Pakistan Journal of Applied Economics, (1987), Vol. VI No.2, (125 – 129)

SHORT-TERM FORECASTING: An Application of Box-Jenkins Methods – A Reply

A. Hafeez SHAIKH* and Asad ZAMAN**

The availability of computer packages has put a variety of econometric techniques within the reach of nearly everyone. The sophisticated mathematics and modern technology required lends a magical quality to these tools. One has the feeling that if only one could pronounce the right magic word (or, use the right regression technique) all econometric problems would disappear. The truth, as one learns from experience, is far from this fantasy. Because strong assumptions are made by econometric theory, it is not at all uncommon for the most sophisticated of techniques to generate the crudest of errors. Therefore it is extremely important to understand:

a) The overall limitations of econometric techniques: what can and cannot be done; and
b) factors not included in the model (such as politics, strikes, outliers) which may influence results.

In the case at hand, we faced the problem of forecasting rice exports for the next five years, given data on the previous twenty four. Before beginning any formal modelling, it is essential to examine the data visually. In fact, this point, often ignored in econometrics texts, cannot be overemphasized: LOOK AT THE DATA! One might say that a picture is worth a thousand regressions. Unfortunately most regression packages make it easier to run regressions than to plot the data. As a consequence, it is entirely typical to find researchers who have run numerous regressions without ever examining the data. There are many data problems which occur frequently and are easily picked up by visual inspection, but do not show up on an output of regression statistics. A simple case is that in which the entire regression depends solely on one or two anomalous observations. One of us recently refereed a paper which claimed a growth rate of 5 per cent for a certain data series. On examining the data series, it became clear that in fact the series was more or less constant over the entire time period, but had a jump near the end. A similar situation arose in our attempt to forecast cotton exports [Shaikh and Zaman (1983), the same paper that Mahmud and Nishat (1987) have attempted to improve upon] A plot of the data clearly revealed that due to the presence of outliers [also called 'leverage points'; see Belsley, Kuh, and Welsch (1980)], usual regression techniques would give misleading results. Thus we were forced to use more ad-hoc techniques to forecast cotton exports.

Outliers are not a serious problem in the rice export data. However a plot of the data does reveal several important and interesting features. The first and obvious one is the clear increasing trend. In the presence of such strong visual evidence, it is unnecessary to run autocorrelation tests to establish nonstationarity.[1] We also remarked in our original paper on the rather high variance around the trend. Indeed the data appears to display heteroscedasticity; the variance increases over time. This preliminary evaluation is extremely important to answering question (a) above: what is the best we can expect to do? If rice exports have behaved erratically over the past, they are clearly not going to stabilize in the future simply to make our forecasts more precise. Because out-of-sample predictions necessarily have higher variance than that of the data (after centering around an estimated model), we expect forecasting errors of at least the same order as the data variance not explained by the model. Judged by these criterion, both the Mahmud-Nishat forecasts and our model forecasts fall within one standard deviation of the actuals and hence are good.

An important question, with a surprising answer, is whether or not additional accuracy (a bull's eye hit) adds validity to the model. Mahmud and Nishat (1987) have been deceived by the apparently obvious idea that the more accurate a forecast, the better the model. We hope to achieve accuracy to within one standard deviation say; we know that greater accuracy can only be achieved accidentally. In fact, the following paradox is sometimes discussed in econometrics courses. Suppose a regression model with $R^2 = 0.8$ generates exact forecasts for say the next five periods. Does this validate the model? The answer is no; the model *cannot* be correct, since the model predicts the error variance to be greater than

[1] Mahmud and Nishat are confused about our 'totally ignoring ... non-stationarity', perhaps because we made no formal tests but simply state that 'Glancing at ... basmati exports, we discover a clear upward trend'. Instead of treating econometric procedures as mysterious black boxes, we strongly recommend researchers to get into the habit of looking at the data and predicting regression results before the computer run. While this is impossible in some situations, it is a very useful tool for developing insight and intuition into the simpler cases.

zero, while it is observed to be zero. If such an event happened it would be necessary to look for other factors beyond the regression to explain why the results were significantly better than predicted!

As just explained, in terms of forecasting performance for two periods, there seems to be no (statistically significant difference between our model and the proposed alternative by Mahmud and Nishat (1987). However, on other grounds, it is quite clear that the Mahmud-Nishat model is highly implausible. We now explain why this is so. First, the process of model selection involved in choosing an appropriate ARMA or ARIMA model causes numerous well-known biases. As a simple illustration, suppose that we try twenty regressions of the type $Y = \alpha+\beta X_i$, where we experiment with twenty different explanatory variables X_i to see which one explains Y the best. Even if all the variables X_i are completely independent of Y, one of these regressions is likely to be significant at the 95 per cent level just by chance. This may mislead us into believing that there is a strong relation between the X_i which yields significant results and Y. This is known as a 'spurious fit'. Although it is common practice, any experimentation with the data in terms of fitting different regressions and searching for a 'best' one causes serious difficulties in interpretation of results, and invalidates the usual regression statistics.[2] For a discussion of the difficulties with model selection, further references, and a new approach to the problem, see Zaman (1984).

Unfortunately, Mahmud and Nishat (1987) did not translate and interpret their estimated model, which is in terms of differences, into the original levels.[3] Had they done so, they would have arrived at the following estimated equation for forecasting:

$$B_t = 17.74 + 0.28B_{t-1} - 0.14B_{t-2} + 0.36B_{t-3} + 0.5B_{t-4} + u_t$$

Here B_t is Basmati exports in period t and u_t is the residual ($u_t = 0.6A_{t-2} + 0.5A_{t-4}$). The meaning of the dynamic structure of this model is clarified by the procedure of back-substitution. At the first stage,

[2] We noted in our original paper that high order ARMA models (like the ones used by Mahmud and Nishat) were likely to generate spurious fits. An amusing illustration of this problem in popular terms is the 'monkeys paradox': A bunch of monkeys randomly pounding on a typewriter will eventually, with probability one, type out the complete works of Shakespeare. That is, sheer luck will generate a good fit/forecast, if we try enough models.

[3] We are inclined to attribute responsibility for this failure to deficient econometrics texts, many of which create the impression that having obtained a regression with a reasonable fit is the end of the econometrician's task, rather than the beginning.

substitute for B_{t-1} its equivalent, i.e. $17.74 + 0.28B_{t-2} - 0.14B_{t-3} + 0.36B_{t-4} + 0.5B_{t-5} + u_{t-1}$. This leads to

$$B_t = 2 \times 17.74 - 0.06B_{t-2} + 0.32B_{t-3} + 0.60B_{t-4} + 0.14B_{t-5} + u_t$$

Similar substitution for B_{t-2} leads to

$$B_t = 3 \times 17.74 + 0.30B_{t-3} + 0.61B_{t-4} + 0.12B_{t-5} - 0.03B_{t-6} + u_t$$

Finally, substituting for B_{t-3} gives us

$$B_t = 4 \times 17.74 + 0.69B_{t-4} + 0.08B_{t-5} + 0.08B_{t-6} + 0.152B_{t-6} + u_t$$

Note that according to these equations, the long run growth trend in Basmati exports is only 17.74 units per period of time. Surely this is much too low an estimate, and in itself sufficient ground to reject the equation. A second peculiarity is revealed by the weights of the leading terms in the four equations, which reveals the full effect of changes in exports at the particular lag value. The first equation shows that the impact of last year's exports is only 28 per cent of current exports. Similarly, exports two years ago enter with a weight of −0.06; a small negative effect. Exports three years ago have a 30 per cent impact on current exports. However, exports four years ago enter with the enormous coefficient of 69 per cent. Effectively, according to this equation, the major portion of current exports is determined by the exports of four years ago! It is not necessary to do diagnostics or undertake forecasting to reject this equation for explaining Basmati exports. The short term good performance of the Mahmud-Nishat model is clearly accidental. We do not in fact have the data available to us, but we can confidently predict that the Mahmud-Nishat model will not work over any serious period of time.

Finally, we would like to conclude by discussing some technical issues about which Mahmud and Nishat have serious misunderstandings. They seem to believe that differencing is necessary in order to estimate autoregressive models for nonstationary time series. In fact, if one is using estimation techniques appropriate to stationary time series (such as Box-Jenkins), it is indeed necessary to remove the trend or otherwise achieve stationarity before beginning analysis. However, even here, differencing is not the only way to do so. In fact, it is appropriate only if the original series has a unit root [see, for example, Bhargava (1986)]. There are numerous reasons, other than a unit root, why a series may be nonstationary. For example, if the series has seasonality or time trends, differencing would be an inappropriate method of removing the trend, and would generate serious problems. The second issue is that autoregressive schemes can be (and indeed are commonly) used for nonstationary series. The theoretical autoregressive time series extends backwards in time to $-\infty$. Such a series must necessarily be stationary. However the simple expedient of permitting the series to start at some finite point in time permits the use of nonstationary autoregressive series. For a simple example see Berenblutt and Webb (1973). A third issue is that it has been learned from experience that high order ARMA models only work well with long time series. In his Ph.D. thesis, Adrian Pagan reports that he was able to discriminate between AR(1) and MA(1) processes reasonably well only with time series of length 100 or more. A time series of 24 observations is much too short for serious applications of Box-Jenkins type techniques.

**University of Karachi and Howard University*
***Columbia University*

References

Belsley, David A., Edwin Kuh, and Roy E. Welsch, 1980, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, Wiley, N.Y.

Berenblutt, I.I., and G.I. Webb, 1973, A new test for autocorrelated errors in the linear regression model, Journal of the Royal Statistical Society, Series-B, 35 : 33-50.

Bhargava, Alok, 1986, On the theory of testing for unit root in observed time series, Review of Economics and Statistics, 53(3): 369-384

Mahmud, F., and M. Nishat, 1987, Short-term forecasting: An application of Box-Jenkins Methods, Pakistan Journal of Applied Economics, VI(1): 61-65.

Pagan, Adrian, 1972, A study of estimation procedures for time series models, Ph.D. Dissertation, Australian National University.

Shaikh, Abdul Hafeez, and Asad Zaman, 1983, Forecasting without theory: Pakistan's export of cotton and rice, Pakistan Journal of Applied Economics, 2 (1) : 65-84.

Zaman, Asad, 1984, Avoiding model selection by the use of shrinkage techniques, Journal of Econometrics, 25 : 73-85.