ECONOMIC RESEARCH REPORTS

THE STATISTICAL PROPERTIES OF DIMENSION CALCULATIONS USING SMALL DATA SETS: SOME ECONOMIC APPLICATIONS

by

James B. Ramsey

Chera L. Sayers

and

Philip Rothman

R.R. #88-10 April 1988

C. V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

March 26, 1988
Revised April 10 1988

THE STATISTICAL PROPERTIES OF DIMENSION CALCULATIONS USING SMALL DATA SETS: SOME ECONOMIC APPLICATIONS

by

James B. Ramsey, New York University

Chera L. Sayers, University of North Carolina - Chapel Hill

Philip Rothman, New York University

March 1988

THE STATISTICAL PROPERTIES OF DIMENSION CALCULATIONS USING SMALL DATA SETS: SOME ECONOMIC APPLICATIONS

I. Introduction[1]

Recently there has been an increase in the resources invested in the study of nonlinear dynamics. Originating in the natural sciences, applications of the theory have spread through various fields including brain research, optics, meteorology, and economics. The revived interest in nonlinear dynamics was sparked by the discovery in the natural sciences of processes characterized by deterministic chaos; that is, highly complex behavior that is generated by relatively simple non-linear functions. Observed time series generated by chaotic processes appear to be random utilizing conventional time series methods such as time series plots, auto-correlation functions, and spectral analysis.

However, while empirical studies in the natural sciences are characterized by large data sets, often numbering in the tens of thousands, data sets in economic applications usually consist of less than one thousand observations. Consequently, statistical procedures designed in the former context may not be appropriate in the latter.

The correlation dimension, a measure of the relative rate of scaling of the density of points within a given space, permits a researcher to obtain

---

1 Technical support by the C.V. Starr Center for Applied Economics at New York University is gratefully acknowledged. The authors wish to thank William Barnett for providing a copy of Fayyad's work and for helpful discussions. The following individuals kindly provided detailed information about their work: Ping Chen, Salam Fayyad, Blake Le Baron.

topological information about the underlying system generating the observed data without requiring a prior commitment to a given structural model. If the time series is a realization of a random variable, the correlation dimension estimate should increase monotonically with the dimensionality of the space within which the points are contained. By contrast, if a low correlation dimension is obtained, this provides an indication that additional structure exists in the time series -- structure that may be useful for forecasting purposes. In this way, the correlation dimension estimates may prove useful to economists wishing to scrutinize uncorrelated time series or the residuals from fitted linear time series models for information on possible non-linear structure. Furthermore, the correlation dimension can potentially provide an indication of the number of variables necessary to model a system accurately.

Three terms that are frequently used in this literature are "attractor", "embedding dimension", and "orbit" and should be at least intuitively defined in this paper. An "attractor" in the context of dynamical analysis is that sub-set of points towards which any dynamical path will converge; that is, the dynamical path is "attracted" to a subspace of the space containing the paths of the dynamical system from any initial condition. "Embedding dimension" is the topological dimension of the space in which the attractor is situated; loosely stated the embedding dimension is the number of axes needed to portray the attractor. Topological dimension specializes in vector spaces to the usual notion of Euclidian dimension. "Orbit" is essentially a synonym for the dynamical path, but also implies the notion that the dynamical path revisits any given part of the attractor infinitely often.

Recent economic empirical applications of the correlation dimension include Barnett and Chen (1986 a,b), Brock (1986), Brock and Sayers (1986), Frank and Stengos (1987 a,b), Gennotte and Marsh (1986), Hsieh (1987), Sayers (1986, 1987) and Scheinkman and Le Baron (1986, 1987). By the standards set in the physical sciences, the sizes of the data sets used for these analyses have been minuscule.

There is increasing concern over the application of the correlation dimension procedure to small data sets, even within the physics community. Furthermore, no finite sample distributional theory exists for the correlation dimension estimator. However, Brock, Dechert and Scheinkman (1987) provide asymptotic results when the observed points are generated by an independently and identically distributed set of random variables.

The work of Ramsey and Yuan (1987), which examines the statistical properties of dimension estimates and their variances, provides some insight into the finite sample properties of correlation dimension estimators. They have estimated empirical relationships between the conditional mean of correlation dimension estimates, the embedding dimension, and the sample size in order to evaluate the small sample biases in these estimators.

The purpose of this paper is to re-evaluate the calculation of dimension utilized in a few recent economic empirical applications in light of the results of Ramsey and Yuan (1987) and to present some cautionary remarks for researchers attempting application of the correlation dimension algorithms. In particular, it will be argued that potential biases are created in the dimension estimation process due largely to small sample size. These biases are dependent on the embedding dimension, the relevant time delay parameter, and the region used to estimate the relative rate of scaling; all of which must be carefully chosen if highly misleading results are not to be obtained.

Section II summarizes the theory of correlation dimension and its estimation, together with a brief summary of the difficulties that are inherent in dimension calculations. Section III presents the models that were examined and the results of our reexamination of previous research. Section IV contains our conclusions.

For the impatient reader, the main conclusion is that while there is abundant evidence for the presence of non-linear stochastic processes, there is virtually no evidence at the moment for the presence of simple chaotic attractors of the type that have been discovered in the physical sciences.

II. Correlation Dimension: Definition and Estimation.

Reviews of the correlation dimension procedures that are written with the economist in mind include Brock(1986), Brock, Dechert, and Scheinkman(1987), Brock and Sayers(1986), Barnett and Chen(1986a,b), and a more detailed evaluation of the details with a guide to the relevant physics literature is Ramsey and Yuan(1987). The basic idea underlying the calculation of dimension is relatively easily stated.

Any sequence of points, $\{x_t\}$, generated by some mechanism, whether random, chaotic, or otherwise, can be transformed into a sequence of d-tuples, $(x_{t1}, x_{t2}, \ldots x_{td})$. These d-tuples, regarded as points in a d-dimensional Euclidian space, can be "plotted" and properties of the cloud of points so created examined. The choice of the value of "d" is the choice of "embedding dimension"; it is the size of the Euclidian space into which the original sequence is being fitted. If the generated points are from observations on a random variable, then as d, the embedding dimension, is increased without bound and assuming an unlimited sample size, the size of the space into which the d-tuples will fit is d for all values of d; that is, random variables are space filling. But if the points are generated by a mechanism that is deterministic, or at least one that produces a shape that requires only "k" dimensions, then as the embedding dimension is increased without limit, the dimension of the points will not increase beyond "k". Imagine, for example, an ellipse, which is an object of dimension 1, that requires at least Euclidian dimension 2 to be observed, but no more; consider embedding an ellipse in a 3 or 4 dimensional space; the dimension of the ellipse is still 1.

Unfortunately, the objects of interest to us involve more complicated structures. The simplest intuitive example is to imagine a mechanism that produces points that are best described as the Cartesian product of the unit interval and a Cantor set; a Cantor set is obtained by deleting middle thirds from the remainder of the unit interval obtained by deleting middle thirds at a previous iteration. This idea can be extended to any number of Cartesian products.

Finally, there exists a set of problems that are particularly severe in a non-experimental discipline like Economics. These problems involve the extended "maintained hypothesis" that is needed in economic analysis, as well as in other non-experimental disciplines. In the problems examined to date in physics and chemistry, the simple dichotomy of: "either an attractor, or the data are merely high dimensional noise" has been considered to be appropriate. But this is not the case in economics. The extended maintained hypothesis must include as alternatives the options that the data come from ARIMA or non-linear stochastic processes.

Even more damaging to a simplistic version of dimension calculation is the realization that often researchers mistakenly perform dimension analysis on data that are highly auto-correlated; this procedure vitiates any conclusions that might possibly be made. This is because dimension is a topological concept and at certain scales of magnification of some stochastic processes, the dimension is in fact quite low; for example, the dimension of a geometric random walk, is at intermediate scales, about 1.1: a geometric random walk can be regarded at such scales as a highly convoluted line, giving it a topological structure that is of dimension slightly higher than that of a line. What is worse is that if the data are generated by a simple ARMA process with long auto-correlation lag, that is, a long period before the auto-correlations die to zero, then dimension calculations with such data will produce, over a range of scaling values, low dimensional results.

The problem for all experimental data, even if there were a perfectly well defined and recoverable attractor, is that at small enough scales the dimension is that of noise. Thus, the practical problem of trying to distinguish between attractors, auto-correlated processes, and non-linear stochastic processes is a real one.

The Grassberger-Procaccia (1983 a,b,c) algorithm will be utilized throughout this paper. Let the ordered sequence $\{X_t\}$, $t = 1,\ldots,N$, represent the observed time series. Then, for a given embedding dimension d, create a sequence of d-histories,

(2.1) $\{(x_t,\ x_{t+\tau}, \ldots\ldots\ x_{t+(d-1)\tau})\}$.

Here, τ stands for the time delay parameter. The sample correlation integral is given by, (2.2) $$C_r^N = N^{-2} \sum_{i,j} \theta(r - |x_i - x_j|),$$

$$r>0, \; x_i = (x_i, x_{i+1\tau}, \ldots, x_{i+(d-1)\tau}).$$

$\theta(.)$ is the Heaviside step function which maps positive arguments into one, and non-positive arguments into zero. Thus, $\theta(.)$ counts the number of points which are within distance r from each other. "r" is the scaling parameter. The calculation of C_r^N is useful because:

(2.3) $$\lim_{\substack{N\to\infty \\ r\to 0^+}} C_r^N \to C_r,$$

and $d\ln C_r/d\ln r = D_2$,

whenever the derivative is defined, Guckenheimer(1984) ; D_2 is a member of a general class of dimensions D_q, $q = \{0,1,2,\ldots\}$, defined by:

$$D_q = -\lim_{r\to 0} K_q(r)/\ln(r)$$

(2.4)

$$K_q = (1-q)^{-1} \ln \Sigma_{i=1}^{N(r)} P_i(r)^q$$

where $P_i(r)$ is the probability of a point of the attractor being within r of the ith point, N(r) is the number of such boxes needed to cover the attractor, and K_q is the Kolmogorov-Sinai (metric) entropy; details are in Badii and Politi, for example. In the rest of the paper, D_2 will be designated dc to stand for correlation dimension. The dc is a measure of pointwise dimension, dp. Pointwise dimension , see for example, Farmer, Ott, and Yorke(1983), measures the relative rate of change in the number of points on the attractor as the diameter of the covering sets is decreased. Pointwise dimension and related concepts differ from capacity and Hausdorff concepts in that they reflect the probability structure of the attractors; purely metric concepts, such as capacity, count all coverings equally, no matter how low the relative frequency of visitation by the orbits.

Correlation dimension is usually estimated from experimental data by a linear regression of the observed values of $\ln C_r^N$ on $\ln r$ over a suitably chosen sub-interval of the range of r, (0,1). The estimated slope coefficient of this regression, designated hereafter as $\hat{dc}$, is the usual estimator of correlation dimension cited in the literature and is the basic variable used in this paper.

However, there are a number of important qualifications to this seemingly simple procedure. First, an important practical issue involves the appropriate choice of the scaling region r actually used to calculate dc. While the theory discusses the properties of C_r as $r \to 0$, the reality is that the range of r used is far from zero and inevitably increases away from zero as embedding dimension is increased. Smaller values for r require substantial increases in sample size at any given embedding dimension in order to be able to determine a logarithmic linear relationship between C_r and r. In fact, the relationship between $\ln C_r$ and $\ln r$ is only approximately linear over a relatively narrow range of values for r. For large values of r, C_r saturates at unity so that the regression of $\ln C_r$ on $\ln r$ is zero. Further, as the value of r declines towards zero even with very large data sets, two complications arise; one is due to the limited precision of the data series and the other is due to the inevitable presence of noise. The former problem sets a practical lower bound on r before C_r collapses to zero and the latter difficulty offsets the decline in values of C_r when r reaches the level of the noise scales. Consequently, the negative slope of $\ln C_r$ on $\ln r$ starts at zero, increases first at an increasing rate, then may remain constant for a short range, before increasing again, and then falls very sharply.

With limited data sets a further problem occurs in that the plot of $\ln C_r$ on $\ln r$ for sufficiently high dimension becomes a step-function for certain ranges of r values, so that the estimation of slope becomes problematic at best. Ranges of r values which produce reliable slope estimates at lower embedding dimensions will yield apparently very low numbers at higher embedding dimensions. A similar phenomenon is observed as r decreases to very low values with fixed embedding dimension and sample size. This phenomenon is caused by the relative scarcity of data points so that C_r remains constant for sizable intervals of r values. This problem is particularly difficult with random models.

The practical implications of this are that the choice of r values with respect to which dc can be usefully calculated is difficult at any given embedding dimension and is doubly so if one is to avoid biases induced by changing the choice of r-region as embedding dimension grows; that is, one needs to avoid moving the r-region too close to the saturation level in response to the above mentioned difficulties. Further, if low scale noise is present, then the appropriate range of r for the calculation of dimension shrinks dramatically.

A subtle aspect of the "choice by eye" of the linear region is that when one is already aware, or believes one is aware, of the true slope coefficient, there is a natural, albeit unconscious, proclivity to choose the "right" value.

Yet another problem is due to an incorrect assumption about the nature of the regression to be used. The traditional model assumes,

(2.5) $\ln C_r = a + b \ln(r) + u$

where b is an estimate of dc, and u is an independently and identically distributed normal error term that is distributed independently of r. Thus, the variance of u generates the variance in estimating the dimension as represented by the coefficient "b". Ramsey and Yuan (1987) present an alternative and more appropriate random coefficient regression model of the following form,

(2.6) $\ln C_r = a(e|N,ED) + b(e|N,ED)\ln(r) + u$

where the variance of u is very small relative to the variance of e for small N and large ED. "e" represents an experimental error term whose distribution is dependent on the sample size, N, and the embedding dimension, ED; e is assumed to be distributed independently of u.

Ramsey and Yuan demonstrate that the conditional mean of the estimate of dc depends both upon the sample size and the embedding dimension in the following manner,

(2.7) $\ln(k+dc) = \gamma_1 + \gamma_2 N^{\gamma_3} + \gamma_4 N^{\gamma_5} [\mathrm{Exp}(\gamma_6/ED^{\gamma_7})) - 1.0]$,

where $\bar{dc}$ is the mean of the estimator $\hat{dc}$.

This equation is not as parameter extensive as it would appear. The parameter k equals zero or one. Its purpose is to improve the approximation for very low dimensional attractors and to allow γ_1 to be strictly positive for attractors. It could easily be set to 0 for all data sets with little serious impact on the overall results;this is the case in this study. The right hand side of equation (2.7) can be considered in two parts:

$$
(2.7') \quad \begin{aligned} &(a) \quad (\gamma_1 + \gamma_2 N^{\gamma_3}) \\ &(b) \quad (\gamma_4 N^{\gamma_5}[\exp(\gamma_6/ED^{\gamma_7}) - 1.0]). \end{aligned}
$$

The expression $(\gamma_1 + \gamma_2 N^{\gamma_3})$ indicates the main effect of small sample size on the expected value of the estimator $\hat{dc}$. For random variables that scale monotonically in ED, $\gamma_1 = 0$ and both γ_2 and γ_3 are positive; that is, the limit towards which the term (2.7'b) is approaching is given by $\gamma_2 N^{\gamma_3}$; the larger N, the larger the limiting value for $\overline{dc}$ expressed as a function of ED.

If, however, one has an attractor, then $\gamma_1 > 0$ and $\gamma_3 < 0$, so that as $N \to \infty$ the small sample bias provided by the term $\gamma_2 N^{\gamma_3}$ goes to zero; one would expect γ_3 to be -1.0 for attractors with low dimension. The asymptote as both N and ED $\to \infty$, but such that $\lim_{N\to\infty} ED/N \to 0$ is given by γ_1.

$$
\lim_{\substack{N\to\infty \\ ED\to\infty}} \ln(1+\overline{dc}) = \gamma_1, \qquad \text{or} \qquad \lim_{\substack{N\to\infty \\ ED\to\infty}} \overline{dc} = e^{\gamma_1}-1.
$$

The second part of the expression shown in (2.7'b) models the bias effect due to the embedding dimension. For both random and attractor generated data, one expects γ_6 to be negative, but γ_7 to be positive so that the approach to the limit is a sigmoid shape. The factor $\gamma_4 N^{\gamma_5}$ modifies the ED bias effect. γ_5 seems to be negative for attractors and zero for random variables. Thus, for low dimensional attractors the effect of a downward bias on $\overline{dc}$ due to small values for ED declines as N increases. γ_4 depends on the units of measurement chosen for ED and the relative weight of the ED effect to sample size, N.

The functional class, defined up to parameter values in equation (2.7), seems to be sufficiently general so as to be able to describe the bias scaling effects of a wide variety of alternative models, including simple attractors and random numbers. The equation would seem to be sufficiently general as to provide at least useful guidance for other, and perhaps less simple, attractors. It is to be hoped that while the coefficient values of equation (2.7) are clearly model specific, the form of the equation is invariant to a wide class of attractors and a wide class of distribution functions.

The authors found that dimension estimates for random signals are downward biased, while for chaotic signals the dimension estimate is upward biased for all, but very small values of ED.

In addition, this relationship contains a test for differentiating between low dimensional processes and random phenomena. If the underlying model is a random variable, then k=0, $\gamma_1 = 0$, and γ_2, $\gamma_3 > 0$. If the observed sequence of observations is being generated by a low dimensional process, then k=1, $\gamma_1 > 0$, and $\gamma_3 < 0$.

One of the implications of equation (2.6) is that the usual formula for the variance of the estimator of the parameter "b" is inappropriate; indeed, the usual estimate is a gross underestimate of the true variance as has been documented by Ramsey and Yuan.The equation that expresses the relationship between the actual standard deviation and the design parameters, N and ED is:

(2.8) $\ln(Std) = \alpha_1 + \alpha_2 \ln N + \alpha_3 \ln ED + \alpha_4 ED/N.$

In general, the standard deviation increases in embedding dimension, and decreases in sample size. However, the relative effect of ED is much greater than that for N; that is, small increases in ED must be offset by proportionately much larger increases in N in order to maintain a given variance. Ramsey and Yuan obtained estimates of the parameters in equation (2.8) by calculating dc for multiple repetitions of a given sample size, N, and for each value of ED. However, given the very small sample sizes involved in the studies examined in this paper we are unable to use this approach. This is a serious deficiency. While the ordinary least squares estimator of the variance is a linear function of the actual variance, the OLS estimate is a small fraction of the actual variance; the ratio varied anywhere from only a fifth to a low of 0.006 for the models studied in Ramsey and Yuan(1987). An idea currently under investigation is to use a Bootstrap procedure to provide a more realistic lower bound to the actual variance. What is very clear is that the OLS estimate is virtually useless.

III. THE ECONOMIC DATA AND THE ESTIMATION OF EQUATION (2.7).

We chose to examine the data sets utilized in Barnett and Chen (1986 a,b), Sayers (1986, 1987), and Scheinkman and Le Baron (1986).

Barnett and Chen assert evidence of chaos in the demand Divisia monetary aggregates. Our efforts concentrated on the demand Divisia M2. This choice was dictated in part by our wish to concentrate on the more important results of Barnett and Chen and in part because we discovered very little difference between the Divisia demand indices for M2 and M3; why there is almost an exact linear relationship between Divisia demand M2 and M3 we have not discovered. The sample period is "weekly" from January 1969 to July 1984, N=807. The data source is Fayyad (1986). The basic procedure that Fayyad seems to have followed is to apply the Divisia procedure to the components of monthly M2, or monthly M3, data series as published by the Federal Reserve, and then to generate a weekly series by spline interpolation at an approximate period of 0.23 to represent a week's fraction of a month. Consequently, only a very few of the original monthly data points are contained in the constructed "weekly" data series; the corresponding number of monthly data points is about 200.

Sayers (1986, 1987) used man-days idle due to work stoppages, monthly figures from January 1928 to December 1981, N=648, as published by The Bureau of Labor Statistics.

Scheinkman and Le Baron utilize Center for Research in Security Prices (CRSP) data. They have indicated finding a dimension of about 4.5. These data are value weighted daily stock returns, with a sample size of 5200 daily returns. Weekly returns were obtained by simple compounding of the daily returns; the details are contained in Scheinkman and Le Baron(1986).This procedure yields 1227 weekly observations.

Our analytical procedure with respect to each data set was as follows. The first step was to replicate each author's published results. Except for minor errors due to differences in algorithms or computer word size, we were able in all cases to duplicate the original results: given the experience of one of the authors of this paper, this is a testimony to the care taken by all of the above mentioned researchers.

The next step was to split each sample into sub-samples of 200, 400, 600, 800, etc. observations in order to attempt to estimate equation (2.7). The original idea was to obtain sub-samples by random sampling of the initial starting value. However, problems that we encountered with the Scheinkman data led us back to a sampling procedure that was more systematic for that data set. The problem that we encountered is of considerable interest and will be discussed below. For each set of sub-sets of data, we estimated dc in the standard manner, for ED = 2,3,...25, if the data set could sustain such a high embedding dimension. In particular cases, the highest sustainable dimension was as low as 8. The final step was to regress the estimated dc on N, sample size, and ED, embedding dimension, as discussed with respect to equation (2.7). The estimated equation was then analyzed in accordance with the discussion above, in order to try to resolve the issue of whether the observed time series indicated the presence of a chaotic attractor. In addition the data were randomized in terms of their time order and the whole procedure repeated with the randomized data.

IV. Results

Figures R.1 to R.3 demonstrate in a striking manner the difference in the relationship between the estimated dimension, sample size, and embedding dimension that was discovered in the context of known attractors and known distributions. These figures will provide a useful benchmark of comparison for the results on the economic data analyzed in this paper.

Sayer's Work Stoppage Data.

Figures S.1 to S.4 and Table S.1 summarize the results of applying the procedures discussed in this paper to the work stoppage data. These data provide the clearest impression of structure that might have low dimension. The dimension calculations were performed at a lag of 5 which represents the lowest lag with zero auto-correlation. However, one must be extremely cautious about any conclusions, because the number of data points is so limited. Nevertheless, the results are intriguing.

A comparison of Figures S.3 and S.4, together with an examination of Table S.1, indicates that there is more evidence of structure in the raw data than in the randomized data. However, with real as opposed to simulated data, especially when using strictly limited data sets, difficulties with estimating the coefficients occur. Examine the estimates for γ_1 and γ_3 for the randomized data. By the analysis in Ramsey and Yuan (1987) γ_1 should be zero and γ_3 should be positive, whereas the estimated values are 4.97 for γ_1 and -0.06 for γ_3. But both estimates have t-ratios that are about .1; that is, the standard deviations are about 10 times the size of the coefficients. In addition, the limited extent of the data produces very elongated confidence ellipsoids. Consequently, the randomized Sayers' data produces an estimate of γ_1 that is very large and offsets that with an estimated value of γ_3 that is negative, instead of the theoretically predicted values of $\gamma_1 = 0$ and $\gamma_3 > 0$.

The results for the raw data indicate an estimated asymptotic value for the dimension of the data of 0.214, but the actual value could easily be as high as 1.68 at an approximate 95% confidence level; a value of 1.6 to 1.8 seems to be a common finding in economic data. But the reader is warned that the estimated standard errors are usually fractions of those indicated by the standard analysis, see Ramsey and Yuan (1987). The remaining coefficient estimates seem to have reasonable values; for example, the power on the additive term in N is approximately -0.5, (actual value is -0.62), which agrees with the usual square root law of asymptotic convergence. The original estimates by Sayers were in a range of about 5.0 to 10.0, depending on sub-sampling size and scaling region used. While these results would indicate the probability at least of a low dimensional result, one cannot conclude from these results alone that one has an attractor. However, one can conclude that these data are neither random, nor simple ARMA processes.

The results produced in this paper can be regarded as a modest improvement in the determination of dimension using these work stoppage data.

Scheinkman and Le Baron Stock Market Data

The results are summarized in Figures Sc.1 to Sc.4 and Table Sc.1. The τ lag used to generate the dimension calculations was 2. While we were able to reproduce the results quoted by Scheinkman and Le Baron, it soon became clear that there were difficulties with the inferences drawn by the authors. First, a careful examination revealed that even with the procedure used by Scheinkman and Le Baron, there is slight, but perceptible, evidence of continued scaling of dimension estimates as embedding dimension increases. Secondly, our initial sub-sampling procedures produced some surprising anomalies that were for some time quite puzzling, until we discovered the key.

Essentially, the authors cited an attractor dimension of less than 6.0. But this conclusion is in fact driven by approximately 25 observations. In the second panel of the time series plot for these data at about 625 observations, the reader will observe that the range of the data suddenly doubles; the data are in fact non-stationary. The effect of the presence of these observations in the series is to lower the entropy of the whole data set dramatically; these data, when normalized to facilitate the analysis, have a long thin tail of values in the histogram plot with a heavy concentration of points in the narrow range defined by the bulk of the data. The implication of low entropy for random data is that it slows, often dramatically, the rate at which estimated dimension increases with embedding dimension. Consequently, such data, especially in the presence of significant auto-correlation, can easily give the appearance of a low dimensional attractor. We discovered that the elimination of 25 consecutive data points in this region led to slightly different estimates for the whole sample, but removed all our anomalous results with the sub-samples. The 25 observations that we dropped were 626 to 650; these observations cover the period from July 1974 to the beginning of January 1975. Figures Sc.3 and Sc.4 were produced with the "purged" data. There is no clear evidence of an attractor in these plots as is confirmed by the regression estimates for the raw and randomized data sets. Even with approximately 1200 observations, there are considerable difficulties with the estimation of the coefficients in equation (2.7). In addition, the reader should be cautioned to remember that the seemingly good fit of each regression of ln Cr on ln r does not provide a good estimate of the true variance of estimate.

Barnett and Chen's Filtered Divisia Money Demand (M2) Data

The results of our efforts are summarized in Figures B.1 to B.5 and Table B.1. Before proceeding, we should warn the reader that the procedures used to generate the 'weekly' data involved spline interpolations from about 200 original data points, not 807 original observations as might be inferred by a careless reader. A greater difficulty with the original approach used by Barnett and Chen is that the first zero of the auto-correlation function for the raw data series is at a lag of approximately 180. As we have seen from the discussion of Ramsey and Yuan (1987), it is crucial to a useful interpretation of dimension calculations that the calculations be carried out at a zero auto-correlation lag, or at least at the first minimum of the mutual information, Fraser (1986). With only 807 splined data points, the standard procedure was clearly impossible.

Our approach was to transform the original data into a stationary series. We did so by subtracting from the original series a 19 point double sided moving average filter in order to eliminate from the series the trend and some of the lowest frequency variation; the corresponding transfer function is shown in Figure B.3. The chosen τ lag was 6.

The raw, but filtered, Barnett data show the least indication of any structure of the three sets of data examined in this paper; compare Figure B.4, the plot of the dimensional estimates on N and ED for the raw filtered data, with Figures R.1 to R.3. While the sign of γ_2 using the randomized data is incorrect, the corresponding t-ratio is less than 0.03, so that little confidence can be placed in the value of its estimate.

Notwithstanding this conclusion, the estimates using the two sets of data tentatively indicate that there was some effect of the order in the original data that was lost when the data were randomized.

V. Conclusions

The main question of burning interest is whether or not there is any evidence of the presence of a strange, or of a chaotic, attractor. The short answer is no. That does not mean that other procedures, or that the use of new data, will produce evidence of attractors; but the current evidence based on minuscule data sets does not provide any indication of an attractor. A possible exception to this general statement is the work stoppage data whose plots show the characteristic shapes exhibited by attractors and their randomized values. However, even this modest claim must be severely hedged by the fact that there are only 648 observations.

There is evidence of varying persuasiveness of non-linearity in all of these data. The effect of randomization is clear, even when applied to data points that are plotted at the zero auto-correlation lags.

The second major pair of lessons is that the calculation of dimension with small data sets is delicate to say the least and that the procedures proposed in Ramsey and Yuan (1987) do provide some relief from the stringency of small sample sizes. By capitalizing on the implicit structure of the bias relationship between the conditional mean of the dimension estimate, sample size, and the embedding dimension, one can improve the quality of one's inferences about the topological structure of non-linear stochastic processes, if not about the topological structure of chaotic attractors.

One final important insight is that economic data seem to show definite signs of non-stationarity, even when in differenced form or when low frequency components have been removed. The typical structure is that of noise components, or seemingly random variation, that are interspersed with periods of very high amplitude, but of low frequency

Table S-1

Regression Results from Fitting

$$\ln(\overline{dc}) = \gamma_1 + \gamma_2 N^{\gamma_3} + \gamma_4 N^{\gamma_5} [\mathrm{Exp}(\gamma_6/ED^{\gamma_7})) - 1]$$

To Sayers "Work Stoppage Data"

Data	(γ)	Coefficient Estimates	Estimated Student-t Ratios	R^2 Value/ Asymptotic $\overline{dc}$ Estimat
Original	1	-1.543	-1.495	$R^2 = 0.983$
	2	4.069	3.772	
	3	-0.621	-4.423	
	4	2.882	4.313	
	5	-0.806	-11.358	60 degrees of freedom (DOF)
	6	-1.066	-4.921	
	7	0.537	4.495	
Randomized	1	4.967	0.084	$R^2 = 0.988$
	2	5.310	0.089	
	3	-0.055	-0.086	
	4	11.872	0.995	
	5	-0.032	-0.601	54 DOF
	6	-1.192	-1.658	
	7	0.298	1.189	

Table Sc.1

$$\ln(\overline{dc}) = \gamma_1 + \gamma_2 N^{\gamma_3} + \gamma_4 N^{\gamma_5}[\mathrm{Exp}(\gamma_6 ED^{\gamma_7})) - 1]$$

To Scheinkman "Computed Stock Return Data"

Data	(γ)	Coefficient Estimates	Estimated Student-t Ratios	R^2 Value
Original	1	4.018	1.585	$R^2 = 0.981$
	2	4.060	0.524	
	3	-0.121	-1.499	
	4	11.255	0.850	
	5	-0.079	-5.269	268 DOF
	6	-0.893	-5.815	
	7	0.160	0.833	
Randomized	1	5.340	8.651	$R^2 = 0.998$
	2	0.646	1.560	
	3	0.267	1.690	
	4	7.690	9.211	
	5	0.032	4.340	165 DOF
	6	-0.715	-19.288	
	7	0.378	8.540	

Table B.1

Regression Results for Fitting

$$\ln(\overline{dc}) = \gamma_1 + \gamma_2 N^{\gamma_3} + \gamma_4 N^{\gamma_5} [\mathrm{Exp}(\gamma_6/ED^{\gamma_7})) - 1]$$

To Filtered Barnett "Money Divisia Demand (M^2) Data"

Data	(γ)	Coefficient Estimates	Estimated Student-t Ratios	R^2 Value
Original	1	3.133	1.669	R^2 = 0.986
	2	3.848	0.750	
	3	3.003	1.156	
	4	7.683	0.916	
	5	1.153	7.478	57 DOF
	6	-1.294	-2.338	
	7	0.293	1.001	
Randomized	1	6.10	0.122	R^2 = 0.982
	2	-1.042	-0.029	
	3	-0.069	-0.034	
	4	8.029	0.138	
	5	0.014	0.216	23 DOF
	6	-0.515	-0.460	
	7	0.310	0.140	

Figure R.1

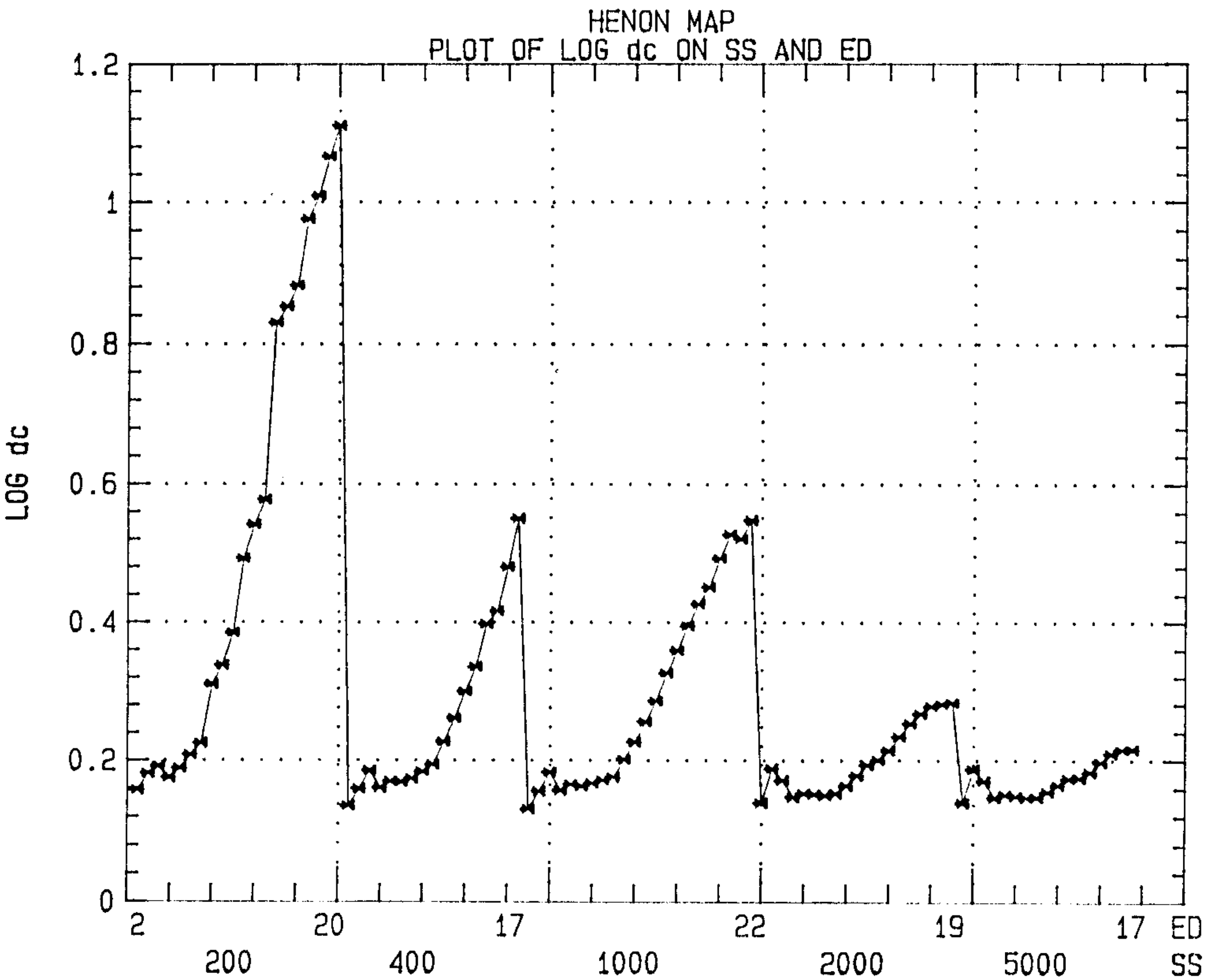

HENON MAP
PLOT OF LOG dc ON SS AND ED
1.2
1
0.8
0.6
0.4
0.2
0
LOG dc
2
20
17
22
19
17
ED
200
400
1000
2000
5000
SS

Figure R.2

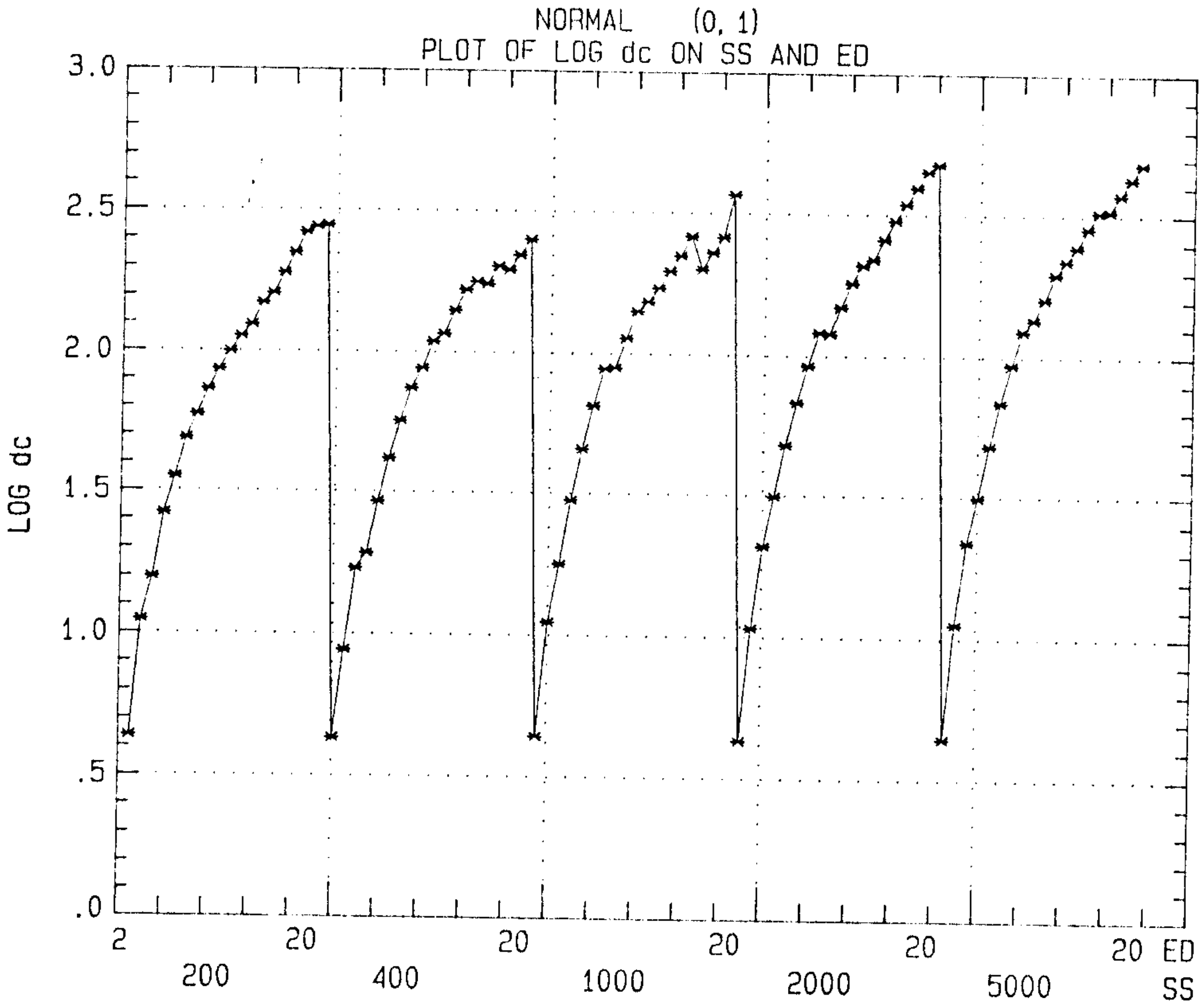

NORMAL (0, 1)
PLOT OF LOG dc ON SS AND ED
LOG dc
3.0
2.5
2.0
1.5
1.0
.5
.0
2
20
20
20
20
20 ED
200
400
1000
2000
5000
SS

Figure R.3

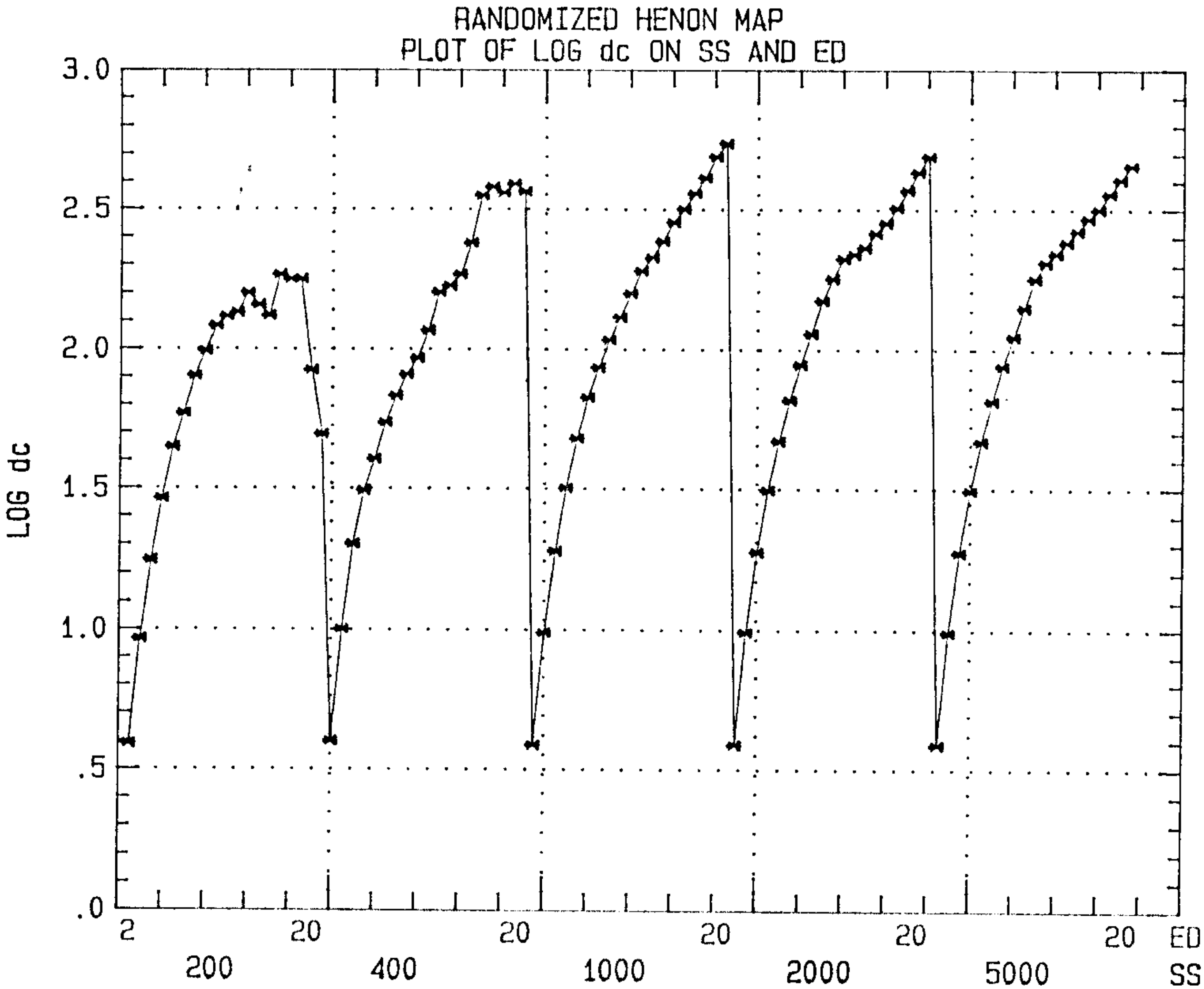

RANDOMIZED HENON MAP
PLOT OF LOG dc ON SS AND ED
3.0
2.5
2.0
1.5
1.0
.5
.0
LOG dc
2
20
20
20
20
20
ED
200
400
1000
2000
5000
SS

Figure S.1

Time Series Plot for Sayers Work Stoppage Data

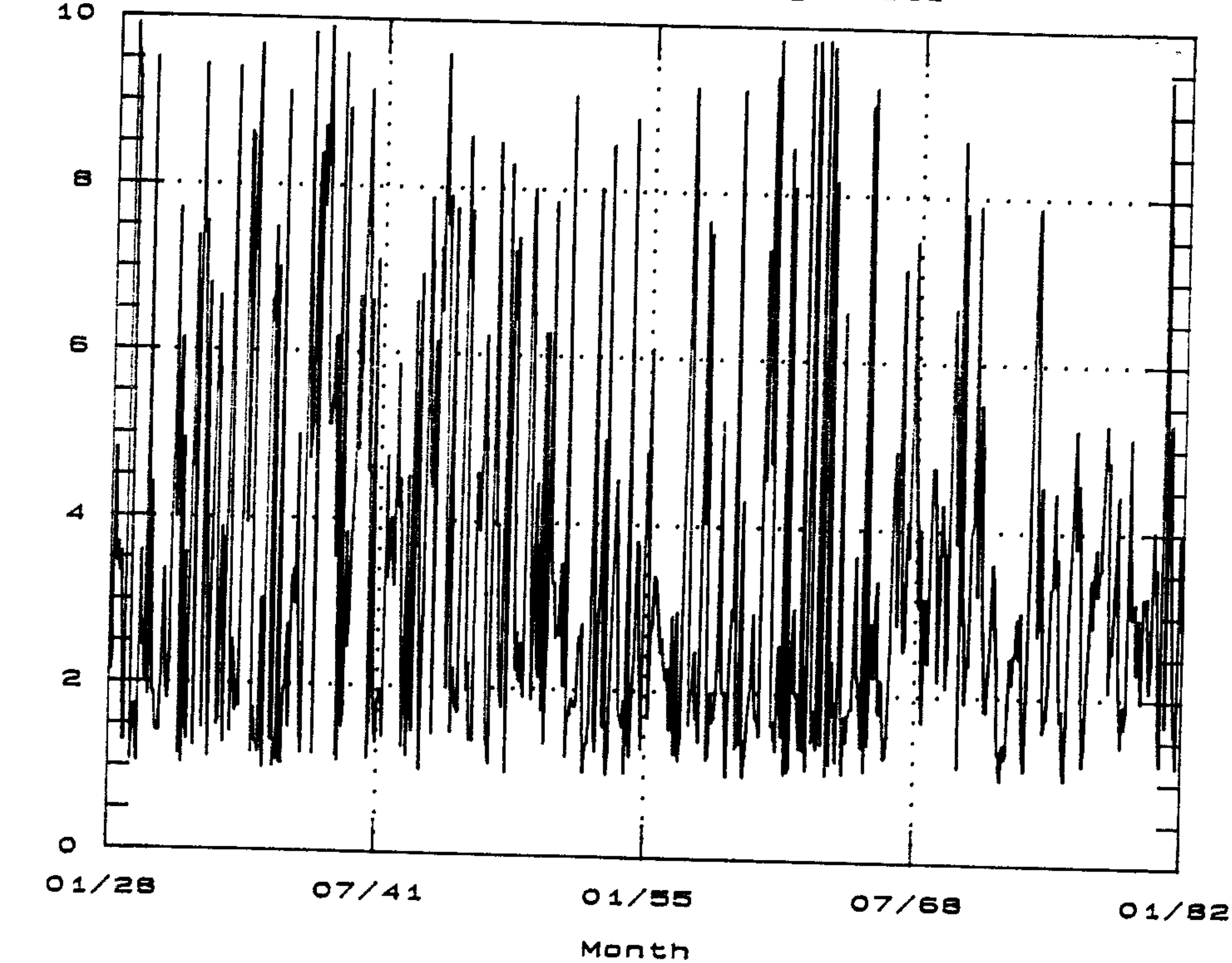


Figure S.2

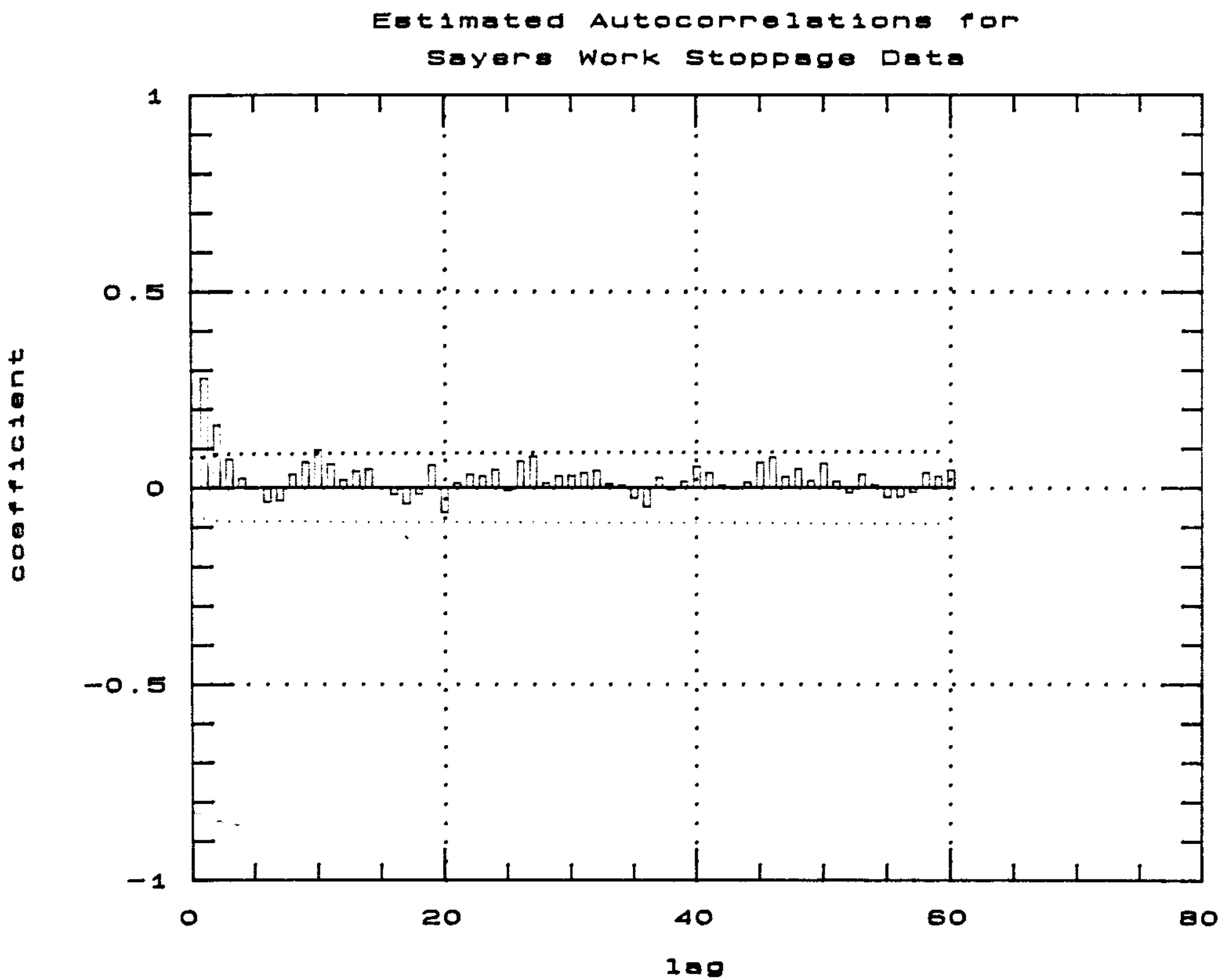

Estimated Autocorrelations for
Sayers Work Stoppage Data
1
0.5
0
-0.5
-1
coefficient
0
20
40
60
80
lag

Figure S.3

Plot of ln (dc) on N and ED for Sayers Work Stoppage Data

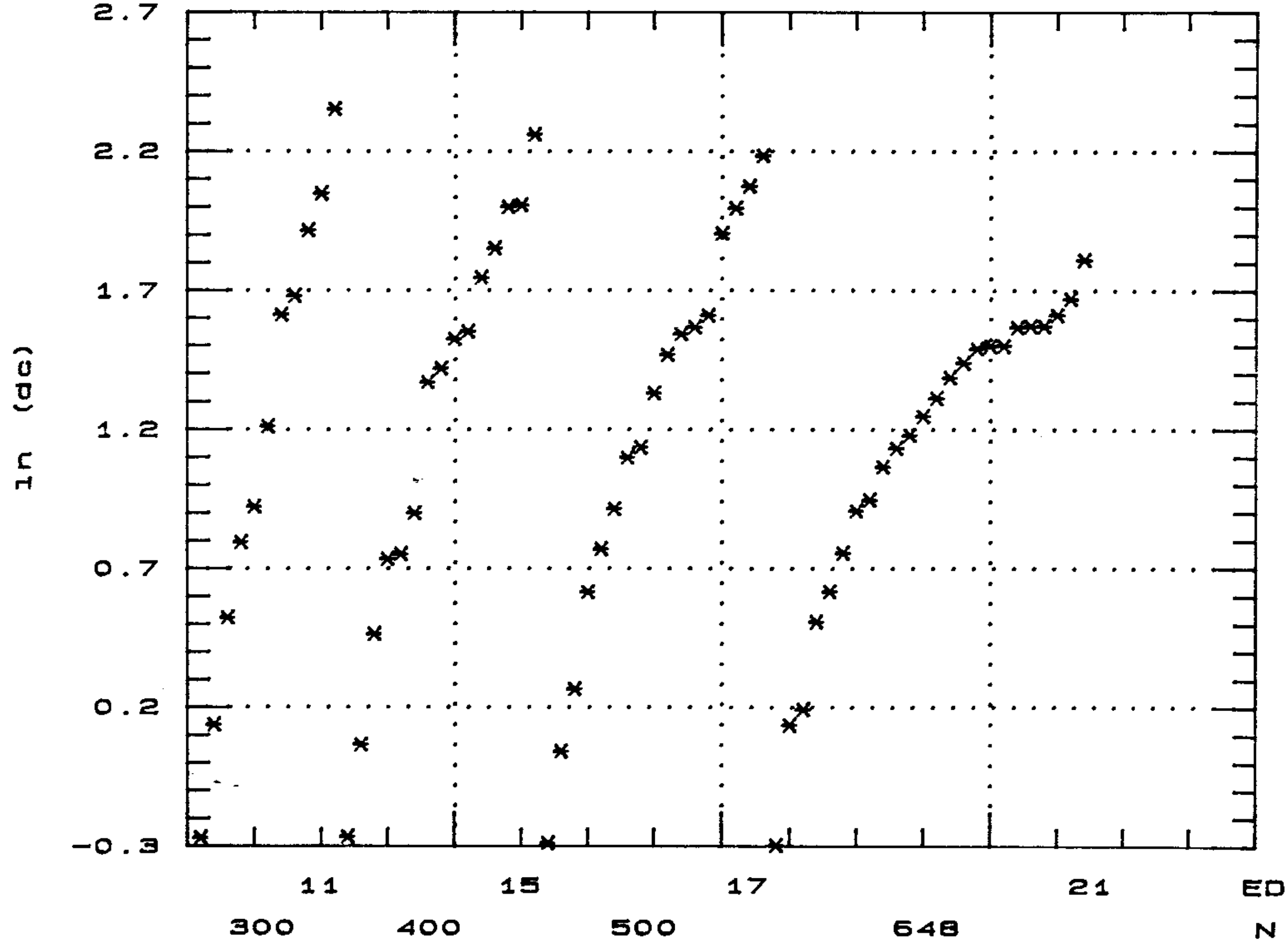


Figure S.4

Plot of ln (dc) on N and ED for Randomized Sayers Work Stoppage Data

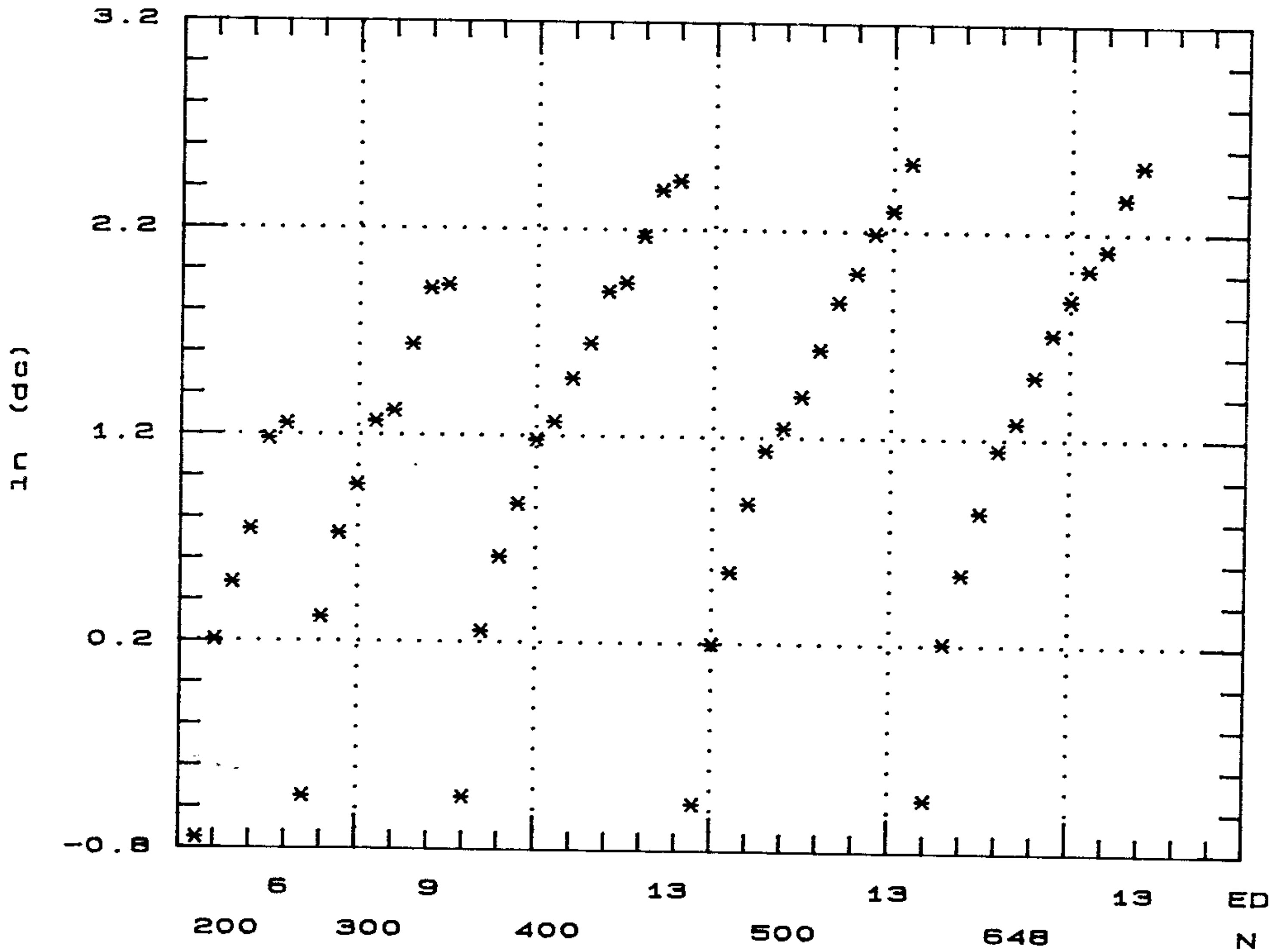


Figure Sc.1a

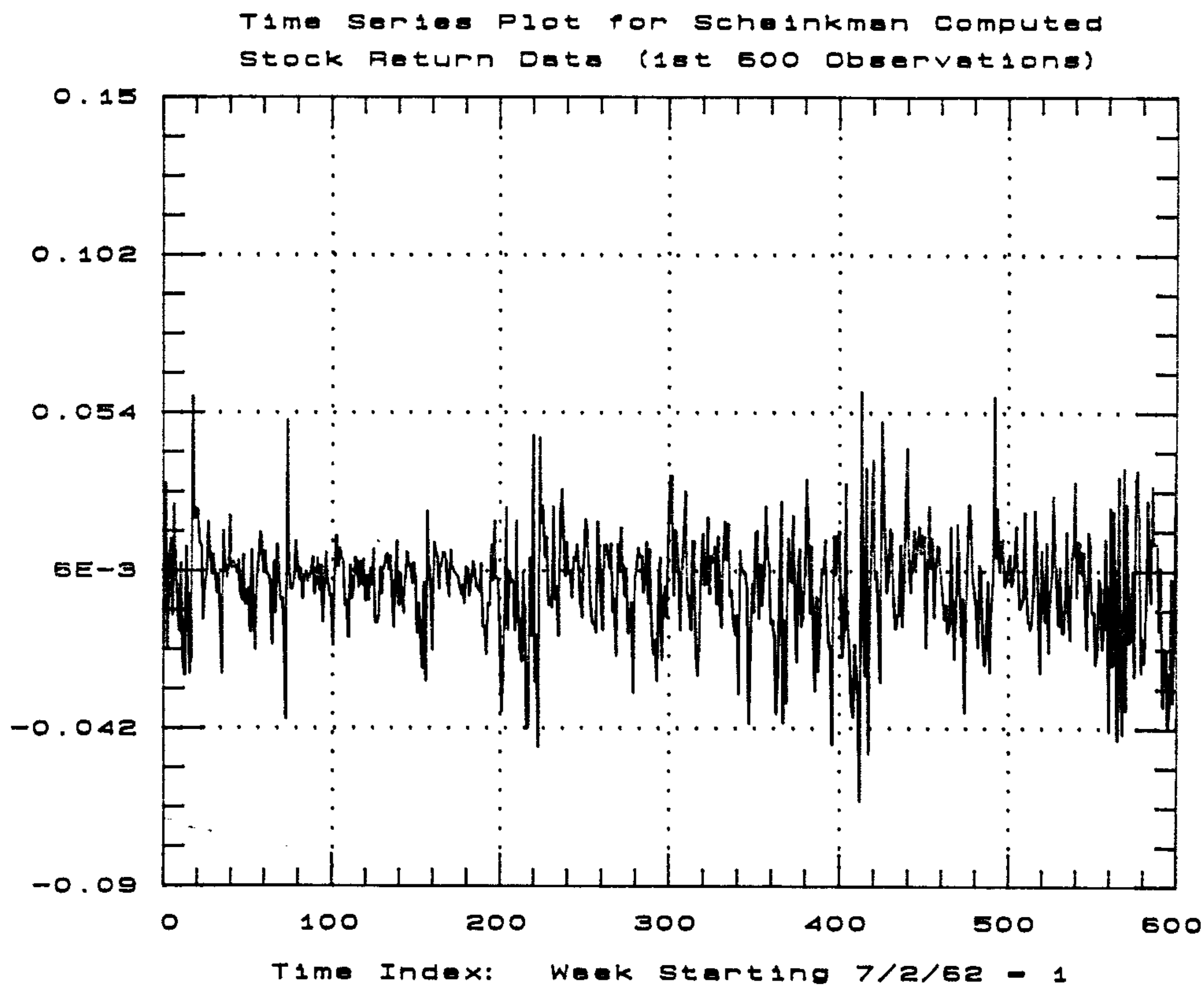

Time Series Plot for Scheinkman Computed
Stock Return Data (1st 600 Observations)
0.15
0.102
0.054
6E-3
-0.042
-0.09
0
100
200
300
400
500
600
Time Index: Week Starting 7/2/62 = 1

Figure Sc.1b

Time Series Plot for Scheinkman Computed
Stock Return Data (Obs. 601-1227)

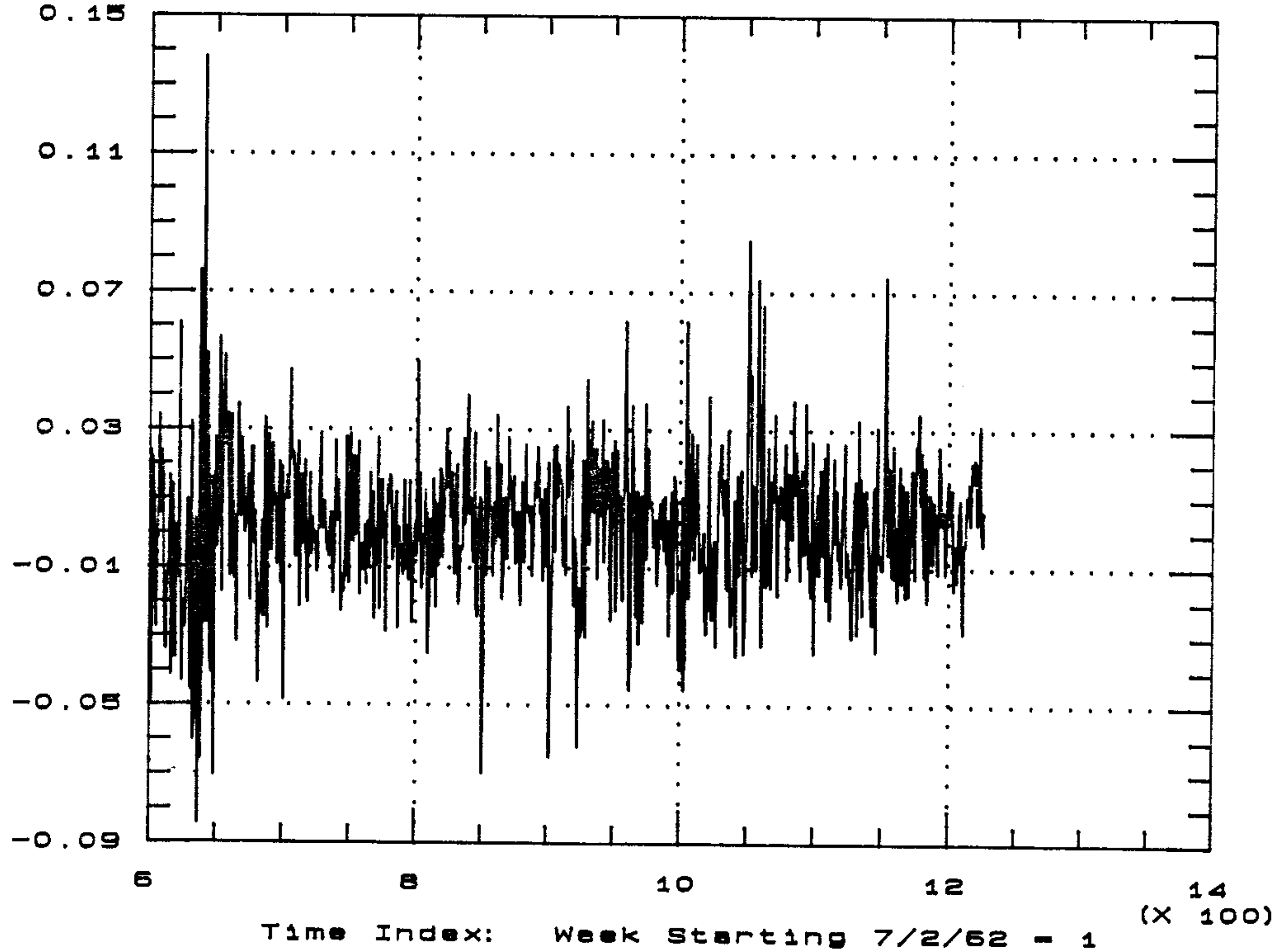
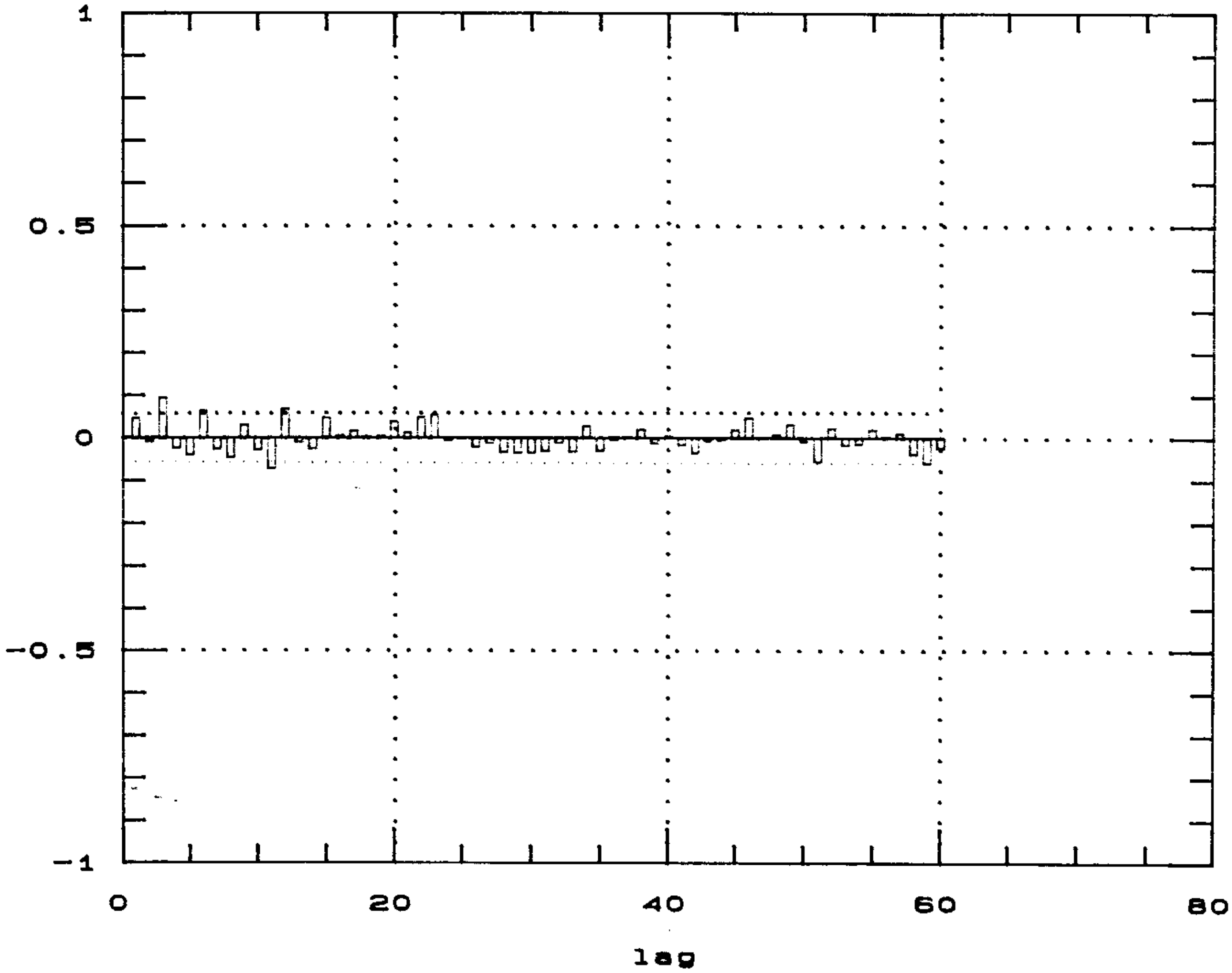


Figure Sc.2

Estimated Autocorrelations For
Scheinkman Computed Stock Return Data



<u>Figure Sc.3</u>

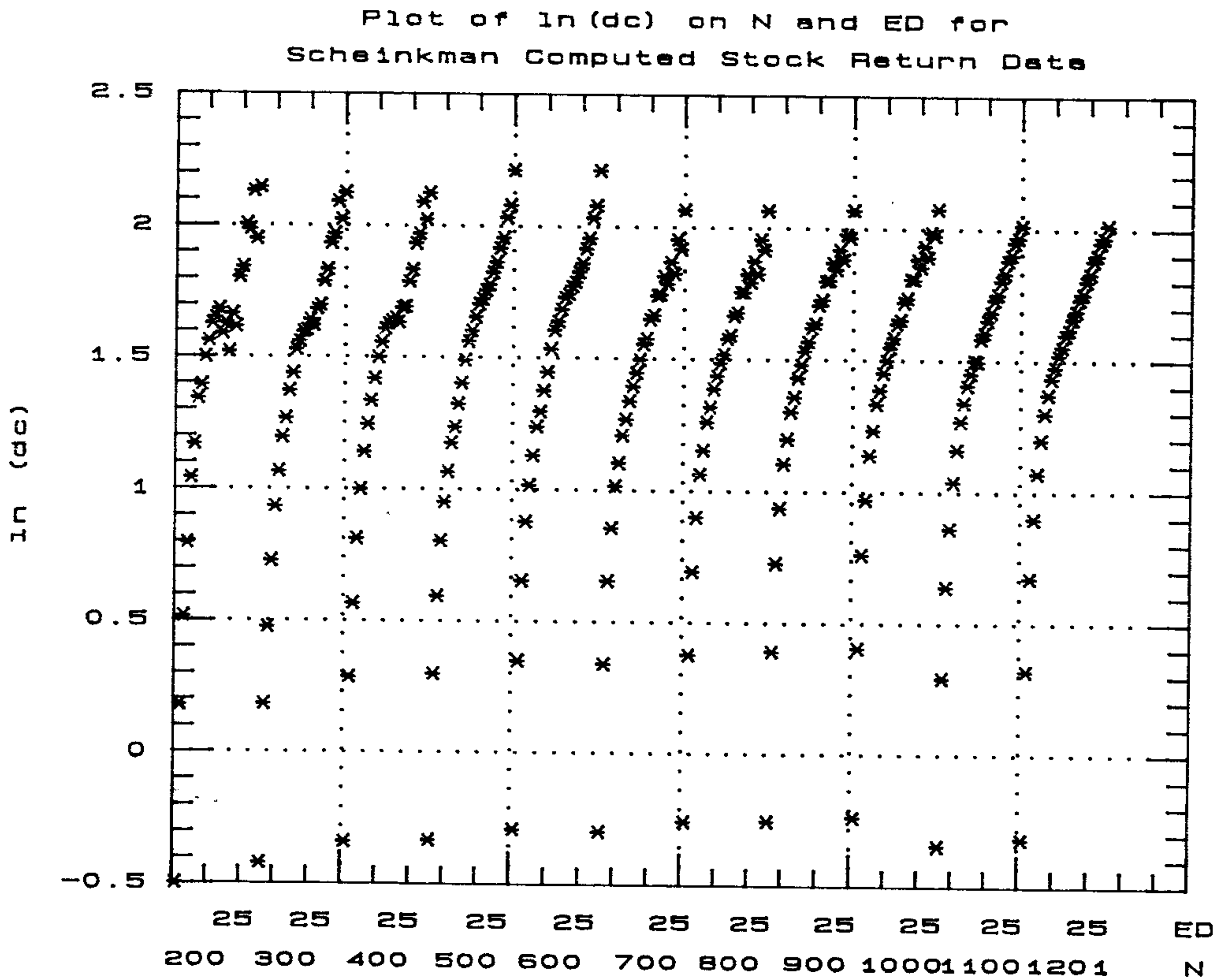


Figure Sc.4

Plot of ln (dc) on N and ED for Randomized Scheinkman Stock Return Data

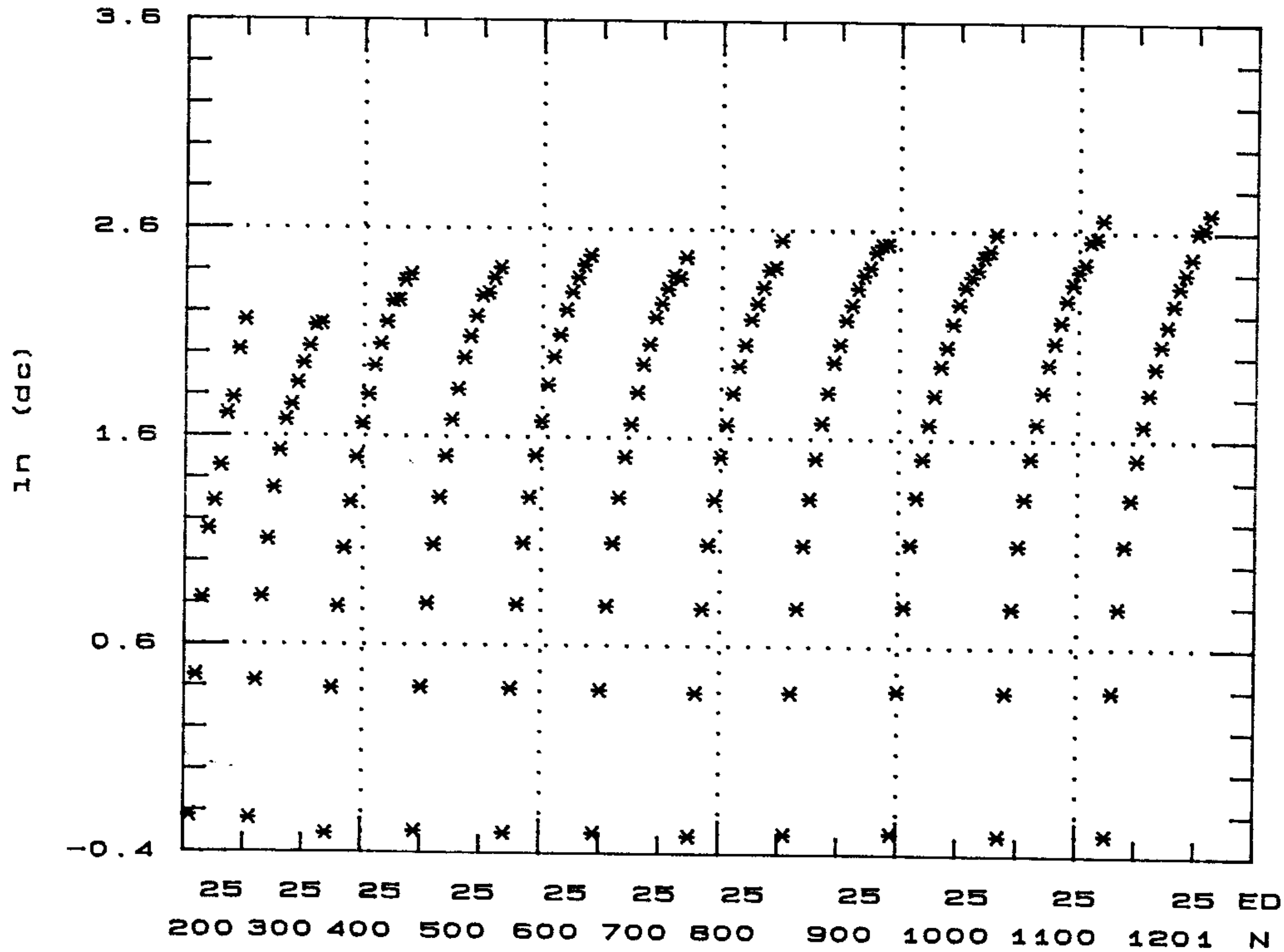


Figure B.1

Time Series Plot for Filtered
Barnett Demand Divisia M2 Data

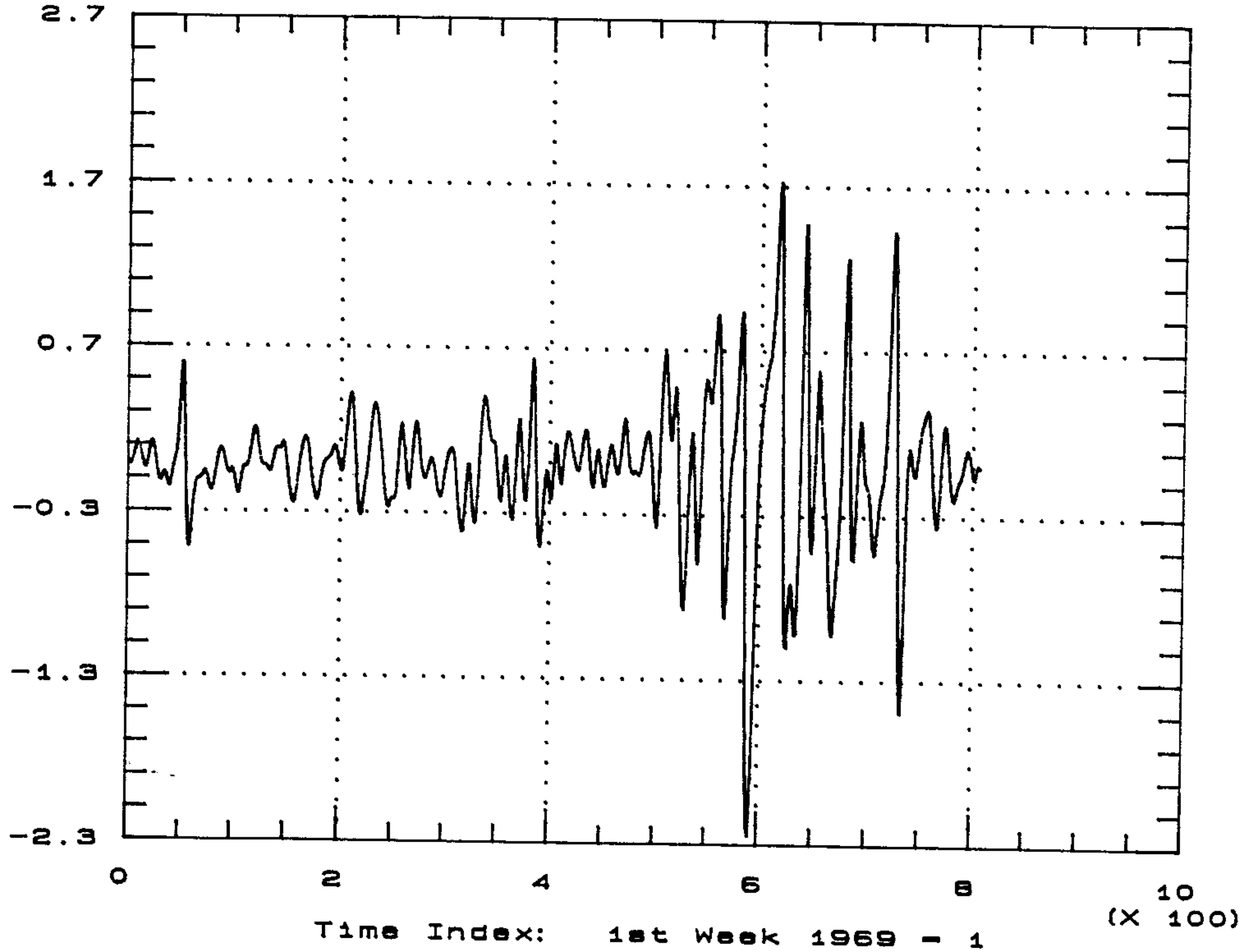


Figure B.2

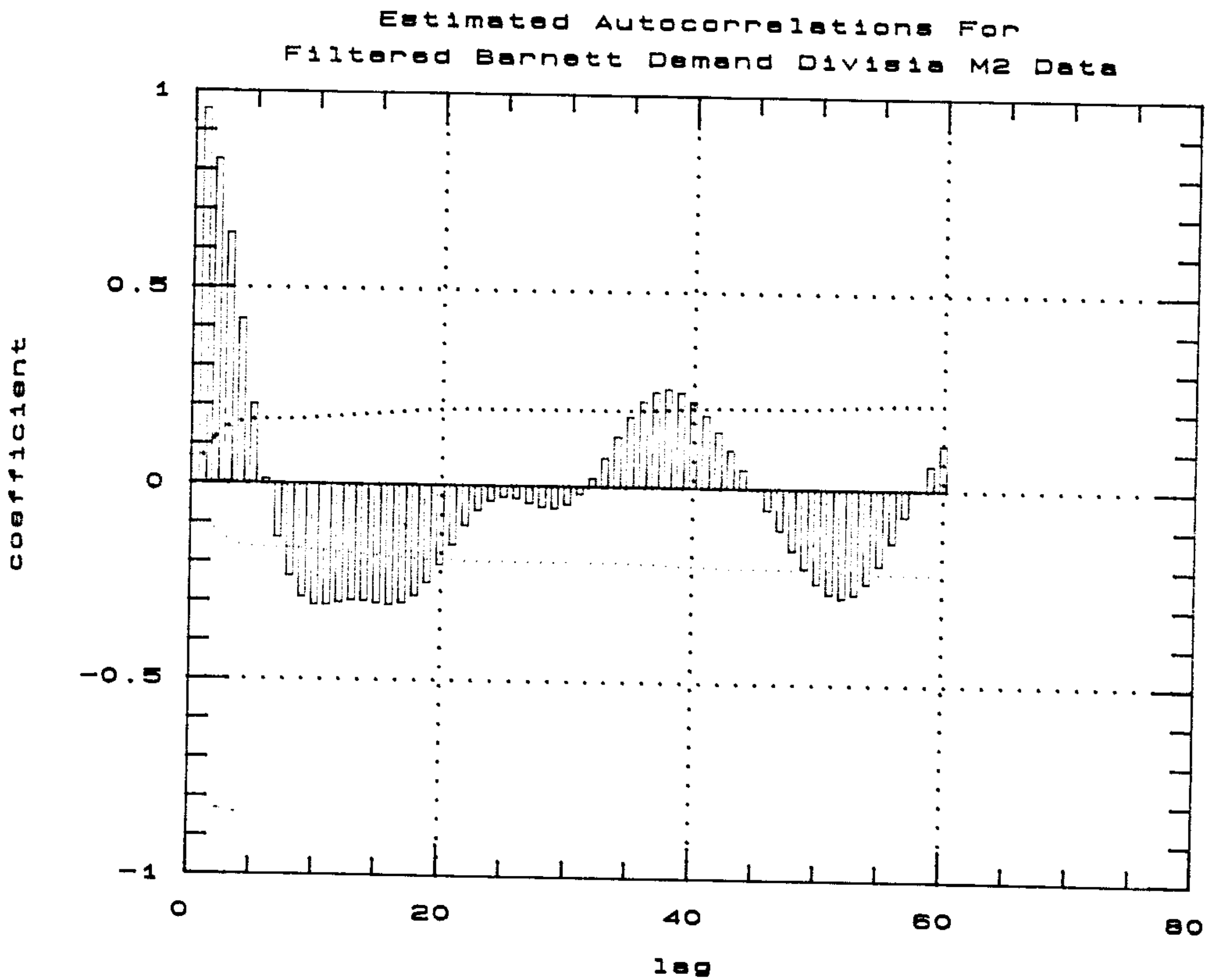

Estimated Autocorrelations For
Filtered Barnett Demand Divisia M2 Data
1
0.5
0
-0.5
-1
coefficient
0
20
40
60
80
lag

Figure B.3

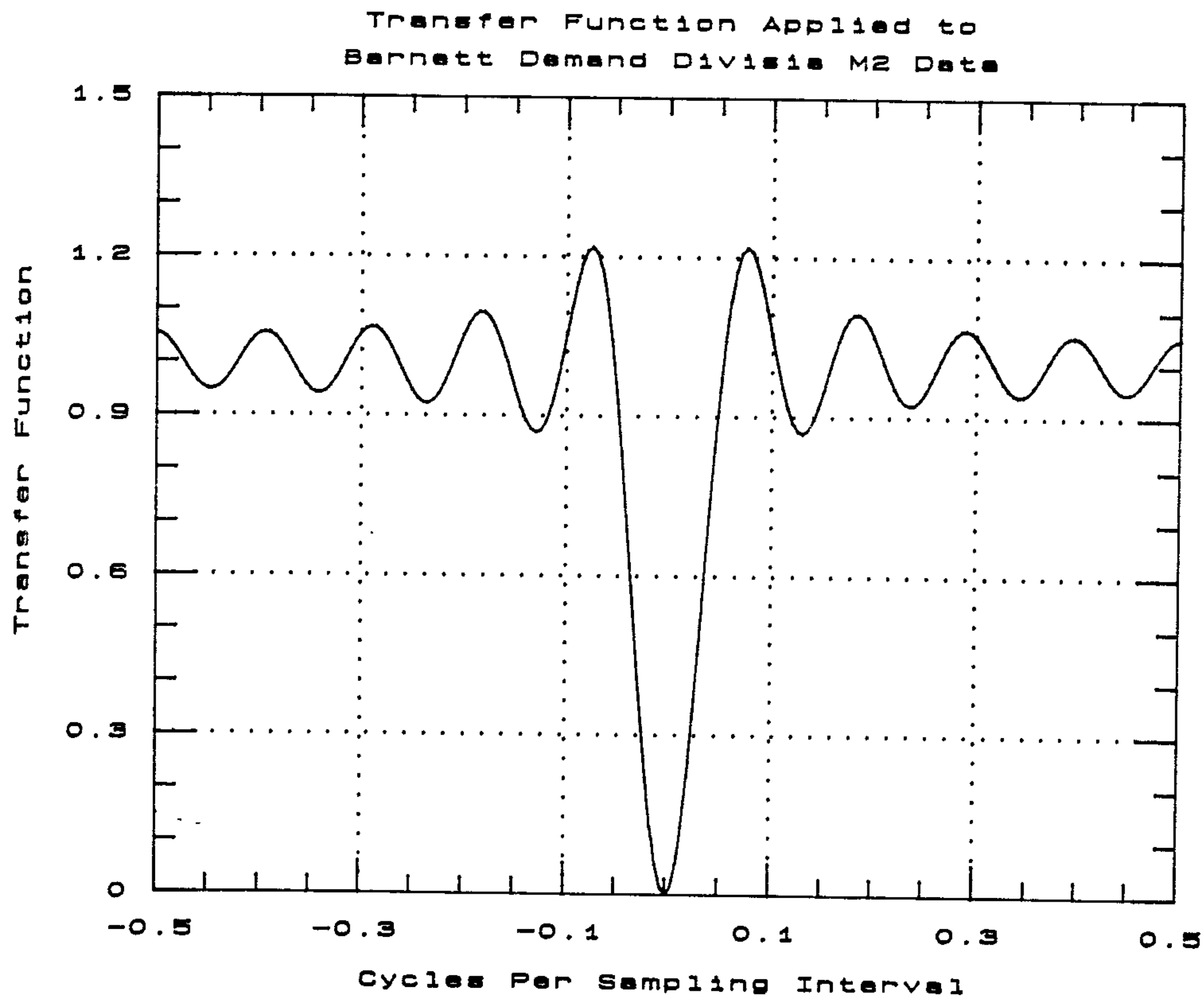

Transfer Function Applied to
Barnett Demand Divisia M2 Data
1.5
1.2
0.9
0.6
0.3
0
Transfer Function
-0.5
-0.3
-0.1
0.1
0.3
0.5
Cycles Per Sampling Interval

Figure B.4

Plot of ln (dc) on N and ED on Filtered Barnett Demand Divisia M2 Data

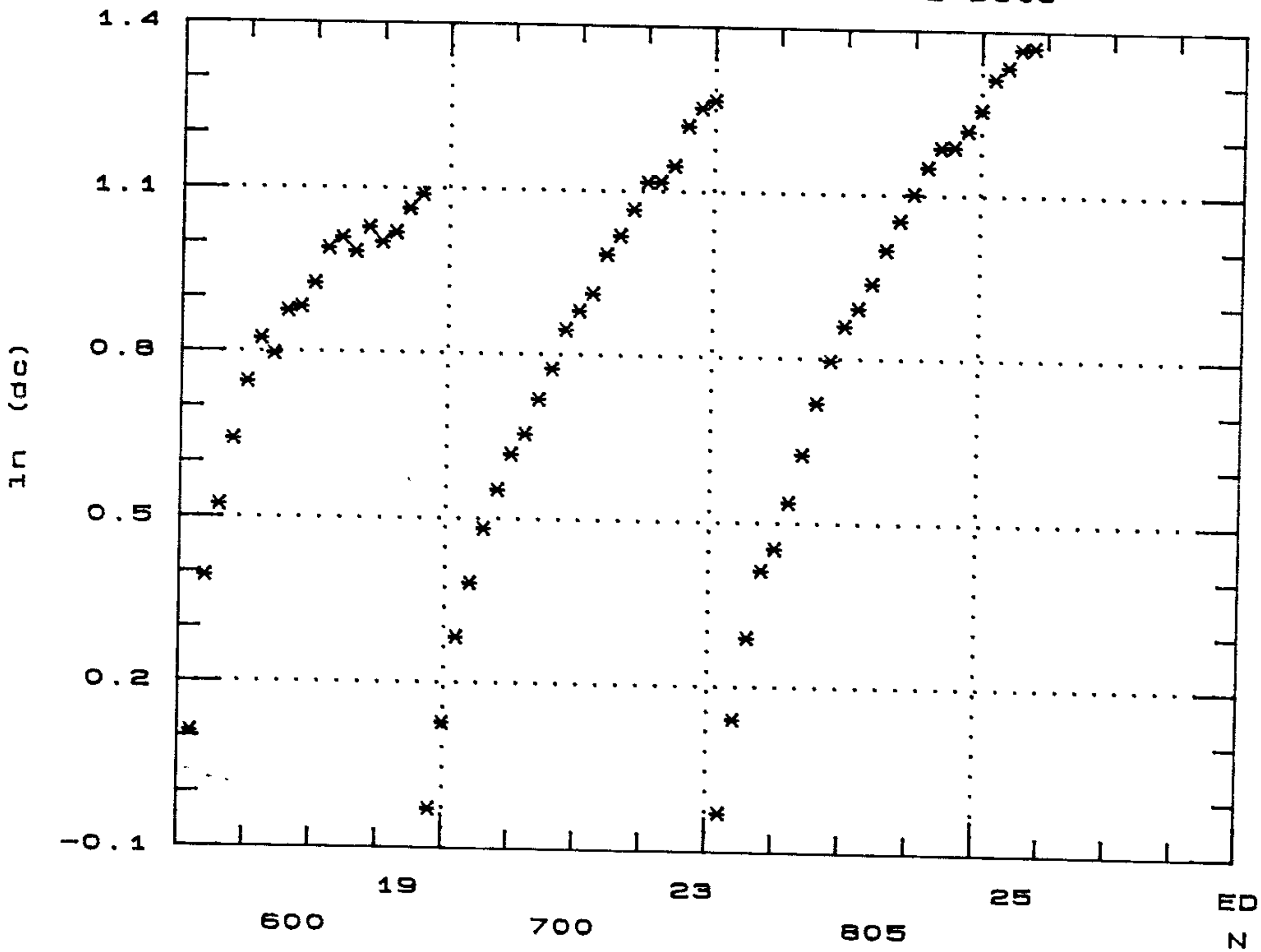
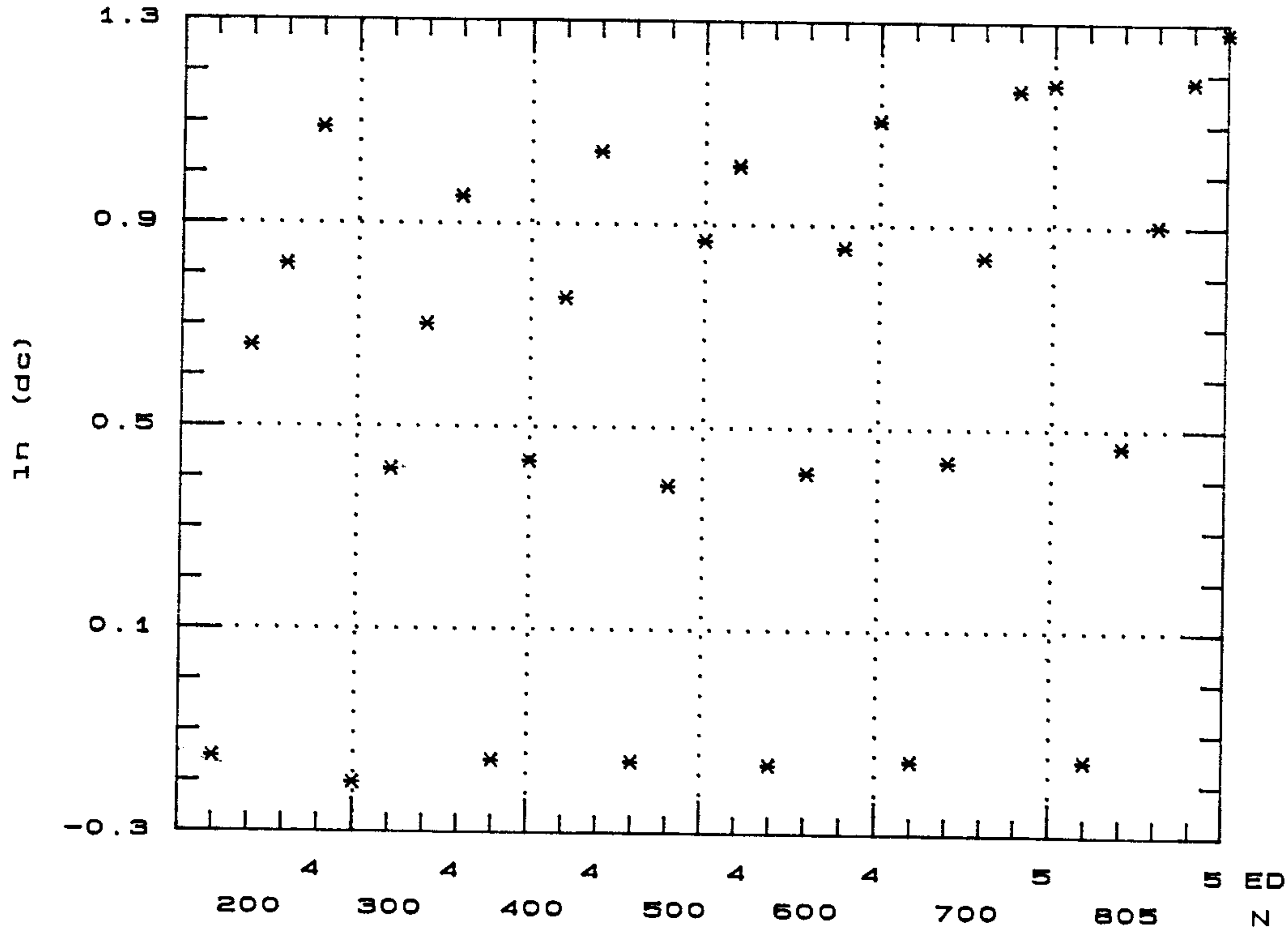


Figure B.5

ln (dc) on N and ED for Randomized Filtered Barnett Demand Divisia M2 Data



REFERENCES

Abraham, N.B., et al. (1986), Calculating the Dimension of Attractors from Small Data Sets," Physics Letters, 114A, 217-221.

Abraham, N.B., et al. (1986), "Definitions of Chaos and Measuring its Characteristics," forthcoming, Optical Chaos, edited by J. Chrostowski and N.B. Abraham, SPIE Proceedings, 667.

Albano, A.M., et al. (1987), "Data Requirements for Reliable EstimatiΔn of Correlton Dimensions," in0Chaotic Biological Systems, edited by A.V. Holden Pergammon Press.

Babloyantz, A. (1986), "Evidence of Chaotic Dynamics of Brain Activity During the Sleep Cycle," in Dimensions and Entropies in Chaotic Systems, Quantification of Complex Behavior, edited by G. Mayer-Kress (Springer-Verlag, Berlin Heidelberg).

Badii, R. and A. Politi (1985), "Statistical Description of Chaotic Attractors: The Dimension Function," J. Stat. Phys., 40.

Barnett, W. and P. Chen (1986a), "The Aggregation-Theoretic Monetary Aggregates are Chaotic and Have Strange Attractors," forthcoming in Barnett, W., E. Berndt and H. White, (eds.) Dynamic Econometric Modelling, Proceedings of the Third Austin Symposium in Economics, Cambridge University Press, Cambridge.

Barnett, W. and P. Chen (1986b), "Determinstic Chaos and Fractal Attractors as Tools for Nonparametric Dynamical Econometric Inference: With a Successful Application to the Divisia Monetary Aggregates," CER Discussion Paper 86-16, University of Texas-Austin.

Baumol, W.J. and J. Benhabib (1987), "Chaos: Significance, Mechanism, and Economic Applications," R.R. *87-16, C.V. Starr Center for Applied Economics, New York Univesity.

Ben-Mizrachi, A., et al. (1984), "Characterization of Experimental (noisy) Strange Attractors," Physical Review A, 29, p. 975.

Brandstater, A. and H.L. Swinney (1986), "A Strange Attractor in Weakly Turbulent Couette-Taylor Flow," University of Texas, Department of Physics and the Center for Nonlinear Dynamics.

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

Brock, W.A. (1987), "Introduction to Chaos and Other Aspects of Nonlinearity," Department of Economics, University of Wisconsin, forthcoming in Brock, W.A. and A.G. Malliaris (1988), Differential Equations, Stability, and Chaos in Dynamic Economics, North Holland, New York.

Brock, W.A., W.D. Dechert and J.A. Scheinkman (1987): "A Test for Independence Based on the Correlation Dimension," SSRI WP #8702, Department of Economics, University of Wisconsin.

Brock, W.A., and C. Sayers (1986), "Is the Business Cycle Characterized by Deterministic Chaos?," SSRI WP #8617, Department of Economics, University of Wisconsin, forthcoming, Journal of Monetary Economics.

Caputo, J.G., et al. (1986), "Determination of Attractor Dimension and Entropy for Various Flows: An Experimentalist's Viewpoint," in Dimensions and Entropies in Chaotic Systems, Quantification of Complex Behavior, edited by G. Mayer-Kress (Springer-verlag, Berlin Heidelberg).

Fayyad, Salam (1986), "Monetary Asset Component Grouping and Aggregation: An Inquiry into the Definition of Money," Doctoral Dissertation, University of Texas at Austin.

Frank, M. and T. Stengos (1987a), "Measuring the Strangeness of Gold and Silver Rates of Return," Department of Economics, University of Guelph.

Frank, M. and T. Stengos (1987b), "Some Evidence Concerning Macroeconomic Chaos," Department of Economics, University of Guelph.

Fraser, A.M.(1986), "Using Mutual Information to Estimate Metric Entropy",in Dimensions and Entropies in Chaotic Systems, edt. by Gottfried Mayer-Kress, Springer-Verlag, New York

Gennottee, G. and T. Marsh (1976), "Variations in Ex ante Risk Premiums on Capital Assets, Berkely Business School, University of California.

Grassberger, P. (1985), "Generalizations of the Hausdorff Dimension of Fractal Measures," Physics Letters, 107A.

Grassberger, p. and I. Procaccia (1983a), "Characterization of Strange Attractors," Physical Review Letters, 50, 346-349.

Grassberger, P. and I. Procaccia (1983b), "Estimation of the Kolmogorov Entropy from a Chaotic Signal," Physical Review, 28.

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

Grebogi, C., et al. (1984), "Strange Attractors that are not Chaotic," Physica, 13D, 261.

Greenside, H.S., et al. (1982), "Impracticality of a Box-counting Algorithm for Calculating the Dimensionality of Strange Attractors," Physical Review A, 25.

Halsey, T.C. et al. (1986), "Fractal Measures and their Singularities: The Characterization of Strange Sets," Physical Review A, 33, No. 2.

Holzfuss, J. and G. Mayer-Kress (1986), "An Approach to Error-free Estimation in the Appliction of Dimension Algorithms," in Dimensions and Entropies in Chaotic Systems, Quantification of Complex Behavior, edited by G. Mayer-Kress (Springer-Verlag, Berlin Heidelberg).

Hsieh, D. (1987), "Testing for Nonlinear Dependence in Foreign Exchange Rates: 1974-1983," Graduate School of Business, University of Chicago.

Kurths, J. and H. Herzel (1987), "An Attractor in a Solar Time Series," Physica 25D, 165-172.

Lovejoy, S. and D. Schertzer (1986), "Scale Invariance, Symmetries, Fractals, and Stochastic Simulations of Atmospheric Phenomena," Bull. Amer. Meteor. Soc., 67.

Priestly, M.B. (1981), Spectral Analysis and Time Series, Volumes I and II, Academic Press, New York.

Ramsey, J.B. and H. Yuan (1987), "The Statistical Properties of Dimension Calculations Using Small Data Sets," R.R. #87-20, C.V. Starr Center for Applied Economics, New York University.

Sayers, C.L. (1987), "Diagnostic Tests for Nonlinearity in Time Series Data: An Application to the Work Stoppages Series," Department of Economics, University of North Carolina - Chapel Hill.

Sayers, C.L. (1986), "Work Stoppages: Exploring the Nonlinear Dynamics," University of Winsconsin.

Scheinkman, J.A. and B. Le Baron (1987), "Nonlinear Dynamics and GNP Data," Department of Economics, University of Chicago, forthcoming, Journal Letters, 114A.

Wolf, A., et al. (1985), "Determining Lyapunov Exponents from a Time Series," Physica, 16D, 285-317.