Journal of Econometrics 16 (1981) 121–130. North-Holland Publishing Company

SOME PROPERTIES OF TIME SERIES DATA AND THEIR USE IN ECONOMETRIC MODEL SPECIFICATION

C.W.J. GRANGER

University of California at San Diego, La Jolla, CA 92093, USA

1. Introduction

It is well known that time-series analysts have a rather different approach to the analysis of economic data than does the remainder of the econometric profession. One aspect of this difference is that we admit more readily to looking at the data before finally specifying a model, in fact we greatly encourage looking at the data. Although econometricians trained in a more traditional manner are still very much inhibited in the use of summary statistics derived from the data to help model selection, or identification, it could be to their advantage to change some of these attitudes. In fact, I have heard rumors that econometricians do data-mine in the privacy of their own offices and I am merely suggesting that some aspects, at least, of this practice should be brought out into the open.

The type of equations to be considered are generating equations, so that a simulation of the explanatory side should produce the major properties of the variable being explained. If an equation has this property, it will be said to be consistent, reverting to the original meaning of this term. As a simple example of a generally non-consistent model, suppose that one has

$$y_t = a + bx_t + e_t,$$

where y_t is positive, but x_t is unbounded in both directions. A more specific example is when y_t is exponentially distributed and x_t normally distributed. The only case when such a model is consistent is when b is zero. Although it would be ridiculous to suggest that econometricians would actually propose such models, it might be noted that two models that appear in the finance literature have

$$D_t = a + bD_{t-1} + cE_t + e_t,$$

and

$$P_t = d + eD_t + fE_t + e_t,$$

0165-7410/81/0000-0000/$02.50

where D_t represents dividends, E_t is earnings and P_t is share price. Note that P_t and D_t are necessarily positive, but that E_t can be both positive and negative, as Chrysler and other companies can testify.

A further example arises from consideration of the question of whether or not a series is seasonal. For the purposes of this discussion, a time series will be said to be seasonal if its spectrum contains prominent peaks round the seasonal frequencies, which are $2\pi j/12$, $j=1,2,\ldots,6$, if the data are recorded monthly. In practice, this will just mean that a plot of the series through time will display the presence of a fairly regular twelve-month repeating shape. Without looking at the data, one may not know if a given series is seasonal or not and economic theory by itself may well not be up to the task of deciding. If now we look at a group of variables which are to be modelled, how does the presence, or lack, of seasonality help with model specification? Considering just single-equation models, which are suitable for the simple point to be made, of the form

$$y_t = a + bx_t + cz_t + e_t, \tag{1.1}$$

then it would clearly be inconsistent to require

(i) if y_t were seasonal, x_t, z_t not seasonal that e_t be white noise (or non-seasonal), or
(ii) generally, if y_t were not seasonal, but just one of x_t *or* z_t was seasonal and that e_t be white noise or AR(1) or any non-seasonal model.

Clearly, we may have information about the time-series properties of the data, in terms of spectral shapes, that will put constraints on the form of models that can be built or proposed. As the point is a very simple one, and ways of dealing with the seasonality are well understood, or are at least currently thought to be so, this case will not be pursued further. There is, however, a special case that is worth mentioning at this point. Suppose that y_t is not seasonal, but that both x_t and z_t are seasonal, then is model (1.1) a possible one, with e_t not seasonal? In general, the answer is no, but if it should happen that the term $bx_t + cz_t$ is non-seasonal, then the model (1.1) is not ruled out. This could only happen if there is a constraint $f = c/b$ such that the seasonal component in x_t is exactly the reverse of f times the seasonal component in z_t. A simple case where this would occur is if the seasonal components is x_t and z_t were identical and f takes the value minus one. This does at first sight appear to be a highly unlikely occurrence, but an example will be given later, in a very different context, where such cancellations could occur.

It is obvious that the spectrum of one side of a generating equation, such as (1.1), must be identical to the spectrum of the other side. If the spectrum of one side has a distinctive feature, it must be reproduced in the spectrum of

the other side; obvious examples being periodic components, such as the seasonal and trend terms. The majority of this paper will be concerned with discussions of this point in connection with a generalized version of the trend term.

2. Integrated series and filters

To proceed further, it is necessary to introduce a class of time series models that have been popular in parts of electrical and hydraulic engineering for some years, but which have so far had virtually no impact in econometrics.

Suppose that x_t is a zero-mean time series generated from a white noise series ε_t by use of the linear filter $a(B)$, where B is the backward operator, so that

$$x_t = a(B)\varepsilon_t, \tag{2.1}$$

with

$$B^k \varepsilon_t = \varepsilon_{t-k}.$$

Further suppose that

$$a(B) = (1-B)^{-d} a'(B), \tag{2.2}$$

where $a'(z)$ has no poles or roots at $z=0$. Then x_t will be said to be 'integrated of order d' and denoted

$$x_t \sim I(d).$$

Further, defining

$$x'_t = (1-B)^d x_t = a'(B)\varepsilon_t,$$

then

$$x'_t \sim I(0)$$

from the assumed properties of $a'(B)$. $a(B)$ will be called an 'integrating filter of order d'. If $a'(B)$ is the ratio of a polynomial in B of order m divided by a polynomial of order l, then x_t will be ARIMA (l, d, m) in the usual Box and Jenkins (1970) notation. However, unlike the vast majority of the literature on ARIMA models, the class of models here considered allow the order of integration, d, to be possibly non-integer. Clearly, not constraining d to be an

integer generalizes the class of models before considered, but to be relevant, the generalization has to be shown to be potentially important. Some earlier accounts of similar models may be found in Hipel and McLeod (1978), Lawrence and Kottegoda (1977), Mandelbrot and Van Ness (1968) and Mandelbrot and Taqqu (1979), although some details in the form of the models are different than those here considered, which were first introduced in Granger and Joyeux (1981).

Some of the main properties of these models may be summarized as follows: The spectrum of x_t, generated by (2.1) and (2.2), may be thought of as

$$f_x(\omega) = (1/|1-z|^{2d})|a'(z)|^2, \qquad z = e^{i\omega},$$

if $\text{var}(\varepsilon_t) = 1$, from analogy with the usual results from filtering theory. It is particularly important to note that for small ω,

$$f_x(\omega) = c\omega^{-2d}. \tag{2.3}$$

It was shown in Granger and Joyeux (1981) that the variance of x_t increases as d increases, and that this variance is infinite for $d \geqq \frac{1}{2}$, but is finite for $d < \frac{1}{2}$. Further, writing

$$x_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j},$$

and denoting

$$\rho_j = \text{correlation}(x_t, x_{t-j}),$$

then, for j large,

$$\rho_j = A_1 j^{2d-1}, \qquad d < \tfrac{1}{2}, \quad d \neq 0,$$

and

$$b_j = A_2 j^{d-1}, \qquad d \leqq 1, \quad d \neq 0,$$

where A_1 and A_2 are appropriate constraints. When $d = 0$, both ρ_j and b_j decrease exponentially in magnitude as j increases, but with $d \neq 0$, it is seen that these quantities decline much slower. Because of this property the integrated series, when $d \neq 0$, have been called 'long-memory'. For long-term forecasting, the low frequency component is of paramount importance and (2.3) shows that if d is not an integer, this component cannot be well approximated by an ARIMA (l, d, m) model with integer d and low order for l and m.

It is not clear at this time if integrated models with non-integer d occur in practice and only extensive empirical research can resolve this issue. However, some aggregation results presented in Granger (1980) do suggest that these models may be expected to be relevant for actual economic variables. It is proved there, for example, that if $x_{jt}, j=1,\ldots,n$, are set of independent series, each generated by an AR(1) model, so that

$$x_{jt}=\alpha_j x_{j,t-1}+\varepsilon_{jt}, \qquad j=1,\ldots,N,$$

where the ε_{jt} are independent, zero-mean white noise and if the α_j's are values independently drawn from a beta distribution on (0, 1), where

$$\mathrm{d}F(\alpha)=(2/B(p,q))\alpha^{2p-1}(1-\alpha^2)^{q-1}\,\mathrm{d}\alpha, \qquad 0\leqq\alpha\leqq 1, \quad p>0, \quad q>0, \tag{2.4}$$

then, if $\bar{x}_t=\sum_{j=1}^{N} x_{j,t}$, for N large

$$\bar{x}_t\sim I(1-q/2). \tag{2.5}$$

The shape of the distribution from which the α's are drawn is only critical near 1 for this result to hold.

A more general result arises from considering x_{jt} is generated by

$$x_{jt}=\alpha_j x_{j,t-1}+y_{j,t}+\beta_j W_t+\varepsilon_{jt}, \tag{2.6}$$

where the series $y_{j,t}$, W_t and ε_{jt} are all independent of each other for all j, ε_{jt} are white noise with variances σ_j^2, $y_{j,t}$ has spectrum $f_y(\omega,\theta_j)$ and is at least potentially observable for each micro-component. It is assumed that there is no feedback in the system and the various parameters α, θ_j, β and σ^2 are all assumed to be drawn from independent populations and the distribution function for the α's is (2.4). Thus, the $x_{j,t}$ are generated by an AR(1) model, plus a independent causal series $y_{j,t}$ and a common factor causal series W_t. With these assumptions, it is shown in Granger (1980) that (i)

$$\bar{x}_t\sim I(d_x)$$

where d_x is the largest of the three terms $(1-q/2+d_y)$, $l-q+d_w$ and $(1-q)/2$, where $\bar{y}_t\sim I(d_y)$, $W_t\sim I(d_w)$, and (ii) if a transfer function model of the form

$$\bar{x}_t=a_1(B)\bar{y}_t+\alpha_2(B)W_t+e_t$$

is fitted, then both $a_1(B)$ and $a_2(B)$ are integrating filters of order $1-q$.

It should be noted from (2.6) that, if $\alpha_j < 1$ then the spectrum of $x_{j,t}$ is

$$f_{x,j}(w) = (1/|1-\alpha_j z|^2)[f_{y,j}(w) + \beta_j^2 f_w(\omega) + f_{e,j}(\omega)],$$

so that if one assumes that $x_{j,t} \sim I(0)$ it necessarily follows that $y_{j,t}$ and W_t are both $I(0)$.

In Granger (1980) it was shown that integrated models may arise from micro-feedback models and also from large-scale dynamic econometric models that are not too sparse. Thus, at the very least, it seems that integrated series can occur from realistic aggregation situations, and so do deserve further consideration.

3. The algebra of integrated series and it's implications

The algebra of integrated series is quite simple. If $x_t \sim I(d_x)$ and $a(B)$ is an integrating filter of order d', then $a(B)x_t$ will be $I(d_x + d')$. Thus, d_x is unchanged if x_t is operated on by a filter of order zero. Further, if $x_t \sim I(d_x)$, $y_t \sim I(d_y)$ then $z_t = bx_t + cy_t \sim I(\max(d_x, d_y))$ in general. This result is proved by noting that the spectrum of z_t is

$$f_z(\omega) = b^2 f_x(\omega) + c^2 f_y(\omega) + 2bc[cr(\omega) + \overline{cr(\omega)}], \tag{3.1}$$

where $cr(\omega)$ is the cross-spectrum between x_t and y_t and has the property that $|cr(\omega)|^2 \leqq f_x(\omega) f_y(\omega)$. For small ω,

$$f_x(\omega) = A_3 \omega^{-2d_x} \quad \text{and} \quad f_y(\omega) = A\omega^{-2d_y},$$

and clearly the term with the largest d value will dominate at low frequencies. There is, however, one special case where this result does not hold, and this will be discussed in the following section.

Suppose now that one is considering the relationship between a pair of series x_t and y_t, and where d_x and d_y are known, or at least have been estimated from the data. For the moment, it will be assumed that d_x and d_y are both non-integer. If a model of the form

$$b(B)y_t = c(B)x_t + h(B)\varepsilon, \tag{3.2}$$

is considered, where all of the polynomials are of finite order, and will usually be of low order, and ε_t is white noise, independent of x_t, then this model is consistent, from consideration of spectral shapes at low frequencies, only if $d_x = d_y$. If one knows that $d_x < d_y$ then to make the model consistent, either $c(B)$ must be an integrating filter of order $d_x - d_x$ or $h(B)$ is an integrating filter of order d_x, or both. In either case, the polynomials cannot

be of finite order. Similarly, if $d_x > d_y$, the necessarily $c(B)$ must be an integrating filter of order $d_y - d_x$, and so cannot be of finite order.

As an extreme case of model (3.2) inconsistency, suppose that $d_x < \frac{1}{2}$, so variance of x_t is finite, but $1 > d_y > \frac{1}{2}$, so variance of y_t is infinite. Using just finite polynomials in the filters, clearly y_t cannot be explained by the model, if variance ε_t is finite, which is generally taken to be true. Similarly if $d_y < \frac{1}{2}$ but $1 > d_x > \frac{1}{2}$, then one is attempting to explain a finite variance series by a infinite variance one. This same problem occurs when the d's can take integer values, of course. Suppose that one knows that change in employment has $d = 0$, and that level of production has $d = 1$, then one would not expect to build a model of the form

$$\textit{Change in employment} = \alpha + \beta\ (\textit{level of production}) + f(B)\varepsilon_t.$$

However, replacing β by $\beta(1-B)$ would produce a consistent model, in the sense of this term being used here. Only with integer d values can a filter, which is a polynomial in B of finite length, be applied to a series to reduce the order of integration to zero.

Naturally, similar constraints can be derived for models involving more than one explanatory variable, although these constraints can become rather complicated if many variables are involved. As an illustration, suppose one has a single-equation model of the form

$$b(B)y_t = c(B)x_t + g(B)Z_t + h(B)\varepsilon_t, \tag{3.3}$$

where ε_t is white noise independent of x_t and z_t and d_x, d_y and d_z are assumed known and non-integer. If all of the polynomials are of finite order, then necessarily $d_y = \max(d_x, d_z)$. If this condition does not hold then, generally, at least one of the polynomials has to correspond to an integrating filter and hence to be of infinite order. When all of the d's are integer, rather simpler rules apply. However, care has to be taken in the model specification so that infinite variance variables are not used to explain finite variance variables, or vice versa. In practice, it is still not uncommon to see this type of misspecification in published research.

4. Co-integrated series

This section considers a very special case where some of the previously stated rules do not hold. Although it may appear to be very special, it also seems to be potentially very important. Start with model (3.3) and ask again, is it possible for $d_y < \max(d_x, d_z)$. For convenience, initially the no-lag case $c(C) = c$, $g(B) = g$ is considered, so that

$$b(B)y_t = cx_t + gz_t + h(B)\varepsilon_t, \tag{4.1}$$

where $d_y > 0$, $h(B)\varepsilon_t$ is $I(d_y)$ and $\text{var}(\varepsilon) = 1$. The spectrum of the right-hand side will be

$$\{c^2 f_x(\omega) + g^2 f_z(\omega) + gc[cr(\omega) + \overline{cr}(\omega)]\} + |h(z)|^2/2\pi, \tag{4.2}$$

where now $cr(\omega)$ is the cross-spectrum between x_t and y_t. The special case of interest has:

(i) $f_x(\omega) = \alpha^2 f_z(\omega)$, ω small, so $d_x = d_z$,
(ii) $cr(\omega) = \alpha f_z(\omega)$, ω small, so that the coherence $C(\omega) = 1$ and the phase $\phi(\omega) = 0$ for ω small.

A pair of series obeying (i) and (ii) will be called *co-integrated.*

If further, $g = -c\alpha$, the part of the spectrum (4.2) inside the main brackets will vanish at low frequencies and so a model of the form (4.1) will be appropriate even when $d_y < \max(d_x, d_z)$ in this special case. It is seen that in this case the difference between two co-integrated series can result in an $I(0)$ series. A slightly more general result arises from considering $x_t = z_t + q_t$, where $d_x = d_x$, $d_q < d_x$, and z_t, q_t are independent, then x_t and z_t will be co-integrated, but the difference $x_t - z_t$ will be $I(d_q)$. It should be noted that if a pair of series x_t, z_t are co-integrated, then so will be $a(B)x_t$, $b(B)z_t$ where $a(B)$, $b(B)$ are any pair of finite lag filters; thus, in particular if x_t and z_t are co-integrated then so will be x_t and z_{t-k} for all k, although the approximation that the phase is zero at low frequencies may become unacceptable for large values of k.

Co-integrated pairs of series may arise in a number of ways, for example:

(i) If x_t is the input and z_t the output of a black box of limited capacity, or of finite memory, the x_t, z_t will be co-integrated, for instance the series might be births and deaths in an area with no immigration or emigration, cars entering and leaving the Lincoln Tunnel, patients entering and leaving a maternity hospital, or houses started and houses completed in some region. For these examples to hold, it is necessary to have $d_x > 0$.
(ii) Series for which a market ensures that they cannot drift too far apart, for example interest rates in different parts of a country or gold prices in London and New York.
(iii) If $f_{n,h}(J_n)$ is an optimal forecast of x_{n+h} based on a proper information set J_n, so that J_n includes x_{n-j}, $j \geqq 0$, then x_{n+h} and $f_{n,h}$ are co-integrated if $d_x > 0$. Thus, if 'unanticipated money supply', $x_{n+1} - f_{n,1}$, is used in a model, this can be appropriate if the variable being explained has $d_x > 0$.

It should be emphasized that for this result to hold $f_{n,h}$ must be on optimal forecast and, if d_x is not an integer, then this means that in theory the forecast has to be based on an infinity of lagged x's. If $1 > d_x > 0$, but an ARIMA (l, d, m) model is used to form forecasts, with integer d, the forecasts will not be optimal and the series and its forecast will not be co-integrated.

There obviously are pairs of economic series, such as prices and wages, which may or may not be co-integrated and a decision on this has to be determined by an appropriate theory or an empirical investigation. It might be interesting to undertake a wide-spread study to find out which pairs of economic variables are co-integrated.

In the frequency domain, the conditions for co-integration of two series state that the two series move in a similar way, ignoring lags, over the long swings of the economy and in 'trend', although the idea of trend is rarely carefully defined and will here mean just the very low frequency component. Although the two series may be unequal in the short term, they are tied together in the long run.

The use of the difference between two series to explain the change in a series has been suggested by Sargan (1964) and Hendry (1978) and implemented in a number of models, particularly in Britain. An example is a model of the form

$$a(B)\Delta y_t = b(B)\Delta x_t + \beta(y_{t-1} - x_{t-1}) + e_t,$$

and the use of the term $\beta(y_{t-1} - x_{t-1})$ has been found in some cases to produce a better model, in terms of goodness of fit. The form of the model has an important property. The difference equation without the innovations e_t,

$$a(B)\Delta y_t = b(B)\Delta x_t + \beta(y_{t-1} - y_{t-1}),$$

is such that if x_t and y_t each tend to equilibrium, so that $\Delta x_t \to 0$ and $\Delta y_t \to 0$ then x_t and y_t tend to the same equilibrium level. When the stochastic elements e_t are present, equilibrium becomes much less meaningful, and is replaced by x_t and y_t tending to having identical means, assuming the means exist. However, if the d values of x_t is greater than d_e, this generating model ensures that x_t and y_t will be co-integrated. They, therefore, will move closely together in the long run, which is possible the property that most naturally replaces the concept of equilibrium for stochastic processes. It is important to note that this property does not hold if $d_x = d_y = d_e$, as then the coherence between x_t and y_t need not be high at low frequencies, depending on the relative variances of e_t and y_t.

5. Conclusion

Having, I hope, made a case for the prior analysis of time series data before model specification inter-relating the variables, it has now to be admitted that the practical implementation of the rules suggested above is not simple. Obviously, one can obtain satisfactory estimates of the spectrum of a series, but it is not clear at this time how d values should be estimated. In the references given earlier, a variety of ways of estimating d are suggested, and a number of sensible modifications to these can easily be proposed, but the statistical properties of these d estimates need to be established. It is possible that too much data is required for practical use of the specification rules or that d values for real economic variables are all integers. Only further analysis, both theoretical and empirical can answer these questions.

References

Box, G.E.P. and G.M. Jenkins, 1970, Time series analysis, forecasting and control (Holden-Day, San Francisco, CA).

Davidson, J.E.H., D.F. Hendry, F. Srba and S. Yeo, 1978, Econometric modelling of the aggregate time-series relationship between consumer's expenditure and income in the United Kingdom, Economic Journal 88, 661–692.

Granger, C.W.J., 1980, Long-memory relationships and the aggregation of dynamic models, Journal of Econometrics 14, 227–238.

Granger, C.W.J. and R. Joyeux, 1981, An introduction to long-memory time series and fractional differencing, Journal of Time Series Analysis V.1.

Hipel, W.H. and A.I. McLeod, 1978, Preservation of the rescaled adjusted range, Part 1, Water Resources Research 14, 491–518.

Lawrence, A.J. and N.T. Kottegoda, 1977, Stochastic modelling of river-flow time series, Journal of the Royal Statistical Society A140, 1–47.

Mandelbrot, B.B. and J.W. Van Ness, 1968, Fractional Brownian motions, fractional noises and applications, Siam Review 10, 422–437.

Mandelbrot, B.B. and M.S. Taqqu, 1979, Robust R/S analysis of long-run serial correlation, Research report RC 7936 (IBM, Yorktown Heights, NY).

Sargan, J.D., 1964, Wages and prices in the United Kingdom: A study in econometric methodology, in: P.E. Hart, G. Mills and J.K. Whitacker, eds., Econometric analysis for national economic planning (Butterworths, London).