CIDE

NÚMERO 14

David Mayer

AN ALTERNATIVE CORRELATION DIMENSION STATISTIC

Introduction

In recent years the application of the Grassberger and Procaccia (G&P) correlation dimension (CD) (1983) in economics has been of interest. This is a statistical procedure applied to time series which originated with the purpose of measuring the fractal dimension of strange attractors and was then used to distinguish between stochastic and deterministic (*e.g.* chaotic) processes.

In a later development, Brock *et al.* introduced a related statistic (DBS statistic) testing for IID (identically and independently distributed) series whose main applications include testing for nonlinearity in stochastic processes. (See for example Sheinkman and LeBaron [1987], and Brock and Malliaris [1989].)

One of the problems plaguing these applications has been the large data requirements of these statistical procedures. Time series numbering 20,000 are common in physical science applications, while results of applications to shorter time series with lengths between 500 and 2,000 —it being difficult enough to obtain such economic data— have suffered from a bias towards low dimensions in the statistical results.

The purpose of this paper is to present a modified statistic which not only calculates dimension and simultaneously tests for the IID null, but is also not biased towards low dimensions and more sensitive to the presence of stochastic structure in the shorter time series mentioned above. Moreover, we give a method to estimate the significance of tests conducted by means of these statistics, based on the reshuffling test and not involving approximation to the normal distribution and the calculation of variance.

We conduct several tests to compare the two statistics, namely the correlation dimension in its usual form (CD for short) and in its modified form (SCD for Statistical Correlation Dimension). The tests consist in comparing, for the CD and SCD,

1) The mean of 2,000 realizations of IID series of length 500 calculated for several stochastic processes.

2) The power of the reshuffling test (described below) applied to distinguish the presence of several processes.

3) Results obtained from applying the reshuffling tests to 103 series of residues obtained from Milton Friedman's Monetary and W. Cox's Treasury Security long-term time series.

4) The dimensions of Brock and Sayers' (1988) original time series.

5) The dimensions of the CRSP daily long-term time series.

The paper contains two parts. First, in sections 1-3, we define the statistical correlation dimension (SCD) and develop some of its properties, as well as obtain significance intervals for the random reshuffling technique. Second, in sections 4-8, we compare results obtained in applications of both statistics.

1. The Grassberger and Procaccia Correlation Dimension

Let us first briefly give a one-paragraph intuitive introduction to the G&P CD and its application, for the sake of readers less familiar with this specialized topic.

Consider any time series $\{a_t : t = 1, \ldots, N\}$. The essential idea behind the concept of correlation dimension is to systematically consider not just single elements of the time series but the sets of m-vectors whose m entries are regularily-lagged members of the time series. This set of vectors, called m-histories, have a graph of points in Euclidean space R^m, which for a large enough m reveals the nature of the time series. Thus, if for sufficiently large m the m-th member of an m-history is a deterministic function (or almost everywhere a locally implicit function) of the preceding elements belonging to the m-vector, the set of m-histories will lie in R^m on some "surface," while if the m-th member is a stochastic function of the previous members, the m-histories will only cluster about some "surface," tending to fill a region of R^m. Now, the dimension of these graphs of points may be measured by the exponent of how fast the number of "close" points grows as the distance allowed between them is increased. In the case of chaotic deterministic processes this number may turn out to be a fraction corresponding to points clustering on "strange" (fractal) subsets of the surfaces mentioned above. In general, deterministic processes yield (with enough data) a finite dimension even as m grows indefinitely, while sufficiently regular stochastic processes will yield (again, with enough data) the largest possible embedded dimension, m, since because of their random nature they tend to fill space. Thus stochastic processes (high dimension) may be distinguished from determinisitic processes (low dimension).

Let us add that in the presence of stochastic structure and finite data we shall get in-between results (*e.g.* a dimension growing with m but less than m) which may be useful for detecting such structure. Clearly a functional transformation of the data should not increase the dimension in the deterministic case (given sufficient data) but in the stochastic case, if the transformation whitens the series, the dimension should rise. This is the basis for the Brock residue test, described below.

Now let us write some of these notions precisely.

1.1. Definition. Let $\mathscr{A} = \{ a_t : t = 1, \ldots, N \}$ be a finite series of real numbers. For each $m = 1, 2, \ldots, N$, $\tau \in \mathbb{N}$ ($\tau = 1$ in the applications in this paper unless otherwise stated) let

$$\mathrm{H}(\mathscr{A}, m, \tau) = \{ (a_t, \ldots, a_{t+(m-1)\tau}) : t = 1, \ldots, N - (m-1)\tau \} \subseteq \mathbb{R}^m$$

be the set of τ-lagged m-histories of $\mathscr{A}$. ■

For the sake of completeness, we mention the following theorem. Consider the following type of deterministic process

$$\frac{dx}{dt} = f(x) \qquad x \in \mathbb{R}^d \tag{1}$$

for which we have a set of measurements

$$a_t = g(x_t) \in \mathbb{R}, \quad t \in \mathbb{N} \tag{2}$$

(We could instead treat a difference equation.) Here d is the dimension of state space.

Takens (1980) proves that generically, for $m \geq 2d + 1$, the set of limit points of $H(\mathcal{A}, m, \tau)$ as N tends to infinity (for any fixed τ) forms a diffeomorphic image of the corresponding set of limit points (attractor) of $\{x_t\}$ which therefore preserves such properties as the dimension mentioned above. Here "generically" means that the measurement function g is non-singular, *i.e.*, that it is sensitive to the state space variables.

Thus any time sequence A of measurements of a process describable by (1), even when the state variables x and/or the evolution function f may be unknown, can, in general, be used to detect its stochastic or deterministic nature by the methods being described.

1.2. Definition. For any finite set $H \subseteq \mathbb{R}^m$ with cardinality $\#H$ let $C_H(\varepsilon) : (0, \infty) \rightarrow [0, 1]$ given by

$$C_H(\varepsilon) = \frac{\#\{\,(x, y) \in H \times H : x \neq y, \|x - y\| < \varepsilon\,\}}{\#H(\#H - 1)}$$

be the relative frequency with which the distance between pairs of vectors in H is less than ε. Here $\|\ \ \|$ is some norm of $\mathbb{R}^m$. ■

Using this function G&P (1983) define the CD.

1.3. Definition. The correlation dimension of a sequence $\mathcal{A} = \{\, a_t : t = 1, \ldots, N\,\}$ is

$$D = \lim_{m \to \infty} d(m)$$

where

$$D(m) = \lim_{\varepsilon \to 0} \frac{\log C_{H(\mathcal{A}, m, \tau)}(\varepsilon)}{\log \varepsilon}. \ ■$$

An application of Takens' theorem (1980) shows that for sequences originated deterministically as in (1), D is the attractor's "limit capacity", a close lower bound of the Hausdorff dimension and is independent of τ, as N tends to infinity. On the other hand, for sufficiently regular stochastic sequences, it may be shown that under the same conditions $D(m)$ tends to m (see for example Brock (1986)). D is also independent of the norm used in R^m.

When $\mathcal{A}$ is generated by (1) and (2), it may be shown that $D \leq d$, the dimension of state space. When $\mathcal{A}$ is the realization of a stochastic process, it may be shown that $p\lim_{N \to \infty} D(m) = m$.

2. *The Statistical Correlation Dimension*

We instead define:

2.1. Definition. The statistical correlation dimension (SCD) of a sequence $\mathcal{A} = \{ a_t : t = 1, \ldots, N \}$ is

$$D' = \lim_{m \to \infty} D'(m)$$

where

$$D'(m) = \lim_{\varepsilon \to 0} \frac{\log C_{H(\mathcal{A}, m, \tau)}(\varepsilon)}{\log C_{H(\mathcal{A}, 1, \tau)}(\varepsilon)}. \blacksquare$$

Thus $D(m) = D'(m)D(1)$. For stochastic processes $D(1) = 1$. For deterministic processes $D(1) = 1$ if $D \geq 1$ while $D'(m) = 1$ if $D \leq 1$. Thus $D'(m)$ distinguishes stochastic from deterministic processes just as $D(m)$, but will yield $D'(m) = 1$ for $D \leq 1$. The theorems which establish the general properties of $D(m)$ also establish those of $D'(m)$.

For IID stochastic processes and the maximum norm

$$E(C_{H(\mathcal{A}, m, \tau)}(\varepsilon)) = E(C_{H(\mathcal{A}, 1, \tau)}(\varepsilon))^m$$

This relationship between the terms in the quotient yields theorems on the distribution of $D'(m)$ (over realizations $\mathcal{A}$ of a stochastic process).

The relationship above shows that, for IID processes, $C_{H(\mathcal{A}, m, \tau)}(\varepsilon)$ behaves as a power of $C_{H(\mathcal{A}, 1, \tau)}(\varepsilon)$ rather than as a power of ε. On the other hand, for finite samples of more general processes, the growth of both functions must decay as ε gets larger and for these reasons, especially for smaller sample sizes, the growth of $C_{H(\mathcal{A}, m, \tau)}(\varepsilon)$ is better compared with $C_{H(\mathcal{A}, 1, \tau)}(\varepsilon)$ than with ε.

Let us mention before proceeding that in practice, the G&P CD of a time series, D, is determined by means of the following procedure. For each m, $D(m)$ is calculated as the coefficient in a linear regression of $\log C_{H(\mathcal{A}, m, \tau)}(\varepsilon)$ with respect to $\log\varepsilon$, in a region of small values of ε where the relation between both variables is approximately linear. Similarily, we calculate $D'(m)$ as the coefficient in a linear regression of $\log C_{H(\mathcal{A}, m, \tau)}(\varepsilon)$ with respect to $\log C_{H(\mathcal{A}, 1, \tau)}(\varepsilon)$.

We now prove some asymptotic properties of $D'(m)$. These follow the application of the technique of U-statistics (see Serfling, 1980), to $C_{H(\mathcal{A}, m, \tau)}(\varepsilon)$ and related statistics by Brock *et al.*

To give a synthetic exposition we begin by quoting two theorems, the first of which summarizes results by Brock and Baek (1991).

2.2. Theorem. Let $\{Z_i\}$ be a strictly stationary process which is absolutely regular. (A definition is ommitted for brevity. There are alternative conditions on the rate of decay of dependence overtime yielding the same result. See Denker and Keller, 1983, p. 507.) Define for each m, n, ε the random variables:

$$C(m, n, \varepsilon) = C_{H(\{Z_1,\dots,Z_{n+m}\}, m, 1)}(\varepsilon)$$

using the maximum norm. Generically (if the asymptotic variance does not tend to zero) there exist numbers $\mu(m, \varepsilon)$, $\sigma(m, \varepsilon)$ for which, as $n \to \infty$,

$$C(m, n, \varepsilon) \quad \text{is} \quad AN(\mu(m, \varepsilon), \frac{1}{n}\sigma(m, \varepsilon))$$

(asymptotically normal). If $\{Z_i\}$ are IID then

$$\mu(m, \varepsilon) = \mu(1, \varepsilon)^m. \ \blacksquare$$

2.3. Theorem. (Serfling, 1980, page 124.) Suppose that the sequence of vector-valued random variables $X_n = (X_n^1,\dots,X_n^k)$ is $AN(\mu, n^{-1}\Sigma)$ (asymptotically normally distributed) with $\Sigma = (\sigma_{ij})$ a covariance matrix. Let $g : \mathbb{R}^k \to \mathbb{R}$ be a real-valued function having a nonzero differential at $x = \mu \in \mathbb{R}^k$. Then

$$g(X_n) \text{ is } AN\left(g(\mu), \frac{1}{n}\sum_{i,j=1}^{k}\sigma_{ij}\left.\frac{\delta g}{\delta x_i}\right|_{x=\mu}\left.\frac{\delta g}{\delta x_j}\right|_{x=\mu}\right). \ \blacksquare$$

The following properties of $D'(m)$ are direct applications of these theorems.

2.4. Theorem. Let $\{Z_i\}$ be a strictly stationary process which is absolutely regular.[1] Define for each n, m, ε for which $C(1, n, \varepsilon) < 1$ the random variables:

$$D'(m, n, \varepsilon) = \frac{\log C(m, n, \varepsilon)}{\log C(1, n, \varepsilon)}$$

using the maximum norm. Generically, $D'(m, n, \varepsilon)$ is asymptotically normally distributed as $n \to \infty$. If $\{Z_i\}$ are IID then

$$D'(m, \varepsilon) = \lim_{n \to \infty} E(D'(m, n, \varepsilon)) = m$$

The asymptotic mean is m. The asymptotic variance $\sigma^2(m, \varepsilon)$ depends on the asymptotic means and covariances of $C(m, n, \varepsilon)$ and $C(1, n, \varepsilon)$. The analogous statement holds if we replace $D'(m, n, \varepsilon)$ with the coefficient for the slope in a linear regression of $\log C(m, n, \varepsilon)$ with respect to $\log C(1, n, \varepsilon)$ (as functions of ε). Then the asymptotic variance $\sigma^2(m)$ is a function of the asymptotic means and covariances of $\{ C(m, n, \varepsilon_j), C(1, n, \varepsilon_j) \}$, where $\{\varepsilon_j\}$ is the set over which the regression is defined. ■

For sequences originated stochastically theorem 2.4 provides, as does theorem 2.2 (DBS statistic), a hypothesis test for the IID null.

Actually calculating the asymptotic standard deviation of the (asymptotically normally distributed) CD, as obtained from the linear regression, would involve an exceedingly large number of computations (of the order of Nn_ε CD computations). An alternative method for obtaining significance intervals is presented below.

Some further considerations follow from theorem 2.4. In practice it turns out that the comparison of $\log C_{H(\mathcal{A}, m, \tau)}(\varepsilon)$ or $\log C_{H(\mathcal{A}, 1, \tau)}(\varepsilon)$ with $\log(\varepsilon)$ requires very large samples. Ramsey, Sayers & Rothman (1990) show empirically, in their criticism of applications of the CD, that in a regression of $\log C_{H(\mathcal{A}, m, \tau)}(\varepsilon)$ against $\log(\varepsilon)$ bias is introduced by m and N (state space dimension and sample size).

By contrast the SCD avoids this problem from its definition, cancelling out the comparison with $\log(\varepsilon)$. Moreover, since $D'(m, \varepsilon) = m$ independently of ε, we have one less limit process requiring large amounts of data. The SCD, besides testing the IID null, is a numerical method for the calculation of the CD for dimensions greater than 1.

[1] There are several definitions which yield the result, which we omit for brevity.

3. The Residue and Reshuffling Tests. Significance Intervals

A few points remain to be made regarding the application of the CD.

As was already mentioned, in the case of stochastic time series, if these are whitened by a functional transformation before applying the CD, a smaller data set will be needed to obtain the correct high dimensional results. On the other hand, it can be shown that such a functional transformation will not raise the dimension in the deterministic case —if in practice there is enough data and the transformation does not involve many variables. This procedure, known as the residue test, was proposed by Brock (1986).

An additional procedure is to randomly reshuffle the series (with replacement). If there is structure, whether stochastic or deterministic, it will be eliminated, so the value of the CD should rise. Thus, given the CD of a time series, one may apply the following significance test for low dimension. Obtain the CD of k independent random reshufflings of the original series and see whether if the original is lower than all of these.

We can find significance intervals for this test which are valid independently of the distribution and which do not require a variance calculation. (Indeed this holds for the same test applied to any statistical procedure with a continuous distribution function.) Suppose that for a given data set the CD of random reshufflings has a distribution well-approximated by an integral

$$P(\mathrm{CD} \in I) = \int_I F(x)dx$$

(as we may for data sets for which the N^N resulting reshuffled CD's are approximately continuously distributed). The probability that a given CD_0 (corresponding to the original series) is smaller than k independent random reshufflings CD_i, $1 \leq i \leq k$ is then well-approximated by

$$P(\mathrm{CD}_0 < \min_{1 \leq i \leq k}(\mathrm{CD}_i)) = \int_{\mathbb{R}} \left(\int_{(x,\infty)} F(y)dy \right)^k F(x)dx$$

$$= \left[-\frac{1}{k+1} \left(\int_{(x,\infty)} F(y)dy \right)^{k+1} \right]_{\mathbb{R}} = \frac{1}{k+1}$$

These confidence intervals are independent of the density function F. In particular, this implies that when applying the random reshuffling test it is unnecessary for F to be close to the normal distribution, or to estimate its variance. Also, we may calculate the CD by means of a regression on a considerable discrete set of ε (which would make the calculation of variance prohibitive) obtaining a much more robust statistic.

4. Comparison of the SCD with the CD

To obtain an informal qualitative comparison of the two statistics first observe figures I, which report $D(m)$ and $D'(m)D(1)$ for the following series of 823 elements.

Stochastic series:
1) A uniformly distributed random variable.
2) Autoregressive series $y_t = 0.9y_{t-1} + w_t$ with w_t uniform on an 1interval.

Chaotic deterministic series:

1) The logistic: $x_{t+1} = 3.57x_t(1 - x_t)$.
2) The Henon series: $x_{t+1} = 1 + y_t - 1.4x_t^2, y_{t+1} = 0.3y_2$.
3) The Lorentz series:
$x_1' = 10(x_3 - x_2), x_2' = 28.0x_2 - x_2 - x_1x_3, x_3' = -2.66x_3 + x_1x_2$
4) The Rossler series: $x_1' = -x_2 - x_3, x_2' = x_1 + 0.2x_2, x_3' = 0.2 + (x_1 - 5.7)x_3$.

As can be observed in the figures, in the case of the chaotic series both statistics gave similar results. However, the results were different in the stochastic case. The proposed method of calculation reduced the bias towards low dimension present in the method of G&P.

We now report comparison of the SCD and CD in the five different ways mentioned in section 1. Let it be stated that both statistics were calculated together by the same computer program, changing only the tables against which the linear regressions of $\log C_{H(\mathcal{A}, m, \tau)}(\varepsilon)$ were calculated from $\log C_{H(\mathcal{A}, 1, \tau)}(\varepsilon)$ to $\log\varepsilon$. In the stochastic case $D(1)$ is very close to 1 in practice, so the graphs show $D(m)$ and $D'(m)$.

The Monte Carlo tests used the following chaotic and stochastic processes. Realizations of the chaotic processes differed in that a uniformly random number of iterations was first discarded, producing realizations with random initial points distributed according to the processes' inherent ergodic distributions.

Chaotic processes (run in double precision and verified to have a chaotic attactor):

1) The logistic $x_{t+1} = 3.57x_t(1 - x_t)$ with $x_0 = 0.255$

2) The Henon series: $x_{t+1} = 1 + y_t - 1.4x_t^2, y_{t+1} = 0.3y_2$ with $x_0 = 0.5, y_0 = 0.5$, (similarily).

Stochastic processes (let u_t be a normal random variable obtained as the normalized average of 12 uniform random variables):

3) The autoregressive process $y_{t+1} = .6y_t + u_t$.
4) The absolute value ARCH process $y_t = |u_{t-1}|u_t$.
5) The same process with autoregression: $y_{t+1} = .6y_t + |u_{t-1}|u_t$.

6) A normal random variable with variance depending on states following a Markov Process:

$$y_t = a_t u_t,\ P(a_{t+1} = i | a_t = j) = A_{ij},\ 1 \leq i, j \leq 2,\ A = \begin{pmatrix} .2 & .8 \\ .6 & .4 \end{pmatrix}.$$

7) The same process with autoregression: $y_{t+1} = .6y_t + a_t u_t$.

The first test obtained the mean of the SCD and CD of 2,000 reshuffles of realizations of length 500 of these stochastic processes. Thus, the dimensions of IID processes with different distributions are obtained. Figures II, in different scales, show that the SCD is slightly upward rather than very downward biased, for IID processes. Also the SCD results for different IID distributions are relatively much less variable than the CD results, which are more dependent on the distribution. In practice, the same result holds for non-IID processes.

5. Power of the Reshuffling Test for the SCD and the CD

How powerful is the random reshuffling test for which significance intervals were obtained in section 3? That is, what is the probability that it will reject the presence of structure when it is indeed present? To test this we applied the random reshuffling test (with $k = 20$) to the seven processes described in the previous section.

We carried out this test for the SCD and CD defined for different values of τ (see definition 1.1), that is, for differently lagged m-histories. The values for τ were 1, 2, 3, 4, 12.

The results are presented in figures III. In the case of the logistic, both statistics obtained a 100% score for each value of t up to 16 dimensions. For the Henon series, the results were very similar (one must recall that with finite precision arithmetic, a chaotic process will present a stochastic component after enough iterations). But in the stochastic case, the SCD performed consistently better than the CD, obtaining power ratings of 95% or above over a much wider range of dimensions and values of τ. The failure of both tests in detecting variance depending on states following a Markov Process points to the large data requirements necessary to detect such fuzzy structure.

6. Summary Statistics of 103 Reshuffling Tests on Residues of M. Friedman's Monetary and W. Cox's Treasury Security Long-term Time Series

Milton Friedman's monetary and W. Cox's Treasury Security long-term time series were examined for the presence of chaos and for nonlinearity. The residue tests consisted in extracting correlation in the mean and variance, as well as in some cases integrated correlation in variance. Evidence was found for these non-linear stochastic structures. The estimation procedures are presented in Appendix A. Altogether, the number of residue series subjected to the random reshuffling test was 103, enough to compare the SCD and CD in the context of practical application.

We report summary statistics for the reshuffling test with $k = 20$ as applied to the macroeconomic time series residues mentioned above (the CD and the SCD of 103 residue series and of 20 reshufflings of each).

Let us describe the summary statistics. For each test and each method we have the following pieces of data: the mean, lowest and highest values of the dimensions obtained in the 20 random reshufflings, and the dimension of the original macroeconomic residue (or normalized) series.

Figures IV.1 and IV.2 contain the means over the 103 tests of each of these pieces of data, for each method respectively. Thus we graph the mean of the unshuffled, lowest, highest and mean dimensions. (The mean mean dimension is a mean over 20 random reshufflings of 103 sets of data.) The downward bias of the G&P method is apparent, with mean dimension not rising above 8 and mean highest dimension not rising above 12, in 64 dimensional phase space. On the other hand, the SCD mean is very close to the theoretically expected mean. It presents a small upward bias due mainly to the asymmetric distribution of $D'(m)$ about m, which is streched upwards (the mode is probably at m). Furthermore , the unshuffled dimensions are much closer to the lowest dimensions in the case of $D(m)$ than in the case of $D'(m)$.

Figure V shows that this last fact, which is a measure of the sensitivity of the statistic to structure, is true not only absolutely but also relatively. Here we normalize each of the pieces of data by dividing by the corresponding mean dimension and comparing the normalized mean lowest dimensions with the normalized mean unshuffled dimensions (which contain non-random and therefore very improbable structure). We find that the unshuffled series can be much more readily distinguished from the shuffled series in the case of the SCD, which is therefore that much more sensitive to structure in the series.

We summarize by observing that not only are meaningful results obtained when removing structure progressively, but these are obtained for a wide range of phase space dimensions. Figure V would suggest that in the reshuffling tests carried out for macroeconomic time series in this paper, the best detection of structure is obtained about $m = 24$. Section 5 above also shows that the power of the reshuffling test may maximize at quite high dimensions. Thus, the SCD seems to preclude the need for data sets of the order of $N = 10^m$ suggested in Ramsey, Sayers & Rothman (1990) for the CD.

7. The SCD and CD of Brock and Sayer's Time Series

To obtain results directly comparable to others in the literature, we applied the SCD and CD to the time series that Brock and Sayers used in their 1988 article (figure VI).

Here we obtained dimensions for the first differences, for autoregressive residues of order 1, and for absolute-value ARCH residues of order 1. These were identical in form to those applied to the monetary series, except that because some of the time series are short (between one and two hundred) we only used one lag (see Appendix A, part 1 for the estimation procedures). A more careful analysis would have to consider a wider array of structures, although in fact most of these series are too short for a good application of dimension estimates.

We reproduce the average of some of the results obtained for the economic series in the following tables.

Table I.1
Average Correlation Dimension Obtained for Residues of Brock and Sayers 1988 Economic Time-Series

	Correlation Dimension					
m	*1*	*2*	*5*	*10*	*20*	*30*
First Difs	0.84	1.63	2.83	3.52	4.42	4.80
AR1	0.99	1.73	2.96	3.65	4.35	4.99
ARCH1	0.99	1.80	3.14	4.05	5.20	6.17

Table I.2
Average Statistical Correlation Dimension Obtained for Residues of Brock and Sayers 1988 Economic Time-Series

	Statistical Correlation Dimension					
m	*1*	*2*	*5*	*10*	*20*	*30*
First Difs	—	1.85	3.78	5.54	9.04	12.0
AR1	—	1.87	3.88	5.95	9.74	13.5
ARCH1	—	1.93	4.21	6.94	11.9	17.6

The main conclusion is that the apparent convergence to a low dimension of 5 or 6, as originally obtained, is spuriously introduced by the CD. The dimensions of the first differences as obtained by the SCD already show consistent growth with phase dimension, pointing to structure of stochastic nature.

Moreover, most of the series' ARCH residues show a higher dimension (including Pig Iron, which was not found to rise by Brock and Sayers), while the AR residues do not rise very much. The case is thus strengthened for their nonlinear, stochastic nature. (If the structure were linear or if it were spuriously induced by the regressions, the AR residue dimensions would rise comparably). Sharp gradient changes in any of the graphs are best interpreted as pointing to the upper limit of phase space dimensions yielding meaningful results.

It is worth commenting that the deterministic sunspot time series used as control by Brock and Sayers, which yields consistently rising dimensions, may not in fact correspond to a low dimensional process, since it is governed by a nonlinear *partial* differential system, to which Takens' theorem does not apply, and which may have infinite or at least very high dimensional states. Moreover, ARCH processes have been linked to diffusion (Nelson, 1990) so that the extraction of ARCH is not meaningless in this context.

8. The SCD and CD of the CRSP Daily Long-term Series

The next point of comparison is a single reshuffling test applied to residues from a CRSP series with length 6409. Three pairs of graphs show the dimensions of the first differences, autoregressive residues, and EGARCH residues as provided by Craig Hiemstra (see Craig Hiemstra, 1992) (figures VII).

It is interesting to note that the CD provides a downward biased and much less stable statistic even for this length of series, yielding a spurious dimension of 5, while the SCD yields smoothly rising dimensions, which are unbiased in the IID case. The graphs also confirm that the SCD provides a more sensitive test for stochastic structure, since the dimensions of the original unshuffled series are much further from the mean than the shuffled series.

The presence of EGARCH is strongly confirmed by both statistics, though the SCD detects additional structure for m-histories with $m > 24$ with a confidence of 5%, which is only vaguely indicated by the CD.

9. Conclusions

Regarding the SCD, the conclusions of this article are the following:

The SCD is a correlation dimension statistic improving the numerical computation of the CD in the stochastic case and its performance for distinguishing deterministic from stochastic and linear from nonlinear processes. In most of the cases, we see meaningful results up to high dimensional phase space, for series with lengths of five hundred or more.

The SCD is a more powerful statistic for the detection of stochastic structure than the CD. Applied to these relatively small-sized data sets, it is unbiased in the stochastic IID case, provides a test of the IID null, reports low dimensions in the chaotic cases, and is very responsive to succesive removal of structure by the residue test method (see appendix A).

These results are confirmed by each of the successive tests applied. Combined with the significance intervals obtained for the reshuffling test, the SCD provides an improved statistic for detecting stochastic structure.

The downward bias of the CD and its low sensitivity to structure when applied to relatively short time series, account for many of the ambiguous results in the Economic literature applying the CD, including spuriously low-dimensional results.

Regarding the macroeconomic time series examined, the conclusions are the following:

Evidence for deterministic chaos was not found in the long-term monetary and Treasury Securities time series examined.

Strong evidence was found for nonlinear stochastic behaviour, especially in the form of structure in the variance. Evidence was found for conditionally autorregresive linear heteroskedasticity in the standard deviation but also evidence for other forms of structure.

In the case of the price index of the Treasury Bills it was found that integrated autoregressive conditional heteroskedasticity may be present in the price variations.

Appendix A. Dimension Tests Applied to M. Friedman's Monetary and W. Cox's Treasury Security Long-term Time Series

A.1. Application to Milton Friedman's Monetary Time Series

Monthly variations between the dates of May 1907 and December 1968 for the following monetary quantities as published by M. Friedman were used: Currency Held by the Public ($M1$), Total Deposits at Commercial Banks, Deposits at Mutual Savings Banks, Total (sum of previous three) ($M3$). (The length of the original series is 740.)

In each case the autocorrelation statistic Q detects the presence of autocorrelation, for the series itself and more strongly for its absolute value, therefore detecting conditional heteroskedasticity.

To apply the residue test the series where transformed estimating several models. These models extracted structure resulting from linear autocorrelation in mean, autocorrelation in variance, and change of regime.

Models extracting linear autocorrelation in the mean were estimated by ordinary least squares (OLS):

$$R_t = \beta_0 + \sum_{i=1}^{L} \beta_i R_{t-i} + \zeta_t$$

where L was set to 1 or 12, and ζ_t are the least square residues.

The residues of these regressions are not autocorrelated, as detected by the Q statistic. However, conditional heteroskedasticity was detected by means of the following OLS regression:

$$|\zeta_t| = \Theta_0 + \sum_{i=1}^{L} \Theta_i |\zeta_{t-i}| + \eta_t$$

$|\zeta_t|$ is a consistent estimator of the standard deviation of R_t (Schwert 1989; Pagan and Schwert 1990). In turn, this condition is sufficient to apply the residue test (Nuissance Parameter Theorem, W. A. Brock and W. D. Dechert [1989]).

Models extracting conditional autoregressive heteroskedasticity were generated by combining the previous estimate of the standard deviation with a new autoregression of the mean estimated by weighted least squares. The dimension of the residue of the standard deviation was also calculated.

Models extracting structure in the form of a change in regime replaced the constant in each regression with two dummies adding to one, representing the periods up to and after World War II respectively.

The resulting SCD may be seen in figures VIII.1, VIII.2, VIII.3, corresponding respectively to the residues of models extracting linear structure, ARCH structure and finally to the residues of the variance estimates. (Abbreviations used in the graphs are explained in Table II.) Each figure includes cases in which lags are chosen at 0, 1 or 12, and the change of regime dummies are included or not. (These are graphs of $D'(m)$; $D(1)$ was consistently very close to 1 for all macroeconomic residues considered in this article.)

Let us first examine the SCD's of the first differences of M1, Deposits at Commercial Banks and M3. In figure VIII.1, these series present consistently low SCD's, reflecting the presence of structure. The SCD for the original first differences yielded results almost identical to the SCD for the linear residues. This means both that the linear autoregression introduced no spurious structure in the data, and that not much of the structure in the data was removed by it. Figure VIII.2 is quite different. We see the SCD consistently rising according to the progressive removal of structure, almost to the expected value for IID series. However, figure VIII.3, which examines the estimated variance residues, shows the presence of structure which has not been taken into account by ARCH. Indeed, the method of extracting the structure attributed to the change of regime (which is not very sophisticated) seems to introduce some spurious structure in the variance, since it lowers the SCD, although it does remove structure from the original series. Thus, we would expect additional or different structure in the variance than that here considered.

Let us now examine the SCD's of Deposits at Mutual Savings Banks. In figure VIII.1 a somewhat-structured stochastic nature is reflected; the change of regime has not much effect upon the conditional means and a certain amount of structure is removed by the linear autoregression. In figure VIII.2, change of regime does not make much difference in any of the cases considered, while a certain amount of ARCH seems to be present in the series. In figure VIII.3 it is apparent that some structure other than those here considered is present in the variance, which however does show some autoregressive structure.

A.2. Application to W. Cox's Treasury Securities Time-Series

Monthly variations between the dates of Jan 1942 and Dec 1984 for the following quantities were used: Market Value of Federal Debt; Price Index of Treasury Debt (any annuities to maturity); Price Index of Treasury Bills (any annuities to maturity); Price Indexes of Treasury Debt (0-1, 1-5, 5-10, 10+ annuities to maturity), as published by W. M. Cox. (The length of the original series is 516.)

The first differences of these series showed low autocorrelation, but their absolute values were highly autocorrelated, therefore detecting conditional heteroskedasticity.

To apply the residue test the series where transformed estimating the variance in several ways. These models extracted structure in the variance resulting from autocorrelation. The original series were then normalized according to the estimated variances. Except for two cases mentioned below, the model estimated was

$$x_t \in N(0, v_t)$$

$$v_t = \delta_0 + \sum_{i=1}^{L} \delta_i v_t + \eta_t$$

where x_t was the first difference of the series being considered and L was set to 2 or 6. Then the dimension was calculated for x_t/v_t and η_t. Autoregressive structure in the means was only detected and extracted in the case of the Market Value of Privately Held Marketable Federal Debt, as for the monetary series above, but with 6 lags.

In the case of the Price Index of the Treasury Bills, it was necessary to consider an autoregressive structure in the first differences of variance before the stochastic nature of the series yielded high CD. The model considered was

$$x_t \in N(0, v_t), \quad v_{t+1} = v_t + \zeta_t$$

$$\zeta_t = \delta_0 + \sum_{i=1}^{\sigma} \delta_i \zeta_t + \eta_t$$

Again X_t was the first difference of the original series, and the dimensin was calculated for x_t/v_t and η_t.

In the case of the Market Value of Privately Held Marketable Federal Debt the series was found to be autocorrelated in the mean, so ARCH was extracted as for the monetary series, but with 6 lags.

The resulting SCD's may be seen in figure IX, which shows in each graph results for a single series in turn. (Again see Table II for abreviations in the graphs.) These are the SCD's for: the first difference of the time series; the first difference normalized according to autocorrelated variances at lags 2 and 6; the residues of these variance estimates. In the second graph for the Price Index of the Treasury Bills the sequence is similar, but the quantities estimated are the first differences of the variance.

The Graph for the Market Value of Privately Held Marketable Federal Debt shows instead dimensions for the residues of the weighted linear regression; for the weighted residues, and for the variance residues, as well as for the first differences normalized without taking into account the autoregressive component of the mean. This last procedure is the only one to introduce spurious structure, as can be seen from the decreased dimension. The other two procedures raise the dimension, with the residues of variance showing similar structure to the residues of the first differences.

In every other case, it may be seen that the successive removal of structure raises the estimated SCD, in some cases from what appeared to be a consistently low dimensional result. In the case of the Price Index of the Treasury Bills this process yielded high SCD only when the conditional autoregressive heteroskedasticity is considered to be an integrated process (conditional autocorrelation in the first difference of variance), though then the residues of the first differences of variance continued to present much structure. Comparison of the SCD of the Price Index of the Treasury Debt at different maturities shows that there is some structure in the shorter term which fades as the length of maturity is increased, different from the structures here considered.

It is interesting that in each case the series and variance residues gave similar SCD's, supporting the hypothesis that most of the structure of these series is in the variance.

Appendix B. Numerical Implementation and Choice of ε

To minimize the computation requirements in the calculation of the CD an algorithm was developed which: *a)* calculates the distribution of distances recursively in the phase space dimension; *b)* mainly uses short integer arithmetic; *c)* has low memory requirements.

This is based on the following. If

$$b_{ij}^{m} = \max_{0 \le r \le m-1} \left| a_{(i+r} - a_{j+r} \right|$$

is the maximum norm distance between the m-histories ($a_i,\ldots,a_{i+m-1}$ and $a_j,\ldots,a_{j+m-1}$)., (We apply here only the case $\tau = 1$) then

$$b_{ij}^{m+1} = \max\,(b_{ij}^{m}, b_{(i+1)(j+1)}^{m})$$

so that, for each difference $j-i$, b_{ij}^{m+1} is obtained recursively from b_{ij}^{m}, requiring in memory only a diagonal vector of norms of the differences.

A set of values of ε is represented as a sequence of integers. Once b_{ij}^{l} has been calculated, the relevant information of its magnitude is therefore an integer, representable in one byte if the number of ε is less than 256.

Since the linear intervals on which the regressions are calculated shift considerably when m is incremented, these were defined, after much experimentation, in terms of the values of $C_{H(st,m,1)}(\varepsilon)$ as a follows:

$$I_m = \{ \varepsilon : 0.02 \leq C_{H(st,m,1)}(\varepsilon) \leq 0.2 \}.$$

These intervals of ε correspond typically (to a factor of 2) to [0.002,0.005] for $m=1$ and [0.02,0.045] for $m=30$. Observe that the lower bound represents a systematic elimination of noise. The optimal bounds must vary with sample size.

The program for this algorithm, implemented in Turbo Pascal for 286 or 486 Intel processors (a mathematical coprocessor proves very useful), is available at the División de Economía, CIDE.

Bibliography

Baek, E. G. and W. A. Brock (1992), "A General Test for Nonlinear Granger Causality: Bivariate Model", preprint.

Barnett, William A. and Ping Chen (1988), "The Aggregation Theoretic Monetary Aggregates are Chaotic and Have Strange Attractors: An Econometric Application of Mathematical Chaos", William Barnett, Ernst R. Berndt and Halbert White (eds.), *Dynamic Economic Modelling*, pp. 199-245.

Barnett, William A. and Seungmook S. Choi (1989), "A Comparison Between the Conventional Econometric Approach to Structural Conference and the Non Each Metric Chaotic Attractor Approach", in William Barnett, John Geweke and Karl Shell (eds.), *Economic Complexity*, pp. 141-212.

Boldrin, Michele and M. Woodford (1990), "Equilibrium Models Displaying Endoge- nous Fluctuations and Chaos. A Survey", *Journal of Monetary Economics*, vol. 25, pp. 182-222.

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

Brock, W. A. and W. D. Dechert (1987), "Theorems on Distinguishing Deterministic from Random Systems", *Social Systems Research 8702*, University of Wisconsin-Madison.

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

Brock, W. A. and A. G. Malliaris (1989), Differential Equations, Stability and Chaos in Dynamic Economics, *Advanced Textbooks in Economics*, North-Holland.

Brock, W. A. and W. D. Dechert (1989), "Statistical Inference Theory for Measures of Complexity in Chaos Theory and Nonlinear Science", in N. B. Abraham *et al.* (eds.), *Measures of Measures of Complexity and Chaos*, Plenum Press, New York.

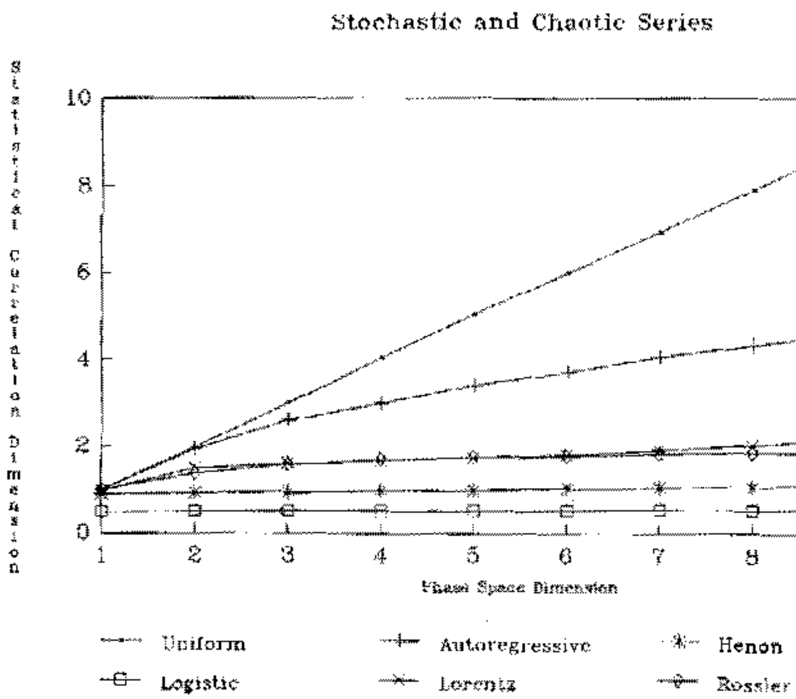


Figure I.1

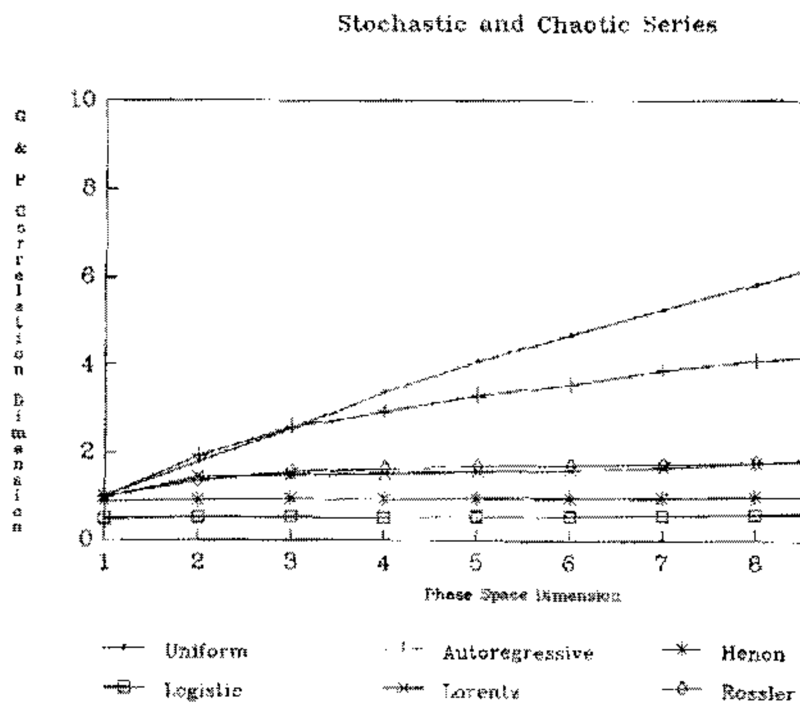


Figure I.2

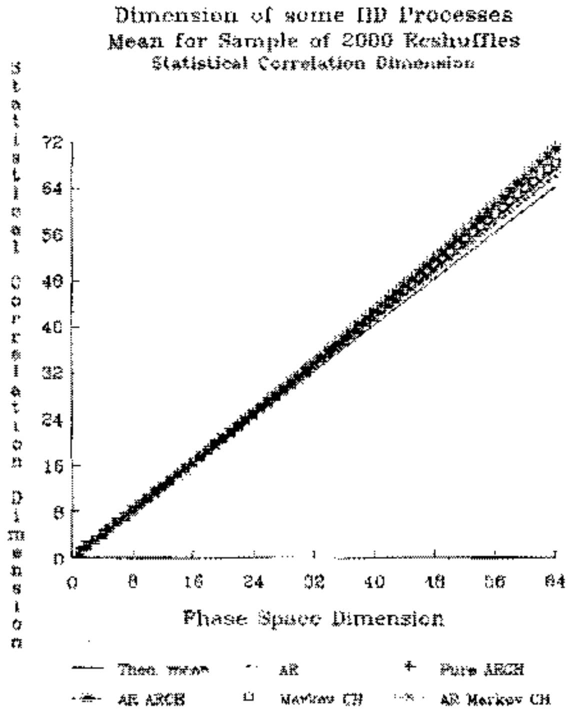


Figure II.1

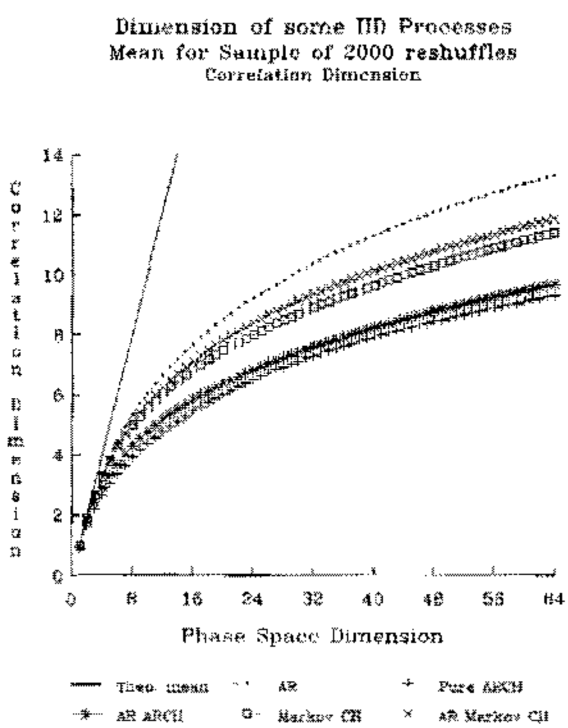
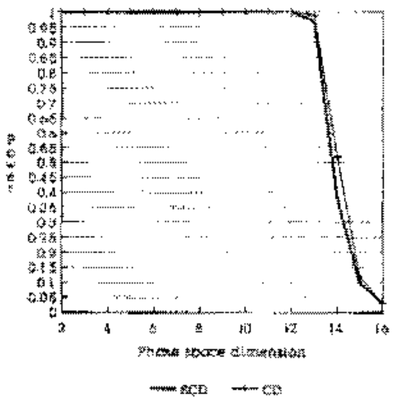


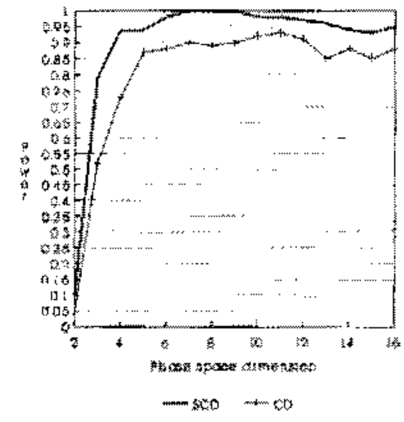
Figure II.2

Power Test 2
Henon Series with
Random Initial Values



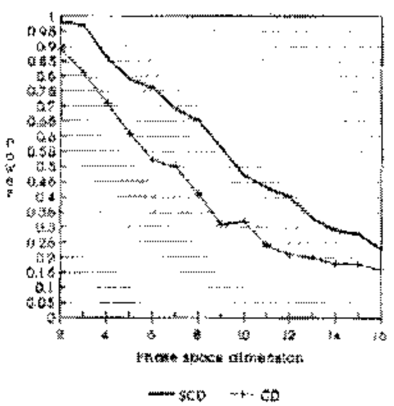
m-histories with lags of 1
Figure III.1

Power Test 3
Linear Autoregressive Series



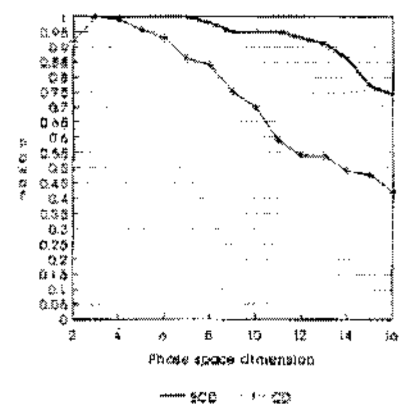
m-histories with lags of 1
Figure III.1

Power Test 4
Absolute Value ARCH
with Non-regressive Mean



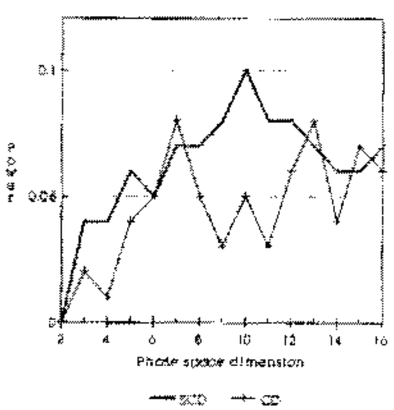
m-histories with lags of 1
Figure III.1

Power Test 5
Autoregressive Series with
Absolute Value ARCH



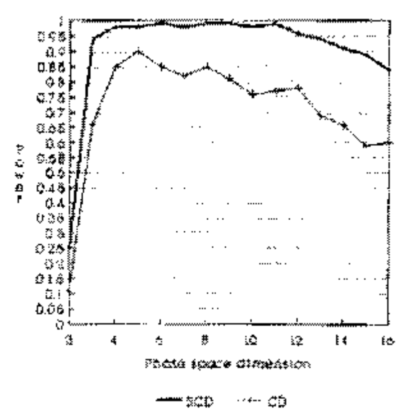
m-histories with lags of 1
Figure III.1

Power Test 6
Normal Random Variable with
Variance following Markov Process



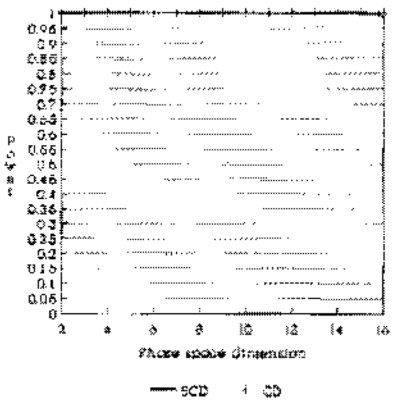
m-histories with lags of 1
Figure III.1

Power Test 7
Autoregressive Series with
Variance following Markov Process



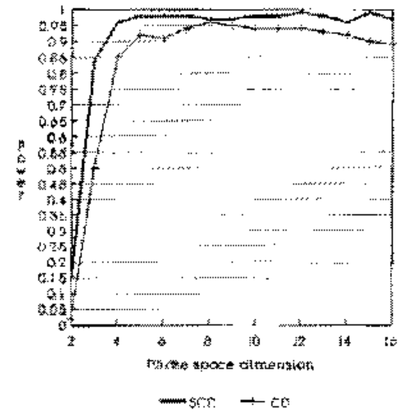
m-histories with lags of 1
Figure III.1

Power Test 2
Henon Series with
Random Initial Values



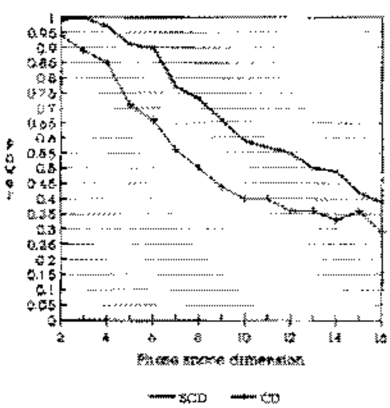
m-histories with lags of 2
Figure III.2

Power Test 3
Linear Autoregressive Series



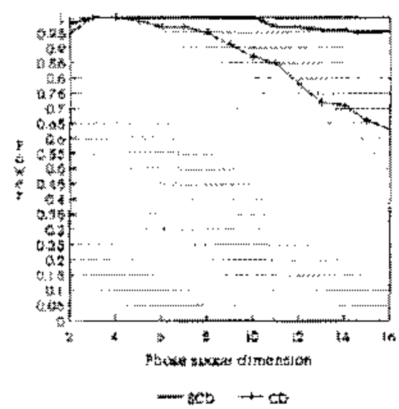
m-histories with lags of 2
Figure III.2

Power Test 4
Absolute Value ARCH
with Non-regressive Mean



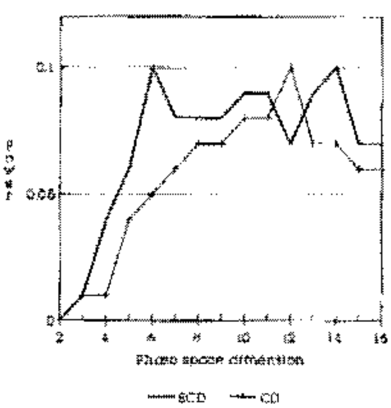
m-histories with lags of 2
Figure III.2

Power Test 5
Autoregressive Series with
Absolute Value ARCH



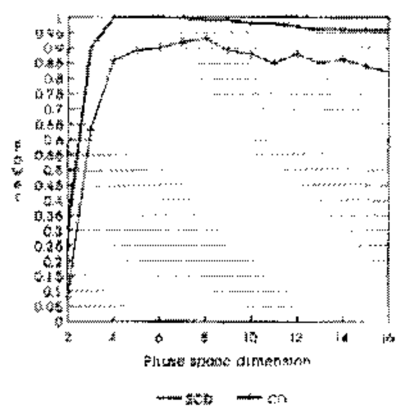
m-histories with lags of 2
Figure III.2

Power Test 6
Normal Random Variable with
Variance following Markov Process



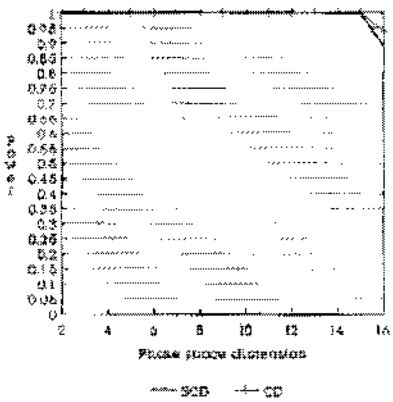
m-histories with lags of 2
Figure III.2

Power Test 7
Autoregressive Series with
Variance following Markov Process



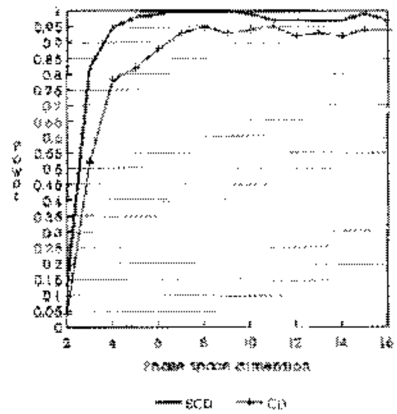
m-histories with lags of 2
Figure III.2

Power Test 2
Henon Series with Random Initial Values



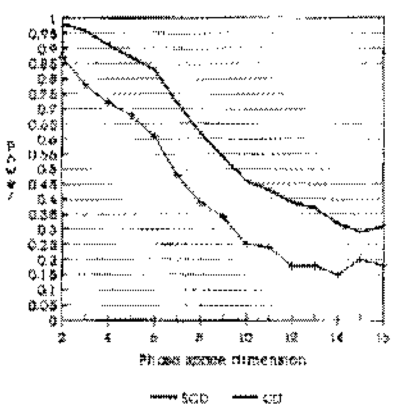
m-histories with lags of 3
Figure III.3

Power Test 3
Linear Autoregressive Series



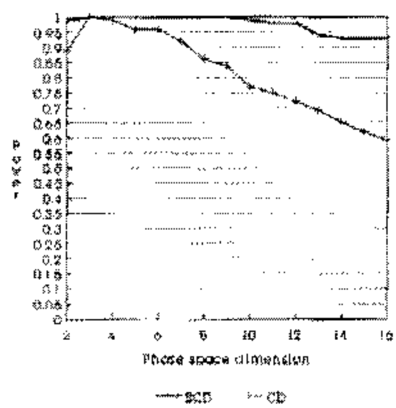
m-histories with lags of 3
Figure III.3

Power Test 4
Absolute Value ARCH with Non-regressive Mean



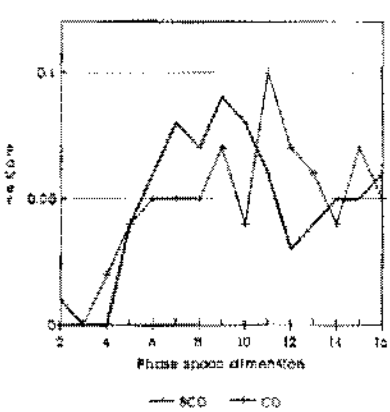
m-histories with lags of 3
Figure III.3

Power Test 5
Autoregressive Series with Absolute Value ARCH



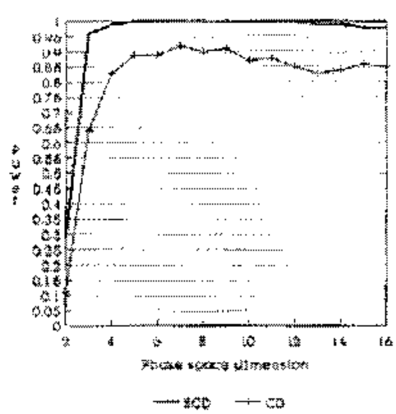
m-histories with lags of 3
Figure III.3

Power Test 6
Normal Random Variable with Variance following Markov Process



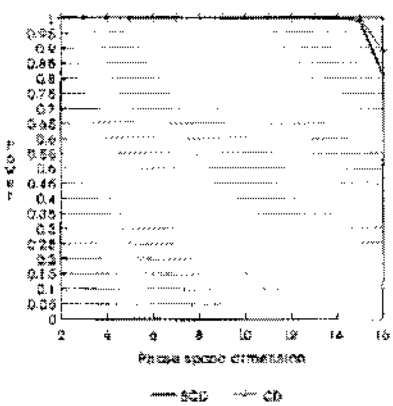
m-histories with lags of 3
Figure III.3

Power Test 7
Autoregressive Series with Variance following Markov Process



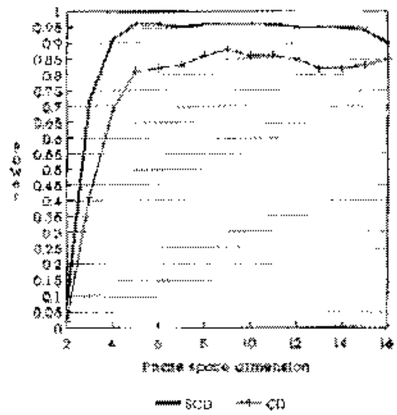
m-histories with lags of 3
Figure III.3

Power Test 2
Henon Series with
Random Initial Values



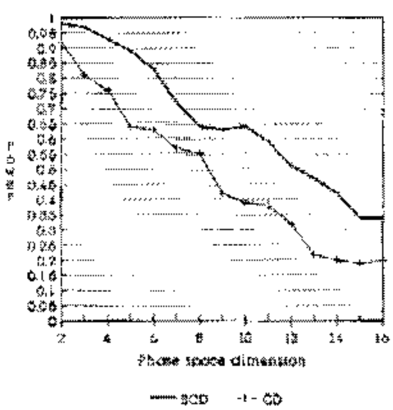
m-histories with lags of 4
Figure III.4

Power Test 3
Linear Autoregressive Series



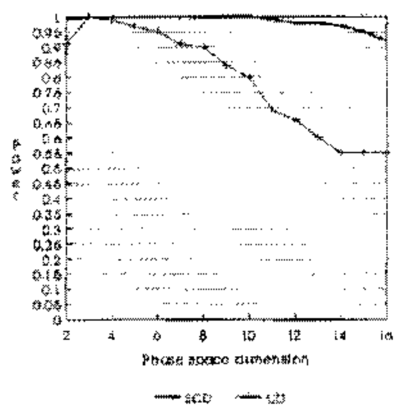
m-histories with lags of 4
Figure III.4

Power Test 4
Absolute Value ARCH
with Non-regressive Mean



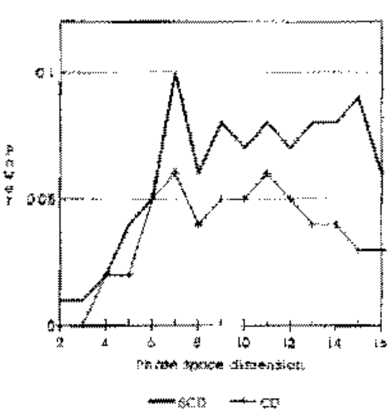
m-histories with lags of 4
Figure III.4

Power Test 5
Autoregressive Series with
Absolute Value ARCH



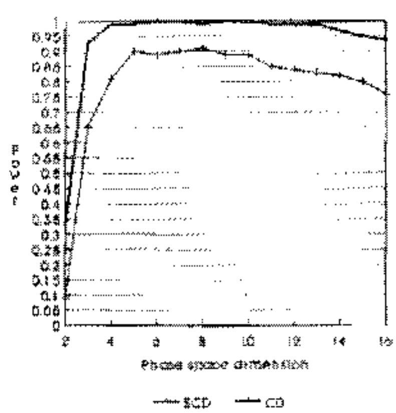
m-histories with lags of 4
Figure III.4

Power Test 6
Normal Random Variable with
Variance following Markov Process



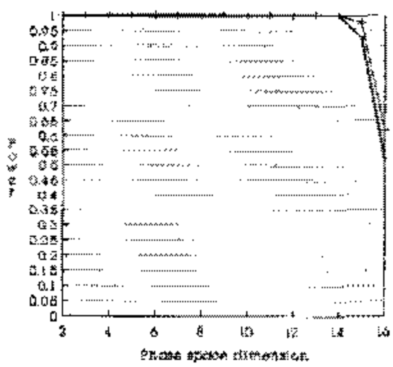
m-histories with lags of 4
Figure III.4

Power Test 7
Autoregressive Series with
Variance following Markov Process



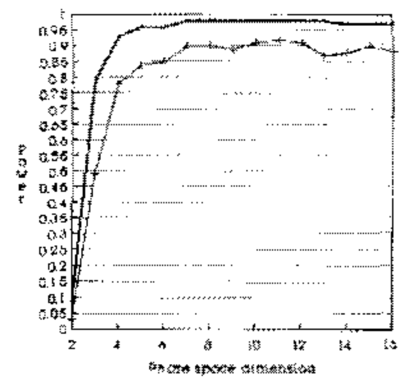
m-histories with lags of 4
Figure III.4

Power Test 2
Henon Series with
Random Initial Values



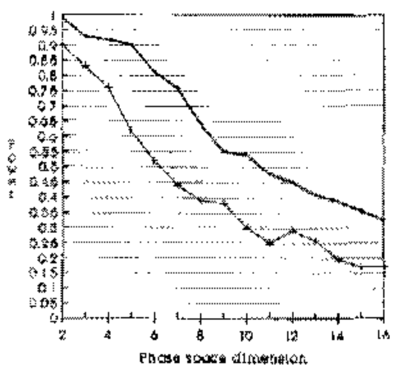
m-histories with lags of 12
Figure III.3

Power Test 3
Linear Autoregressive Series



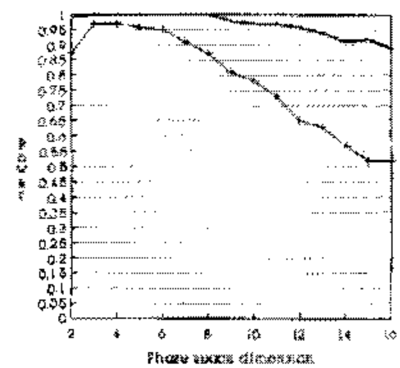
m-histories with lags of 12
Figure III.4

Power Test 4
Absolute Value ARCH
with Non-regressive Mean



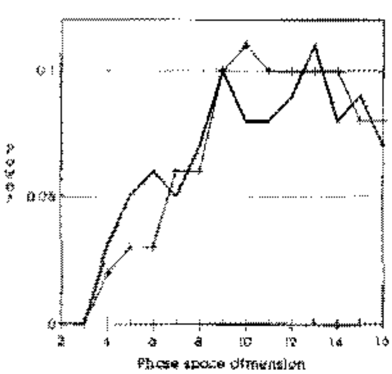
m-histories with lags of 12
Figure III.5

Power Test 5
Autoregressive Series with
Absolute Value ARCH



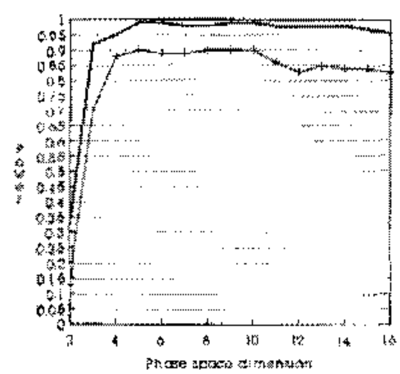
m-histories with lags of 12
Figure III.6

Power Test 6
Normal Random Variable with
Variance following Markov Process



m-histories with lags of 12
Figure III.7

Power Test 7
Autoregressive Series with
Variance following Markov Process



m-histories with lags of 12
Figure III.8

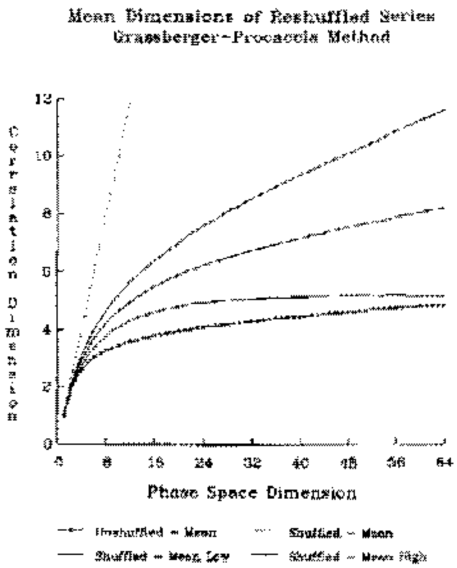


Figure IV.1

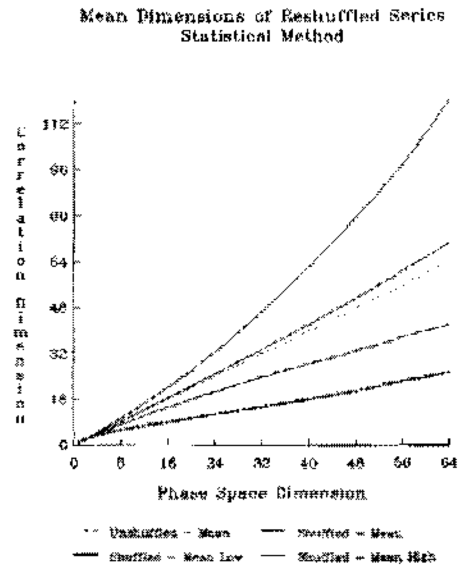
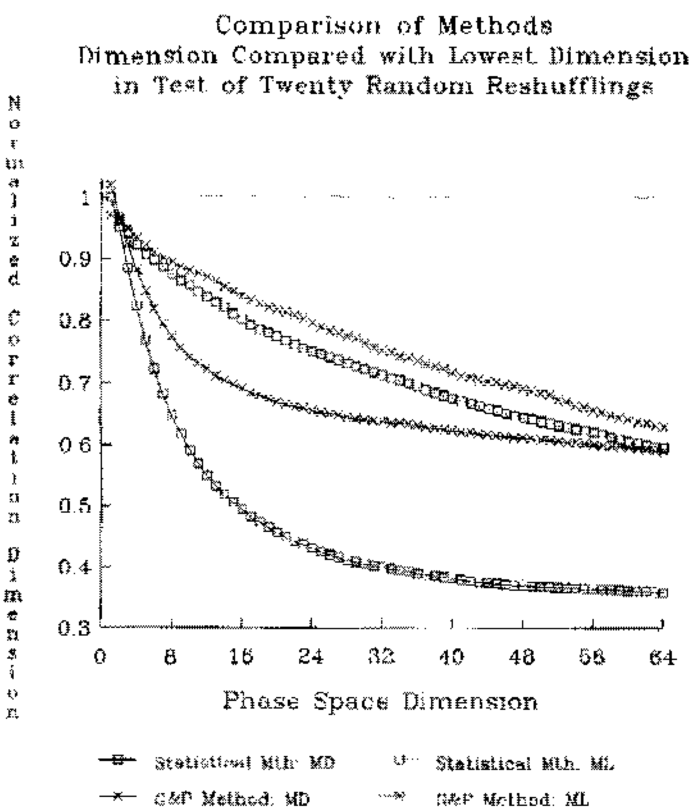


Figure IV.2



MD: Mean Dimension (103 series)
ML: Mean Lowest Dimension (103 tests)

Figure V

Brock and Sayers' Series
Dimensions of First Differences

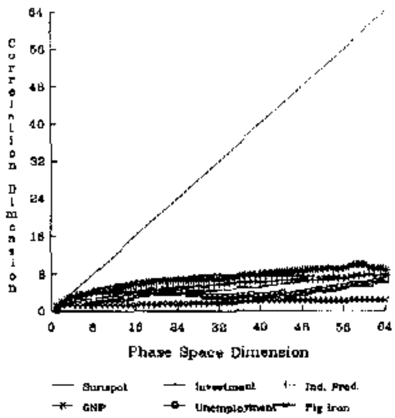


Figure VI.1

Brock and Sayers' Series
Dimensions of First Differences

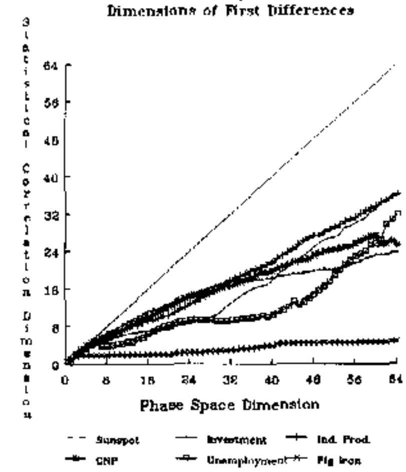


Figure VI.2

Brock and Sayers' Series
Dimensions of Autoregressive residues

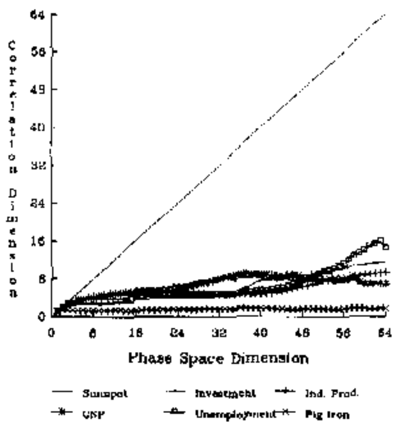


Figure VI.3

Brock and Sayers' Series
Dimensions of Autorregresive Residues

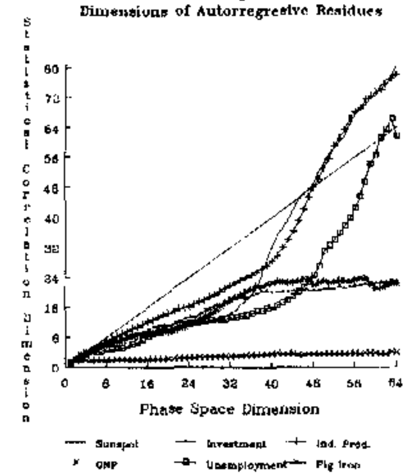


Figure VI.4

Brock and Sayers' Series
Dimensions of ARCH residues

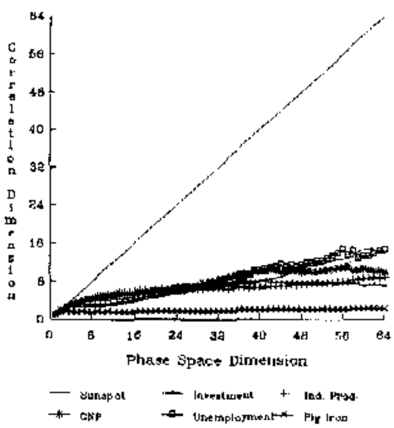


Figure VI.5

Brock and Sayers' Series
Dimensions of ARCH Residues

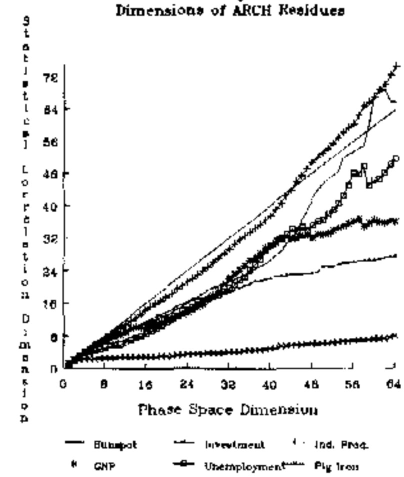


Figure VI.6

CRSP
Dimension of First Differences

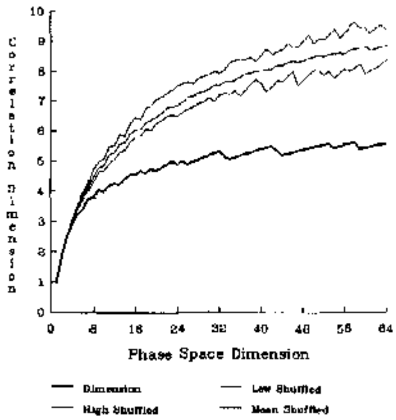


Figure VII.1

CRSP
Dimension of First Differences

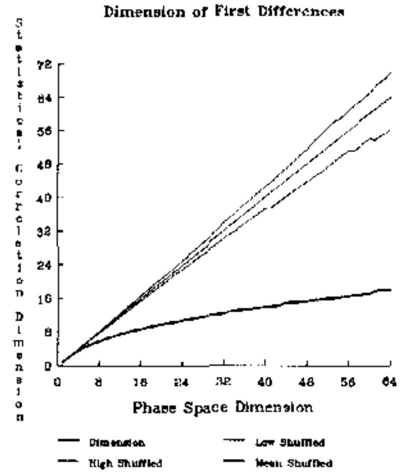


Figure VII.2

CRSP
Autoregressive Residue Dimensions

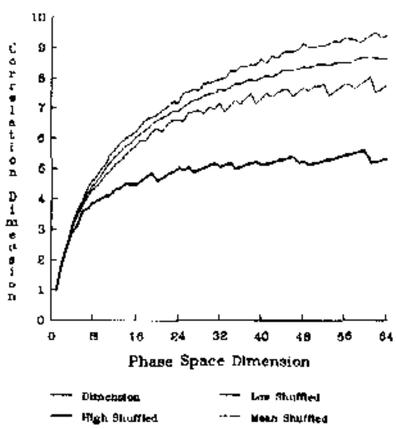


Figure VII.3

CRSP
Autoregressive Residue Dimensions

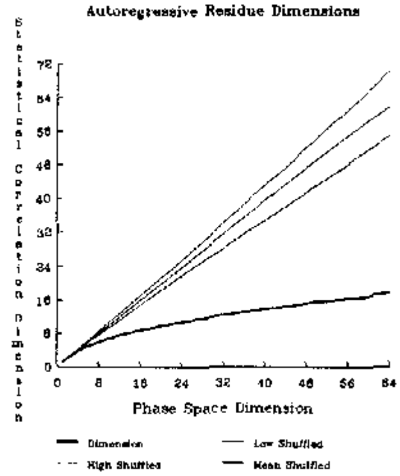


Figure VII.4

CRSP
Egarch Residue Dimensions

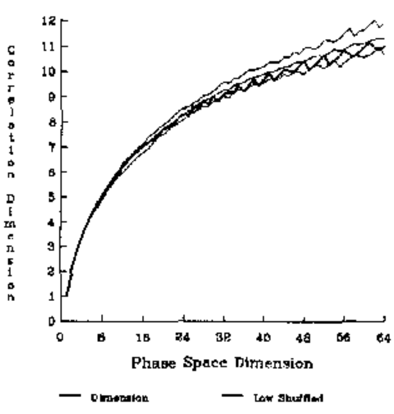


Figure VII.5

CRSP
Egarch Residue Dimensions

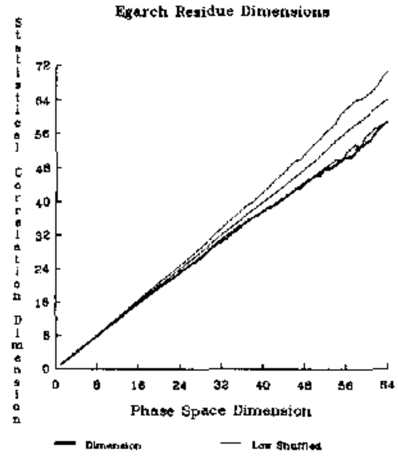


Figure VII.6

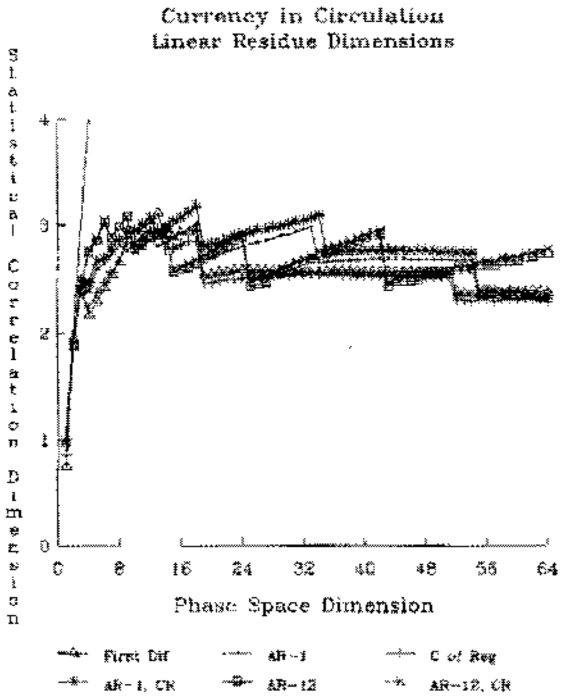


Figure VIII.1.1

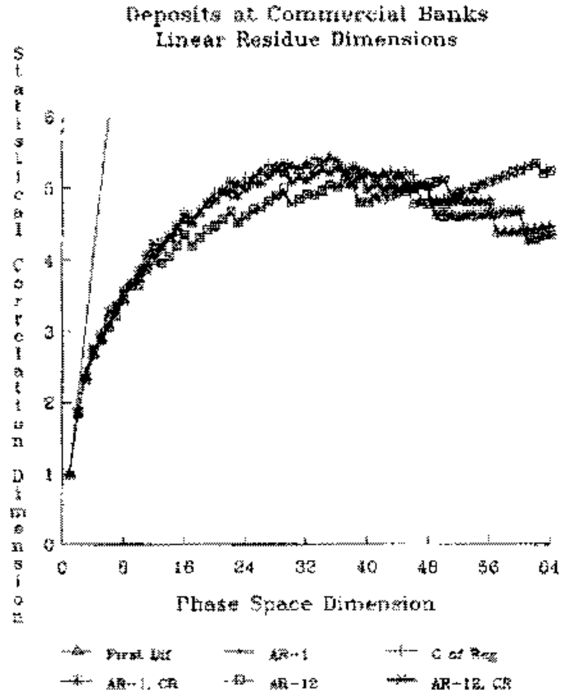


Figure VIII.1.2

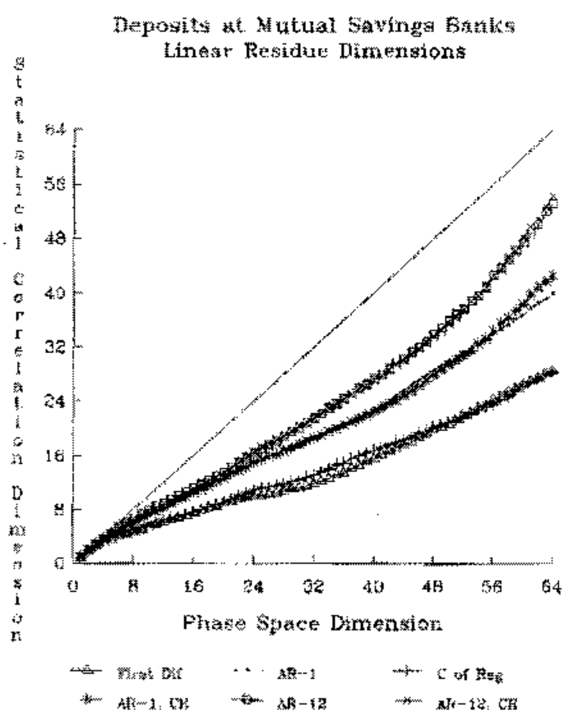


Figure VIII.1.3

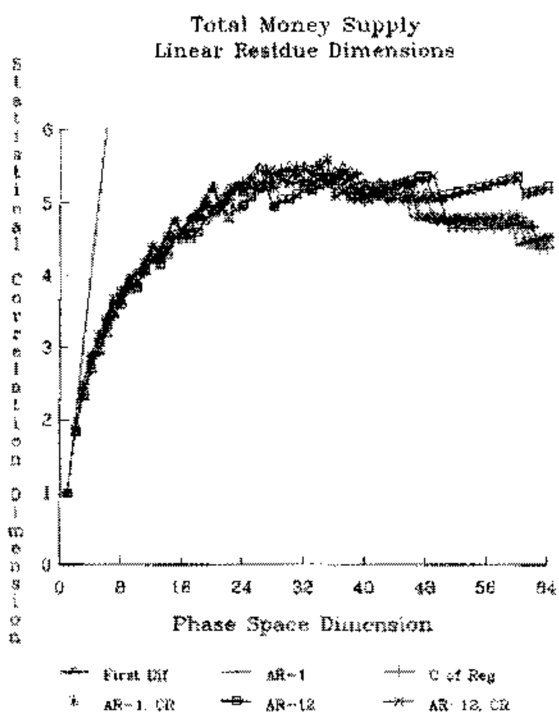


Figure VIII.1.4

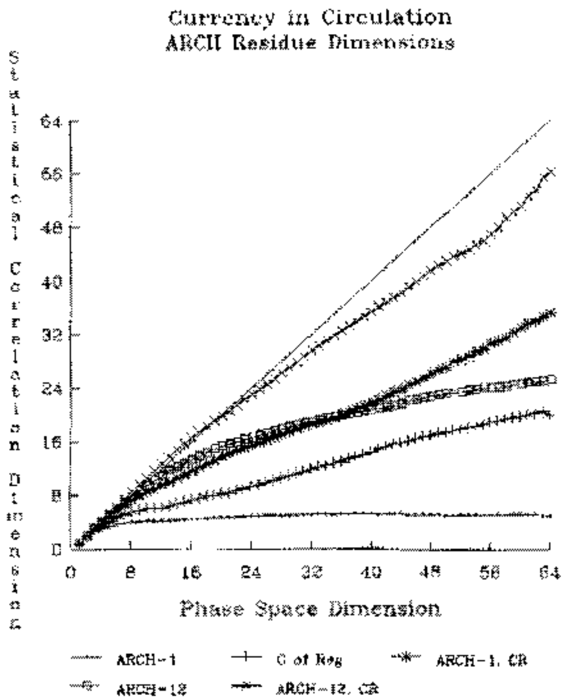


Figure VIII.2.1

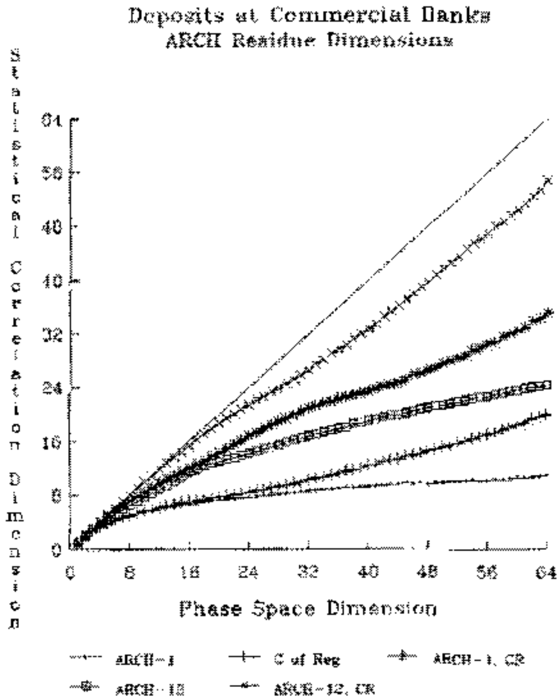


Figure VIII.2.2

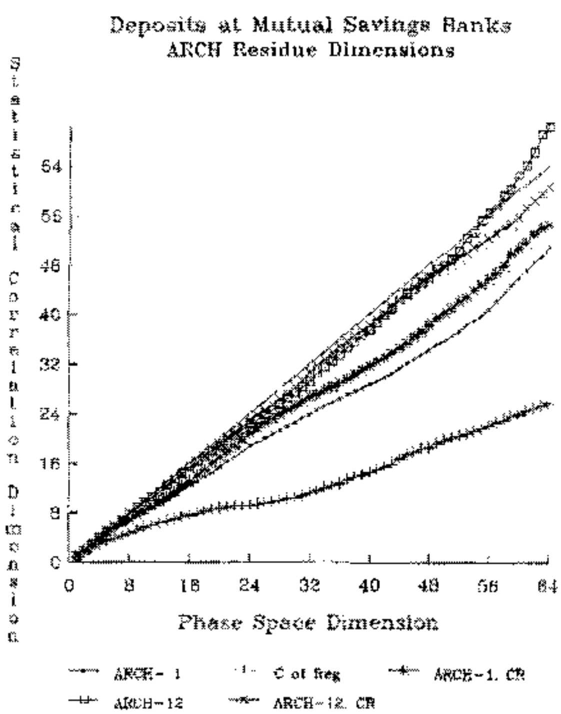


Figure VIII.2.3

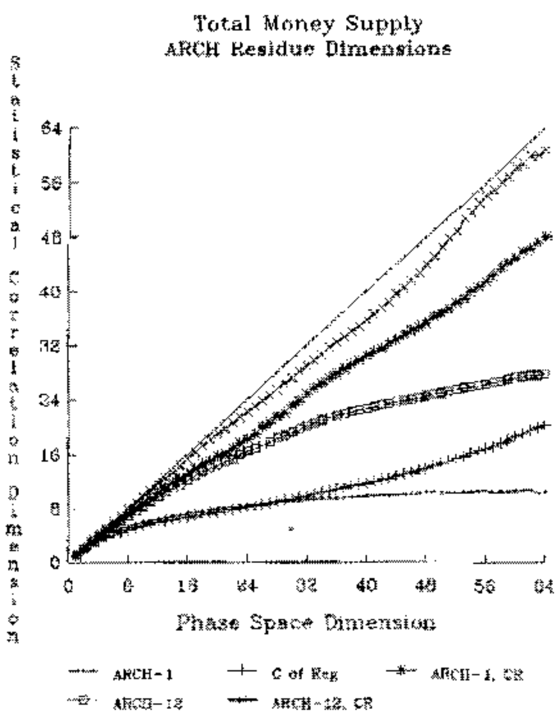


Figure VIII.2.4

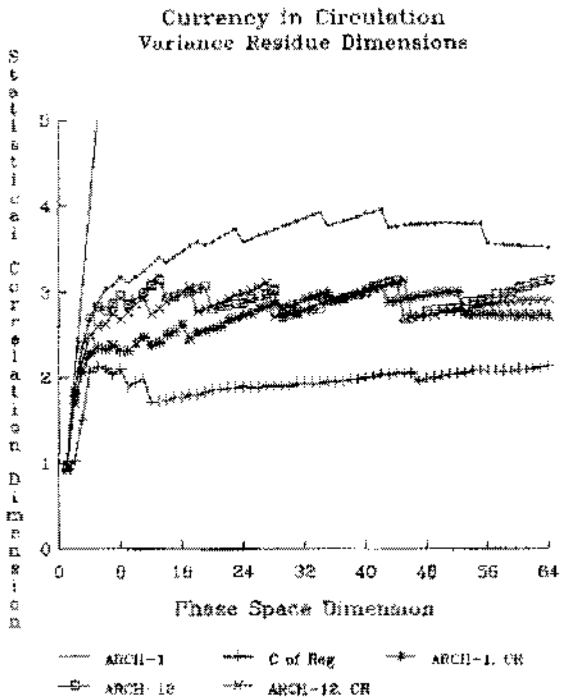


Figure VIII.3.1

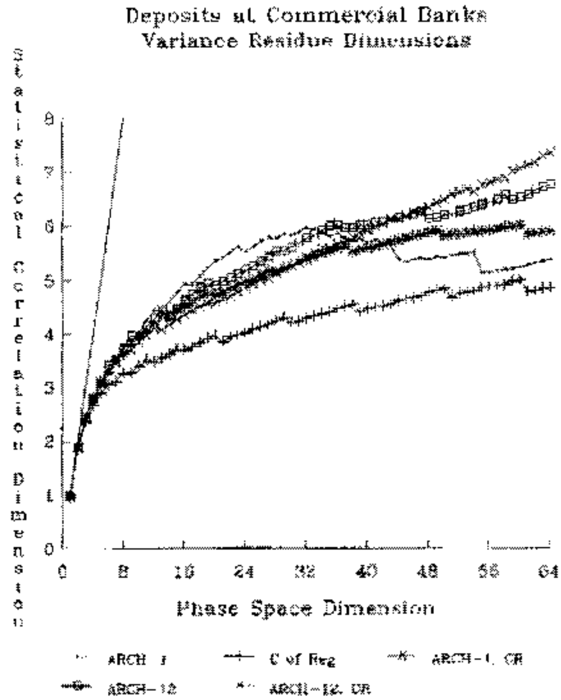


Figure VIII.3.2

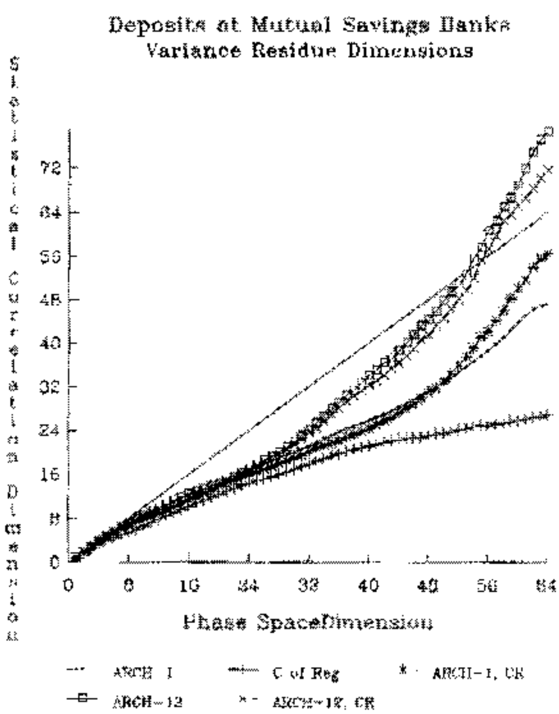


Figure VIII.3.3

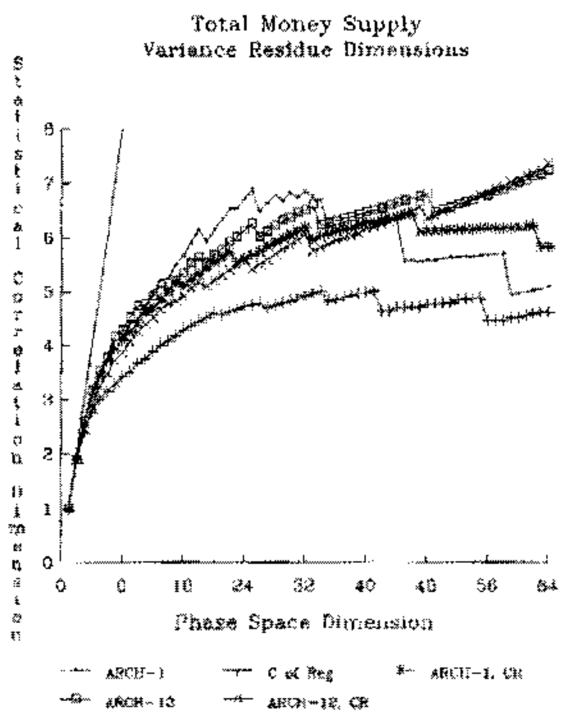


Figure VIII.3.4

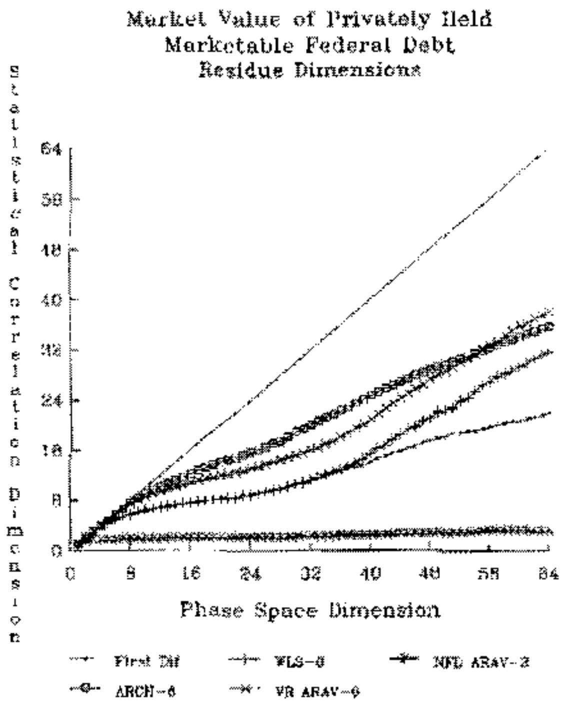


Figure IX.1

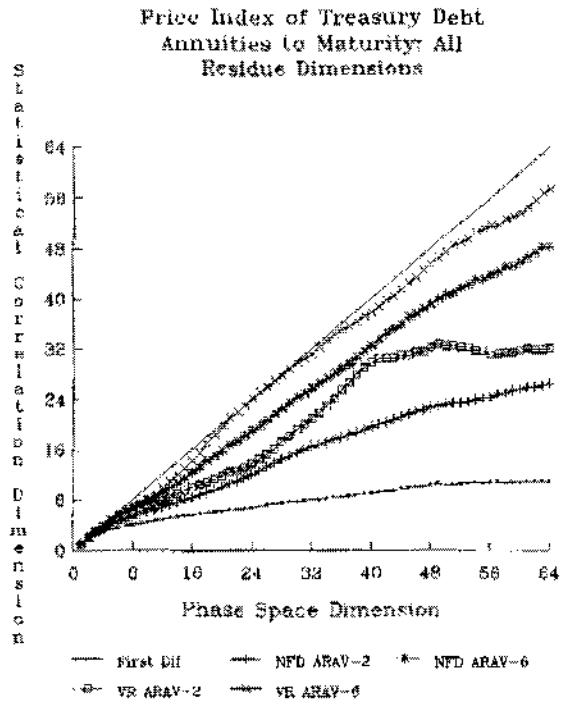


Figure IX.2

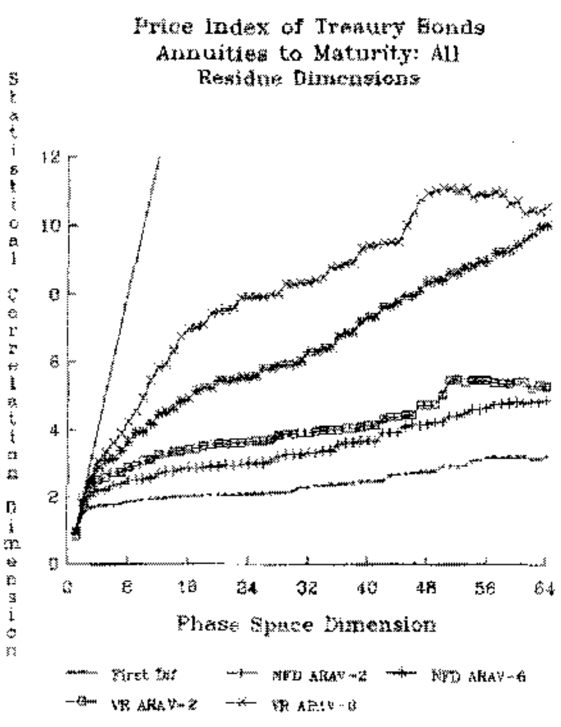


Figure IX.3

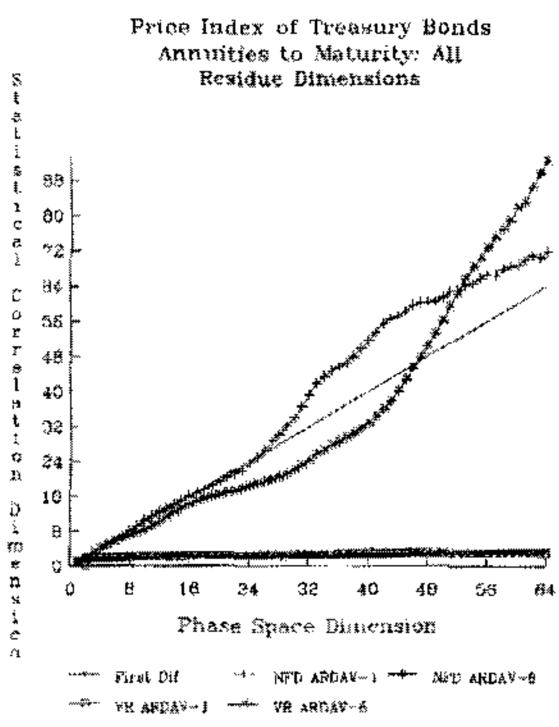


Figure IX.4

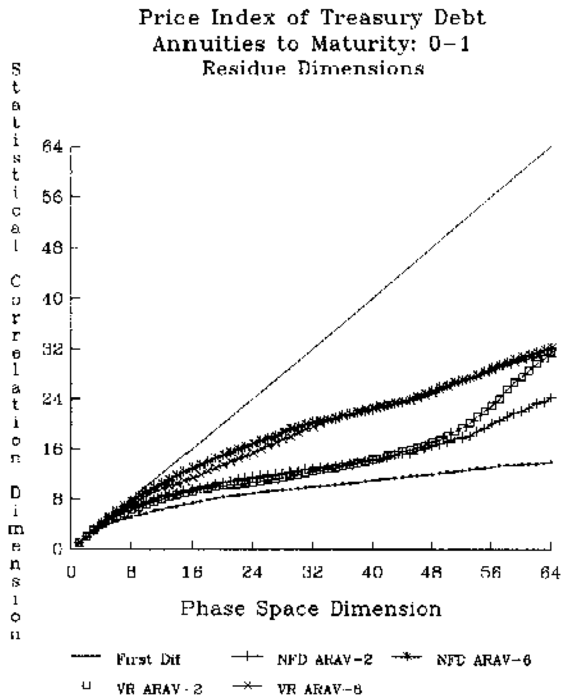


Figure IX.5

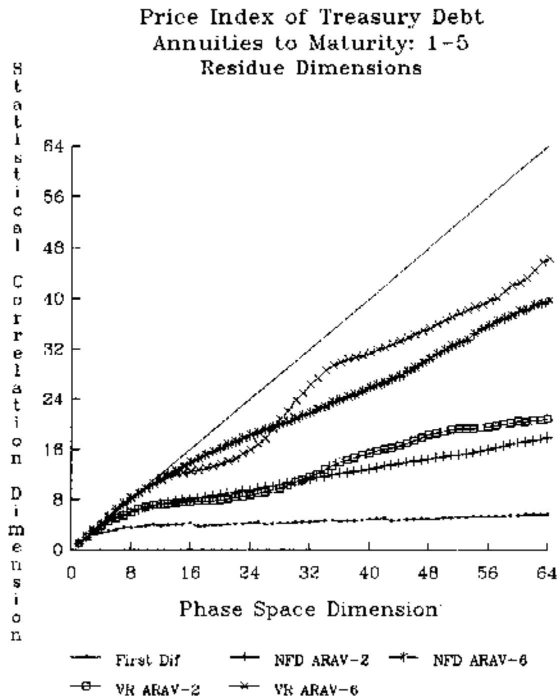


Figure IX.6

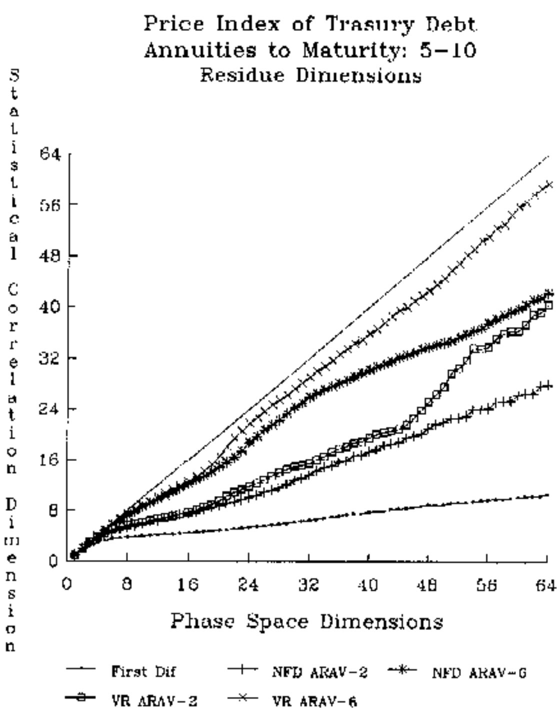


Figure IX.7

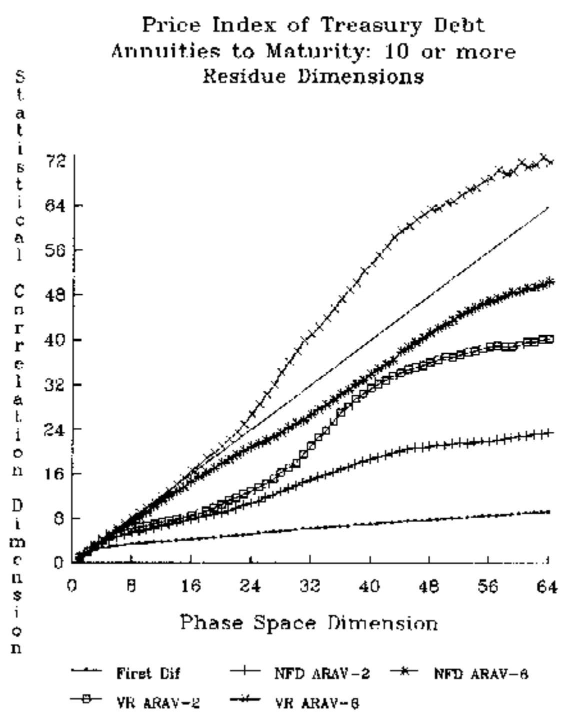


Figure IX.8