Some Comments on a Paper of Coen, Gomme and Kendall
Author(s): George E. P. Box and Paul Newbold
Source: *Journal of the Royal Statistical Society. Series A (General)*, Vol. 134, No. 2 (1971), pp. 229-240
Published by: Wiley for the Royal Statistical Society
Stable URL: http://www.jstor.org/stable/2343873

Some Comments on a Paper of Coen, Gomme and Kendall

By George E. P. Box[1] and Paul Newbold[2]

University of Wisconsin

[Received July 1970. Revised February 1971]

Summary

The method of analysis used in a recent paper on economic forecasting is reviewed. Evidence is presented that what were believed to be highly significant relationships making possible the forecasting of the *Financial Times* share index arise because of the inflexibility of the assumed error structure.

1. Introduction

In a recent publication, by Coen, Gomme and Kendall (1969) which for convenience we refer to as the C.G.K. paper, the forecasting of the *Financial Times* ordinary share index using various other lagged series is discussed. It is sufficient for illustration to consider relation (7) of the C.G.K. paper which we write as

$$Y_t = \alpha + \beta_1 X_{1,t-6} + \beta_2 X_{2,t-7} + n_t \tag{1}$$

which is projected to obtain forecasts. In this expression the "output" Y_t is the *Financial Times* ordinary share index. The two "inputs" are $X_{1,t}$, United Kingdom car production, and $X_{2,t}$, the *Financial Times* commodity index, and n_t is an error term. Quarterly data were employed yielding time series containing 51 successive observations. Of course this is only one of a number of such relationships which the authors postulate. However, it is their methods which we are doubtful about and our reservations would apply equally to their other analyses.

The authors built their model (1) by cross-correlating the detrended series which in this instance were the residuals remaining after fitting linear least-squares regressions on time to each series. For example, Fig. 1(a) shows a plot of the sample cross-correlation function† between the *Financial Times* share index (Y_t detrended) and United Kingdom car production ($X_{1,t}$ detrended). The authors display this cross-correlation for positive lags only (Y leading X) and note that a moderately large cross-correlation occurs near lag 5 or 6. The choice of lags for the independent variables in the linear regrassion was made by initially including among the regressors an independent variable at several different lags close to the value where the sample cross-correlation was a maximum in absolute value. A stepwise regression program was used to determine which lags should be included in the final equation. Values of Student's t-statistic of 14·1 and $-9{\cdot}9$ were computed for the estimates of the parameters β_1 and β_2 in (1) and these appeared to be very highly significant. The authors were thus led to believe, for example, that car production six quarters previously and the *Financial Times* commodity index seven quarters previously could be used to forecast

[1] At University of Essex in 1971.

[2] Now at University of Nottingham.

† Our calculated cross-correlations differ somewhat from those given in the C.G.K. paper; however, the general pattern is similar.

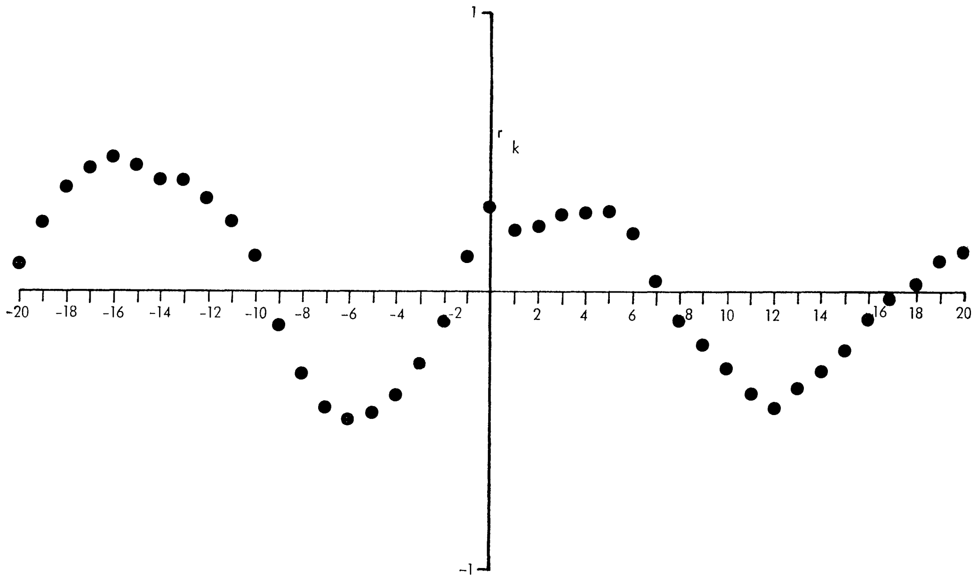

Fig. 1(a). Sample cross-correlations r_k between the *Financial Times* share index (detrended) and lagged values of United Kingdom car production (detrended). Fifty-one pairs of observations.

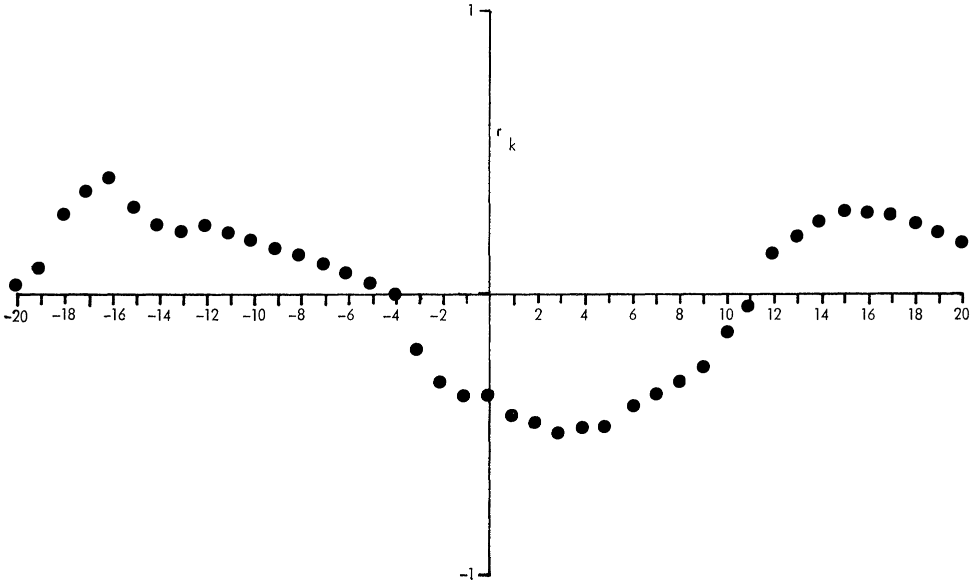

Fig. 1(b). Sample cross-correlations r_k between two unrelated detrended random walks generated from the first two columns of Wold's table of random normal deviates. Fifty pairs of observations.

the *Financial Times* ordinary share index. Because it was felt that the method of detrending itself might give rise to a cross-correlation effect the paper includes an appendix by E. M. L. Beale which shows that *if* individual series were of the form

$$X_t = \alpha + \beta t + a_t, \tag{2}$$

where a_t are assumed to be independent identically distributed random deviates, then the serial correlations of the residuals obtained when (2) is fitted by ordinary least squares would not differ much from those of a_t unless the sample size was small, thus indicating that sample cross-correlations such as those of Fig. 1 were not created by the detrending procedure. They, furthermore, conducted sampling experiments in which cross-correlations of residuals from detrended *random* series were plotted. They remark that the resulting diagrams are sufficiently unlike the *smooth* curves (see Fig. 1) which characterized the cross-correlations of the economic series as "to indicate that the latter are not artefacts created by the trend removal process, at least so far as concerns random residuals".

We believe that the authors were right to suspect that the apparent lagged relationships which they found might be produced by an artefact. The object of this report is to present evidence that this is so, and to show that this happens because of the inappropriateness of the error structure chosen.

2. EXAMINATION OF THE PROPOSED MODEL

In this paper we suppose throughout that $\ldots a_t, a_{t-1}, a_{t-2}, \ldots$ are a sequence of *independent* identically distributed random variables having means equal to zero. We shall subsequently call this a *white noise* process.

Let us consider equation (1). Allowing for detrending, the model may be written

$$Y_t = \alpha + \beta_0 t + \beta_1 X_{1,t-6} + \beta_2 X_{2,t-7} + n_t. \tag{3}$$

A critical assumption on which the C.G.K. analysis hangs is that the n_t's, representing error or "noise", are independent identically distributed random variables. Thus, by employing ordinary least squares, it is tacitly assumed that $n_t = a_t$.

However, it must surely be rare that the noise structure for economic models of this kind can be so represented. One might suspect instead that the n_t's were dependent and possibly best represented by some stable non-stationary noise model such as one in which the first difference of the noise was represented by a stationary process. A particular noise structure which has often proved useful in applications of this kind is the integrated moving average or "noisy random walk".

$$n_t = (1-\theta) \sum_{j=1}^{\infty} a_{t-j} + a_t, \quad -1 < \theta \leqslant 1, \tag{4}$$

where the a_t's are independent but for $\theta \neq 1$ the n_t's can be highly dependent.

Characteristics of the integrated moving average model of special interest are:

(i) It may alternatively be written as an infinite autogressive process in which the autogressive weights diminish exponentially

$$n_t = (1-\theta)(n_{t-1} + \theta n_{t-1} + \theta^2 n_{t-2} + \ldots) + a_t$$

with θ between zero and unity. The model thus exponentially discounts past information.

(ii) It follows, as was first shown in Muth (1960), that the model produces the widely used exponentially weighted average as an optimal forecast.

(iii) The model can be written in the convenient alternative form

$$\nabla n_t = a_t - \theta a_{t-1}$$

which implies that the first difference $\nabla n_t = n_t - n_{t-1}$ of the noise is a first-order moving average process and is therefore readily identifiable.

(iv) In the special case $\theta = 1$ we obtain the noise structure assumed in the C.G.K. paper.

(v) In the special case $\theta = 0$ we obtain a noise model which is a pure random walk.

(vi) For intermediate values we have the discounted disturbance structure of (i) which often provides a satisfactory representation of noise in economic series models.

We may now substitute the alternative noise structure (4) in (3) and regard θ as a parameter to be estimated along with the other parameters. If the estimate of θ is close to unity then the simpler error structure assumed in the C.G.K. paper will be vindicated.

After substituting the augmented error structure and differencing we obtain

$$y_t = \beta_0 + \beta_1 x_{1,t-6} + \beta_2 x_{2,t-\iota} + a_t - \theta a_{t-1},$$

where

$$y_t = Y_t - Y_{t-1}, \quad x_{1,t} = X_{1,t} - X_{1,t-1},$$

$$x_{2,t} = X_{2,t} - X_{2,t-1} \quad \text{and} \quad a_t - \theta a_{t-1} = n_t - n_{t-1}.$$

Now the model may be written in the form

$$a_t = \theta a_{t-1} + y_t - \beta_0 - \beta_1 x_{1,t-6} - \beta_2 x_{2,t-7},$$

whence we may recursively compute the quantities $a_t = (a_t \mid \theta, \beta_0, \beta_1, \beta_2)$ from $t = 8$ onwards for any given choice of parameters. It will make little practical difference if the starting value of a_7 in this recursion is set equal to its unconditional expected value of zero, or if it too is treated as a parameter to be estimated. In either case the least-squares estimates of the parameters obtained by minimizing $\sum (a_t \mid \theta, \beta_0, \beta_1, \beta_2)^2$, where summation extends over the whole sample, will closely approximate maximum-likelihood estimates (Barnard *et al.*, 1962).

The least-squares estimates with their approximate standard errors obtained from an iterative nonlinear squares fit are as follows:

$$\hat{\theta} = -0{\cdot}06 \pm 0{\cdot}15, \quad \hat{\beta}_0 = 1{\cdot}78 \pm 2{\cdot}71,$$

$$\hat{\beta}_1 = 0{\cdot}00016 \pm 0{\cdot}00009, \quad \hat{\beta}_2 = -1{\cdot}16 \pm 1{\cdot}18.$$

The analysis is remarkably revealing.

(1) The value of $\hat{\theta}$ is close to zero and not to unity implying that the noise structure is very different from that assumed in the C.G.K. paper and is in fact like a random walk.

(2) While one of the four estimates ($\hat{\beta}_1$) is $1{\cdot}78$ times its standard error, this can hardly be regarded as unusual. Thus with the less restrictive error structure there is no real evidence of any relation at all between the output on the one

hand and the two lagged inputs on the other hand. Thus among the class of models considered there is in fact little reason to question the unsophisticated model

$$\nabla Y_t = a_t,$$

which implies that Y_t is approximately a random walk and agrees with the frequently confirmed conclusion of Bachelier (1900) concerning the behaviour of stock prices.

(3) It would follow in particular if this model were appropriate that an observation at time $t+l$ could be expressed in terms of that at time t by the equation

$$Y_{t+l} = Y_t + \sum_{j=t+1}^{t+l} a_j.$$

Now (see, for example, Whittle, 1963) the minimum mean-square error (m.m.s.e.) forecast $\hat{Y}_t(l)$ of Y_{t+l} made at origin t for lead time l is given by taking expected values conditional on available knowledge at the time origin t. Hence the m.m.s.e. forecast for l steps ahead would be independent of the inputs $X_{1,t-6}$ and $X_{2,t-7}$ and would be simply the current value of the output

$$\hat{Y}_t(l) = Y_t. \tag{5}$$

The more general error structure proposed above might of course still be unduly restrictive. This could be checked in two ways:

(i) by considering other error structures;
(ii) by examining the behaviour of residuals from the fitted models.

A noise structure which presents a plausible alternative to the integrated moving average (4) is a low-order autoregressive process such as the second-order process

$$n_t = \phi_1 n_{t-1} + \phi_2 n_{t-2} + a_t. \tag{6}$$

It is to be noted that with $\phi_1 = 1$ and $\phi_2 = 0$ this coincides with the random walk model $\nabla n_t = a_t$. Thus the two classes of models represented by (4) and (6) intersect at the random walk.

Table 1 summarizes the results from fitting various models and it will be seen that the calculations confirm that the noise model is approximately a random walk. Furthermore, pronounced residual autocorrelations from the C.G.K. model are evident with a highly significant value for the Durbin and Watson (1951) statistic but, no such evidence of inadequacy is found for the other models.

3. Cross-Correlation Patterns of Random Walks

The apparent relationships in the C.G.K. paper between economic series have a number of puzzling aspects. For example, we see from our Fig. 1(a) that when cross-correlations with negative as well as positive lags are plotted one finds even larger cross-correlations existing at negative lags than those found in the C.G.K. paper at positive lags. This might suggest on the reasoning of that paper that the stock prices might be used to forecast car production instead of vice versa. And *a priori* this seems at least equally plausible.

The question thus arises of how cross-correlation patterns of this kind can have arisen. The answer involves the structure of the individual series which are not adequately represented by the classical regression model

$$X_t = \alpha + \beta t + a_t,$$

but rather require a model closer to the random walk form

$$\nabla X_t = \beta + a_t,$$

that is,

$$X_t = X_{t-1} + \beta + a_t$$

or equivalently

$$X_t = \alpha + \beta t + \sum_{j=0}^{\infty} a_{t-j}.$$

TABLE 1

Estimates and standard errors of coefficients in equation (3) for various noise structures

Assumed structure for noise in Y_t	*White noise (C.G.K.)*	*Integrated moving average*	*Second-order autoregressive*	*First-order autoregressive*	*Random walk*
	$n_t = a_t$	$n_t = a_t + (1-\theta)\sum_{j=1}^{\infty} a_{t-j}$	$n_t = \phi_1 n_{t-1} + \phi_2 n_{t-2} + a_t$	$n_t = \phi_1 n_{t-1} + a_t$	$n_t = \sum_{j=0}^{\infty} a_{t-j}$
α	653 ± 57		306 ± 108	318 ± 106	
β_0		$1·78 \pm 2·7$	$2·31 \pm 1·0$	$2·04 \pm 1·1$	$1·74 \pm 2·6$
β_1	$0·00047 \pm 0·00004$	$0·00016 \pm 0·00009$	$0·00017 \pm 0·00009$	$0·00018 \pm 0·00009$	$0·00017 \pm 0·00008$
β_2	$-6·13 \pm 0·62$	$-1·16 \pm 1·18$	$-1·76 \pm 1·22$	$-1·87 \pm 1·18$	$-1·27 \pm 1·17$
θ		$-0·06 \pm 0·15$			
ϕ_1			$0·93 \pm 0·16$	$0·82 \pm 0·10$	
ϕ_2			$-0·14 \pm 0·16$		
σ_a^2	497	321	299	298	315
Durbin–Watson statistic	Significant at 1 per cent	Not significant	Not significant	Not significant	Not significant

To gain preliminary insight into the behaviour of cross-correlations between series generated by models of this latter kind, a sampling experiment was performed as follows.

Each column of Wold's table of random normal deviates contains 50 entries. The first five columns on the first page of the table could therefore be used to generate five random walks of 50 observations by computing cumulative sums of column entries. The five independent random walk series so obtained were detrended as in the C.G.K. paper and the sample cross-correlations between each pair computed. Fig. 1(b) shows the cross-correlation function between the series generated from the first two columns of the random deviates. It has a pattern typical of those found for the other pairs.

Now persuasive features of the cross-correlation patterns in the C.G.K. paper, to which the authors have drawn attention are:

(i) their smoothness;
(ii) the large absolute magnitude of the biggest cross-correlation.

But it is exactly these features which are displayed by the cross-correlations of the independent random walks. In particular the largest in absolute value of the cross-correlations found and, in brackets, the lag at which this correlation appeared is shown in Table 2.

TABLE 2

Largest cross-correlations (with lag in brackets) found between independent random walks generated from the first five columns of Wold's table of random normal deviates

Column \ Column	2	3	4	5
1	−0·48(5)	−0·42(5)	0·50(−9)	−0·48(8)
2		0·51(1)	0·53(5)	0·54(4)
3			0·49(4)	−0·35(−16)
4				−0·79(18)
5				

Suppose we now treat one detrended random walk series as the detrended "input" X'_t and another as the detrended "output" Y'_t and following the C.G.K. paper fit by ordinary least squares the usual regression model

$$Y'_t = \alpha + \beta X'_{t-j} + a_t,$$

where j is chosen to give the maximum cross-correlation. Then applying the standard t-test it is readily confirmed that *every one* of the ten pairs of series yields a regression "significantly different" from zero *at least* at the 5 per cent point, even though the series are in fact independent.

In the C.G.K. paper the authors mention that much of the early work was done by graphing the series on transparencies to a roughly comparable scale, superposing them and sliding along the time axis to see whether there was any fairly obvious coincident variation. Their Fig. 1 is such a superposition which seemingly shows a remarkably close relationship. Where high cross-correlation is found at the particular lag we would expect to be able to visually demonstrate the relationship by this graphical technique and vice versa. It may be asked then whether a visual impression of a relationship is obtained when two unrelated random walks are treated in this way. Fig. 2 shows the plot for independent random walks 2 and 4 obtained from Wold's random deviates detrended, comparably scaled, and at lag 5 where maximum correlation is produced. The apparent relationship is partly due to the flexibility allowed in what is treated as similar—we can in effect adjust for location, spread, trend and lag before we need find similarity, partly due to the comparative smoothness of what is to be compared—to find a correlation only *a few* detrended rescaled and suitably lagged bumps have to roughly match, and partly due to selection process—among n series there are $\frac{1}{2}n(n-1)$ pairs of series that could show such an apparent relationship.

The reason for the smoothness and large absolute magnitudes of the cross correlations is readily explained from a theoretical viewpoint as follows.

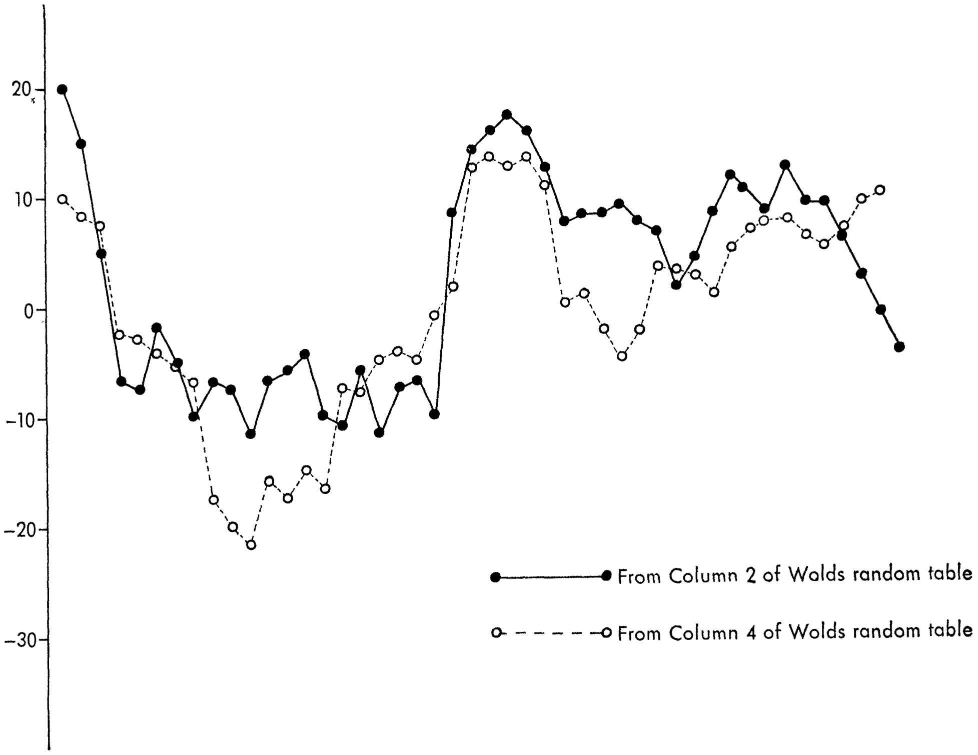

FIG. 2. A plot of comparably scaled and detrended independent random walks at lag 5.

Smoothness

Suppose we have a series of n values generated by the random walk process

$$(X_t - X_{t-1}) - \beta_1 = U_t$$

or

$$x_t - \beta_1 = U_t$$

and N values of Y generated similarly by

$$(Y_t - Y_{t-1}) - \beta_2 = V_t$$

or

$$y_t - \beta_2 = V_t$$

where U_t and V_t are independent white noise processes and let us define the cross-covariance between the differences x_t and y_t as

$$C_k^* = (N-1)^{-1} \sum_{t=2}^{N-k} (x_t - \bar{x})(y_{t+k} - \bar{y}).$$

Now, suppose we postulate trend relationships of the form

$$X_t = \alpha_1 + \beta_1 t + e_{1,t}$$

and

$$Y_t = \alpha_2 + \beta_2 t + e_{2,t},$$

and let $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$ and $\hat{\beta}_2$ be *any* estimators of α_1, α_2, β_1 and β_2. The detrended series are then

$$X'_t = X_t - \hat{\alpha}_1 - \hat{\beta}_1 t,$$

$$Y'_t = Y_t - \hat{\alpha}_2 - \hat{\beta}_2 t,$$

and the sample cross-covariances for the detrended series are

$$C_k = N^{-1} \sum_{t=1}^{N-k} (X'_t - \bar{X}')(Y'_{t+k} - \bar{Y}'). \tag{7}$$

It is shown in the appendix that for moderate or large sample sizes to a close approximation

$$\nabla^2 C_{k+1} = -C^*_k. \tag{8}$$

Now (see, for example, Bartlett, 1955) the cross-covariances C^*_k between two independent white noise processes are independently distributed about zero with constant variance. Thus writing $e_k = -C^*_{k-1}$ the e_k's form a white noise process and the C_k's satisfy the difference equation

$$\nabla^2 C_k = e_k$$

the solution of which may be written

$$C_k = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} e_{k-i-j}.$$

Thus on the assumption made the cross-covariances C_k themselves follow a highly *non-stationary* stochastic process—the cumulative sum of a cumulative sum of random deviates. The appearance of any particular series of cross-covariances and hence of the corresponding cross-correlations is bound therefore to be smooth. Thus with the assumptions made, even though X and Y are generated by independent processes, their cross-covariances and hence their cross-correlations will wander about in a smooth pattern peculiar to each generating set of random numbers, in much the same way as was found for the economic series in the C.G.K. paper. This will be so irrespective of whether, or in which way, the series are detrended.

Size

The above can explain the smooth appearance of the cross-correlations, there remains the question of their large size. The variance of the sample cross-correlations r_k between two independent normal sequences X and Y is given (Bartlett, 1955) approximately by

$$\operatorname{var}(r_k) = (n-k)^{-1} \sum_{\nu=-\infty}^{+\infty} \rho_{xx}(\nu)\, \rho_{yy}(\nu),$$

where $\rho_{xx}(\nu)$ and $\rho_{yy}(\nu)$ are the theoretical autocorrelations. This variance is $(n-k)^{-1}$ for series which are not autocorrelated. However, it can be substantially inflated for correlated sequences. For example, suppose the series under study can be represented by unrelated first-order autoregressive processes each with parameter ϕ.

Then substitution yields

$$\text{var}(r_k) = (n-k)^{-1}(1+\phi^2)/(1-\phi^2).$$

The "inflation factor" $(1+\phi^2)/(1-\phi^2)$ becomes large as ϕ approaches unity and as the sequences approximate to the random walks we are considering.

Furthermore, as is noted, for example, by Hannan (1960), the variance of the regression coefficient between two such unrelated autoregressive sequences is inflated approximately by the same factor. The standard errors of the regression coefficients quoted in the C.G.K. paper could thus easily be underestimated by an order of magnitude. This possibly accounts for the high levels of significance obtained.

It is seen then that the observed cross-correlation phenomena are to be expected from unrelated but autocorrelated sequences.

Actual performance of the forecasts

To compare the C.G.K. forecasts with those obtained using equation (5) which totally ignores the inputs X_1 and X_2, forecasts were compared from one step, to six steps, ahead. The forecasting process was begun from the origin 1963/4—that is, the fourth quarter of 1963. All previous data were used to obtain forecasts

(i) from model (1) (C.G.K. forecast),
(ii) from model (5) (present price is forecast price)

The origin was then moved forward to 1964/1 and the whole process was repeated. The origin was moved forward one step at a time to 1967/3, thus producing 16 pairs of forecasts made one step ahead, 15 pairs two steps ahead, and so on. The averages for the squared errors of these forecasts are shown below

	Equation (1)	*Equation* (5)
One step ahead	969	386
Two steps ahead	1,164	894
Three steps ahead	1,264	1,301
Four steps ahead	1,279	1,270
Five steps ahead	1,274	739
Six steps ahead	1,500	375

Comparison of these results verifies that for these data equation (5) usually provides better forecasts.

CONCLUSIONS

Coen, Gomme and Kendall end their paper with the conclusion that their method deserves serious consideration for short-term economic forecasting. We have written this paper because on the contrary we believe this method should not be employed because of an innate and insidious capacity to mislead which we have discussed in some detail. The criticisms we have made are in the spirit of a recently published book (Box and Jenkins, 1970) which is, in turn, based on a number of previous reports and papers there referenced. These latter authors regard the process of model construction as involving first the consideration of an adequately flexible and theoretically sensible family of models followed by the iterative use of the sequence: model identification—model fitting—model diagnostic checking.

In that context we believe we have shown in this paper that:

(i) the *class of* C.G.K. *models* considered—linear multiple regression on lagged input variables with uncorrelated errors—is a demonstrably inadequate family. Adequacy would mean that transformations of the data of the kind

$$n_t = Y_t - \alpha - \beta_0 t - \beta_1 X_{1,t-s_1} - \beta_2 X_{2,t-s_2}$$

for suitable choice of α, β_0, β_1, s_1 and s_2 could produce uncorrelated noise n_t and this has not been found to be so;

(ii) the process of *identification* involving superposition of the highly autocorrelated time series backed by cross-correlation analysis invites the discovery of spurious relationships;

(iii) the process of *fitting* by ordinary least squares with implied uncorrelated errors is inappropriate and could lead to t values inflated by an order of magnitude;

(iv) no diagnostic *checking*, such as analysis of residuals, which would have pointed to these inadequacies, seems to have been attempted.

ACKNOWLEDGEMENT

We are grateful to Larry Haugh for carrying out additional calculations and for his help in the revision of this paper. This work was supported by the Air Force Office of Scientific Research and by the U.S. Army Research Office, Durham.

APPENDIX

CROSS-CORRELATION PROPERTIES OF DETRENDED RANDOM WALKS

We may establish the approximate relation between C_k and C_k^* of equation (8) as follows:

$$(N-1)\,C_k^* = \sum_{t=2}^{N-k}\left[\left\{X_t' - (N-1)^{-1}\sum_{t=2}^{N} X_t'\right\} - \left\{X_{t-1}' - (N-1)^{-1}\sum_{t=1}^{N-1} X_t'\right\}\right]$$
$$\times\left[\left\{Y_{t+k}' - (N-1)^{-1}\sum_{t=2}^{N} Y_t'\right\} - \left\{Y_{t+k-1}' - (N-1)^{-1}\sum_{t=1}^{N-1} Y_t'\right\}\right]$$

and approximately

$$(N-1)^{-1}\sum_{t=2}^{N} X_t' = (N-1)^{-1}\sum_{t=1}^{N-1} X_t' = \bar{X}';$$

$$(N-1)^{-1}\sum_{t=2}^{N} Y_t' = (N-1)^{-1}\sum_{t=1}^{N-1} Y' = \bar{Y}'.$$

Thus

$$(N-1)\,C_k^* \simeq \sum_{t=2}^{N-k}[(X_t' - \bar{X}') - (X_{t-1}' - \bar{X}')]\,[(Y_{t+k}' - \bar{Y}') - (Y_{t+k-1}' - \bar{Y}')],$$

therefore

$$(N-1)\,C_k^* \simeq \sum_{t=2}^{N-k}(X_t' - \bar{X}')(Y_{t+k}' - \bar{Y}') + \sum_{t=1}^{N-k-1}(X_t' - \bar{X}')(Y_{t+k}' - \bar{Y}')$$
$$-\sum_{t=2}^{N-k}(X_t' - \bar{X}')(Y_{t+k-1}' - \bar{Y}') - \sum_{t=1}^{N-k-1}(X_t' - \bar{X}')(Y_{t+k+1}' - \bar{Y}').$$

For moderate or large samples, on dividing by N the approximate relation (8) is now obtained.

REFERENCES

BACHELIER, L. (1900). Théorie de la spéculation. *Ann. Sci Éc. Norm. Sup., Paris*, Series 3, **17**, 21–86.

BARNARD, G. A., JENKINS, G. M. and WINSTEN, C. B. (1962). Likelihood inference and time series. *J.R. Statist. Soc.* A, **125**, 321–352.

BARTLETT, M. S. (1955). *Stochastic Processes*. Cambridge: Cambridge University Press.

BOX, G. E. P. and JENKINS, G. M. (1970). *Time Series Analysis Forecasting and Control.* San Francisco: Holden-Day.

COEN, P. G., GOMME, E. D. and KENDALL, M. G. (1969). Lagged relationships in economic forecasting. *J.R. Statist. Soc.* A, **132**, 133–152.

DURBIN, T. and WATSON, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, **38**, 159–178.

HANNAN, E. J. (1960). *Time Series Analysis*. London: Methuen.

MUTH, G. F. (1960). Optimal properties of exponentially weighted forecasts of time series with permanent and transitory components. *J. Amer. Statist. Ass.*, **55**, 299–306.

WHITTLE, P. (1963) *Prediction and Regulation by Linear Least-squares Methods*. London: English Universities Press.