Comments on Modeling and Interpreting Economic Relationships

by David A. Pierce

I found the concept of "structure" brought forth in Professor Sims' paper to be very helpful in clarifying my ideas on a word about which I have often been confused. It seems that the words "structure" and "structural" are used in two different ways, and I found a sentence in Chris Sims' paper that illustrates this point very well: "No characteristic of a relation's structure can guarantee that the relation is structural." I think this distinction does reflect the way we ordinarily use these words. For in referring to an econometric model as a structural model, we might have in mind the way the model is put together or we might be referring to the way that the model is intended to reflect reality.

One of the uses to which Sims puts this concept of a structural relation is to provide a framework for assessing whether observed causal orderings ("causal" in Clive Granger's sense) are spurious. Sims discusses two general situations where spurious orderings are possible: one due to a common association with a third variable and the other due to the possibility of feedback control by a policy maker. His conclusion that a common dependence of two variables on a third does not in a natural way tend to give spurious orderings seems to me to strengthen the usefulness of those methodologies – such as that employed by Granger and Paul Newbold – which are most suited to bivariate relationships. One should reasonably expect, based on these results, to be able to analyze the pairwise groupings of a set of variables as a first step toward building an interactive model for them – whatever mixture of time series and econometric methodologies is employed in doing so.

However, the result on feedback control seems to me to be much more limiting. Professor Sims has given some fairly general conditions under which any variable which the fiscal or monetary authorities are capable of reacting to, and therefore influencing, is vulnerable to a spurious ordering. For example, if interest rates are raised to curtail money growth, if money growth is curtailed with the aim of lowering

inflation, or if money growth or the size of the government budget (or the government deficit) are raised with the aim of raising GNP or lowering unemployment, then we are in danger of "finding" spurious causal orderings.

Thus, perhaps we ought to consider whether the fact that policy makers have indeed made policy during virtually the entire postwar period places a rather severe limitation on the value of postwar macroeconomic data in building structural models or, for that matter, in building time series models whose usefulness is to be something more than forecasting under the assumption of unchanged policy rules. So I gather I am more pessimistic than is Chris Sims regarding the implication of this particular result for the empirical modeling of what otherwise would be potentially structural relationships.

One further comment I have on the subject of structural relations is that the empirical presence of *instantaneous* Granger causality places a further limitation on our ability to infer anything concerning a system's "structuralness" or even concerning statements that a variable is exogenous. On the other hand, it would seem that the presence of *simple* causality only could constitute a more direct verification of the notion of exogeneity.

Turning to the paper by Professors Granger and Newbold, quite a number of interesting issues have been discussed; and while I think that some of the more flagrant methodological abuses illustrated are not as prevalent as they were even a few years ago, I think it is still worthwhile having their consequences brought to our attention. However, the sensible and direct procedures discussed and employed in the paper have been and continue to be surrounded by some controversy — more so than they should be in my opinion — and what I want to do is to take a further look at three of these procedures: first, series differencing; second, bivariate model building through examining univariate residual cross correlations; and third, empirically measuring the strength of economic relationships.

One recommendation that is made by Granger and Newbold is, in the absence of contrary empirical evidence, to use first differences of the data. I personally think this is a good idea, as the evidence in most macroeconomic data (exceptions include excess reserves) is overwhelmingly in favor of the hypothesis of nonstationarity or near-nonstationarity in series levels. I say near-nonstationarity because it is probably impossible to distinguish whether, for example, in a first-order AR process, the parameter is actually 1 or something close to but less than 1, say .95.

On the other hand, differencing is not necessary in these situations. Simply because of this indistinguishability, a prefilter of $(1 - .95L)$ would be expected to have about the same effects as the differencing filter $(1 - L)$, and in fact, the prefilters used by John Geweke in his paper, which are of the form $(1 - aL)^2$, may well approximate filters including a first difference operation together with a further autoregression. But I think that first differencing, if warranted, is on the whole easier to work with and to interpret.

In any event, those cases where first differencing is clearly contra-

indicated can almost always be discovered empirically. For example, if an already stationary series is differenced, the resulting series' moving average representation is noninvertible; thus, an estimated moving average on differenced data with a root close to 1 may indicate overdifferencing. Alternatively, an autoregression fitted to the *un*differenced series might contain a factor $1 - \alpha L$ for $\alpha \approx 1$, if differencing *is* appropriate. I would be curious to know whether the fifth order *AR*'s in the paper by Tom Sargent and Sims ever showed something like this. (As a further comment concerning their prewhitening, I would be interested in seeing tests of fit on these *AR* models, for example, based on the residual autocorrelations, as with quarterly data I have usually found moving average or mixed *ARMA* models to be appropriate, which in some cases would require longer than fifth-order autoregressions to approximate.)

Another procedure employed by Granger and Newbold is in the use of the cross correlations of residuals from fitted univariate time series models in modeling multivariate, particularly bivariate, relationships. A problem with this procedure, of which they are aware in referencing a paper by Sims where the subject is discussed, is that significance is underestimated in hypothesis tests of, say, feedback when one-way causality is present if those tests are constructed under the pretention that the residuals are calculated from true rather than estimated univariate models. Thus, a different procedure needs to be used for a significance test of feedback, and several such tests are available. A point noted by Granger and Newbold in their unemployment-industrial production index example is that an asymptotically valid test has been obtained for the fitted transfer function model. But at the model-*building* stage, the issue seems to be less important. The residual cross correlations are consistent estimates of the population cross correlations, the asymptotic bias being in their variance only; and since this is a *downward* bias, we have a tighter *upper* bound to their magnitudes as well as a tighter *lower* bound. What this means is that if we are more suspicious that a feedback relationship exists, we are also more confident that this relationship is a weak one.

But perhaps this raises the issue of whether a weak relationship is something we should go after. My view on this is that if we were to forecast a variable Y and if modeling its relationship with another variable X would reduce the standard error of such a forecast by, say, 1 or 2 percent, then this relationship would be a weak one and hardly something to get excited about – even though in a sufficiently large sample it may be highly statistically significant.

This brings me to my final point, concerning measuring economic relationships. Granger and Newbold have noted that the usual R^2 measure often tends to be inflated as a result of improperly accounting for autocorrelation problems. I have proposed [118][†] carrying this notion to what I think is its logical conclusion: that is, any variation in the dependent

†Numbers in [ ] correspond to reference list, p. 219.

variable Y reported as having been explained by X, but which could equally well have been explained by past Y, should be eliminated from the R^2-measure. Thus, a "net" R^2 can be defined as (our estimate of) the percentage reduction in the *innovation* variance of Y, achievable from modeling the X-Y relationship. There are variations on this—for example, whether variance reduction achievable from knowing present X, or present and future X, is included in the measure — but the idea is always not to measure as part of an *inter*-variable relationship anything that is actually an *intra*-variable relationship.

When we apply these calculations, however, the resulting net R^2's often turn out to be rather low. For example, in the unemployment-industrial production index example of Granger and Newbold this net R^2 is .22. (In their analysis of the St. Louis Model, Granger and Newbold occasionally found sharp decreases in forecast variance possible. It would be interesting to see how their smallest forecast variances compared with those from *ARIMA* models fitted to the endogenous variables themselves.) Moreover, my experience [119] is that even this is higher than what we're likely to see with many economic time series, even with those that are supposed to be strongly related. One way of stating this is that once own-series effects are taken into account, inter-series relations tend to be weak, a phenomenon that is also observed in the paper by Sargent and Sims. There are several possible explanations of why this tendency toward weak relationships in macroeconomic data should exist [119], but perhaps the failure of economic relations to be structural in the sense of Sims' paper in this volume is an important factor. In this regard I think that Lucas's argument, summarized in Sims' paper, is a very interesting one; for if fixed-parameter systems are inappropriate, then both the time series and econometric approaches are possibly in need of overhauling.