecology and biology, [26], climatology, [28], and so on. Prigogine [31] gives a strong argument in general for nonlinear modelling in science.

The literature cited above relies heavily on mathematical literature on

* Research support of NSF Grant SES-8420872 and the Wisconsin Graduate School is gratefully acknowledged. I thank the participants at the Brussels and Paris conferences for their stimulating comments on this work. J. M. Grandmont, G. Laroque, and M. Kurz made especially helpful comments. I especially thank W. Dechert, Don Hester, C. Sayers, and J. Scheinkman. James Ramsey of NYU informed me that he is testing for chaos in financial data. W. Barnett and P. Chen of the University of Texas, Austin have recently shown that monetary aggregates have low correlation dimension. Most of all I thank J. M. Grandmont for sending me the long letters of careful and helpful suggestions many of which I incorporated in this paper. Furthermore I would like to thank the Centre for Policy Studies at Monash University, Australia for providing an excellent environment in which some of the work on this paper was done. I also thank my friend, Alice Wilcox, for typing N revisions of this thing where N is large. The usual exemption applies.

0022-0531/86 $3.00

"chaos" and nonlinear dynamics. See [14, 24] for overviews of the relevant mathematical literature.

The main reason for this recent explosion of interest in nonlinear dynamics by the applied sciences is that the trajectories generated by some nonlinear difference equations look completely random to the naked eye. A particularly dramatic example was given by Sakai and Tokumaru [35]. They show that most trajectories of the difference equation $x_{t+1} = F(x_t)$, x_0 given where

$$F(x) \equiv x/a, \qquad x \in [0, a], \; F(x) \equiv (1-x)/(1-a), \; x \in [a, 1], \; 0 < a < 1 \tag{1}$$

generate the same autocorrelation coefficients as the first-order AR process

$$v_{t+1} = (2a-1)v_t + u_{t+1}, \; \{u_t\} \text{ i.i.d.} \tag{2}$$

Brock and Chamberlain [11] show that given any spectral measure G there is a deterministic overlapping generations economy whose equilibrium trajectory generates an empirical spectrum that approximates G. Hence linear time series methods (spectral analysis and autocovariance functions) may not be able to observationally distinguish between deterministic and random systems. We come to the subject of this article.

The subject of this article is to discuss tests that are potentially capable of distinguishing between certainty and uncertainty. We will argue that such tests can be usefully applied to economic data.

This paper is organized as follows. Section 1 contains the introduction. The second section gives tests for deterministic chaos that can be applied to a time series of data. Section 3 reports applications of the tests to postwar U.S. business cycle data and to the Wölfer sunspot series. Section 4 is a short summary. Finally an expanded version of this paper is available in [13].

2. Testing Time Series for Deterministic Chaos

The setup is as follows. One observes a time series of real numbers $\{a_t\}_{t=1}^{\infty}$ that looks "erratic" or "random" to the naked eye. We are interested in discovering a few implications of the hypothesis that the time series $\{a_t\}$ is generated nevertheless by a deterministic, nonlinear, "chaotic" dynamical system, and to see how these implications can be confronted with data.

Before we can discuss methods of testing time series for deterministic chaos we must set up a precise framework, give a precise definition of "deterministic chaos", and state a precise maintained hypothesis to be tested. We need

DEFINITION 2.1 [41, 42]. The time series of real numbers $\{a_t\}_{t=1}^{\infty}$ has a *smoothly* (i.e., at least C^2) *deterministic* explanation if there exists a *system* (h, F, x_0) such that $h: R^n \to R$, $F: R^n \to R^n$ are smooth and

$$a_t = h(x_t),\ x_t = F(x_{t-1}), \qquad t = 1, 2,...;\ x_0 \text{ given.} \tag{2.1}$$

There is an analogue of this definition for continuous time systems where "$x_t = F(x_{t-1})$" is replaced by "$dx/dt = F(x)$" but for space reasons, we concentrate on discrete time systems in this paper. Also [41, pp. 316–317] only requires that h, F be Lipchitz. We will discuss later how much differentiability we need.

Think of "$x_t = F(x_{t-1})$" as an unknown law of motion of an unknown state variable. Nature knows F but the observer, the scientist, does not know F. However, the scientist observes measurements $a_t = h(x_t)$ where h is a measuring apparatus. Can the scientist by observing $\{a_t\}_{t=1}^{\infty}$ uncover and reconstruct the hidden dynamics F or can the scientist adduce evidence from $\{a_t\}$ for the coarser claim that the unknown law of motion F is deterministic?

One wishes to study the implications for the observed time series of the hypothesis that the unknown law of motion F is "chaotic". A crucial observation in this respect was made by Takens [41]. He showed that if one considered "m-histories"

$$a_t^m \equiv (a_t,..., a_{t+m-1}) = (h(x_t),..., h(F^{m-1}(x_t)) \equiv J_m(x_t)$$

the map J_m from R^n to R^m was generically a smooth diffeomorphism when m is large enough. The consequence is that for large m, the dynamical behavior of m histories will typically mimic the behavior of the unknown trajectory $\{x_t\}$. In particular, if the unknown law of motion is "chaotic", this will translate into a similar property of the trajectory followed by m-histories, for m large, which can be, in principle, confronted by the data-provided one has enough observations. Specifically, we wish to test the hypothesis that F, and thus observed m-histories a_t^m, have a *low dimensional "chaotic" attractor.*

To define the notion of deterministic chaos precisely we need a definition of chaotic attractor. This turns out to be a knotty and controversial issue that seems to be unsettled in the literature [24, p. 256]. From this point on we write "Guckenheimer and Holmes" as "GH." We are not going to be able to shed any light on the most useful definition of a chaotic or strange attractor in this paper. The definition that seems to work best for the work we want to do is that of GH [24, p. 256]:

DEFINITION 2.2. An *attractor* is an indecomposable, closed, invariant set Λ with the property that, for any $\varepsilon > 0$, there is a set U of positive Lebesgue measure in the ε-neighborhood of Λ such that $x \in U$ implies that the ω-limit set of x is contained in Λ, and the forward orbit of x is contained in U. Following [44], the attractor is *chaotic* if the largest Lyapunov exponent (to be defined below) is positive.

The above definition of attractor is not strong enough for our purposes. Since we get to observe only one time series we need existence of an orbit that densely covers the entire attractor. This property will be needed later to apply the Wolf, *et al* [44] algorithm to estimate the largest Lyapunov exponent. This extra property is given in

DEFINITION 2.3 [24, p. 237]. A closed invariant set Λ is *topologically transitive* if F has an orbit which is dense in Λ.

We will also need, in order to position ourselves for use of the [29] Multiplicative Ergodic Theorem that our attractor Λ and dynamical system F come equipped with a unique ergodic invariant measure ρ.

HYPOTHESIS 2.1. *The time series* $\{a_t\}_{t=1}^{\infty}$ *has a* C^2 *deterministic explanation* (h, F, x_0) *where* F *has a unique compact attractor that is topologically transitive and is equipped with a unique ergodic invariant measure* ρ *that has a continuous density* $\rho(dx) \equiv j(x)\,dx$. *Furthermore, the forward orbit* $\{x_t\}_{t=0}^{\infty}$ *determined by* x_0 *lies on the orbit which is dense in* Λ. *The largest Lyapunov exponent (to be defined below) is positive.*

Let us try to explain the ingredients of the maintained Hypothesis 2.1.

The subject of this section is to discuss methods for using observations of $\{a_t\}$ to uncover knowledge about (h, F, x_0) even though (h, F, x_0) is unknown and even the dimension of the vector x_0 is unknown. We need

DEFINITION 2.4 [24, pp. 283–284]. Let $F\colon R^n \to R^n$ define a discrete dynamical system. Assume F is C^1. Fix $x \in R^n$. Let $F^i(x)$ denote the ith iterate of F. Suppose that there are subspaces $V_i^{(1)} \supset V_i^{(2)} \supset \cdots \supset V_i^{(n)}$ in the tangent space at $F^i(x)$ and numbers $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$ that depend on x with the properties that

(a) $DF(V_i^{(j)}) = V_{i+1}^{(j)}$, $i = 0, 1,\ldots$

(b) $\dim V_i^{(j)} = n + 1 - j$, $i = 0, 1,\ldots$

(c) $\lim_{N\to\infty}(1/N) \ln \|D_x F^N(v)\| = \mu_j$ for all $v \in V_0^{(j)} - V_0^{(j+1)}$, $\|v\| = 1$.

Then the μ_j are called the *Lyapunov exponents* of F at x. They depend upon x. Lyapunov exponents are a generalization to general forward orbits of eigenvalues of $D_{\bar{x}}F$ at a fixed point $\bar{x}$.

Let us attempt to give meaning to the spaces $V_i^{(j)}$. For example, when $n = 2$ if $\bar{x}$ is a fixed point and $D_x F = \left[\begin{smallmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{smallmatrix}\right]$ with $\lambda_1 > \lambda_2$ then pick $V^{(1)} \equiv$

span$\{(1, 0), (0, 1)\} = R^2$ and $V^{(2)} \equiv \text{span}\{(0, 1)\}$. It is easy to check that $V_i^{(1)} = V^{(1)}$, $V_i^{(2)} = V^{(2)}$, $i = 0, 1, \ldots$ satisfy properties (a) and (b); and that $\mu_j = \ln \lambda_j$, $j = 1, 2$. Note that, in general, $V_0^{(1)} - V_0^{(2)}$ consists of all vectors which grow at the fastest possible rate, $V_0^{(2)} - V_0^{(3)}$ consists of all vectors which grow at the next fastest rate, etc. To put it another way V_0^1 is the direct sum of the "eigenspaces" corresponding to μ_i, $i = 1, 2, \ldots, n$. V_0^2 is the direct sum of the eigenspaces corresponding to μ_i, $i = 2, 3, \ldots, n$, where $\mu_1 > \mu_2$ and so on. See [2] for this interpretation and see [29, 33], for the general development. For later use we need the Oseledec multiplicative ergodic theorem.

THEOREM 2.1 [29, 24, p. 284; 2]. *Let F be C^2. Let F and its attractor A possess an ergodic invariant* [24, p. 280] *measure ρ. Then there is a ρ-measurable set $A_1 \subseteq A$ such that $\rho(A_1) = \rho(A)$ such that for all $x \in A_1$, Lyapunov exponents exist.*

Remark. One can get by with F being $C^{1+\theta}$ for some $\theta > 0$. See [33].

Remark. Since $\dim V_0^{(1)} = n + 1 - 1 = n$ therefore, as pointed out by [2, p. 2339], if one chooses the vector v in (c) "at random" one may expect to find $\mu_j = \mu_1$. Ergodicity of ρ implies [25, p. 138] that μ_1 is *independent* of x. Hence we may (and will) speak of *the* largest Lyapunov exponent of (F, A, ρ).

EXAMPLE 2.1. The tent maps (1.1) have as invariant measure $\mu(dx) = j(x)\,dx$, $j(x) = 1$, $x \in [0, 1]$. The logistic map $x_{t+1} = 4x_t(1 - x_t)$ has invariant measure $\rho(dx) = j(x)\,dx$, $j(x) = 1/(\pi(x(1-x))^{1/2})$. The largest Lyapunov exponent for these maps is given by

$$\lambda = \int_0^1 \ln |F'(x)|\, \mu(dx) \tag{2.2}$$

so that for the tent map $F(x) = 2x$, $x_t \in [0, \frac{1}{2}]$, $F(x) = 2(1 - x)$, $x \in [\frac{1}{2}, 1]$, we get $\lambda = \ln 2$. Strictly speaking Theorem 2.1 does not apply to the tent maps because F is not differentiable at $x = \frac{1}{2}$. However, Bennettin, *et. al* [3, p. 13] point out that the Raghunathan improvement of the Oseledec theorem covers cases where F is differentiable at all but a finite number of points and where F may not be invertible.

With this preparation we may explain the two tasks that the researcher must do to test for deterministic chaos in time series data: (1) show that the dimension of the time series is low, (2) show that the largest Lyapunov exponent is positive. It turns out that these quantities are invariants that may be calculated from the sequences $\{(a_t, \ldots, a_{t+m-1})\}_{t=1}^{\infty}$ under Hypothesis 2.1.

Calculating Invariants from a Time Series

In order to motivate notions of invariants suppose that $\{a_t\}_{t=1}^{\infty}$ has a deterministic explanation by the system $(h. F, x_0)$ and look at the m-history starting at t:

$$a_t^m \equiv (a_t, a_{t+1}, \ldots, a_{t+m-1}) = (h(x_t), \ldots, h(F^{m-1}(x_t))) \equiv J_m(x_t). \quad (2.3)$$

Hence $J_m: A \subseteq R^n \to R^m$. In order to grasp what is coming look at the logistic. Here $A = [0, 1]$ and $j(x) = \pi^{-1}(x(1-x))^{-1/2}$, $\pi \equiv 3.14\ldots$ Thus J_m: $[0, 1] \to R^m$, x_t is distributed according to ρ which has a continuous density on $(0, 1)$ so that the dimension of $J_m(A) \equiv \{J_m(x), x \in A\}$ is *one* for any sensible notion of dimension. This is so because $J_m(A)$ is a one-dimensional arc embedded in R^m provided that J_m is a smooth map. Notice that it is not enough to only require that J_m be continuous because you might get a space-filling curve. We need to insure that J_m carries an r-manifold A onto an r manifold $J_m(A)$ for m large enough. Furthermore Takens [42] has shown that generically J_m is 1–1 from A to $J_m(A)$ if $m \geqslant 2n+1$. The same reasoning applies whatever the dimensions of A. Calculate the dimension D_m of $J_m(A)$ for all "embedding dimensions" m and find $\lim_{m \to \infty} D_m$.

We need a precise statement of Taken's Theorem which establishes that, for $m \geqslant 2n+1$, the dynamics of "$x_t = F(x_{t-1})$" on A is equivalent to a dynamics of m-histories a_t^m on $J_m(A)$. Here "smooth" means, as always, "at least C^2".

THEOREM 2.2 [42, p. 369]. *Let $\mathcal{N}$ be a compact manifold of dimension n. For pairs (h, F), where $F: \mathcal{N} \to \mathcal{N}$ is a smooth diffeomorphism and where $h: \mathcal{N} \to R$ is a smooth function, it is a generic property that the map $J_m: \mathcal{N} \to R^m$ is an embedding provided that $m \geqslant 2n+1$. That is J_m is* 1–1 *onto $J_m(\mathcal{N})$.*

Takens gives an analogue of this theorem for flows. The importance of this theorem for us here is that the dynamics of $x_{t+1} = F(x_t)$ on $A \subseteq R^n$ are generically equivalent to the dynamics of $a_{t+1}^m = J_m \circ F \circ J_m^{-1}(a_t^m) \equiv \psi_m(a_t^m)$ for $m \geqslant 2n+1$. Here "$\circ$" means composition. To put it another way properties of the unknown dynamical system F that governs the unknown state x that are invariant through conjugacy can be, generically, transposed into properties of the dynamical system $a_{t+1}^m = \psi_m(a_t^m)$ for $m \geqslant 2n+1$. We state this as the

INVARIANCE PRINCIPLE. *Generically, conjugacy invariants of F and attractor A are preserved by J_m for $\psi_m \equiv J_m \circ F \circ J_m^{-1}$ and attractor $J_m(A)$, when $m \geqslant 2n+1$.*

There are many invariants that one could examine. Indeed Takens [41] discusses a number of such dynamical invariants which include topological entropy, various notions of dimension, and Lyapunov exponents. We are interested in notions of dimension here.

In particular it is intuitively obvious that the image $J_m(\mathcal{N})$ of an n dimensional manifold under the smooth map,

$$J_m(x) \equiv (h(x), h \circ F(x), \ldots, h \circ F^{m-1}(x)), \tag{2.4}$$

cannot have a larger dimension (for any sensible notions of dimension) than n. This follows directly from manifold theory.

Notice that continuity of J_m is not enough because "space filling curves," i.e., continuous maps J_m, [21, p. 133] can be constructed where, for example, $\mathcal{N} = [0, 1]$ and $J_m(\mathcal{N}) = [0, 1] \times [0, 1]$. Space filling curves increase "dimension" even for "sensible" notions of dimension. Turn now to notions of dimension. We shall concentrate on two notions of dimension: (a) the *limit capacity* [41, 42], and (b) the [23, 41] *correlation dimension.* The theory and properties of the limit capacity are developed by Takens [41]. This measure of dimension preserves our intuition of what "dimension" should be. But it is expensive to compute. Turn now to the practical problems of dimension computation.

The Correlation Dimension

We now face the practical problem of calculating the dimension of $J_m(A)$ for each m from a finite data set $\{a_t\}_{t=1}^{N}$. After much experimentation discussed in [13] the natural science community seems to have settled on the Grassberger–Procaccia [23] correlation dimension α_m as one of the most useful dimension measures. It is defined for $\varepsilon > 0$ by

$$\alpha_m \equiv \lim_{\varepsilon \to 0} \ln C_m(\varepsilon)/\ln \varepsilon \tag{2.5}$$

$$C_m(\varepsilon) \equiv \lim \# \{(i, j)/\|a_i^m - a_j^m\| < \varepsilon, 1 \leqslant i \leqslant N_m, 1 \leqslant j \leqslant N_m\}/N_m^2 \tag{2.6}$$

$$N_m \equiv N - (m-1). \tag{2.7}$$

Here $\# A$ denotes the cardinality of set A and ln is the natural logarithm.

Let us explain the meaning of the correlation dimension and its relationship to the more intuitive notion of limit capacity mentioned above.

The limit capacity of A measures the *rate* of growth of the cardinality of a minimal ε-separated set that "covers" A as $\varepsilon \to 0$. If A is an n dimensional manifold then the cardinality of a minimal ε-separated set that covers A will increase like ε^{-n} as $\varepsilon \to 0$. To put it another way if the minimal number of r dimension hyper cubes of side length ε needed to cover A grows like ε^{-D} as $\varepsilon \to 0$ then D is the limit capacity. Notice that D is independent of r for r large enough.

Grassberger and Procaccia [23] (GP) cite studies that show that D is very costly to compute. Therefore they and Takens [43] proposed the correlation dimension α as an alternative. We explain it in two steps. First we discuss the meaning of the correlation dimension of the underlying system (A, F) then we discuss the problem of calculating the correlation dimension of (A, F) from the observable data $\{a_t\}_{t=1}^{\infty}$ under Hypothesis 2.1.

DEFINITION 2.5. The *correlation dimension* α of (F, A, x_0) is given by (when the limits below exist)

$$C(\varepsilon) = \lim_{N \to \infty} \#\{(x_i, x_j) | \, \|x_i - x_j\| < \varepsilon, 1 \leqslant i, j \leqslant N\}/N^2 \tag{2.8}$$

$$\alpha = \lim_{\varepsilon \to 0} \ln C(\varepsilon)/\ln \varepsilon \tag{2.9}$$

where $x_i \equiv F^i(x_0)$, $i = 1, 2,...$; $x_j \equiv F^j(x_0)$, $j = 1, 2,...$

GP indicate that (i) $\alpha \leqslant D$, (ii) α is approximately 0.5 for the logistic $x_{t+1} = ax_t(1 - x_t)$ at the point $a_\infty = 3.5699456...$ that marks the end of the Feigenbaum cascade of period doubling bifurcations enroute to chaos; in contrast to $D =$ about 0.538, [23, p. 193], (iii) GP calculate α and compare it to D for the Hénon, Kaplan–Yorke, *et al.* maps and find that $\alpha \leqslant D$ in all cases. (iv) They introduce a notion of "information dimension" σ (which we shall not use in this paper) and indicate that $\alpha \leqslant \sigma \leqslant D$, (v) they show that $\alpha = \sigma = D = 1$ for the logistic map $x_{t+1} = ax_t(1 - x_t)$ with $a = 4$, and they show how to extract the correlation dimension of (F, A, x_0) from observations $\{a_t\}$ under the hypothesis of a smoothly deterministic explanation of $\{a_t\}$. The correlation dimension has nice properties. We have not found formal statements and proofs of the properties developed below for the correlation dimension in the natural science literature. We do this here. First there is

THEOREM 2.4. *The correlation dimension is independent of any two norms* $\|\cdot\|_1$, $\|\cdot\|_2$, *provided there is a constant* $K > 0$ *such that for all* $x \neq 0$

$$K \|x\|_1 \geqslant \|x\|_2 \geqslant K^{-1} \|x\|_1. \tag{2.10}$$

Proof. It is easy to see that for $1 \leqslant i, j \leqslant N$ we have

$$\begin{aligned} \#\{(i, j) \mid K \|x_i - x_j\|_1 < \varepsilon\} &\leqslant \#\{(i, j) \mid \|x_i - x_j\|_2 < \varepsilon\} \\ &\leqslant \#\{(i, j) \mid K^{-1} \|x_i - x_j\|_1 < \varepsilon\} \end{aligned} \tag{2.11}$$

Take $\lim_{N \to \infty} N^{-2}$ of both sides to get

$$C(\varepsilon/K, 1) \leqslant C(\varepsilon, 2) \leqslant C(\varepsilon K, 1) \tag{2.12}$$

where $C(\cdot, i)$ denotes the quantity C calculated under norm i, $i = 1, 2$. It is easy to see that $K > 0$ implies

$$\begin{aligned}\lim_{\varepsilon \to 0} (\ln C(\varepsilon, 1)/\ln \varepsilon) &= \lim_{\varepsilon \to 0} (\ln C(\varepsilon/K, 1)/\ln \varepsilon) \\ &\leqslant \lim_{\varepsilon \to 0} (\ln C(\varepsilon, 2)/\ln \varepsilon) \\ &\leqslant \lim_{\varepsilon \to 0} (\ln C(K\varepsilon, 1)/\ln \varepsilon) \\ &= \lim_{\varepsilon \to 0} (\ln C(\varepsilon, 1)/\ln \varepsilon). \quad \text{Q.E.D.}\end{aligned} \tag{2.13}$$

The second nice property of the correlation dimension is given by

THEOREM 2.5. *Under Hypothesis* 2.1, α_m *is (generically) independent of* m *for* $m \geqslant 2n + 1$.

Proof. Notice that $a_i^m = J_m(x_i)$, $a_j^m = J_m(x_j)$ implies

$$C_m(\varepsilon) = \#\{(i, j) \mid \|J_m(x_i) - J_m(x_j)\| < \varepsilon\}/N_m^2. \tag{2.14}$$

Put

$$\|x_i - x_j\|_2 = \|J_m(x_i) - J_m(x_j)\|, \qquad \|x_i - x_j\|_1 = \|x_i - x_j\| \tag{2.15}$$

and copy the previous proof. Use the fact that, under Hypothesis 2.1, J_m is a Lipchitz homeomorphism from Λ onto $J_m(\Lambda)$. Hence there is $K > 0$ such that for all $x_i \in \Lambda$, $x_j \in \Lambda$ we have

$$K\,\|x_i - x_j\| \geqslant \|J_m(x_i) - J_m(x_j)\| \geqslant K^{-1}\,\|x_i - x_j\|. \tag{2.16}$$

This ends the proof.

A third property of the correlation dimension is

THEOREM 2.6. *Assume Hypothesis* 2.1. *If* Λ *is an n-dimensional manifold (with boundary) that possesses an invariant measure* ρ *with* $\rho(dx) = j(x)\,dx$, $j(x) > 0$, *for all* $x \in \Lambda$, j *continuous, then* $\alpha_m = n$, *generically, for* $m \geqslant 2n + 1$.

Proof. (*See* [12]).

Fourth, a collection of other properties of the correlation dimension are discussed in [12]. In sum the correlation dimension squares well with what our intuition says dimension should be.

Natural scientists (e.g., [39] and references) construct Grassberg–Procaccia dimension plots of $\ln C_m(\varepsilon)$ against $\ln \varepsilon$ and attempt to measure the slopes $\hat{\alpha}_m$ of these G–P plots for each embedding dimension m. After constructing these plots they look to see if $\hat{\alpha}_m$ levels off to some $\hat{\alpha}$ as $m \to \infty$.

This procedure requires skill and judgment for two reasons. First since $\ln C_m(\varepsilon) = 0$ for large ε you must "estimate" the slope of the G–P plot over a range of ε of moderate to small size. Second, Brock and Dechert [12] prove that if there is noise in the data of positive variance even if the variance is small then for each embedding dimension m, $\alpha_m = m$ almost surely. Therefore the slopes $\hat{\alpha}_m$ of the dimension plots must be estimated over ε's larger than the scale $\underline{\varepsilon}$ of any noise that is present in the data set.

In practice a range of ε's over which the slope $\hat{\alpha}_m$ of the G–P plot appears to the "stable" is chosen by "eyeballing". The noise level $\underline{\varepsilon}$ is "estimated" by hunting for an $\underline{\varepsilon}$ small enough so that $\hat{\alpha}_m \cong m$ when $\hat{\alpha}_m$ is estimated over ε in $(0, \underline{\varepsilon}]$ [5]. This procedure is applied to data in Section 3 of this article. After discussing the calculation of Lyapunov exponents we will say a few words about noisy systems at the end of this section of the paper.

Calculating the Largest Lyapunov Exponent

Calculation of the largest Lyapunov exponent μ_1 is based upon the formulae (a)–(c) of Definition 2.4 and the Oseledec Theorem 2.1. We briefly explain the Wolf, *et al.* [44] algorithm which we use in our empirical work.

For each embedding dimension m we use the time series $\{a_t\}_{t=1}^{N}$ to form a time series $\{a_t^m\}_{t=1}^{N_m}$ of m-histories. Start the algorithm by locating the nearest neighbor $a_{t_1}^m \neq a_1^m$ to the initial m-history a_1^m. Let $d_1^{(1)} = \|a_{t_1}^m - a_1^m\|$. Note that $d_1^{(1)}$ is the smallest *positive* distance $\|a_t^m - a_1^m\|$. Select a positive integer q and set $d_2^{(1)} = \|a_{t_1+q}^m - a_{1+q}^m\|$ and store $g_1(q) = d_2^{(1)}/d_1^{(1)}$. We shall call q an "evolution time." This ends the first iteration. We are now ready to enter the main program loop.

Ideally, in order to start the second iteration we would like to find a new m-history $a_{t_2}^m$ near a_{1+q}^m whose angle $\theta(a_{t_2}^m - a_{1+q}^m, a_{t_1+q}^m - a_{1+q}^m)$ is close to zero. In this way we mimic the definition 2.4(c) of Lyapunov exponent as closely as possible with $a_{t_1}^m - a_1^m$ determining v through $J_m(\cdot)$. Definition 2.4(b) shows that except for hairline cases $\lim(1/N) \ln \|D_x F^N(v)\| = \mu_1$, the *largest* Lyapunov exponent, because the set $V_0^{(1)} - V_0^{(2)}$ has full Lebesque measure.

Motivated by this strategy we choose t_2 to minimize the penalty function

$$p(a_t^m - a_{1+q}^m, a_{t_1+q}^m - a_{1+q}^m) \equiv \|a_t^m - a_{1+q}^m\| + \hat{w}\,|\theta(a_t^m - a_{1+q}^m, a_{t_1+q}^m - a_{1+q}^m)| \tag{2.17}$$

subject to the nondegeneracy requirement $a_t^m \neq a_{1+q}^m$. Here $\hat{w}$ is a penalty weight on the deviation $|\theta|$ from zero. Store

$$\begin{aligned} g_2(q) \equiv d_2^{(2)}/d_1^{(2)}, \qquad d_1^{(2)} &\equiv \|a_{t_2}^m - a_{1+q}^m\|, \\ d_2^{(2)} &\equiv \|a_{t_2+q}^m - a_{1+2q}^m\|. \end{aligned} \tag{2.18}$$

This ends iteration two. Continue in this manner.

For iteration k store

$$\begin{aligned} g_k(q) \equiv d_2^{(k)}/d_1^{(k)}, \qquad d_1^{(k)} &\equiv \|a_{t_k}^m - a_{1+(k-1)q}^m\|, \\ d_2^{(k)} &\equiv \|a_{t_k+q}^m - a_{1+kq}^m\|, \end{aligned} \tag{2.19}$$

where t_k minimizes

$$p(a_t^m - a_{1+(k-1)q}^m,\, a_{t_{k-1}+q}^m - a_{1+(k-1)q}^m)$$

subject to $a_t^m \neq a_{1+(k-1)q}^m$. Continue until $k = K$ where K solves $\max\{k \mid 1 + kq \leqslant N_m\}$. Set

$$\hat{\lambda}_q \equiv \frac{1}{K} \sum_{k=1}^{K} [\ln(d_2^{(k)}/d_1^{(k)})/q]. \tag{2.20}$$

It is possible to show that, generically, Hypothesis 2.1 implies that an idealized version of the Wolf, *et al.* algorithm converges to the largest Lyapunov exponent μ_1 of (F, Λ, ρ) ρ-almost everywhere. In order to see this look first at the quantity

$$\|a_{t_1(\xi)+kq}^m - a_{1+kq}^m\| / \|a_{t_1(\xi)}^m - a_1\| \equiv R_m(\xi, w, r) \tag{2.21}$$

where

$$a_{t_1(\xi)}^m - a_1^m = \xi w + o(\xi),\ r \equiv kq \tag{2.22}$$

for a fixed direction vector w. Here $o(\xi)$ is a function of ξ such that $o(\xi)/\xi \to 0$, $\xi \to 0$. The idea we are trying to capture is that of taking a sequence of "nearest neighbors" $a_{t_1(\xi)}$ converging to $a_1^m \equiv J_m(x_1)$ along direction vector w in the tangent space $T_{a_1^m}$ to $J_m(\Lambda)$ at $J_m(x_1)$. By the definition of $T_{a_1^m}$ we may find vectors $a_{t_1(\xi)}^m$ satisfying (2.22) provided that $\{a_t^m\}_{t=1}^{\infty}$ is dense in $J_m(\Lambda)$. Density of $\{a_t^m\}_{t=1}^{\infty}$ in $J_m(\Lambda)$ follows by $\{x_t\}$ lying on the orbit who is dense in Λ by topological transitivity (cf. Hypothesis 2.1). We may now state the proof of

THEOREM 2.7 [12].
For $m \geqslant 2n+1$ under Hypothesis 2.1, for Lebesque almost all $w \in T_{a_1^m}$, for ρ-almost all x_1,

$$\lim_{r \to \infty} (\lim_{\xi \to 0} \ln(R_m(\xi, w, r)))/r = \mu_1 \equiv \lim_{r \to \infty} (\ln \|D_{x_1} F^r(v)\|/r). \tag{2.23}$$

Proof (*See* [12]).
In practice, given a finite data set $\{a_t^m\}_{t=1}^{N_m}$, the Wolf *et al.* algorithm calculates an approximation to μ_1. First the nearest neighbor a_{t_1} determines to an approximation a direction w through (2.22) which in turn determines v. This is yet still another approximation. Second for evolution time q the vector $a_{t_1+q}^m - a_{1+q}^m = J_m(F^q(x_{t_1})) - J_m(F^q(x_1))$ can be expected to become a progressively poorer approximation to $D_{x_1} J_m \circ F^q(v)$ as q increases. For this reason the Wolf, *et al.* algorithm tries to find a new nearest neighbor $a_{t_2}^m$ to a_{1+q}^m on the line segment connecting $a_{t_1+q}^m$ and a_{1+q}^m. In practice of course the distance of $a_{t_2}^m$ from this line segment must be traded off against the distance of $a_{t_2}^m$ from a_{1+q}^m. We do this by minimizing a penalty function like (2.17) where the analyst may weight the angle θ by his choice of $\hat{w}$. This induces yet another approximation.

Finally for a finite data set the average in (2.20) is necessarily a finite average. The following theorem of Brock and Dechert [12] shows that an idealized form of the Wolf, *et al.* algorithm converges to μ_1.

THEOREM 2.8 [12]. *Assume Hypothesis 2.1. Fix q and for each $k = 1, 2, \ldots$ select a new vector $a_{t_{k+1}(\xi)}$ in the line segment connecting $a_{t_k(\xi)+q}$ and a_{1+kq}. Form the ratios*

$$g_k(q, \xi) = d_2^{(k)}(\xi)/d_1^{(k)}(\xi) \tag{2.24}$$

and the sum $\hat{\lambda}_q(\xi, k)$ as in (2.19) and (2.20), respectively. Let the sequence $a_{t_1(\xi)}^m$ be determined as in Theorem 2.7. That is

$$a_{t_1(\xi)}^m - a_1^m = \xi w + o(\xi). \tag{2.25}$$

Then

$$\lim_{k \to \infty} \lim_{\xi \to 0} \lambda_q(\xi, k) = \mu_1. \tag{2.26}$$

Proof. See Brock and Dechert [12].

Remark. [2] prove this theorem for the case where x is observable. Our statement and proof seem to be the only one available for the case where x can only be observed indirectly through an "observer" function $h(x)$.

In the applications to actual time series that we shall discuss in Section 3 there will always be a natural alternative hypothesis for each time series besides deterministic chaos. For example we shall see later in Section 3 that detrended real U.S. GNP, $\{ex_t\}$, is described very well by the AR(2) model

$$ex_t = 1.36ex_{t-1} - 0.42ex_{t-2} + \delta x_t$$

where δx_t denotes the AR(2) residual at date t. But if the series $\{ex_t\}_{t=1}^{\infty}$ has a deterministic explanation (Hypothesis 2.1) one can show that the series $\{\delta x_t\}_{t=1}^{\infty}$ has the same largest Lyapunov exponent and the same dimension as the series $\{ex_t\}_{t=1}^{\infty}$. This leads to a simple test for deterministic chaos that will be used in Section 3. In general we have

RESIDUAL TEST THEOREM. *Let $\{a_t\}_{t=1}^{\infty}$ have a deterministic explanation that satisfies Hypothesis 2.1. Fit a linear time series model with a finite number of lags to $\{a_t\}_{t=1}^{\infty}$ i.e.,*

$$a_t + \gamma_1 a_{t-1} + \cdots + \gamma_L a_{t-L} = \varepsilon_t, \qquad t = L+1, \ldots$$

where ε_t is the residual at time t and $\gamma_1, \ldots, \gamma_L$ is the sequence of estimated coefficients. Then, generically, the dimension and the idealized Wolf algorithm estimate (the L.H.S. *of* (2.23)*) of the largest Lyapunov exponent of $\{a_t\}$ and $\{\varepsilon_t\}$ are the same.*

Proof. This is a simple application of the invariance principle. Since $\{a_t\}$ has a deterministic explanation $a_t = h(x_t), \ldots, a_{t-L} = h(x_{t-L})$ so that

$$\varepsilon_t = h(F^L(x_{t-L})) + \gamma_1 h(F^{L-1}(x_{t-L})) + \cdots + \gamma_L h(x_{t-L}) \equiv M(x_{t-L}).$$

Since $\varepsilon_t = M(x_{t-L})$ is a smooth function of x_{t-L}, therefore Taken's Theorem 2.2 implies that, generically, the map

$$J_m^*(y) = (M(y), M(F(y)), \ldots, M(F^{m-1}(y)))$$

is an embedding for $m \geqslant 2n+1$. Hence Theorem 2.6 implies that the correlation dimension of $\{\varepsilon_t\}$ is the same as the correlation dimension of $\{a_t\}$. Theorem 2.7 show that for $m \geqslant 2n+1$

$$\lim_{r \to \infty} (\lim_{\xi \to 0} (R_m(\xi, w, r)))/r = \mu_1,$$

regardles of whether x is "observed" through the "observer" function h or the "observer" function M. Q.E.D.

Remark. The same proof works for a generic class of smooth time series models of the form

$$\psi^*(a_{t+L_1},\dots, a_{t-L_2}; \gamma) = \varepsilon_t,\ t = L_2 + 1,\dots$$

where γ is a vector of estimated parameters and ε_t is the residual of the estimated model at date t.

Remark. The residual test is easy to apply: Find the best fitting time series model ψ^*, estimate the dimension, and the largest Lyapunov exponent of the residuals $\{\varepsilon_t\}$ and compare these estimates with those computed from $\{a_t\}$. We will use this procedure in Section 3.

Before we turn to empirical applications there are two major problems that have not been treated here that we must discuss. The first problem is the case where the initial condition, x_0, is only required to be in the basin of attraction, $B(\Lambda)$, of the attractor Λ. The second problem is the case of noise. We take these up in turn.

The Case $x_0 \in B(\Lambda)$, $x_0 \notin \Lambda$

DEFINITION 2.6. The *basin* $B(\Lambda)$ or *Domain of Attraction* [24, p. 34, pp. 111–114] of the attractor Λ is defined by

$$B(\Lambda) \equiv \{ y_0 \mid F^i(y_0) \to \Lambda, i \to \infty \}. \tag{2.27}$$

Our techniques for testing time series for deterministic chaos involve computing the correlation dimension α_m and the largest Lyapunov exponent μ_1. We have shown above that α_m, and μ_1 are built up from long run time averages and we assumed Hypothesis 2.1 to insure that these time averages exist and are independent of the initial condition $x_0 \in \Lambda$. What one needs to push the theory through for $x_0 \in B(\Lambda)$ is that the same limiting time averages exist and be independent of x_0. The appropriate tool seems to be the Sinai–Bowen–Ruelle ergodic theorem [24, p. 282] which, under the assumption of hyperbolicity of attractor Λ, asserts that there is a unique invariant probability measure ρ on Λ such that for Lebesque almost all $y_0 \in B(\Lambda)$ long run time averages of continuous functions exist. Ergodicity of ρ would imply that time averages are independent of initial conditions. Hence there would seem to be no problem in pushing the analysis of this section through for the more realistic case $x_0 \in B(\Lambda)$. However hyperbolicity of Λ [24, p. 238] is a strong assumption that is difficult to establish in examples [24, p. 238] and requires F^{-n} to exist. That is F must be globally invertible, i.e., a diffeomorphism. This excludes attractive examples like logistic chaos.

It is beyond the scope of this paper to do anymore with the case where the initial condition is not in Λ. Turn now to the next problem.

The Case of Noise

Noise can enter in two basic ways: (i) through the measuring apparatus or observer h, (ii) through the unknown (to the scientist) law of motion F. To be specific we need

DEFINITION 2.8. $\{a_t\}$ has a smooth system explanation with noisy observer (deterministic observer) and noisy law of motion (deterministic law of motion) if there is a C^2 function $h: R^n \times R \to R$ (if there exists a C^2 function $h: R^n \to R$) and if there exists a C^2 function $F: R^n \times R \to R^n$ (if there exists a C^2 function $F: R^n \to R^n$) and if there exists two independently and identically distributed stochastic processes $\{\eta_t\}$, $\{v_t\}$, each possessing a continuous, nondegenerate, density with compact support such that for all t

$$a_t = h(x_t, \eta_t), \qquad (a_t = h(x_t)) \tag{2.28}$$

$$x_t = F(x_{t-1}, v_t), \qquad (x_t = F(x_{t-1})). \tag{2.29}$$

For convenience of exposition we will run time from $-\infty$ to ∞. We need

HYPOTHESIS 2.2. *The stochastic processes* (2.28), (2.29) *are "regular" enough so that $\{a_t\}$ is strictly stationary and metrically transitive* (MTSSS). *Furthermore the joint density $P(a_t^m = u_0^m) \equiv g_m(u_0^m)$ is assumed to exist and to satisfy*

$$\int_{z \in R^m} g_m^2(z)\, dz > 0,$$

and to be bounded for each $m = 1, 2, \ldots$

There are two cases of particular interest to economists. First is the case where F is chaotic and satisfies Hypothesis 2.1 for $\{v_t\}_{t=-\infty}^{\infty}$ deterministic at mean value $\bar{v}$, where $\bar{v}$ equals the mean of v_t. We will call this case *noisy chaos*. There is endogenous instability in this case. Second is the case where F has a unique globally asymptotically stable fixed point $\bar{x}$ when $\{v_t\}_{t=-\infty}^{\infty}$ is deterministic at value $\bar{v}$, and the sequence of distribution functions of $\{x_t\}$ converges to a unique limit distribution function independently of the initial distribution (cf., [10] and references). The latter case is the case typically considered in the stochastic growth theory and modern macro-finance literature. We shall call this case a stable stochastic system. It corresponds to a system that is globally asymptotically stable if it were not buffeted by exogenous shocks. This system is inherently stable.

How do we tell from the data $\{a_t\}_{t=-\infty}^{\infty}$ which case is present or whether some other type of system generates the data? Dimension tests are useless.

THEOREM 2.9. *Under Hypothesis* 2.2, $\alpha_m = m$ *almost surely.*

Proof. By Hypothesis {2.2} $\{a_t\}$ is a stationary stochastic process with nondegenerate density at each date t. Hence $\{a_t^m\}$ fills m cubes for each m. See [12] for the details.

Remark. Intuitively speaking a theorem like Theorem 2.9 should be obvious. After all we just need a condition to insure that the sequence $\{a_t\}_{t=1}^{\infty}$ fills m-dimensional cubes for each $m = 1, 2,...$ But making this rigorous involves some mathematics, and gets rather intricate. For example, we needed the mean ergodic theorem to get a handle on

$$C_m(\varepsilon) \equiv \lim_{N \to \infty} \#\{(i, j) |\ \|a_i^m - a_j^m\| < \varepsilon, 1 \leqslant i, j \leqslant N\}/N^2.$$

The MTSSS property gives us the mean ergodic theorem. Let us give a brief idea how to show $\alpha_m = m$. The mean ergodic theorem implies

$$C_m(\varepsilon) = \int_{u,v \in B_\varepsilon(u)} g_m(u)\, g_m(v)\, du\, dv$$

$$= \int_u g_m(u) \int_{v \in B_\varepsilon(u)} g_m(v)\, dv$$

If

$$\sup g_m \equiv \bar{g}_m < \infty$$

then

$$C_m(\varepsilon) \leqslant \left(\int_u g_m(u)\, du \right) \bar{g}_m V_m(\varepsilon) = \bar{g}_m V_m(\varepsilon)$$

where $V_m(\varepsilon)$ is the m-dimensional Lebesque volume of the m-dimensional ε ball, $B_\varepsilon(u)$, centered at u. Hence

$$\alpha_m = \lim_{\varepsilon \to 0} \ln C_m(\varepsilon)/\ln \varepsilon \geqslant \lim_{\varepsilon \to 0} \ln V_m(\varepsilon)/\ln \varepsilon = m$$

See [12] for a precise statement, discussion, and the rest of the proof. Hypothesis 2.2 is stronger than necessary to get $\alpha_m = m$ for each $m = 1, 2,...$ But we do not have the space to discuss improvements here. See [12] for the details as well as a discussion of what assumptions are needed on h, F in order to get $\alpha_m = m$, $m = 1, 2,...$.

Since dimension calculations are useless when observer noise is present what can the analyst do to uncover information about (F, h, x_0) from her data?

The only thing we can offer the reader here is rough empirical suggestions that may be formalizable.

If the support or the variance of the noise in the observer is small relative to the variation of the data then we would expect information about the unknown deterministic law of motion F and the unknown deterministic attractor Λ_F to be contained in the functions $\{C_m(\varepsilon)\}^\infty$ and the Wolf, *et al.* [44] numbers $g_k(q)$, λ_q. Suppose (F, Λ_F) is chaotic with correlation dimension α and largest Lyapunov exponent $\mu > 0$.

To be specific suppose that the noise in the observer is additive and uniformly distributed with mean zero over the interval $[-\underline{\varepsilon}, \underline{\varepsilon}]$. Ben–Mizrachi *et al.* [5] indicate that in cases like this, provided the variance of x_t is large relative to $\underline{\varepsilon}$, there will be a level δ such that $C_m(\varepsilon)$ is approximately m for $\varepsilon > \delta$ and $C_m(\varepsilon)$ is approximately α for an interval of ε where $\varepsilon > \delta$. They indicate that δ can be estimated empirically and could be used to indicate the level of noise in the system. They do not present a rigorous argument and proof of this claim but they adduce experimental evidence and heuristic argument to support the claim. The same analysis is done in the case where "small" (relative to the variation in x_t) noise exists in the law of motion F.

Obviously if the noise is large enough relative to the variation of x_t across the attractor Λ_F then the structure of Λ_F will be obscured and we will have $C_m(\varepsilon)$ approximately equal to m at all scales ε that potentially contain information about Λ_F. Recall that $C_m(\varepsilon) = 1$ for ε large enough so that large ε's generate no information about Λ_F. We refer the reader to Ben–Mizrachi, *et al.* for the details. Turn now to the estimation of Lyapunov exponents in the presence of noise.

Wolf, *et al.* [44] contains a detailed heuristic discussion, sample calculations for well-known attractors, and presentation of problems that noise presents for calculation of Lyapunov exponents from finite data sets. We divide discussion into two cases (i) noise in the measurement process, i.e., the observer, and (ii) noise in the law of motion. For concreteness we consider uniform i.i.d. additive noise with support $[-\underline{\varepsilon}, \underline{\varepsilon}]$.

In case (i) the system possesses well defined Lyapunov exponents that are potentially recoverable. Strictly speaking in case (ii) Lyapunov exponents are not defined. But Wolf, *et al.* [44] mention approaches that generate numbers that are the Lyapunov exponents for the noise free system averaged over the range of noise induced states. More discussion on what information is extractable from empirical estimates of Lyapunov exponents in the presence of noise is in the empirical section of this paper. Turn now to empirical applications.

3. Empirical Application of These Ideas

Empirical calculation of the Grassberger–Procaccia [23] correlation dimension α and the largest Lyapunov exponent λ is the procedure typically used in natural science to test for the presence of chaos in time series data [39, 44]. For some low dimensional systems the dynamics $\psi_m = J_m \circ F \circ J_m^{-1}$ are "reconstructed". Economists must deal with time series much shorter than the 10,000–30,000 observations typically used in natural science work and, furthermore, economic time series are probably noisier. The problem is especially acute in business cycle analysis. For example, in some of the natural science studies reported by Swinney [39, 40] explicit reconstruction of the attractor is possible. Here we apply the residual test to test the hypothesis of deterministic chaos against the alternative of a linear model.

Brock and Sayers [9] test for nonlinearities in quarterly data on U.S. real GNP and U.S. real gross private domestic investment by (a) calculating the Grassberger–Procaccia [23] correlation dimension and estimating the largest Lyapunov exponent for various embedding dimensions; (b) calculating measures of asymmetry, and measures of skewness and kurtosis, and (c) applying the residual test for determinatistic chaos. For lack of space, we only describe some of the results for U.S. quarterly data from 1947 : 1 to 1985 : 1 (1972 = 100), and for Wölfer's sunspot numbers 1749–1924. Wölfer's sunspot numbers is a well known asymmetric series to contrast with our economic time series.

Let x_t = real GNP at quarter t and y_t = real gross private investment at quarter t. The data was detrended by the following OLS regressions:

$$\begin{gathered} \log x_t = \underset{(982.1)}{2.681} + \underset{(119.4)}{0.003671}t + ex_t, \\ R^2 = 0.990,\ \mathrm{RSS} = 0.043,\ \mathrm{TSS} = 4.065, \\ \mathrm{SE} = 0.017,\ \mathrm{nobs} = 153 \end{gathered} \tag{3.1}$$

$$\begin{gathered} \log y_t = \underset{(223.3)}{1.851} + \underset{(403)}{0.003765}t + ey_t, \\ R^2 = 0.915,\ \mathrm{RSS} = 0.393,\ \mathrm{TSS} = 4.625, \\ \mathrm{SE} = 0.051,\ \mathrm{nobs} = 153 \end{gathered} \tag{3.2}$$

Here the numbers in parentheses denote t statistics, RSS = residual sum of squares, TSS = total sum of squares, SE = standard error of estimate, and nobs = number of observations. In view of the well-known result that autoregressive models of order two (AR(2)) fit detrended U.S. real GNP well we fit two AR(2) models:

$$ex_t = \underset{(18.2)}{1.36}\, ex_{t-1} \underset{(-5.6)}{-0.42}\, ex_{t-2} + \delta x_t,$$

$$R^2 = 0.933,\ \mathrm{RSS} = 0.003,\ \mathrm{TSS} = 0.042, \qquad (3.3a)$$

$$\mathrm{SE} = 0.025,\ \mathrm{nobs} = 151$$

$$\sigma_{ex} = 0.0167,\ \sigma_{\delta x} = 0.0043,\ \sigma_{ex}/\sigma_{\delta x} = 3.88,$$

$$sk_{ex} = -0.244,\ sk_{\delta x} = -0.086, \qquad (3.3b)$$

$$k_{ex} = 2.19,\ k_{\delta x} = 4.05.$$

$$ey_t = \underset{(14.4)}{1.12}\, ey_{t-1} \underset{(-4.0)}{-0.31}\, ey_{t-2} + \delta y_t,$$

$$R^2 = 0.760,\ \mathrm{RSS} = 0.094,\ \mathrm{TSS} = 0.392, \qquad (3.4a)$$

$$\mathrm{SE} = 0.025,\ \mathrm{nobs} = 151$$

$$\sigma_{ey} = 0.051,\ \sigma_{\delta y} = 0.025,\ \sigma_{ey}/\sigma_{\delta y} = 2.03,$$

$$sk_{ey} = -0.46,\ sk_{\delta y} = -0.57, \qquad (3.4b)$$

$$k_{ey} = 2.78,\ k_{\delta y} = 5.14.$$

Have σ_A, sk_A, k_A denote standard deviation of A, skewness of A, kurtosis of A. The "t-statistics" are reported even though the "significance" of the second coefficient changes a lot when we fit AR(3) and AR(4) to this data. Lags greater than two did not appear to be "significant" however.

For $\{ex_t\}$, $\{ey_t\}$, $\{\delta x_t\}$, $\{\delta y_t\}$ we calculated, for embedding dimension d,

$$C(\varepsilon) \equiv \frac{1}{n_{d^2}} \#\{(i, j) \mid \|a_i^d - a_j^d\| < \varepsilon\},$$

$$C^*(\varepsilon) \equiv \frac{1}{n_d^*} \#\{(i, j) \mid i \neq j,\ \|a_i^d - a_j^d\| < \varepsilon\} \qquad (3.5)$$

$$\alpha(\varepsilon) \equiv \ln C(\varepsilon)/\ln \varepsilon,\ \alpha^*(\varepsilon) \equiv \ln C^*(\varepsilon)/\ln \varepsilon, \qquad (3.6)$$

$$SC(\varepsilon_k, \varepsilon_{k-1}) \equiv (\ln C(\varepsilon_k) - \ln C(\varepsilon_{k-1}))/(\ln \varepsilon_k - \ln \varepsilon_{k-1}) \qquad (3.7)$$

$$SC^*(\varepsilon_k, \varepsilon_{k-1}) \equiv (\ln C^*(\varepsilon_k) - \ln C^*(\varepsilon_{k-1}))/(\ln \varepsilon_k - \ln \varepsilon_{k-1}) \qquad (3.8)$$

and estimated the largest Lyapunov exponent λ. Here $n_d^* \equiv n_d^2 - n_d$, $\#S$ denotes the cardinality of set S; $n_d = n - (d-1)$ is the number of d-histories $a^d = (a_t^d, \ldots, a_{t+d-1})$ constructed from the sample of length n, and $1 \leqslant i, j \leqslant n_d$.

Three difficulties emerged. First a standard of comparison was needed so that meaning could be attached to a "large" or "small" dimension or Lyapunov exponent. To take care of this difficulty, we calculated the same measures of dimension and Lyapunov exponent for normal pseudo random numbers of the same mean and standard deviation for each series.

We remark parenthetically that there is a problem in finding a good source of "random" numbers. We used the Gauss [20] software package to generate "random" numbers. The Gauss generator, like all computer "random" number generators, is a deterministic mechanism. Since I do not know how to generate true random numbers at a reasonable cost I looked to see if the correlation dimension of Gauss's random numbers was large. The evidence for embedding dimension 20 in Table I is typical in my

TABLE I

Dimension Calculations for Detrended U.S. Real GNP
U.S. Real GNP 1947 : 1–1985 : 1

ex_t, $n_d = 134$						$\tilde{e}x_t$, $n_d = 134$					
ε	S_{20}	α_{20}	α^*_{20}	SC_{20}	SC^*_{20}	ε	S_{20}	α_{20}	α^*_{20}	SC_{20}	SC^*_{20}
0.9	9582	5.96	6.02	2.11	1.97	0.9	7246	8.61	8.72	1.54	1.56
0.9^2	7674	4.03	3.99	2.45	2.65	0.9^2	6166	5.07	5.14	1.59	1.63
0.9^3	5936	3.50	3.55	3.17	3.26	0.9^3	5216	3.91	3.97	2.19	2.25
0.9^4	4248	3.42	3.48	3.35	3.46	0.9^4	4144	3.48	3.54	3.13	3.25
0.9^5	2990	3.40	3.48	3.41	3.59	0.9^5	2982	3.41	3.48	2.62	2.75
0.9^6	2090	3.40	3.50	3.14	3.39	0.9^6	2266	3.27	3.36	2.32	2.49
0.9^7	1502	3.36	3.48	2.87	3.19	0.9^7	1774	3.14	3.23	3.54	3.88
0.9^8	1112	3.30	3.44	3.13	3.65	0.9^8	1224	3.19	3.32	2.91	3.32
0.9^9	800	3.28	3.47	2.38	2.95	0.9^9	902	3.15	3.32	2.61	3.13
0.9^{10}	622	3.19	3.41			0.9^{10}	686	3.10	3.30		

δx_t, $n = 132$						$\tilde{\delta} x_t$, $n = 132$					
ε	S_{20}	α_{20}	α^*_{20}	SC_{20}	SC^*_{20}	ε	S_{20}	α_{20}	α^*_{20}	SC_{20}	SC^*_{20}
0.9	5636	11.00	10.87	6.03	6.21	0.9	4966	11.91	12.10	6.93	7.22
0.9^2	2992	8.50	8.54	6.02	7.30	0.9^2	2392	9.42	9.66	8.16	8.95
0.9^3	1590	7.67	8.12	7.16	7.24	0.9^3	1012	9.00	9.42	9.05	11.65
0.9^4	750	7.54	7.90	6.68	8.98	0.9^4	390	9.02	9.98	7.13	15.20
0.9^5	372	7.36	8.12	4.25	7.76	0.9^5	184	8.64	11.02	2.46	12.19
0.9^6	238	6.84	8.06	2.55	7.13	0.9^6	142	7.61	11.79	0.69	∞
0.9^7	182	6.23	7.93	2.49	17.39	0.9^7	132	—	∞	0	∞
0.9^8	140	5.76	9.11	0.56	∞	0.9^8	132	—	∞	0	∞
0.9^9	132	5.14	∞	0	∞	0.9^9	132	—	∞	0	∞
0.9^{10}	132	4.62	∞	0	∞	0.9^{10}	132	—	∞	0	∞

Note. Embedding dimension $d = 20$, $n_d = 132$, 134, $S_{20} = \#\{(i, j) \mid \|a_i^d - a_j^d\| < \varepsilon\}$.

simulations. In contrast I get small dimensions for short series like that of Tables I–II for the Henon and logistic maps. See [9] for more discussion. The Gauss generator seems to generate a usably large dimension. Also the behavior of the Lyapunov exponents reported in Table III is typical in my simulations. This behavior appears tolerably random.

Second, the utility of the measure $\alpha(\varepsilon)$ (suppressing obvious subscripts for ease in typing) is based on "the power law conjecture".

$$C(\varepsilon) \cong K(\varepsilon)\, \varepsilon^{\alpha(\varepsilon)}, \qquad \alpha(\varepsilon) \to \alpha,\ K(\varepsilon) \to K,\ \varepsilon \to 0. \tag{3.9}$$

so that

$$\ln C(\varepsilon)/\ln \varepsilon = \ln K(\varepsilon)/\ln \varepsilon + \alpha(\varepsilon) \to \alpha,\ \varepsilon \to 0. \tag{3.10}$$

With a time series of length 10,000, $\ln C_d, n_d(\varepsilon)/\ln \varepsilon$ may approximate $\ln C_{d,\infty}(\varepsilon)/\ln \varepsilon$ even for small ε but a time series of length 100–200 is a different matter. Here we use subscripts d, n_d when we want to emphasize the dependence upon d, n_d. For example let $\underline{\rho} = \min\{\|a_i^d - a_j^d\| > 0, 1 \leqslant j \leqslant n_d\} > 0$, be the smallest positive distance. For $\varepsilon < \underline{\rho}$, $C_{d,n_d}(\varepsilon) = 1/n_d$, $\alpha_{d,n_d}(\varepsilon) = -\ln(n_d)/\ln \varepsilon \to 0$, $\varepsilon \to 0$. Note also that the slope estimator $SC_{d,n_d}(\varepsilon_k, \varepsilon_{k-1}) = 0$ for $\varepsilon_k, \varepsilon_{k-1} < \underline{\rho}$. This is true even if $\{a_t\}$ is a sequence of random numbers for which the theoretical value of α and SC is d. Hence α and SC are poor estimators of the underlying dimension especially for small data sets.

Since α^*, SC^* give values much closer to the theoretical value of d for random numbers therefore we calculated the measures α^*, SC^* as well.

TABLE II

Dimension Calculations for Wolfer Sunspot Numbers
Wölfer Sunspot Numbers: 1749–1924 from [1]

x_t, $n_d = 157$						δx_t, $n_d = 155$					
ε	S_{20}	α_{20}	α_{20}^*	SC_{20}	SC_{20}^*	ε	S_{20}	α_{20}	α_{20}^*	SC_{20}	SC_{20}^*
0.9	2627	21.25	21.77	2.98	3.21	0.9	1543	26.06	27.00	5.00	5.74
0.9^2	1919	12.12	12.49	2.93	4.52	0.9^2	551	11.94	16.37	4.79	6.16
0.9^3	1409	9.05	9.83	2.98	2.16	0.9^3	313	10.30	12.97	5.37	8.73
0.9^4	1029	7.54	7.91	4.61	5.75	0.9^4	197	9.12	11.91	4.39	12.58
0.9^5	633	6.95	7.48	3.18	4.51	0.9^5	179	7.75	12.04	0.91	5.31
0.9^6	453	6.32	6.99	2.81	4.72	0.9^6	159	6.80	10.92	1.13	17.01
0.9^7	337	5.82	6.66	2.00	4.17	0.9^7	155	5.98	11.79	0.24	∞
0.9^8	273	5.34	6.35	1.18	3.07	0.9^8					
0.9^9	241	4.88	5.99	0.77	2.89						
0.9^{10}	223	4.47	5.62								

Note. $x_t \equiv s_t - \bar{s}$, $x_t = 1.336x_{t-1} - 0.65x_{t-2}$, $R^2 = 0.802$, $n = 174$, $n_d = 155, 157$.

TABLE III

Lyapunov Exponents and Other Instability Measures for Wölfer's Sunspot Numbers

Sunspot numbers

q	K	$\hat{\lambda}$	g	σ_g
1	100	0.64	3.22	5.67
2	50	0.57	8.22	22.81
3	45	0.38	5.96	7.67
4	35	0.35	7.78	8.68
5	25	0.28	8.52	10.71
6	24	0.21	6.45	8.20
7	21	0.25	7.93	5.45
8	18	0.15	4.31	2.55
9	16	0.16	5.78	4.54

AR(2) Residuals of Sunspot Numbers

q	K	$\hat{\lambda}$	g	σ_g
1	100	1.45	6.91	8.08
2	50	0.97	11.26	10.11
3	45	0.62	9.40	7.88
4	35	0.48	10.46	9.24
5	25	0.34	8.05	6.92
6	24	0.30	8.93	8.30
7	21	0.27	12.14	12.45
8	18	0.20	8.54	9.51
9	16	0.19	6.97	6.94

3 Runs on pseudo random numbers to compare with sunspot numbers

q	K	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\bar{g}_1$	$\bar{g}_2$	$\bar{g}_3$	σ_{g1}	σ_{g2}	σ_{g3}
1	100	1.01	0.89	1.01	4.06	4.51	3.95	4.26	10.57	3.77
2	50	0.64	0.68	0.57	5.88	6.32	4.67	6.80	6.64	5.49
3	45	0.47	0.47	0.49	5.84	6.07	8.07	6.01	6.22	10.31
4	35	0.36	0.33	0.35	6.13	6.49	6.55	4.92	10.90	6.01
5	25	0.28	0.24	0.30	5.79	5.21	6.09	3.89	5.02	4.35
6	24	0.29	0.19	0.25	10.66	4.19	7.53	12.52	2.95	9.05
7	21	0.19	0.18	0.23	5.95	6.79	8.16	8.62	9.49	8.24
8	18	0.17	0.15	0.21	5.50	17.87	7.42	4.37	60.42	6.12
9	16	0.17	0.15	0.16	7.07	4.88	8.94	6.16	2.88	13.12

These measures attempt to remove the distortion in a possible power law caused by the fact $C_{d,n_d}(\varepsilon) = 1/n_d$ on $[0, \rho)$.

There is the further problem of how to report the results of dimension calculations for the small data sets used in business cycle analysis. That is to say what constitutes a "confidence interval" and "rejection region" under a given null hypothesis for an "estimator" where we know very little about the sampling distribution and where generating an empirical distribution under various null hypotheses is expensive? We shall handle this problem as it is handled in the natural science literature that we have seen to date. That is, to try to report our findings in such a way that the reader may make her own judgment as to the significance or stability of any estimate. Due to lack of space we can only report the highlights of the results of Brock and Sayers [9] here.

Third, even less is known about the small sample properties of our estimate of the largest Lyapunov exponent. Hence we calculated a measure of *cumulative* spreading between initially close trajectories for (a) our data set, (b) random numbers of the same mean and variance as our data set, and (c) a time series generated by the tent map (1.1) with $a=\frac{1}{2}$.

In a nutshell our results indicate the following: there is not enough information in the 1947 : 1–1985 : 1 data set for the Grassberger–Procaccia type dimension measures to reject the null hypothesis that detrended real GNP and real GPDI are generated by an AR(2) process as estimated in (3.3) and (3.4). To put it another way we need more data to establish that U.S. real GNP and real GPDI are generated (in the main) by a low dimension chaotic deterministic dynamical system.

It may be helpful to future researchers in this area to describe the chain of thought and the sequence of events that lead us to this conclusion. First we calculated measures of dimension for $\{ex_t\}$, $\{ey_t\}$ and got evidence of low dimension as revealed in the tables below.

Second, our colleague Don Hester suggested that AR(2) models like (3.3) and (3.4) might generate a low dimension estimate if they fit the data well. Although Theorem 2.9 assures us that the dimension of an infinite data set generated by (3.3) or (3.4) must be infinity, a data set of length 153 may be a long way from infinity especially if the rate of convergence of dimension estimates is slow. Since our time series was too short to reconstruct the attractor along the lines of Swinney [39, 40] and since an AR(2) model is the received alternative hypothesis, therefore we used the residual test to test the hypothesis of deterministic chaos against the AR(2) hypothesis.

Recalling the logic of the residual test it is easy to prove that if, say, $\{ex_t\}$ has a smoothly deterministic explanation

$$ex_t=h(z_t), \qquad z_t=F(z_{t-1}),\ z_0 \text{ given}, \tag{3.11}$$

then, from (3.3a),

$$\delta x_t=h(F^2(z_{t-2}))-1.36h(F(z_{t-2}))+0.42h(z_{t-2})\equiv M(z_{t-2}). \tag{3.12}$$

Hence the dimension of $\{\delta x_t\}$ should be the same as the dimension of $\{ex_t\}$ if $\{ex_t\}$ has a deterministic explanation. In fact Table I shows evidence that the dimension of $\{\delta x_t\}$ is slightly (?) smaller than the "large" dimension of a sequence $\{\tilde{\delta} x_t\}$ of the same length of normal pseudo-random numbers with the same mean and variance. This is evidence against the chaotic, low, dimensional, deterministic, dynamical, system hypothesis.

Third, faced with this mixed evidence we generated sequences $\{\tilde{\delta} x_t\}$, $\{\tilde{\delta} y_t\}$ of normal pseudo-random numbers with the same mean and variance as $\{\delta x_t\}$, $\{\delta y_t\}$ and generated two simulated AR(2) time series $\{\tilde{e}x_t\}$, $\{\tilde{e}y_t\}$ of the same length as $\{ex_t\}$, $\{ey_t\}$ using the fitted values of the AR(2) models in (3.3) and (3.4). Grassberger–Procaccia measures of dimension were calculated for $\{\tilde{e}x_t\}$ and $\{\tilde{e}y_t\}$ and compared with those calculated for $\{ex_t\}$ and $\{ey_t\}$. There appears to be no significant difference to the naked eye. See Table I for real GNP. A qualitatively similar result holds for real GPDI [9].

Fourth, the same battery of tests was applied to the Wölfer sunspot series [1, 30] which displays asymmetries in that the average gradient of rise from trough to peak is greater than the average gradient of fall from peak to trough across "cycles" [30, p. 882]. Furthermore, an AR(2) model fits the sunspot series quite well although a better fit is obtained with a bilinear model according to the Akaike criterion (Priestley [30, p. 884]). The story is similar to that of real GNP and real GDPI. See Tables II, III.

Fifth, the same procedure was applied to the Lyapunov exponent calculations. Look at Table 3 which contains calculations for Wolfer's sunspot numbers. We calculated $\hat{\lambda}_q$ from (2.20) for $q = 1, 2, \ldots, 9$ as well as the mean $\bar{g}$ and the standard deviation σ_g of $\{g_k(q)\}_{k=1}^{K}$. Notice that if $\{a_t\}$ is a sequence of i.i.d. random numbers, then from (2.19) $\ln(d_2^{(k)}/d_1^{(k)})/q \to 0$, $q \to \infty$. Hence $\hat{\lambda}_q \to 0$, $q \to \infty$ by (2.20). But, in contrast, if there is deterministic chaos present in the data we should see a tendency of $\bar{g}(q)$ to grow with q. This is so because $\bar{g}(q)$ is a measure of spreading of nearby trajectories.

Look at Table III where three typical runs with pseudo-random numbers were done in order to get a perspective. Notice the absence of any tendency of $\bar{g}(q)$ to grow with q and the tendency of $\hat{\lambda}_q$ to fall with q for each of the three runs. Now turn to the same calculations with the sunspot numbers and their AR(2) residuals. If there are instabilities, i.e., deterministic chaos present in the sunspot data we should see a tendency of $\bar{g}(q)$ to grow with q and $\hat{\lambda}_q$ should not fall with q. Our own naked eye sees little difference from the three runs of random numbers. The same pattern emerged in Lyapunov exponent calculations from the AR(2) residuals from real GNP and real GPDI.

Conclusion. Dimension tests and Lyapunov exponent calculations have a hard time rejecting a linear model where there is a lot of variation of the data "*within* the regression plane" relative to the variation of the data "*around* the regression plane," when the data set is small. Our findings reported above should not be constructed as negative to the attempt to find evidence of significant nonlinearities in economic data.

This is so for several reasons. First we used aggregate data. Hence many nonlinearities at the microlevel may have been "washed out". Second the data set 1947 : 1–1985 : 1 is a period where severity of recessions has fallen (Zarnowitz and Moore [46, pp. 12–15]) and growth cycles [27, Table 8]) look quite symmetric throughout this period. Therefore, our results square with DeLong and Summers [18] who find little evidence of significant skewness of growth rates of GNP and industrial production for the postwar period. The failure to find large kurtosis agrees with Blanchard and Watson [6]. This does not contradict Blatt because he looks at different series over different periods for example pig iron production data from Burns and Mitchell [7, p. 231].

Third Brock and Sayers [9] have examined US unemployment data and have found tentative evidence for nonlinear structure in that data, as well as in other data.

We summarize the main message for empirical applications of these methods. It is *not* enough when you are working with short data sets to report low dimension and positive Lyapunov exponents to make the case for deterministic chaos in your data. Plausible alternative hypotheses (such as a linear model for aggregative macroeconomic time series data) must be tested. This can be done by fitting a plausible model to the data (such as an AR_2 model used here) and calculating the dimension and Lyapunov exponents of the residuals. This is the residual test.

We have not seen the residual test used before. It appears to be new. In any event we believe, in the absence, of explicit reconstruction of the dynamics $a^m_{t+1} = J_m(F(J^{-1}_m(a^m_t)))$ that the residual test could be used to test claims of chaos in the natural sciences as well as in economics.

4. Summary

In this paper we briefly reviewed methods devised by the natural sciences to test for the presence of low dimensional nonlinear chaos in time series data. The tests consist of two parts. First calculate some notion of dimension and show that it is small. Second calculate an estimate of the largest Lyapunov exponent and show that it is positive.

In this paper we amended and adapted these tests to maximize the information available in the short data sets available in business cycle analysis. We added the residual test. Our findings indicate that there is not enough information available in U.S. real GNP, real gross private domestic investment, and Wölfer's sunspot series for the three part test discussed here to reject the null hypothesis that the series under scrutiny was generated by an AR(2) process.

Finally a more extensive discussion of these methods is available in [13] and a more extensive discussion of empirical applications to business cycles and labor statistics is in [9, 36]. Scheinkman and LeBaron's [37] paper calculates dimension estimates for stock market data. They find strong evidence of *non* independence of stock returns.

More references that I have found useful in my literature search as well as references to other studies that I have found are in [13].

We close with a final caveat: The empirical results reported here are tentative and should be viewed as illustrative calculations only. For example, we are not satisfied with our present algorithm to calculate Lyapunov exponents for the small data sets that we are analyzing.

Also the residual test may mis-identify deterministic chaos as random noise in a short data set. As an example generate a short time series $\{x_t\}$ comparable in length to the real GNP series studied in this paper from the tent map (1.1), fit an AR(2), and compare the dimension of the residuals with the dimension of the original series. The estimated dimension of the residuals may be larger than the estimated dimension of the original series because the map J_m^* on m-histories induced by the AR(2) residuals may be more "irregular" than the map J_m induced by the original series. Our group is extending and updating our findings reported here.

References

1. T. W. Anderson, "The Statistical Analysis of Time Series," Wiley, New York, 1971.
2. G. Benettin, et al, Kolmogorov entropy and numerical experiments, *Phys. Rev. A* (3) **14**, (1976), 2338–2345.
3. G. Benettin, et al, Lyapunov characteristic exponents for smooth dynamical systems: A method for computing all of them, *Meccanica* **15**, (1980), 9–20.
4. J. Benhabib, and R. H. Day, A characterization of erratic dynamics in the overlapping generation model, *J. Econ. Dynam. Control* **4**, (1982), 37–55.
5. A. Ben-Mizrachi, et al. Characterization of experimental (noisy) strange attractors, *Phys. Rev. A* (3) **29**, (1984), 975–977.
6. O. Blanchard and M. Watson, "Are All Business Cycles Alike?" Department of Economics, MIT Press and Harvard Univ. Press, Cambridge, Mass. Jan. 1984.
7. J. M. Blatt, "Dynamic Economic Systems," (M. E. Sharpe, Ed.), Armonk, New York, 1981.
8. M. Boldrin and L. Montrucchio, On the Indeterminacy of Capital Accumulation Paths, *J. Econ. Theory* **40** (1986), 26–39.
9. W. Brock and C. Sayers, "Is the Business Cycle Characterized by Deterministic Chaos?" Department of Economics, Univ. of Wisonsin–Madison, Madison, Wisc. November 1985.
10. W. Brock, Asset prices in a production economy, "Economics of Information and Uncertainty", (J. J. McCall, Ed.), pp. 1–43, Univ. of Chicago Press, Chicago 1982.
11. W. Brock, and G. Chamberlain, "Spectral Analysis Cannot Tell a Macro-Econometrician, whether his Time Series came from a Stochastic Economy or A Deterministic Economy," SSRI W.P. 8419, Depart. of Econ. Univ. of Wisc., Madison, Wisc. 1984.
12. W. A. Brock, and W. D. Dechert, "Theorems on Distinguishing Deterministic and Random Systems," Depart. of Econ. Univ. of Wisconsin, Madison, Wisconsin and Depart. of Econ. Univ. of Houston, Houston, Texas, 1985.
13. W. Brock, "Distinguishing Random and Deterministic Systems: An Expanded Version," Depart. of Econ., Univ. of Wisconsin–Madison, Wisc. August, 1985.
14. P. Collet, and J. Eckmann, "Iterated Maps on the Interval as Dynamical Systems," Birkhäuser, Basel, 1980.
15. R. A. Dana and P. Malgrange, The dynamics of a discrete version of a growth cycle

model, J. P. Ancot, Ed.) *in* "Analyzing the Structure of Econometric Models," Nijhoff, The Hague, 1984.
16. R. H. DAY, Irregular growth cycles, *Amer. Econ. Rev.* **72**, (1982), 406–414.
17. R. H. DAY AND W. J. SHAFER, "Keynesian Chaos," Depart. of Econ., Univ. of Southern Cal., Los Angeles, 1983.
18. J. DELONG, AND L. SUMMERS, "Are Business Cycles Symmetric?" Harvard Inst. of Econ. Res. Discussion Paper 1076, August, 1984.
19. R. DENECKERE AND S. PELIKAN, Competitive chaos, *J. Econ. Theory* **40** (1984), 13–25.
20. L. EDLEFSEN AND S. JONES, "Gauss," Applied Technical Systems, Kent, Washington, 1984.
21. B. GELBAUM AND J. OLMSTED, "Counterexamples in Analysis," Holden Day, San Francisco, Calif., 1964.
22. J. M. GRANDMONT, On endogenous competitive business cycles, *Econometrica* **53** (1985), 995–1045.
23. P. GRASSBERGER AND I. PROCACCIA, Measuring the strangeness of strange attractors, *Phys. 9D* (1983), 189–208.
24. J. GUCKENHEIMER AND P. HOLMES, "Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields," Springer-Verlag, New York, 1983.
25. A. KATOK, Lyapunov exponents, entropy, and periodic orbits for diffeomorphisms, *Publ. Math. I.H.E.S.* (1980), 137–174.
26. R. MAY, Simple mathematical models with very complicated dynamics, *Nature* **261**, (1976), 459–467.
27. G. MOORE AND V. ZARNOWITZ, "The Development and Role of the National Bureau's Business Cycle Chronologies," NBER Working Paper 1394, Cambridge, Mass. July 1984.
28. C. NICOLIS AND G. NICOLIS, Is there a climatic attractor? *Nature* **311** (1984), 529–532.
29. V. OSELEDEC, A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems, *Trans. Moscow Math. Soc.* **19**, (1968), 197–231.
30. M. PRIESTLY, "Spectral Analysis of Time Series," Vol. I, II, Academic Press, New York, 1981.
31. I. PRIGOGINE, "From Being to Becoming," Freeman, New York, 1980.
32. D. RUELLE AND F. TAKENS, On the nature of turbulence, *Comm. Math. Phys.* **20** (1971), 167–192.
33. D. RUELLE, Ergodic theory of differentiable dynamical systems, *Publ. Math. I.H.E.S.* **50** (1979), 27–58.
34. D. RUELLE, Small random perturbations of dynamical systems and the definition of attractors, *Comm. Math. Phys.* **82** (1983), 137–151.
35. H. SAKAI AND H. TOKUMARU, Autocorrelations of a Certain Chaos, *IEEE Trans. Acoust. Speech Signal Process.* **V.I. ASSP-28**, No. 5, Oct., (1980), 588–590.
36. C. SAYERS, "Work Stoppages: Measuring the Attractor Dimension," Univ. of Wisc. Madison, Depart. of Econ. April, 1985.
37. J. SCHEINKMAN AND B. LEBARON, "Nonlinear Dynamics and Stock Returns," Dept. of Econ. Univ. of Chicago Press, Chicago, 1986.
38. M. STUTZER, Chaotic dynamics and bifurcations in a macro model, *J. Econ. Dynam. Control* **2** (1980), 353–376.
39. H. SWINNEY, Observations of complex dynamics and chaos (Cohen E.G.D. Ed.) *in* "Fundamental Problems in Statistical Mechanics, VI," Elsevier, North-Holland, Amsterdam, 1985.
40. H. SWINNEY, Observations of order and chaos in nonlinear systems, *Phys. D*, No. 5, (1985), 1–3, 3–15.
41. F. TAKENS, Distinguishing deterministic and random systems, (G. Borenblatt, G. Iooss and D. Joseph, Eds.), *in* "Nonlinear Dynamics and Turbulence, 315–333, Pitman, Boston, 1985.

42. F. TAKENS, Detecting strange attractors in turbulance, (D. Rand, L. Young, Eds.), *in* "Dynamical Systems and Turbulence," 366–382, Warwick 1980, Lecture Notes in Mathematics No. 898: Springer-Verlag, Berlin.
43. F. TAKENS, "On The Numerical Determination of the Dimension of an Attractor," unpublished, manuscript, 1984.
44. A. WOLF, J. SWIFT, H. SWINNEY, AND J. VASTANO, Determining Lyapunov exponents from a time series, *Phys. 16D* (1985), 285–317.
45. V. ZARNOWITZ, "Recent Work on Business Cycles in Historical Perspective: Review of Theories and Evidence," Univ. of Chicago and NBER, Chicago, October 1984.
46. V. ZARNOWITZ AND G. MOORE, "Major Changes in Cyclical Behavior," NBER Working Paper 1395, Cambridge, Mass., July 1984.