DISTRIBUTIONS OF CORRELATION COEFFICIENTS IN ECONOMIC TIME SERIES

EDWARD AMES AND STANLEY REITER*
Purdue University

This paper presents results, mainly in tabular form, of a sampling experiment in which 100 economic time series 25 years long were drawn at random from the *Historical Statistics for the United States*. Sampling distributions of coefficients of correlation and autocorrelation were computed using these series, and their logarithms, with and without correction for linear trend. We find that the frequency distribution of autocorrelation coefficients has the following properties:

(a) It is roughly invariant under logarithmic transformation of data.
(b) It is approximated by a Pearson Type XII function.
(c) It approaches a rectangular distribution symmetric about 0 as the lag increases.

The autocorrelation properties observed are not to be explained by linear trends alone.

Correlations and lagged cross-correlations are quite high for all classes of data. E.g., given a randomly selected series, it is possible to find, by random drawing, another series which explains at least 50 per cent of the variances of the first one, in from 2 to 6 random trials, depending on the class of data involved. The sampling distributions obtained provide a basis for tests of significance of correlations of economic time series. We also find that our economic series are well described by exact linear difference equations of low order.

1. INTRODUCTION

SINCE 1913, when H. L. Moore first undertook to measure the elasticity of demand for agricultural commodities, economists have been trying to extract information about economic relationships from time series data. Since that time, a variety of statistical models have been employed to analyze time series data. Despite the considerable sophistication of the available statistical models of time series, we cannot yet claim to know the class of statistical models appropriate to economic time series. An accumulation of systematic information about statistical properties of economic data would improve our ability to decide the relevance of alternative statistical models, and suggest the lines of development of new models, if that should seem desirable. This paper is a step toward systematizing information regarding statistical properties of observed economic time series.

This paper presents frequency distributions of coefficients of correlation and autocorrelation computed from annual economic data. We chose the correlation properties of economic time series for our first step, because of the relevance of these statistics for current theoretical and empirical work. Regression models are perhaps those most widely used for estimating economic relationships. In these models, the correlation coefficients obtained indicate the amount of association observed. The significance of such association is usually evalu-

* This work has been in part supported by the Office of Naval Research.

ated by computing the probability of a correlation as large or larger than the observed correlation on the hypothesis of no more than chance association. The significance test appropriate for such applications should be based on the probability distribution of correlations among randomly selected economic variables rather than on the distribution of correlation coefficients among variables from another uncorrelated population. Thus, the frequency distributions of correlation coefficients presented in this paper provide a more appropriate basis for tests of significance of correlations among economic variables than does the distribution of the correlation coefficient under the usual null hypothesis of statistical theory that the expected correlation is zero.

It is known that estimates of the coefficients in linear autoregressive models may be obtained by correlation techniques; and that properties of such models may be formulated in terms of the roots of their characteristic polynomial. Hence, a knowledge of the empirical distribution of autocorrelation coefficients points the way to a study of autoregressive structures which represent economic processes. In this sense, our sytematization of this knowledge could assist ultimately in the selection of stochastic models suitable to describe economic processes.

2. DESCRIPTION OF THE DATA USED

We took a random sample of 100 economic time series from the *Historical Statistics of the United States*,[1] using the RAND random digits to select them. The volume contains 2,994 columns of data, of which roughly half are annual series covering the period 1929–1945. These we extended through 1953, using the annual *Statistical Abstract*. We thus accepted the definition of "economic data" used by the Census Bureau. This hundred series for the 25-year period 1929–1953 form the raw material for our work.

Each series was then considered in four possible representations:

1. In the units originally measured, called natural numbers;
2. As residuals about a linear regression function $x=a+bt$, where t, the integers 1, · · · , 25, stands for "time," i.e., "corrected for trend";
3. As logarithms;[2] and
4. As residuals about a linear regression function $\log_{10} x=a+bt$, where t is as before, i.e., "logarithms, corrected for trend."

The use of "trend-corrected data" requires comment. There is a view in the economics profession that economic processes may be additively separable into "long-term" processes, which depend upon slowly-changing factors such as population, technology, resource endowment, etc., and "short-term" processes, including primarily the factors loosely describable as "the business cycle."[3] A part of the economic literature associates discussion of "the business cycle" with trend-corrected empirical data, and for this reason, we have felt it would be instructive to include "trend-corrected" data in this study.

---

[1] U. S. Bureau of the Census, Washington, 1946.

[2] Two of the series chosen contained negative numbers, and had to be replaced by other series for the third and fourth representations.

[3] This problem is discussed by Edward Ames, "A Theoretical and Statistical Dilemma," *Econometrica*, 1948.

3. DESCRIPTION OF THE CALCULATIONS

Each of these representations was then processed as follows:

(1) The coefficient of autocorrelation was computed for each series, for lags of 1, 2, 3, 4, 5 years.

(2) Successive pairs of series were correlated (in the random order of selection); the second series in each pair was treated as the independent variable in a linear regression equation, and lagged 0, 1, 2, 3, 4, 5, years behind the first (dependent) variable.

(3) Since these various coefficients involved different degrees of freedom, all of the coefficients were made comparable by means of the transformation

$$\bar{r} = \left[1 - (1 - r^2)\frac{n-1}{n-2}\right]^{1/2}$$

where

$\bar{r}$ is the corrected correlation coefficient;
r is the uncorrected correlation coefficient;
$n = 25 - L$, where L is the lag, and 25 the number of years for which data were collected.

(4) In the process of deriving the trend-corrected data, correlation coefficients of the natural numbers and the logarithms with "time" (the integers 1, 2, · · · , 25) were computed.

(5) Frequency distributions of the corrected coefficients of correlation and autocorrelation were prepared, for each lag and each representation of the data.

(6) The mean; the second, third, and fourth moments about the mean; and the coefficients β_1 and β_2[4] were computed for each lag and each representation of the data. These were computed from the individual observations and not from the frequency distributions, so that no corrections for grouping were required.

4. SUMMARY OF THE FINDINGS

The frequency distributions resulting from our calculations, and moments of these distributions, are given in the Appendix. We may summarize these tables as follows:

4.1 *Autocorrelation* is clearly very great in unadjusted economic series (both natural numbers and logarithms). Some 70 per cent of the variance of such data is "accounted for" on the average, if we correlate any series with itself, with lag of one year (Tables A-I, A-III). The same series, correlated with the integers 1, 2, · · · , 25 give U-shaped distributions of correlation coefficients, with modes around +.90 and −.70 (Table B-I). Although the average coefficient of autocorrelation declines as the lag increases, it is still .48 for a lag of 5 years. A series having autocorrelation close to +1 for lag of one year will be approximated closely by a straight line. Likewise, a series correlated with "time" will tend to increase by a constant amount each

[4] If μ_i is the ith moment about the mean, $\beta_1 = \mu_3^2/\mu_2^3$ and $\beta_2 = \mu_4\mu_2^2$.

year, i.e., its value in any year will roughly equal the value of the preceding year plus a constant. In this sense, trend and autocorrelation analysis measure the same thing.

To determine other properties of this distribution, we made use of the coefficients β_1 and β_2 and the Pearson system of curve-fitting.[5] Both β_1 and β_2 decrease as the lag increases. Their values are (a) rather similar for the distribution of the natural numbers and of the logarithms; and (b) with one exception,[6] all satisfy the condition

$$1.125\beta_1 + 1.5 < \beta_2 < 1.250\beta_1 + 2.0.$$

Distributions meeting this requirement are of a simple, "Type XII" form,

$$f(x) = \frac{1}{b\Gamma(1+m)\Gamma(1-m)}\left[\frac{a+x}{b-x}\right]^m$$

where the constants, a, b, and m are calculated from the moments. In particular, when $d=0$, this distribution is the rectangular distribution; if $\beta_1=0$, then $m=0$; and it will be noted that, within our range of observation, β_1 declines from 5.2 to .3 for the logarithms, and from 3.9 to 1.4 for the natural numbers. Calculated values of the functions are given in Table 1.

Thus, our calculations indicate that the autocorrelation properties of economic time series may be described by a frequency distribution which

(i) is roughly invariant when the data are subjected to a logarithmic transformation;
(ii) is approximated by a Pearson Type XII function; and
(iii) approaches a rectangular distribution with mean 0, as the lag increases.

Economic series, whether in natural numbers or logarithms, have autocorrelation properties more complicated than can be explained by the simple linear trend model. If, apart from "trend," there were no tendency for an economic series to be "influenced by its own past," the trend-corrected data would show small autocorrelations. Actually, even these show average autocorrelations of .65 and .68 for lag 1. Moreover, although these averages decline rapidly toward zero, the possibility is not excluded that, over a range of lags longer than 5, the average might be a periodic function of the lag (perhaps damped), as is the case with some autocorrelation functions observable in the natural sciences.

A large body of empirical work by the National Bureau of Economic Research has pointed to a persistent short-term fluctuation three to four years in length, on the average, for most economic series. Such a fluctuation does not appear to affect our autocorrelation coefficients, since nothing like a corresponding period appears in them. If the average autocorrelation does have a period, its length would appear to be substantially greater than the period which most writers have ascribed to business fluctuations,

The distributions obtained for the trend-corrected data have fairly constant

[5] William P. Elderton, *Frequency Curves and Correlation*, Fourth Edition. Washington, D. C. 1953, provides a useful collection of methods and criteria.

[6] Natural numbers, lag 1, fail by a little to meet this requirement.

TABLE 1. FITTED FREQUENCY FUNCTIONS OF COEFFICIENTS OF AUTOCORRELATION IN 100 RANDOMLY SELECTED ECONOMIC SERIES*

Lag in Years	Autocorrelation in Series** Expressed as	
	Natural Numbers	Logarithms
1	$\frac{1}{.885}\left[\frac{.091+r}{.968-r}\right]^{.754}$	$\frac{1}{1.362}\left[\frac{.410+r}{.962-r}\right]^{.797}$
2	$\frac{1}{1.071}\left[\frac{.287+r}{.963-r}\right]^{.617}$	$\frac{1}{1.062}\left[\frac{.389+r}{.947-r}\right]^{.614}$
3	$\frac{1}{1.148}\left[\frac{.495+r}{.966-r}\right]^{.498}$	$\frac{1}{1.046}\left[\frac{.351+r}{.970-r}\right]^{.408}$
4	$\frac{1}{1.211}\left[\frac{.572+r}{.955-r}\right]^{.466}$	$\frac{1}{1.230}\left[\frac{.581+r}{.963-r}\right]^{.388}$
5	$\frac{1}{1.023}\left[\frac{.615+r}{.959-r}\right]^{.355}$	$\frac{1}{1.214}\left[\frac{.550+r}{.948-r}\right]^{.318}$

* The constants in these functions give bounds to the values of r for which non-zero frequencies exist. Thus, for natural numbers, lag 1, $-.091 < r < .968$. The calculations follow Elderton, *op. cit.*, p. 111.

** Not corrected for trend.

variance for all lags, and (for lags greater than 1) are almost symmetrical. They do not seem well approximated by the normal distribution, since β_2 (which in the case of the normal distribution equals 3) is as low as 1.76 (roughly a rectangular distribution) and as high as 4.8, with no great concentration of values near 3.0. It does not appear reasonable to consider that the mean of the values of any moment for the various lags has any particular meaning, although if a moment fell in a particular small range for all lags, it might lead us to explore this element of similarity in the distributions. No great stability in β_2 can be indicated, however, and we cannot now say much about the distribution of autocorrelation coefficients in *trend-corrected* economic data, except that

(i) The autocorrelations from the trend-corrected data have distributions unlike those from uncorrected data. The distributions obtained from the trend-corrected natural numbers and from trend-corrected logarithms appear to be rather similar.

(ii) For lag 1, the mean is substantially above zero, but declines rapidly to zero as the lag increases. The mean autocorrelation might be a periodic function of the lag, but we have not computed enough lags to be sure.

(iii) The variance of the distributions is roughly constant, and (for lags greater than 1) the distributions are almost symmetrical.

(iv) Other features of the distribution do not seem to us to be very regular.

4.2 *Serial correlation coefficients*, like autocorrelation coefficients, are distributed in roughly the same way for logarithms as for natural numbers. Like the distributions of autocorrelation coefficients, they differ according to whether original data or deviations from trend are considered. All distributions are close to symmetrical: in no case is β_1 over .167. Moreover, all the means are small: for the uncorrected data, the means are between .10 and .16, and for the corrected data, they are in most cases between $\pm .1$ and in no case as high as .16. Variances tend to be smaller for larger lags (in autocorrelations, they were constant or increased with the lag), and are several times as large as the autocorrelation variances for the same lag.

Trend-correction makes for one important difference in the serial correlation distributions. In the uncorrected data, β_2 is almost always less than 1.8;[7] in the corrected data, β_2 ranges from 1.9 to 3.2. It will be recalled that the Pearson distribution corresponding to $\beta_1=0$, $\beta_2=1.8$ is rectangular. In symmetrical ($\beta_1=0$) Pearson distributions, those for which $\beta_2<1.8$ are bimodal, and those for which $\beta_1>1.8$ are unimodal. Thus, it would be correct to say that for the uncorrected data, the serial correlation coefficient is almost rectangularly distributed for all lags; but where it deviates from the rectangular, it is in the direction of bimodality rather than normality.

On the other hand, serial correlation coefficients from corrected data tend to have a unimodal distribution. If β_2 were symmetrically distributed about 3.0, one might be tempted to use the normal distribution as an approximation to these distributions. Only in one case out of twelve, however, does β_2 exceed 3.0, and in nine, β_2 is below 2.6. This means that, within any given interval about the mean (expressed in multiples of the standard deviation), there would be *fewer* observations in these distributions than in a normal distribution with the same standard deviation. Thus, a significance test, assuming a normal distribution whose mean and variance were given by the present estimates, would lead to overestimation of the significance of any serial correlation coefficient actually obtained.

Two tentative experiments were carried out on the serial correlation coefficients in an early stage of our work. These are reported here despite their inexactness,[8] because the gross findings are of some interest. Suppose that it is posited, on theoretical grounds, that variable x_1 may influence x_2, and it is considered possible that the effect of x_1 may be felt after some delay. We might calculate serial correlation coefficients for these variables for lags 0, 1, 2, etc., years. Having selected the lag for which the coefficient is maximized, we obtain (a) an estimate of the closeness of the relationship, and (b) an extimate of the speed with which the independent variable affects the dependent variable. This sort of procedure is, in fact, sometimes followed in empirical work.

The results we obtained by similar operations on our randomly selected data are given in Tables D-I, D-II, and D-III. Two different hypotheses were used: (1) The cause must not succeed its effect; and (2) the cause must strictly precede its effect. For the first, we selected the maximum correlation coefficient between each pair of variables over lags 0, 1, 2, 3, 4, 5; for the second, the maximum over lags 1, 2, 3, 4, 5. Thus, for the first hypothesis, we note that for the natural numbers in 46 per cent of all cases ($|r|\geq .7$), it was possible to select a lag such that the "independent variable" would account for half or more of the

[7] It is 1.87 for the logarithms, lag 5.

[8] They used correlation coefficients which were *not* corrected for changes in the number of observations. Thus, two correlations of .5, one for lag 0 and one for lag 5 in the independent variables, were treated as equal, although the latter is actually smaller, being based on 20 rather than 25 observations. These experiments were not repeated after the correction, since the computer program used did not print out the data needed.

variance in the "dependent variable"; for the logarithms, we could do as well in 60 per cent of all cases.

In fact, if we use data not corrected for trend, on any of the following hypotheses:

(1) the independent variable has an immediate effect;
(2) the independent variable has an effect within 5 years;
(3) the independent variable has a delayed effect, but the delay is of not more than 5 years;

then we can, on the average, find an independent variable capable of "explaining" 50 per cent or more of the variance in *any* randomly selected economic series in three random trials or less (Table D-IV).[9] If we used data corrected for trend, it may take as long, on the average, as twelve tries; but if we omit the case with zero lag, we can succeed even here in four to six tries, on the average.

5. CONCLUSIONS

Our results do not yet provide very clear specifications for delimiting the class of statistical models appropriate for analysis of annual economic time series. They do, however, permit some comments about earlier work. In particular, we shall discuss some of the economic literature on linear stochastic difference equations.

Tinbergen,[10] Orcutt,[11] and Gartaganis[12] have made studies anticipating some of our findings. Like us, they have studied autocorrelation, but whereas we selected economic series at random, the first two writers used series suggested by the theory of business cycles, and the third used production series from Arthur Burns' classic study *Production Trends*. Thus, a portion of the differences between our results and theirs may result from their selection techniques, which prefer only a sub-population of economic series.

Second, all three writers have studied a short period (the first two studied 1919–32; the third, 1914–29), and the third studied a non-overlapping long period (1870–1913). The longer period shows quite different autocorrelation properties from the shorter periods, but it is not clear whether the differences result from the difference in the length of the period for which autocorrelations were studied, or whether they are due to the fact that non-overlapping periods have different autocorrelations. The evidence suggests that autocorrelation (for any lag) would increase with the number of observations, but we have not tested this conjecture.

Tinbergen's original contribution was to prepare a business cycle model involving a number of variables of theoretical interest. He showed that he could compute multiple regression curves which would closely approximate the time paths of the individual variables in the system (for the period 1919–32), and

[9] For example: 46 per cent of the pairs of variables (natural numbers not corrected for trend) showed a correlation with absolute value at least .7 for at least one lag in the set 0–5 (Table D-I). Thus, on the average, 1/.46, or slightly over two random tries, would be required to find as good a correlation as this.

[10] Tinbergen, "Statistical Testing of Business Cycle Theories," Vol. II, *Business Cycles in the United States of America, 1919–1932*, League of Nations, Geneva, 1939.

[11] Orcutt, "A Study of the Autoregressive Nature of the Time Series Used for Tinbergen's Model of the Economic System of the United States, 1919–1932," *Journal of the Royal Statistical Society*, Vol. X, Series B (1948), 1.

[12] Gartaganis, "Autoregression in the United States Economy, 1870–1929," *Econometrica*, 22 (1954), 228.

that correlation coefficients large enough to pass ordinary significance tests were generated. He also showed that he could, by a process of successively eliminating variables, end up with a single "explanatory" variable (corporate profits); that all the other variables in his system could be expressed as a function of corporate profits; and that corporate profits could be closely approximated by a fourth-order linear difference equation.

At the time Tinbergen wrote, less was known than now about the properties of such systems. A decade later, Orcutt undertook a comparison of the autocorrelation present in Tinbergen's series (for lags of 1–4 years) with that in a Monte Carlo experiment, using data generated by autoregressive functions containing a random term. He concluded that the autocorrelation in Tinbergen's series did not resemble the autocorrelation to be expected from fourth-order linear autoregressive equations, but instead they resembled drawings from a population of series generated by a single second-order linear stochastic difference equation, the elements of which had different non-homogeneous terms.

Our results differ from Orcutt's. The average autocorrelation in our data is consistently higher than in his (for lag of one year, his mean is .597, while ours (Table A-I) is .837; for lag 4, his mean is −.125, ours .525). Our variance and his are about the same for lag of one year (ours .040, his .048), but ours rises steadily with the lag (for lag 4, our variance is .158, his .035), and his does not.

Gartaganis is in part concerned with differences in the autocorrelation properties of production series belonging to different sectors of the economy and to different periods. He finds that there are, in general, notable differences between autocorrelation in 1870–1913 and in 1914–1929. Within each period, average correlation for a given lag does not differ greatly among agriculture, mining, and industry. Whereas the distributions in the first two sectors are homogeneous, in the last, they appear to be messy, and Gartaganis does not treat industry in as much detail as the other sectors.

It is, of course, clear that a series' autocorrelation in a subperiod will not necessarily equal the autocorrelation in an entire period; and that if a series has stable autocorrelation properties, it will still vary from one interval to the next for sampling reasons.

Our results show, on the average, less autocorrelation in the 25-year period 1929–53 than Gartaganis' in the 44-year period 1870–1913; They show more, on the average, than Gartaganis' in the 15-year period 1914–1929.

Disregarding the fact that Gartaganis uses different (and not randomly selected) series from ours, we might conjecture *either* that the longer the time period studied, the greater the observed autocorrelation in economic series as would follow from the properties of the autocorrelation coefficient as a function of the number of observations, *or* that the interwar period was a disturbed one in which pre-1914 economic properties were disrupted, and post-depression properties had not been established. On the second conjecture, 16-year autocorrelations for the period before 1914 would be greater than 16-year autocorrelations for the period since 1932; and these, in turn, would be greater than those Gartaganis saw for 1914–29. Such a view would certainly match the intuition of many economic historians and be consistent with a certain amount of sta-

TABLE 2. DETERMINANTS OF AUTOCORRELATION COEFFICIENTS FOR 100 RANDOMLY SELECTED ECONOMIC TIME SERIES

(Natural Numbers, not corrected for trend)

Number of Determinant Values Less Than	Order of Autocorrelation Function[a]			
	2	3	4	5
.001	19	46	60	68
.01	46	64	74	83
.1	73	82	91	94

(Logarithms, not corrected for trend)

Number of Determinant Values Less Than	Order of Autocorrelation Function[a]			
	2	3	4	5
.001	15	43	59	69
.01	43	65	73	83
.1	72	85	90	93

(Natural Numbers, corrected for trend)

Number of Determinant Values Less Than	Order of Autocorrelation Function[a]			
	2	3	4	5
.001	2	15	29	40
.01	12	30	48	56
.1	42	61	74	81

(Logarithms, corrected for trend)

Number of Determinant Values Less Than	Order of Autocorrelation Function[a]			
	2	3	4	5
.001	0	8	27	42
.01	6	30	52	65
.1	44	60	76	85

[a] To an autocorrelation function of nth order corresponds a determinant of order $n+1$. The squares of the first degree functions are discussed in the text.

tistical evidence.[13] We have no evidence as to whether either of these is correct.

We have mentioned the discussion by these writers of autocorrelation functions involving several lags in the variables involved. Such functions generate determinants whose components are lagged autocorrelation coefficients of the type we have discussed. Our data permit some further remarks on the extent to which economic data resemble those generated by higher-order linear autocorrelation functions.

[13] See Ames, "Trends, Cycles, and Stagnation in U. S. Manufacturing Since 1860," *Oxford Economic Papers*, Vol. XI, 1959, p. 270.

In correlation analysis, it is shown that the determinant of a matrix of correlation coefficients vanishes when one of the variables is a linear combination of the others. The value of this determinant, moreover, can be interpreted as the percentage of the variance of one of the variables which cannot be explained by changes in the other variables. We can associate with a linear difference equation of order n a matrix of order $n+1$, whose entries are autocorrelation coefficients, and whose determinant vanishes.[14] Since we have autocorrelation coefficients for lags 1–5, we may readily find out how many of these autocorrelation matrices of order 1–5 actually approach singularity.[15] Such data would indicate the proportion of time series that it would be possible to approximate to the desired degree of precision by an exact linear difference equation. Our results are given in Table 2.

It is clear that the data not adjusted for trend can, in large part, be approximate by linear difference equations of low order; trend-corrected data require higher-order equations to achieve a given degree of approximation.

These findings suggest that our series more nearly satisfy a linear deterministic process than they do a linear stochastic one. However, the findings still allow our series to satisfy, for example, a linear difference equation with nonconstant coefficients, provide these coefficients change slowly over time. Thus, the coefficients might be generated by a "slow" stochastic process, or perhaps change discontinuously at relatively long intervals.[16] If these findings and interpretations are correct, they also suggest the possibility that shock models (disturbances in the equations) may not fit annual economic data very well.

---

[14] Wold, *A Study in the Analysis of Stationary Time Series*, Second Edition, Uppsula, Almquist & Wiksell, 1954, p. 45, Theorem 2.

[15] Even if economic time series were actually generated by linear difference equations, errors of measurement alone would prevent us from observing empirically the vanishing of autocorrelation matrices.

[16] This point arose in a discussion with R. Radner.

TABLE A-I. FREQUENCY DISTRIBUTION OF THE CORRECTED AUTOCORRELATION COEFFICIENT IN 100 RANDOMLY SELECTED ECONOMIC TIME SERIES, FOR VARIOUS LAGS

(Natural Numbers, not corrected for trend)

First Digit of r, unrounded	Lag of 1 year	Lag of 2 years	Lag of 3 years	Lag of 4 years	Lag of 5 years
.9	56	38	21	16	10
.8	17	17	24	18	20
.7	08	08	09	16	11
.6	06	10	08	02	08
.5	06	05	03	09	04
.4	03	07	04	02	06
.3	00	03	09	08	07
.2	02	03	05	07	08
.1	00	04	08	07	06
.0	02	01	00	03	02
−.0	00	02	04	04	04
−.1	00	02	03	04	10
−.2	00	00	02	02	00
−.3	00	00	00	00	03
−.4	00	00	00	02	00
−.5	00	00	00	00	01
−.6	00	00	00	00	00
−.7	00	00	00	00	00
−.8	00	00	00	00	00
−.9	00	00	00	00	00
Mean	.837	.705	.599	.525	.453
Moments about Mean					
Second	.040	.091	.134	.158	.180
Third	−.016	−.038	−.049	−.054	−.050
Fourth	.012	.035	.053	.068	.074
β_1	3.947	1.842	.991	.743	.433
β_2	7.120	4.152	2.946	2.742	2.281

TABLE A-II. FREQUENCY DISTRIBUTION OF THE CORRECTED AUTOCORRELATION COEFFICIENT IN 100 RANDOMLY SELECTED ECONOMIC TIME SERIES, FOR VARIOUS LAGS

(Natural Numbers, corrected for linear trend)

First Digit of r, unrounded	Lag of 1 year	Lag of 2 years	Lag of 3 years	Lag of 4 years	Lag of 5 years
.9	16	00	00	00	00
.8	22	02	01	01	01
.7	23	11	00	00	00
.6	12	12	00	00	00
.5	07	12	04	00	00
.4	04	11	11	00	00
.3	05	08	12	00	00
.2	04	21	24	20	20
.1	06	14	13	17	04
.0	01	03	11	15	08
−.0	00	03	09	16	22
−.1	00	02	12	20	28
−.2	00	01	02	05	07
−.3	00	00	01	03	05
−.4	00	00	00	03	02
−.5	00	00	00	00	02
−.6	00	00	00	00	00
−.7	00	00	00	00	01
−.8	00	00	00	00	00
−.9	00	00	00	00	00
Mean	.675	.376	.146	−.033	−.012
Moments about Mean					
Second	.056	.070	.061	.055	.063
Third	−.014	−.007	−.002	.003	.009
Fourth	.010	.013	.010	.011	.017
β_1	1.106	.158	.024	.054	.302
β_2	3.090	2.751	2.638	3.722	4.348

TABLE A-III. FREQUENCY DISTRIBUTION OF THE CORRECTED AUTOCORRELATION COEFFICIENT IN 100 RANDOMLY SELECTED ECONOMIC TIME SERIES, FOR VARIOUS LAGS

(Logarithms, not corrected for trend)

First Digit of r, unrounded	Lag of 1 year	Lag of 2 years	Lag of 3 years	Lag of 4 years	Lag of 5 years
.9	55	28	18	15	10
.8	15	25	24	15	14
.7	14	08	10	15	15
.6	04	07	07	08	07
.5	04	08	05	05	05
.4	00	09	06	05	04
.3	01	05	02	07	07
.2	00	03	08	05	09
.1	05	00	10	06	10
.0	01	02	04	02	03
−.0	00	02	02	07	09
−.1	01	03	03	08	02
−.2	00	00	01	00	01
−.3	00	00	00	00	03
−.4	00	00	00	02	00
−.5	00	00	00	00	01
−.6	00	00	00	00	00
−.7	00	00	00	00	00
−.8	00	00	00	00	00
−.9	00	00	00	00	00
Mean	.823	.689	.579	.491	.437
Moments about Mean					
Second	.056	.093	.121	.169	.168
Third	−.031	−.038	−.033	−.506	−.040
Fourth	.026	.036	.036	.066	.066
β_1	5.215	1.812	.600	.532	.338
β_2	8.083	4.250	2.468	2.325	2.330

TABLE A-IV. FREQUENCY DISTRIBUTION OF THE CORRECTED AUTOCORRELATION COEFFICIENT IN 100 RANDOMLY SELECTED ECONOMIC TIME SERIES, FOR VARIOUS LAGS

(Logarithms, corrected for linear logarithmic trend)

First Digit of r, unrounded	Lag of 1 year	Lag of 2 years	Lag of 3 years	Lag of 4 years	Lag of 5 years
.9	06	00	00	00	00
.8	25	00	00	00	00
.7	29	07	00	00	00
.6	13	09	00	00	00
.5	06	10	03	00	00
.4	05	11	05	00	00
.3	04	08	07	01	00
.2	02	18	23	13	11
.1	05	21	18	13	05
.0	03	05	04	12	09
−.0	00	05	17	25	28
−.1	02	04	18	21	21
−.2	00	01	05	02	15
−.3	00	00	00	07	04
−.4	00	01	00	05	04
−.5	00	00	00	01	02
−.6	00	00	00	00	00
−.7	00	00	00	00	01
−.8	00	00	00	00	00
−.9	00	00	00	00	00
Mean	.652	.293	.064	−.104	−.178
Moments about Mean					
Second	.065	.075	.059	.052	.049
Third	−.025	−.008	—[a]	−.001	.001
Fourth	.020	.017	.006	.006	.008
β_1	2.259	.158	.001	.007	.006
β_2	4.770	2.982	1.763	2.314	3.314

[a] Absolute value less than .0005.

TABLE B-I. FREQUENCY DISTRIBUTION OF THE COEFFICIENT OF CORRELATION BETWEEN TIME (INTEGERS 1, 2, · · · , 25) AND EACH OF 100 RANDOMLY SELECTED ECONOMIC TIME SERIES

First Digit of r, unrounded	Natural Numbers	Logarithms
.9	22	28
.8	23	16
.7	8	10
.6	5	3
.5	5	5
.4	3	3
.3	2	2
.2	3	2
.1	2	4
.0	2	2
−.0	2	3
−.1	1	1
−.2	0	2
−.3	2	0
−.4	1	0
−.5	1	3
−.6	6	3
−.7	7	6
−.8	1	2
−.9	4	5
Mean	.397	.392
Moments about Mean		
Second	.418	.425
Third	−.264	−.262
Fourth	.042	.043
β_1	.956	.894
β_2	2.380	2.362

TABLE C-I. PERCENTAGE DISTRIBUTION OF THE CORRECTED COEFFICIENT OF CORRELATION IN 50 PAIRS OF RANDOMLY SELECTED ECONOMIC TIMES SERIES, FOR VARIOUS LAGS IN THE INDEPENDENT VARIABLE

(Natural Numbers, not corrected for trend)

First Digit of r, unrounded	Lag of 0 years	Lag of 1 year	Lag of 2 years	Lag of 3 years	Lag of 4 years	Lag of 5 years
.9	08	06	04	06	06	04
.8	06	08	10	04	04	06
.7	08	12	08	10	04	08
.6	12	06	08	04	14	06
.5	02	06	06	06	02	06
.4	06	00	06	10	04	04
.3	06	06	00	02	06	06
.2	06	04	08	02	12	02
.1	06	08	10	08	04	10
.0	04	02	02	06	04	06
−.0	02	04	00	00	08	08
−.1	04	08	06	10	04	10
−.2	02	02	04	04	06	04
−.3	02	06	10	08	02	01
−.4	08	06	04	04	10	10
−.5	08	08	04	04	02	00
−.6	06	04	08	10	08	08
−.7	02	04	02	00	00	00
−.8	02	00	00	00	00	00
−.9	00	00	00	00	00	00
Mean	.145	.122	.126	.107	.113	.109
Moments about Mean						
Second	.352	.343	.326	.300	.281	.271
Third	−.059	−.029	−.036	−.020	−.022	−.010
Fourth	.203	.185	.171	.148	.138	.132
β_1	.080	.020	.038	.015	.022	.005
β_2	1.641	1.573	1.607	1.642	1.749	1.795

TABLE C-II. PERCENTAGE DISTRIBUTION OF THE CORRECTED COEFFICIENT OF CORRELATION IN 50 PAIRS OF RANDOMLY SELECTED ECONOMIC TIME SERIES, FOR VARIOUS LAGS IN THE INDEPENDENT VARIABLE

(Natural Numbers, corrected for linear trend)

First Digit of r, unrounded	Lag of 0 years	Lag of 1 year	Lag of 2 years	Lag of 3 years	Lag of 4 years	Lag of 5 years
.9	02	00	00	00	00	00
.8	06	02	00	00	00	00
.7	06	08	04	00	00	06
.6	00	06	10	04	04	00
.5	10	10	04	08	06	02
.4	08	04	02	06	04	00
.3	06	08	04	02	08	02
.2	08	12	16	14	22	28
.1	14	08	06	14	04	10
.0	08	12	12	08	06	04
−.0	10	10	10	16	14	14
−.1	08	10	10	10	12	18
−.2	06	02	10	02	02	04
−.3	04	04	00	08	08	04
−.4	00	02	04	06	10	06
−.5	00	00	02	02	00	02
−.6	00	02	02	00	00	00
−.7	04	02	00	00	00	00
−.8	00	00	00	00	00	00
−.9	00	00	00	00	00	00
Mean	.158	.137	.046	.018	.012	—[a]
Moments about Mean						
Second	.178	.171	.151	.121	.123	.113
Third	−.016	−.014	−.001	−.001	.001	.010
Fourth	.082	.071	.049	.031	.029	.034
β_1	.047	.040	—[a]	.001	—[a]	.073
β_2	2.605	2.416	2.154	2.098	1.929	2.618

[a] Absolute value less than .0005.

TABLE C-III. PERCENTAGE DISTRIBUTION OF THE CORRECTED COEFFICIENT OF CORRELATION IN 50 PAIRS OF RANDOMLY SELECTED ECONOMIC TIME SERIES, FOR VARIOUS LAGS IN THE INDEPENDENT VARIABLE

(Logarithms, not corrected for trend)

First Digit of r, unrounded	Lag of 0 years	Lag of 1 year	Lag of 2 years	Lag of 3 years	Lag of 4 years	Lag of 5 years
.9	04	02	02	00	04	08
.8	14	10	12	12	08	06
.7	06	10	08	04	06	06
.6	06	14	12	12	08	10
.5	08	02	02	10	08	08
.4	04	06	08	04	06	04
.3	06	02	04	02	00	00
.2	00	06	04	10	08	08
.1	14	06	06	04	10	16
.0	02	06	04	04	04	02
−.0	05	02	02	02	06	02
−.1	04	08	06	06	04	04
−.2	00	00	00	04	02	00
−.3	00	00	06	02	02	06
−.4	06	06	04	02	06	04
−.5	04	00	02	10	06	06
−.6	06	12	10	04	04	04
−.7	06	04	06	06	06	04
−.8	04	04	02	02	02	02
−.9	00	00	00	00	00	01
Mean	.127	.125	.120	.102	.108	.155
Moments about Mean						
Second	.371	.371	.371	.356	.339	.325
Third	−.093	−.089	−.081	−.065	−.056	−.068
Fourth	.246	.235	.225	.214	.205	.198
β_1	.167	.154	.128	.093	.082	.135
β_2	1.781	1.706	1.640	1.689	1.786	1.868

TABLE C-IV. PERCENTAGE DISTRIBUTION OF THE CORRECTED COEFFICIENT OF CORRELATION IN 50 PAIRS OF RANDOMLY SELECTED ECONOMIC TIME SERIES FOR VARIOUS LAGS IN THE INDEPENDENT VARIABLE

(Logarithms, corrected for linear logarithmic trend)

First Digit of r, unrounded	Lag of 0 years	Lag of 1 year	Lag of 2 years	Lag of 3 years	Lag of 4 years	Lag of 5 years
.9	00	00	00	00	00	00
.8	00	02	00	00	02	00
.7	06	04	02	02	00	02
.6	02	00	02	00	02	00
.5	04	04	06	06	00	02
.4	06	02	04	00	02	04
.3	06	08	02	02	08	04
.2	08	14	10	10	14	22
.1	14	16	20	18	08	10
.0	16	08	02	08	12	08
−.0	16	12	18	16	18	12
−.1	16	12	14	16	16	10
−.2	02	06	04	04	02	00
−.3	00	06	04	10	08	08
−.4	02	02	04	04	04	04
−.5	00	02	04	02	02	04
−.6	00	00	04	02	02	00
−.7	00	02	00	00	00	00
−.8	02	00	00	00	00	00
−.9	00	00	00	00	00	00
Mean	.066	.028	−.035	−.070	−.052	−.024
Moments about Mean						
Second	.119	.132	.128	.105	.105	.103
Third	—[a]	−.001	.003	.006	.009	—[a]
Fourth	.046	.049	.040	.028	.032	.026
β_1	—[a]	—[a]	.004	.029	.078	—[a]
β_2	3.233	2.796	2.439	2.560	2.934	2.476

[a] Absolute value less than .0005.

TABLE D-I. PERCENTAGE DISTRIBUTION OF THE MAXIMUM ABSOLUTE VALUE OF THE CORRELATION COEFFICIENT OVER LAGS 0–5, FOR 50 PAIRS OF RANDOMLY SELECTED ECONOMIC TIME SERIES

First Digit of r, unrounded	Natural Numbers		Logarithms	
	Not corrected for trend	Corrected for linear trend	Not corrected for trend	Corrected for linear logarithmic trend
.7 to .9	32	24	36	10
.2 to .6	30	34	26	30
−.1 to .1	2	2	4	10
−.2 to −.6	20	34	10	42
−.7 to −.9	14	2	24	8

TABLE D-II. PERCENTAGE DISTRIBUTION OF THE MAXIMUM ABSOLUTE VALUE OF THE CORRELATION COEFFICIENT OVER LAGS 1–5, FOR 50 PAIRS OF RANDOMLY SELECTED ECONOMIC TIMES SERIES[a]

First Digit of r, unrounded	Natural Numbers		Logarithms	
	Not corrected for trend	Corrected for linear trend	Not corrected for trend	Corrected for linear logarithmic trend
.7 to .9	30	20	30	8
.2 to .6	30	32	32	30
−.1 to .1	2	6	4	12
−.2 to −.6	24	38	12	40
−.7 to −.9	12	4	22	8

[a] In some cases, the total may be less than 100 per cent. This is because two correlations were of absolute value equal to the maximum, but of opposite sign, and could not be entered in this table.

TABLE D-III. AVERAGE ABSOLUTE VALUE OF THE CORRELATION COEFFICIENT IN 50 RANDOMLY SELECTED PAIRS OF ECONOMIC TIME SERIES, FOR VARIOUS LAGS

Lag in Years	Natural Numbers		Logarithms	
	Not corrected for trend	Corrected for linear trend	Not corrected for trend	Corrected for linear logarithmic trend
0	.571	.385	.573	.283
1	.559	.386	.583	.303
2	.548	.362	.583	.312
3	.518	.327	.558	.284
4	.508	.306	.545	.283
5	.499	.290	.541	.269

TABLE D-IV. AVERAGE NUMBER OF RANDOM TRIALS NEEDED TO FIND AN ECONOMIC VARIABLE "EXPLAINING" HALF OR MORE OF THE VARIANCE IN A GIVEN VARIABLE

	Natural Numbers		Logarithms	
	Not corrected for trend	Corrected for linear trend	Not corrected for trend	Corrected for linear logarithmic trend
Lag 0	3	5 plus	2.5	12
Lags 0–5 max.	2 plus	4 minus	1.7	5 plus
Lags 1–5 max.	2.5	4 plus	2 minus	6