THE ECONOMIC AND SOCIAL RESEARCH INSTITUTE

ECONOMETRIC FORECASTING FROM LAGGED RELATIONSHIPS

M. G. KENDALL

Sixth Geary Lecture, 1973

Copies of this paper may be obtained from
The Economic and Social Research Institute. (Limited Company No. 18269).
Registered Office: 4 *Burlington Road. Dublin* 4. *Ireland.*

Price £0.50

(Special rate for students £0.25)

Maurice Kendall was formerly Professor of Statistics at the University of London, Director of Scientific Services and Chairman of Scientific Control Systems Ltd., and is now Director of the World Fertility Survey. This Paper has been accepted for publication by ESRI, which is not responsible for either the content or the views expressed therein.

Printed by Cahill & Co. Limited, East Wall Road, Dublin 3.

Econometric Forecasting From Lagged Relationships

1. My distinguished predecessor in contributing to this series of lectures, Jan Tinbergen, included in the title of his lecture the word "interdisciplinary". I follow him in spirit, for the problem of forecasting is truly interdisciplinary, calling as it does on the combined skills of the economist, the statistician and the mathematician as well as the commonsense of the practitioner. And I cannot think of any more suitable name on whose honour a lecture of this kind should be given than that of Roy Geary, himself an interdisciplinarian if ever there was one, equally at home in all these subjects, and among his many distinctions possessing one which I think is unique, that of being the only former head of a Government Central Statistical Office whose name is attached to a mathematical theorem, has acted at the Abbey Theatre and has been offered a job as a professional footballer.

2. Several years ago two colleagues and I (Coen, Gomme and Kendall, 1969) presented a paper to the Royal Statistical Society on forecasting from lagged relationships. Our general argument was that if cause precedes full effect by a substantial lapse of time, and if one can measure the causal variables, then it ought to be possible to forecast the effect variable: and we presented some numerical material to illustrate the point, among other things by forecasting the *Financial Times* quarterly index of industrial share prices. We did not claim that the work was conclusive, but we thought it raised some possibilities which were worth examining.

3. Apparently nobody else thought so, and the paper came in for some rather rough handling, both from the econometricians and from the statisticians. The econometricians' criticisms I have never

really understood. Apart from the fact that they themselves have given a lot of attention to the problem of distributed lags, it is standard practice in setting up an econometric model to express the (endogenous) variables of interest as linear functions of other (exogenous) variables and of lagged values of the endogenous variables, the latter two groups being known as "pre-determined" because they are assumed known at any particular time-point when a forecast from the model is to be made. In matrix terms the model is usually written as

$$\boldsymbol{A}\boldsymbol{u} + \boldsymbol{B}\boldsymbol{v} = \boldsymbol{\Gamma} \tag{1}$$

where $\boldsymbol{u}$ is a vector of m endogenous variables, $\boldsymbol{A}$ is an $m \times m$ matrix of structural coefficients, $\boldsymbol{v}$ is a vector of e exogneous variables, $\boldsymbol{B}$ is a $m \times e$ matrix of structural coefficients and $\boldsymbol{\Gamma}$ is an m-column vector of "error" terms. In this form the model may be unidentifiable in the sense that we cannot estimate $\boldsymbol{A}$, however large the sample, and some attention has to be given to modifying it if necessary to achieve identifiability. However, if $\boldsymbol{A}$ is not degenerate we can write

$$\boldsymbol{u} = -\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{v} + \boldsymbol{A}^{-1}\boldsymbol{\Gamma} \tag{2}$$

This, in general, is a set of linear relationships known as the "reduced form" in which the u's are expressed in terms of the pre-determined variables (which, it will be remembered, may include lagged values of the u's). In short, the equations are, in general, mixed auto-regression models with errors in equations but not errors in variables.

4. Now this is precisely the kind of model which we were proposing for consideration, with three possible exceptions:

(*a*) We considered one equation {one component of (2)} at a time.

(*b*) Whereas the econometrician can specify the variables $\boldsymbol{u}$ from economic theory, or hopes to be able to do so, the statistician is more inclined to set up working empirical relationships by treating (2) as a regression, putting in all the variables which he thinks may be relevant and leaving a statistical routine to reject the non-contributory variables.

(*c*) To be useful for forecasting all the variables on the right in equation (2) must be predetermined.

It does not seem to me that there need be any gap between the econometric and the statistical approach. The economist, admittedly, may have a much deeper insight into what variables ought to appear in the equations, but the statistician is not without knowledge on the subject too; and we know so little about the outward and visible signs of the inward and invisible workings of an economy that there are advantages in a collateral approach to the problem of structure without imposing preconceived notions on the analysis.

5. A more formidable antagonist has appeared in the form of Professor Box and one of his colleagues (Box and Newbold, 1971). Box maintains that the relationships which we found in the 1969 paper are illusory, and that just as good forecasts can be obtained from the F.T. series itself regardless of information which might be contributed from concomitant series. The method of forecasting which Box and Jenkins have tried to popularise recently relies on what I call auto-projection, the prediction of the future course of a series from past values. The method, in fact, is one of auto-regression but differs from the more classical type in two respects: it works on the differences of the original series (the order of difference being taken far enough to induce relative stationarity in the result), and it allows for an error term which may be a moving average of independently occurring random errors.

6. There is clearly a lot of ground for argument here. On the face of it, one might expect that a method which brings in concomitant variables would give better results than one which relies on the series itself, there being, in some sense, more information brought into play. This is certainly true of some situations; for example, a forecaster who was predicting daily load on an electric grid network would soon find himself in difficulties if he ignored the weather. On the other hand, if the additional information is subject to observational error (and what economic information is not?) the importation of new variable may, in the engineer's terminology, introduce more noise into the system than the effort is worth; and in any case correlations between concomitant series, especially short ones, may

be temporary evanescent efforts creating a spurious (or at least an impermanent) degree of dependence which will frustrate attempts to predict. I do not think that there is any simple way of resolving conflicts of opinion on such matters as these. I will admit that there are many more problems in the method of lagged relations than I originally expected, but I am not yet ready to discard it in those regions where it has been successfully applied in the last few years.

7. But let me, at this stage, make two general points. The first is that it seems unlikely that any one method will be preferable to all others for all kinds of economic series—there is, in standard statistical terminology, no uniformly most powerful forecasting method. The point has been well brought out by Reid (1971) who subjected 117 economic series (mostly British, but including a number of USA series) to several auto-projective methods, and showed that the effectiveness of different techniques varied according to the nature of the series, for example whether or not seasonal variation was present, whether or not there was a large irregular component, whether the series was short or long. Most practitioners, I think, try a number of different methods and, so far as I can judge from talking to them, there is no general consensus favouring any one regardless of circumstances.

8. The second point, arising from the first, is that, even with the aid of an electronic computer, which indeed is essential in this class of work, the amount of labour required to compare different methods over a wide range of types of series is very large and apt to be expensive in computer time. I referred earlier to conflicts of opinion about the relative points of different forecasting methods. In science we like to think that there are no conflicts of opinion which are not resolvable by appeal to the facts. However, in the behavioural sciences we are not always able to wait long enough for our methods to be fully validated and compared. A more regrettable feature of the situation is that some of the facts do not come to light—one can hardly expect a professional forecaster to publish his failures. Inquests and post-mortems are no more popular in statistics than elsewhere in the world. What ought to be objective appraisals of scientific method have shown a slight tendency to become promotional campaigns. I think we are working towards much greater objectivity—even in econometrics, truth will out—but we still have some way to go.

9. For some time the theory of regression has been regarded as one of the least understood parts of statistical theory. But the more one goes into it, the more unsolved problems appear. The one which bedevils the whole of econometrics is that of collinearity among the regressand* variables, the situation in which they are highly inter-correlated among themselves, their covariance matrix is accordingly ill-conditioned and hence estimates of the coefficients are highly unstable. (I may remark in passing that in my opinion interpolation of a coefficient in a regression as a measure of elasticity is in general unsound.) But there are other difficulties; for example regressions which change direction at an unknown point; the discarding of redundant variables; the analysis of sets of equations in which the errors may be correlated from one equation to another. The subject is far from being fully explored, and there is one particular difficulty affecting time-series.

10. If we fit an ordinary regression to p regressors on n observations one would think of the residual as depending on the $n—p$ degrees of freedom which ought to give it considerable sampling stability. Unfortunately such a conclusion is not necessarily sound. Successive values of the regressors are not independent, nor are the regressors independent *inter se*. The "length" of a time-series is not to be reckoned by the number of points at which it is observed. In ordinary statistical theory we are accustomed to talk of "amount of information" in a sample as proportional to the sample size. Personally, I think this is an objectionable use of the word "information" but it is undoubtedly true that in ordinary random samples the precision of estimates derived from them, as measured by the sampling variance, is proportional to the reciprocal of the sample size. However, a realisation of a time-series is not a collection of random observations. The precision of the estimates which we base on it depends on the internal structure of the series, and this is in general not known. It is therefore often impossible to say whether, in the case of an economic series we have enough residual degrees of

*What I call the "regressand" is in the older terminology, the "dependent" or "explained" variable, the one on the left of the regression equation as usually written. The "regressors" are those on the right, the so-called "independent" or "explanatory" variables. I dislike the older terminology on the grounds that the "independent" variables are hardly ever independent and the "explanatory" variables may not be explanatory.

freedom to be able to lean heavily on the predictive formula. It may have worked in the past; we do not know its reliability for the future. And this is independent of another objection frequently, and not unfairly, lodged against all forecasting methods in econometrics, namely that some human agency or Nature herself may alter the rules of the game without much warning or without much further announcement.

11. One further objection may be advanced against equations of the type of equation (2). Long lags look suspicious. In some work on forecasting economic series by autoregressive methods I have sometimes obtained equations in which the lag appears at three or four years. On the face of it, I should think, not very plausible. However, we must remember two things at least.

(*a*) In a highly interconnected system such as a country's economy, where most of the variables are correlated either because they react on one another or because they are moving concomitantly through time, the fact that a rejection routine throws out some variables does not mean that they are unimportant or non-causal; only that they are so highly correlated with the retained variables that the latter are, so to speak, acting on their behalf. Much as has been written on the problem of collinearity in multivariate analysis, both within the econometric field and outside, I do not think that it has been satisfactorily solved.

(*b*) In an economy exhibiting oscillatory movements, whether about a trend or not, there may be a pseudo-periodicity. If, for example, there were a four-yearly cycle in the F.T. Index the appearance of F_{t-15} (t being measured in quarters) in an autoregression would be understandable; it would correspond to F_{t+1}, and the apparent relation to history four years ago would merely reflect the fact that history repeats itself every four years. In economics, as has been known for at least 50 years, oscillations in the economy are not strictly cyclical and behave much more like stochastic processes. However, over a relatively short range the way in which disturbed series get out of step may not have had time to exert its full effect. It certainly looks, from some UK series, as if there has been a swing of about four years period during the last twenty years. The same is true of the USA and Australia (and for all I know, of other countries). It has even been suggested that, so far as the UK is concerned, this is a human artifact, dependent on the semi-regular intervals at which the British elect their House of Commons. I am a sceptic about cycles in any strict sense when no underlying rhythmic cause can be identified, and I would certainly not use them for medium-term forecasting; but for short-term movements up to, say, a year ahead, they may be contributing something to the lagged terms.

12. The more one tries to penetrate into the setting up of mathematical relationships in economics the more unsolved problems appear. Econometric model-building has grown into a subject of its own since Tinbergen's pioneer work before the war. For the time being we are compelled to make forecasts from unreliable data with unreliable methods. In the remainder of this lecture I want to discuss the lines along which research should proceed in order to improve the situation. It all adds up, I fear, to a very large programme of research; but perhaps we may derive some consolation from the fact that we are at least in a position to ask some of the right questions.

13. For the most part econometricians have concerned themselves with systems which are assumed to be expressible in the form of equation (1). Here the variables, whether endogenous or exogenous, are assumed to be measured without error. This is asking rather a lot of some economic material which we have, but the plain fact is that the alternative approach in which there are admitted to be errors in the variables raises some difficult problems of estimation to which Geary himself has contributed a good deal, but which remain for the most part unsolved. Apart from a paper by Konijn (1962) little seems to have been done recently on this topic, the general hope being that the errors are not large enough to invalidate the conclusions we reached by ignoring their existence. I think this may be true for models involving one or two equations; I doubt it very much for models involving ten or more. The situation may be compared to a cocktail party where, as soon as the numbers present exceed a threshold value, there is so much noise that nobody can hear what is being said. This does not usually matter at a cocktail party, but it matters a great deal in econometric model building.

14. The model of equation (1) is generally known as that of errors-in-equations. The "errors" may be errors of mis-specification, indicating that possibly some variable has been left out of the model, or they may be regarded as exogenous shocks to the system. In either case, to make much headway, we have to assume something about their nature. The most popular assumption is to assume that they behave like a random variable in the same sense that values from one time-point to another are independent; and it is usually further assumed that the distribution of errors follows the Gaussian (normal) form.

15. There are grounds for doubting whether either of these assumptions is legitimate in practice. We are therefore faced with three important questions: how do we know whether they are to be complied with? How far can we suppose that our analysis is robust under departures from the assumptions? And if we cannot so suppose, what methods of analysis are required?

16. As to the randomness of the "error" term, the situation might, perhaps, be thought amenable to treatment by calculating the residuals and examining them. This is always a good thing to do, if only because it throws up outlying observations which may be unduly influencing the estimation. But it does not directly answer the question, because the process of fitting and estimating itself will generate serial correlations in the observed residuals, even if the real residuals are uncorrelated. For an ordinary regression the problem has been extensively studied by Durbin and Watson (1971 and earlier papers) who provide a test of independence based on observed residuals. Apart from the fact that the test is not easy to apply in some cases, it depends on normality in the real residuals, is an asymptotic test and cannot be applied with complete confidence to small samples, does not apply even asymptotically to autoregressive or mixed regression-autoregressive systems, and has not been generalised to the simultaneous estimation in more than one equation. I do not say this in depreciation of the Durbin-Watson test, which is based on skilful and powerful mathematics, but it will be clear that there remains an enormous amount of research to be done to deal with the problem of dependence in residuals.

17. Economic variables are not usually normally distributed. Even when the frequency distributions are relatively symmetrical and unimodal, there are more values in the tails than the normal distribution should contain. In fact, there are some strong reasons for supposing that a number of distributions in the social sciences have infinite variance. The point has been effectively made in a number of papers by Mandelbrot (e.g. 1963) who maintains that economic variables can be better represented by a family of distributions known as stable. The distinguishing feature of the family is that the sum of a number of variables distributed according to a stable law is itself distributed according to the same law. The Gaussian distribution is, of course, stable in this sense but it is a very special case.

18. Until recently the study of stable laws was restricted by the fact that, although their characteristic function can be written down explicitly, it cannot be inverted to give a closed form for the frequency distribution. However, with the aid of the computer, numerical conversion of the integral is now possible and in an unpublished thesis Bartels (1972) has successfully obtained some of the distributions in a form which enables their shape to be studied. We can no longer ignore them on the ground that they are too difficult to handle.

19. Two important consequences follow. If the residual terms in the errors-in-equation model are more likely to be of the stable type than Gaussian, the whole of the least-square fitting requires reconsideration. For some time statisticians have speculated whether it would be better to fit theory to observation by minimising the sum of absolute deviations rather than the sum of their squares, the latter assigning undue importance to the large deviations. But least-squares can be handled theoretically with relative ease whereas least-absolute-deviations cannot. The computer has restored the absolute-deviation to its rightful place and there is some growing evidence that it gives better fits, on occasion, than least-squares. This opens up the whole question as to how we should fit econometric relations, and in particular how we should judge the efficiency of a forecasting formula.

20. A second consequence of adopting the non-Gaussian type of law is that new effects appear in the estimation and significance-testing of our constants, and particularly of those required to study the internal constitution of a time-series, the autocorrelations. It has been known for some time that a serial correlation estimated from a stationary process is subject to downward bias in small samples, but that for normally distributed errors the distribution of the serials tends asymptotically to normality. Some experiments by Bartels in the thesis under reference lead to the conclusion that the usual tests of significance no longer apply when the distributions are non-normal stable. Once again our basic theory seems to require extensive re-examination. This is true even when we are concerned with one equation and, presumably, *a fortiori* when a simultaneous set of equations are concerned.

21. I turn now to two other topics which seem to me to require much further research: the problem of distributed lags and the problems raised by differencing a series to get rid of trend or, more generally, to reduce series to stationarity.

Some time-series exist at all points of time, like temperature or price. Others exist only by aggregation, like rainfall or production. In economics we do not often have much choice in the interval of observation or in the period over which aggregation takes place. Much of the data we handle is given by official sources; and even when we can get inside an economic organism, there are usually practical reasons why the accounts department or some other provenant agency cannot produce figures at intervals which are inconvenient to them. (Incidentally, this raises the problem of how we analyse time-series which *are* observed at unequal intervals.) However, the present point is that the intervals of observations may be relatively far apart compared with the period during which economic reactions take place. We are compelled, for the most part, to deal with the mathematics of the situation by the calculus of differences rather than the calculus of differential coefficients. The calculus of differences offers no particular mathematical difficulty but we do not know what are the precise effects, at least in the forecasting context, of approximating to a continuously existent series by observations at sparse intervals.

22. If we conceive of a time-series as a relatively long-term systematic component plus short-term fluctuation, there are grounds for hoping that we might be able to remove the longer component by taking successive differences of the series. The argument is that if the longer component is smooth and can be represented by a polynomial, at least locally, the successive differences will obliterate it and leave the other components behind for separate study. This thought lies at the basis of what is known as the Variate-Difference method. Unfortunately, the process does not work in practice as well as one might expect. There are at least two reasons for this: one is that if the interval of observation is relatively short compared with the oscillatory effects the method may, so to speak, mistake these for trend and partially remove them; the other is that unsystematic random effects are enhanced by differencing so that the original noise in the series is increased to the point when it begins to drown systematic effects even if such still remain. I have given elsewhere two examples of this: (Kendall and Stuart, 1968, vol. 3, p. 375). A series was constructed from the formula

$$u_t = (t-26) + \tfrac{1}{10}(t-26)^2 + \tfrac{1}{100}(t-26)^3 + \epsilon_t \qquad (3)$$

for $t=1$ to 51 and a random rectangular variable ϵ_t ranging from 0 to 99. The series varied for -119 at $t=1$ to 244 at $t=51$. The variate-differences for $k=1$, 2, etc. and dividing by $\binom{2k}{k}$, this being the factor by which the variance of a random term is inflated, and then observing at what point the quotient settles down. On this series the method indicates that the series is effectively linear, and the reason is easy to see. Consider, for example, the cubic term in equation (3). In the original series this varies from $-156{\cdot}25$ to $+156{\cdot}25$, where the random element ranges (about its mean) from $-49{\cdot}5$ to $49{\cdot}5$. When first differences are taken the cubic term becomes

$$\{3(t-26)^2 + 3(t-26) + 1\}/100$$

which ranges from $-18{\cdot}01$ to $+19{\cdot}51$ whereas the random term may range from -99 to $+99$.

23. In an earlier study (Kendall, 1946) I considered the variate-difference method applied to a stationary autoregressive series of the second order. The variate difference method suggested that the series was a cubic plus a random residual with a variance of about 9 per cent of the real variance of the error term.

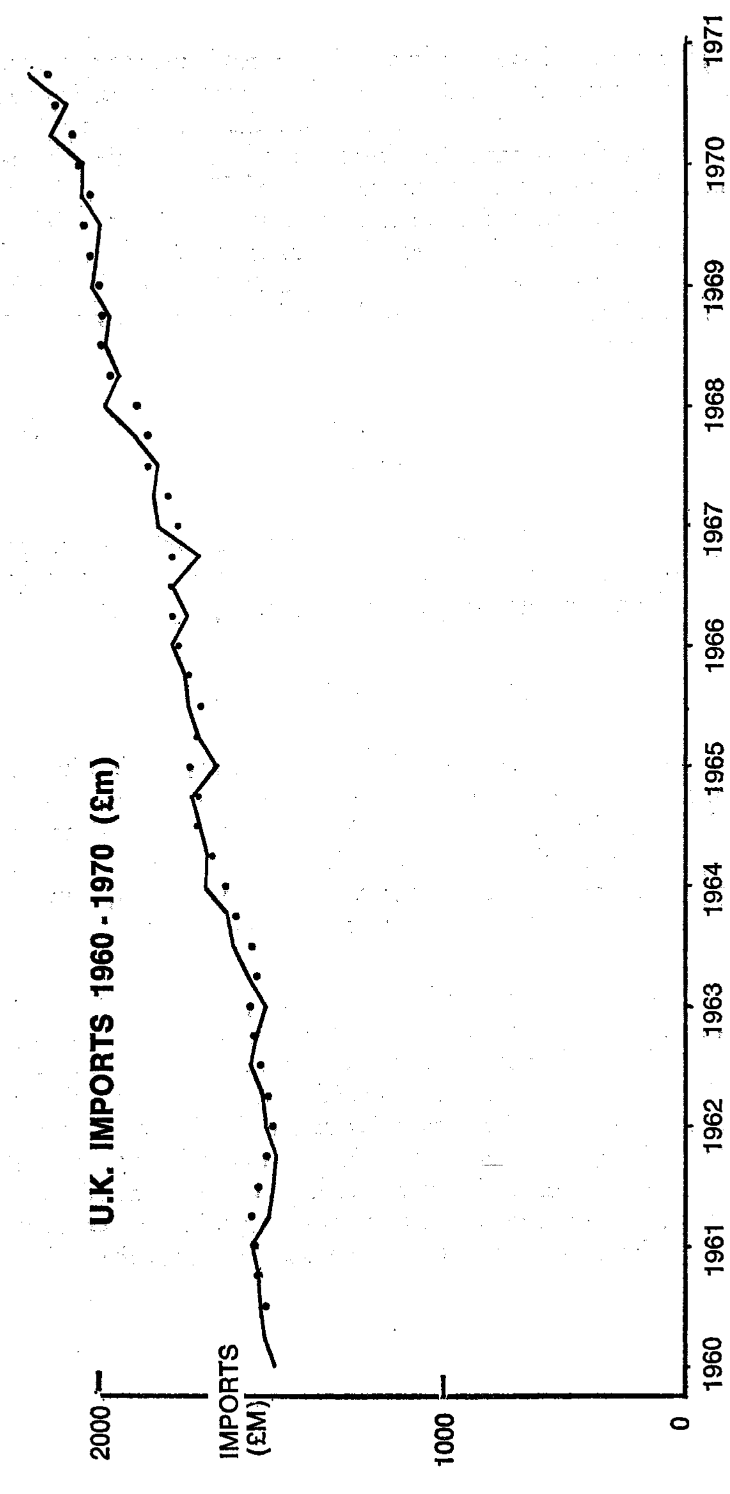
24. It seems to me, therefore, that one must proceed with differencing rather cautiously, especially when the random component of a series is substantial. Not only is the random variance inflated by differencing, but substantial correlations are generated in the residuals of the differenced series. Now the auto-projection method of forecasting advocated by Box and Jenkins attempts to deal with departures from stationarity in a series by taking differences until the resultant series appears to be stationary. It is, I think, a fair question to ask why stationarity is considered so important, but one can see the point of getting rid of trend by some means or other so as to study residual effects in isolation. The question is whether differencing does so without creating its own kind of difficulty. The only safe rule, I think, is not to be dogmatic. In the present state of knowledge an appeal to empirical success seems to be the best way of justifying the use of a model.

25. An example may illustrate the point in the context of forecasting from lagged relationships. I considered the problem of forecasting UK imports (on a quarterly basis) and took the imports for each quarter of 1960 through 1970, corrected for seasonality. There are other variables which may plausibly be supposed to influence imports, namely stocks, consumer durable expenditure and fixed investment. The imports on one quarter were regressed on those for the previous four quarters and similar lagged values of the other three variables, 16 regressors in all. It turned out, however, that the optimal regression routine rejected as non-contributory all the variables except imports at lags one and 2, the equation being

$$I_t = -44 \cdot 4579 + 0 \cdot 52751 I_{t-1} + 0 \cdot 51485 I_{t-2} \qquad (4)$$

This is not a *stationary* autoregressive series, but there is no reason why it should be. The relation between "prediction" and "actual" is shown in Figure 1. R^2 is 0·96.

Figure 1


U.K. IMPORTS 1960 - 1970 (£m)
IMPORTS (£M)
2000
1000
0
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971

26. Now here we have a fairly clear linear trend, and the question therefore arises whether we should not do better by forecasting the change from one quarter to the next, i.e. the first difference. It turned out that even with the differences of all the series as regressands the predictor was poor, the value of R^2 being 0·40 if we calculated it as the maximum likelihood estimator equal to $1-S_R/S_T$ where S_R is the sum of squares of residuals and S_T the sum of squares of the regressand about its mean. In point of fact, this estimate is biased. If we "correct" for degrees of freedom by computing R^2 as

$$1-\frac{S_a}{n-p}\cdot\frac{n}{S_T} \qquad (5)$$

we may get a negative variance! (This raises still one more problem, on which I have no space to dwell). If we retain merely the first difference, so that we are regressing on ΔI_{t-1} we find

$$\Delta I_t = 26{\cdot}7333 - 0.47281\ \Delta I_{t-1}$$

which is, in fact, the same as equation (5). The maximum likelihood estimator of R^2 is 0·20, despite the equivalence of the equations, stems from the fact that in the former case the residual variance is expressed as proportional to the variance of I_t whereas in the second case it is proportional to the variance of ΔI_t—yet one more caution in the interpretation of measures of goodness of fit.

There is, however, one hopeful sign. With the aid of the computer we can generate an unlimited number of series, as long as we like, whose constitution is known and test our methods on them. It has always been a surprise to me that statisticians have not been quicker to take advantage of this fact—there is a remarkable lack of literature concerned with the validation of methods on artificially generated material, at least in time-series analysis. I hope that econometricians, most of whom now have access to a computer, will become more alert to this fact and not become too despondent about theoretical problems.

27. This paper raises more problems than it solves. Indeed, that was part of my purpose, to call attention to some of the outstanding topics in time-series analysis on which further research is needed. So far as concerns forecasting from lagged relationships, while I freely admit that empirically determined relationships are not entirely satisfactory, especially for projection into the future, I would still claim that they should not be discarded until they have been shown to be worse than other forecasting methods.

REFERENCES

Bartels, R. H. (1972). Stable distributions in economics. *Thesis presented to the University of Sydney. Australia.*

Box, G. E. P. and Newbold, P. (1971). Some comments on a paper of Coen, Gomme and Kendall, *Journal of the Royal Statistical Society*, A, 134, 299.

Coen, P. G., Gomme, F. D. and Kendall, M. G. (1969). Lagged relationships in economic forecasting. *Journal of the Royal Statistical Society*, A, 132, 133.

Durbin, J. and Watson, G. S. (1971). Testing for serial correlation in least-squares regression. *Biometrika* 58, 1.

Kendall, M. G. (1946). *Contributions to the study of oscillatory time-series.* Cambridge University Press.

Kendall, M. G. and Stuart, A. (1968). *The Advanced Theory of Statistics.* vol. 2. Charles Griffin & Co., London.

Konijn, H. S. (1962). Identification and estimation in a simultaneous equations model with errors in the variables. *Econometrica.* 30, 79.

Mandelbrot. B. (1963). The variation of certain speculative prices. *Journal of Business*, 36, 394.

Reid, D. J. (1971). Forecasting in action: a comparison of forecasting techniques in economic time-series. *Joint Conference of Operational Research Society's Group on Long Range Planning.*

Tinbergen, Jan (1972). *An Interdisciplinary approach to the measurement of utility or welfare.* Economic and Social Research Institute, Dublin.