.

Estimating Short and Long Run Relationships: A Guide to the Applied Economist

B. Bhaskara Rao

rao_b@usp.ac.fj

University of the South Pacific, Suva (Fiji)

Abstract

Many applied economists face problems in selecting an appropriate technique to estimate short and long run relationships with the time series methods. This paper reviews three alternative approaches viz., general to specific (GETS), vector autoregressions (VAR) and the vector error correction models (VECM). As in other methodological controversies, definite answers are difficult. It is suggested that if these techniques are seen as tools to summarize data, as in Smith (2000), often there may be only minor differences in their estimates. Therefore a computationally attractive technique is likely to be popular.

JEL Classification: C230, C510, C520, C590.

Keywords: General to Specific Models, Cointegration, VAR Models, Alternative Methods, Short and Long Run Equations.

1. INTRODUCTION

Methodological debates are not new in economics. In all such debates it is difficult to reach definite conclusions because empirical verifications seldom consistently favour one particular approach. For example, there is a considerable volume of empirical evidence to support theories and policies based on the free market frictionless models. In contrast, there is also a large body of evidence to validate models based on a variety of imperfections. These alternative approaches have different policy implications. Consequently, it may be said that our preference for a paradigm is perhaps influenced by our preference for its policy implications, although we warn at the outset that theories should be evaluated on their theoretical and empirical merits and not on the basis of their policy implications. Currently a similar methodological debate exists on the relative merits of the statistical techniques for the estimation of the time series models to forecast and estimate theoretical relationships. Although this new debate has more positive aspects, it appears that it is also difficult to evaluate their relative merits.

In examining this new methodological debate, it is useful to draw a distinction between three stages of a research programme, as suggested by Smith (2000).

1. Purpose (or objective)
2. Summary of facts and
3. Interpretation of facts.

Within this three fold classification, statistical techniques are seen as tools to develop credible summaries of the observed facts. Purpose and interpretation are left to the analyst according to his/her prefered economic theories. Therefore, in evaluating the relative merits of the alternative statistical techniques, it is important to ask how good a particular technique is for summarizing the observed facts. It is unlikely that these alternative techniques provide conflicting summaries, at least qualitatively, with the same set of observed facts. They may differ, however, in their precision and perhaps only marginally. If so, as Granger (1997) has observed, eventually a computationally simpler technique with an acceptable degree of precision will be widely used. The more demanding and the high precision techniques will also be used in specialized applications. A typical example is the frequently used Cobb-Douglas production function, in comparison to the translog function, in a large number of studies on technical progress.

Since there are at least three statistical technique to estimate long and short run relationships in economics, it is important to ask: Which is simpler and how good is its precision. At times it is also necessary to check whether alternative techniques yield conflicting summaries of facts by using two or three different techniques. Seen this way, the new controversy is, predominantly, a positive issue. But, at the end of the day, the applied economist is likely to select a technique that is simple. The outline of this paper is as follows. Section 2 presents some conflicting positions of the proponents of the alternative approaches to highlight the philosophical nature of this controversy.[1] Section 3 briefly reviews the three alterna-

---

[1] This way of looking at this controversy, from a philosophical perspective, can be justified with the following observation by Granger (1997).

> "⋯ the actual economy appears to be very complicated, partly because it is the aggregation of millions of non-identical, non-independent decision-making units, ⋯ A further practical problem is that the observation period of the data does not necessarily match the decision making periods (temporal aggregation). It can be argued that even though the quantity of data produced by a macroeconomy is quite large, it is still quite insufficient to capture all of the complexities of the DGP. The modeling objective has thus to be limited to providing an adequate or satisfactory approximation to the true DGP. Hopefully, as model-

tive approaches and in section 4 we consider some guidelines for selection. Section 5 concludes.[2]

2. A FLAVOUR OF THE DEBATE

The three frequently used techniques are:

1. The LSE general to specific approach (GETS) is based on the pioneering works of Saragan and Hendry. It is also referred to as the LSE approach; see Hendry, Pagan and Saragan (1984) and Hendry (1987). This approach, in spite of its simplicity and a good methodological basis, is not widely popular and often criticized for some weaknesses.
2. Following Johansen (1988) this approach insists on pretesting for unit roots, testing for the existence of cointegration vectors and using the causality tests to estimate scaled-down VAR models. For convenience, this approach is referred to as the cointegrating vector error correction model (VECM). In contrast to GETS, this is very popular.

---

ing technology improves and data increases, better approximations will be achieved, *but actual convergence to the truth is highly unlikely."* p.169, my italics. In light of this observation it may be said that no matter how complicated is our technique, we may never know the truth. Therefore, a simpler method with good and acceptable degree of precision is likely to be widely used.

[2] At the outset a limitation of this article should be noted. The widely used current techniques viz., GETS, VAR and VECM are all based on autoregressive (AR) formulations. This may not satisfy some specialist time series econometricians who favour ARIMA or ARMA formulations; Harvey (1997). For a good survey of the historical developments in various econometric techniques see Section V in Pesaran and Smith (1992). They also briefly comment on the problems in using ARIMA formulations in a systems framework.

3. The vector autoregression (VAR) approach follows the pioneering work of Sims (1980) and is the dominant approach in the USA. There is also a more sophisticated sub-group, known as the structural VAR (SVAR) group, which is becoming prominent and has implications for how economic theories should be developed. The early VAR models were used mainly to forecast variables and not for interpretation. Therefore, VAR is criticized as atheoretical. The SVAR models are developed to identify the implied structure of the underlying theoretical models, but as yet these models are in an early stage of development. VAR Modeling is highly popular with the North American researchers, especially for forecasting of economic and financial variables.

It is hard to evaluate the relative merits of these three approaches because each proponent is critical of others. For example, Sims (2000, p.238), in his comments on VECM, says that the power of the unit root tests is low. Therefore, it is dangerous to insist on pre-testing the data for unit roots and then develop models of summaries to satisfy the traditional, but now naive, Neymann-Pearson test procedures to test the significance of the estimated coefficients, e.g. the coefficients in the cointegrating vectors. Instead he advocates the use of the likelihood functions of the alternative models to determine their relative importance.

Hendry (2000, p.241), the proponent of GETS says on the Sims VAR approach that:

> "Now let's get on to Sims's idea that the Sims $\cdots$ theorem allows one only to throw all the garbage into a model and leave it hanging around like that and then make inference. Now we know that these things do indeed converge a bit faster, so sometimes it is not so serious. But in fact, multicollinearity, in the old-fashioned sense, occurs appallingly in models in which you have eight or nine levels-co integrated variables. You just get uninterpretable effects. Now

> one of the beauties of cointegration is that certain restrictions on certain parts of the cointegration space, $\cdots$, can have very profound implications for restrictions that you are trying to put on elsewhere, and sometimes you simply cannot put them on at all."

Commenting further on the use of cointegration, Hendry (ibid) says that

> " I actually thought cointegration was so blindingly obvious that it was not even worth formalizing it. You could get the mortgage stock going off to infinity, the deposits of the building societies going off to infinity, but the ratio between them would stay essentially constant at 90 per cent. I still think, in that sense, it is completely trivial but it is very, very interesting because these are the things that, $\cdots$, are equilibria in the sense that (a) they are the targets agents are trying to achieve and (b) when they get there they will stay there and when they are not there they will try to move there. They are certainly not error-correction mechanisms, as they are still often called in the literature. They have the property that, in forecasting models, for example, if you have built them in, and if the equilibrium shifts, your forecasts are appalling for ever."

Now consider what the creator of VAR approach has to say on the concept of cointegration and therefore on CIVAR/VECM and by implication GETs. Sims (2000, pp.238-9) says that:

> "The approach of bringing in theory in a casual and naive way, $\cdots$ is really a bad approach. We see, for example, a cointegrating relationship between M, P, Y, and r, presented and then the possibilities discussed of treating it as a money-demand relation or, if the sign on the coefficient on r is wrong, as a money supply relation. This is really no different from saying, 'We are going to regress p on q or q on p and if the coefficient comes out negative we will call it a demand curve and if it turns out positive we will call it a supply curve. If somebody did that, we would all recognize that this is a fallacy, a naive way to proceed. There is an identification problem, and once there is an identification problem probably a regression of q on p is neither demand nor supply, and if there is a situation where

> you have a doubt you cannot solve it just by looking at the sign of a coefficient. This is just as true in a cointegration relationship as in an ordinary regression. *Cointegration analysis is of no help in identification.* That is my view." (my italics).

Hendry is also critical on the claims that the SVAR models can identify, with any degree of accuracy, the underlying structural parameters of the models. He says that

> "Where this gets to its most ludicrous, in my view, is using identified VARs trying to interpret shocks. You could have a completely structural model of the economy and the shocks not be structural and never interpretable. I think identification problems occur in that literature because you are trying to identify the unidentifiable. Shocks are made up of everything that is missing from the model. They are derived from the properties of the specification, from the measurement structure of the data, from the information you have used and the restrictions you have put on, and in my view will never be structural and never identifiable. "

Although Hendry defends the LSE approach against the US VAR and VECM approaches, the LSE approach, which can be also interpreted to be consistent with VECM, is criticized because no formal cointegrating tests are conducted on the variables in the error-correction part of GETS. To further complicate the problems of evaluations, some established model builders, using the Cowles Commission approach, have started reestimating their models with the standard classical methods and by altogether ignoring the significance of the unit roots revolution. Ray Fair (2004, p.5), for example, says that

> "The econometric assumption is made that all variables are stationary around a deterministic trend. If this assumption is wrong, the estimated asymptotic standard errors may be poor approximations to the true standard errors. One way to examine the accuracy of asymptotic distributions is to use a bootstrap procedure $\cdots$".

These three alternative methods are based on the autoregressive (AR) formulations and some time series specialists take an entirely skeptical view of their usefulness. Instead they suggest using autoregressive integrated moving average (ARIMA) formulations. Harvey (1997, p. 199), for example, says that

> "··· many applied economists resort to fitting a vector autoregression. Such a model is usually called a VAR. To many econometricians, VAR stands for 'Very Awful Regression'. (I am indebted to Arnold Zellner for introducing me to this delightful terminology.) ··· However, they become a little more respectable if modified in such a way that they can embody cointegration restrictions which reflect long run equilibrium relationships. The vector error correction mechanism (VECM) has been very influential in this respect as it enables the researcher to make use of the procedure devised by Johansen (1988) to test for the number of co-integrating relationships. ···
>
> "There are a number of reasons why my enthusiasm for VAR-based cointegration methods is somewhat muted. The fundamental objection is that autoregressive approximations can be very poor, even if a large number of parameters is used. One of the consequences is that co-integration tests based on autoregressive models can, like unit root tests, have very poor statistical properties. ··· However, casting these technical considerations aside, what have economists learnt from fitting such models? The answer is very little. 1 cannot think of one article which has come up with a co-integrating relationship which we did not know about already from economic theory."

Given such handline methodological positions of the leading econometricians, it is hard to evaluate their relative merits with a non-controversial procedure. Consequently, often there are appeals to *the authority* to defend one's methodological choice. A pragmatic choice, therefore, is to use more than one technique to summarise the data and examine if there are any major conflicts in their findings.

3. ALTERNATIVE TECHNIQUES

In this section we briefly explain, with simple examples, the aforesaid three widely used approaches based on autoregressive (AR) formulations. However, it should be also noted that there are a few relatively less frequently used techniques viz., the Engle-Grager (1987) two-step procedure, the Phillips-Hansen (1990) fully modified OLS estimators and the bounds test of Pesaran and Shin (1995) and Pesaran, Shin and Smith (2001). These are also based on AR formulations. It would be interesting to examine how well they all perform at the empirical level, but we limit the scope of this article to an examination of GETS, VAR and VECM.

At this stage of development, there do not seem to be any well developed and widely usable techniques, based on ARIMA formulations, to test economic propositions and eventually develop structural models. However, ARIMA equations are widely used for forecasting purpose. We first explain GETS and then VAR and VECM. GETS is explained in some detail because some of that will be also useful to understand VECM and VAR. Furthermore, GETS is the earliest approach to address the methodological conflicts between the equilibrium nature of economic theory and the data generated from the disequilibrium real world.

3.1 GETS Approach

The LSE GETS approach is based on a long standing tradition and was developed before the present developments in time series econometrics had much impact. GETS does not conflict with the traditional Cowles Commission approach, because it is essentially an alternative method of estimation of the dynamic structural equations. Econometricians at the LSE were concerned with a methodological conflict between the static equilibrium nature of economic theories and the data used to test and estimate the theoretical propositions. There is a conflict because the data used to estimate the theoretical relationships are generated in the real world that is seldom in a state of equilibrium and economic theory seldom yields any guidelines on the dynamic adjustments in the theories. This conflict was patched up in the past with arbitrary lag structures like partial adjustments and Almon lags. The late Professor Saragan took the view that it is appropriate to determine the dynamic adjustment structure using the data so that it is consistent with the underlying data generating process (DGP). This is a pragmatic solution to reconcile an irreconcilable methodological conflict between theory and data and attracted a lot of interest. GETS can be illustrated with a simple example.[3] Suppose, the theory implies that there is a relationship between consumption (C) and income (Y) i.e., say

$$C = \alpha_0 + \alpha_1 Y. \tag{1}$$

Since this is an equilibrium relationship, the DGP consistent dynamic adjustment equation can be searched by starting first with a very general and elaborate specification. Subsequently, this initial general specification is termed as the general unrestricted model (GUM). A good GUM for the consumption equation could be as follows:[4]

[3] We are using a version that was developed after the unit roots revolution.

[4] This is also known as the unrestricted auto-regressive distributed lag model (ARDL). In the ARDL approach the lag structure is determined by an optimal lag search procedure. This procedure can be also used to determine the lag length in the GUM. Equation (2) in the text is based on transforming the following equation in the levels of the variables. For simplicity, only one lag is used for C.

$$\Delta C_t = \sum_{i=1}^{n} \beta_{ci} \Delta C_{t-i} + \sum_{i=0}^{m} \beta_{yi} \Delta Y_{t-i} + \lambda(C_{t-1} - \alpha_0 - \alpha_1 Y_{t-1}) \tag{2}$$

Depending on the number of observations available the total number of lagged variables to be included, on the right hand side, may be large. However, it is important to ensure that enough lagged variables are included so that there is no serial correlation in the residuals of the GUM. Note that the equilibrium theoretical consumption relationship can be recovered from (2) by imposing the equilibrium condition that all the changes in the variables are zero, i.e., from the term in the second line in (2) since in equilibrium (2) will be

$$0 = 0 + 0 + \lambda(C_{t-1} - \alpha_0 - \alpha_1 Y_{t-1})$$
$$\text{therefore,} \quad C^* = \alpha_0 + \alpha_1 Y^* \tag{2A}$$

The expression in the lagged level variables in (2) is known as the error correction term and models that include it are known as the error correction models (ECM). It implies that departures from the equilibrium position in the immediate past period will be offset in the current period by λ propor-

---

$$C_t = a_0 + b_1 C_{t-1} + g_0 Y_t + g_1 Y_{t-1} \tag{A}$$

This can be written as:

$$C_t - C_{t-1} = b_0 + b_1 C_{t-1} - b_1 C_{t-1} + g_0 Y_t - g_0 Y_{t-1} + g_1 Y_{t-1} - (1 - b_1) C_{t-1} + (g_0 + g_1) Y_{t-1}$$
$$\Delta C_t = b_0 + b_1 \Delta C_{t-1} + g_0 \Delta Y_t - \lambda(C_{t-1} - \gamma Y_{t-1}) \tag{B}$$

where $\lambda = (1 - b_1)$ and $\gamma = (g_0 + g_1)/(1 - b_1)$. Note that if (A) is solved for equilibrium, the coefficient of Y is the same as this. To obtain the equilibrium equation from (A), set $C_t = C_{t-1}$ and $Y_t = Y_{t-1}$ and solve. A trend variable may be also added to the ARDL.

tion. Note that λ should be negative and its absolute value need not be always less than unity, implying that, at times, overshooting is a possibility.[5]

Next, a parsimonious version of equation (2) is developed, by deleting the insignificant variables and imposing constraints on the estimated coefficients. GETS is thus a highly empirical approach, needing the judgment of the investigator at each stage, until a final parsimonious specification is obtained. The computational effort in searching for a parsimonious equation can now be reduced by using a recent automatic model selection software PcGets of Hendry and Krozig (2001); see also Hendry and Krolzig (2005), Owen (2004), Hoover and Perez (1999) and Rao and Singh (2005c).[6]

A standard criticism against GETS is that it does not pretest for the order of the variables. In general most macroeconomic variables are I(1) in their levels and I(0) in their first differences. Therefore, (2) in which both the I(0) and I(1) variables are present is not a balanced equation and may give spurious results. In answering this criticism Hendry repeatedly stated that if the underlying economic theory is correct, then the variables in the levels part must be cointegrated and therefore a linear combination of the I(1) levels of the variables

[5] It is of interest to note that ECM is very much a LSE concept and its usefulness has been illustrated in several applications. Professor W.B. Phillips seems to have first developed this concept to determine the adjustment needed in the policy instruments to maintain a target variable close to its desired value. It was later used in other applications, including in GETS, by Saragan, Hendry, Mizon and also Engle and Granger (1987) in developing their concept of cointegration; for a survey of ECM and its different interpretations see Alagoskoufis and Smith (1991).

[6] A judicious application of the variable deletion test in the standard software like Microfit also gives good parsimonious equations. When the sample size is small, relative to the number of explanatory variables, using these variable deletion tests is a more practical option because PcGets is good when the GUM is adequately elaborate.

must be I(0). In general, if the underlying economic theory is inadequate, the coefficient estimates in the levels part of the equation would be contrary to expectation, for example a_1 in (2) could be more than unity or insignificant or even negative yielding an implausible value for MPC. Furthermore, some arbitrary selection procedures that might be necessary in using GETS are no more serious than similar procedures in the alternatives.

A variant of GETS was developed by Pesaran and Shin (1995) and Pesaran, Shin and Smith (2001), known as the ARDL approach or the bounds test. The manual for Microfit has a good exposition. Like GETS, it does not need pre-testing for the order of the variables. However, unlike in GETS, it is possible to test for the cointegration of the levels of variables, irrespective of their order, using a modified F test. A limitation of this approach is due to a large uncertain rage for its test statistic and only the critical values for samples of 500 and 20,000 are in the manual for Microfit. Therefore, at times, the null of no cointegration can be neither accepted nor rejected and Shin, in a personal communication, informed that the critical values, adjusted for small samples, if any, will be larger making the rejection of the null even more difficult. It is also difficult to accept that variables of different orders are cointegrated. For a recent application of this approach to estimate the demand for money for several developing Asian countries see Bahmani-Oskooee and Rehman (2005).

The Phillips-Hansen FMOLS is also simple to use and preferable when one or more explanatory variables are endogenous. The Microfit manual has a good exposition of FMOLS. However, unlike VECM and GETS, it is less flexible because it is not possible to estimate the coefficients of restricted intercept and trend. Furthermore, it is also not possible to include I(0) variables like shift dummies in the cointegratiing equation although these may be added to the dynamic adjustment equations.

3.2 VAR Models

VAR models were developed as an alternative to the large scale econometric models based on the Cowles Commission approach. Sims (1980) argued that the classification of variables into endogenous and exogenous, the constraints implied by the traditional theory on the structural parameters and the arbitrary dynamic adjustment mechanisms used in the large scale models are all arbitrary and too restrictive. At about this time, forecasts from the large scale models also were also found to be unsatisfactory and thus lending further support to Sims' criticisms. Furthermore, Lucas was also critical of the methodology of the policy models based on the Cowles Commission approach. The Lucas critique argued that if expectations are formed rationally, economic agents change their behaviour to take into account the effects of policies. Therefore, the structure of conventional econometric models will change. These attacks on the Cowles Commission methodology virtually wiped out policy model building activity since the 1980s. VAR models became popular for forecasting purpose and in fact these models yielded better forecasts than large scale econometric models.

VAR models treat all the relevant variables as endogenous. We shall use an example from Smith (2000) and ignore the nominal rate of interest in the following model of money and income. Suppose that the objective is to model the relationship between the logs of real money and real income. Assume that, for simplicity ignoring the rate of interest, the underlying theory of the goods and money markets implies that these equations are given by:[7]

$$m_t = a_{11} + a_{12} y_t + a_{13} T + \epsilon_{1t} \tag{3}$$

$$y_t = a_{21} + a_{22} m_t + a_{23} T + \epsilon_{2t} \tag{4}$$

[7] A real balance effect could link the goods and money markets.

In the past it was not unusual to estimate, for example, (3) with OLS and interpret it as the demand for money by treating y_t as an exogenous variable. The rate of interest is ignored for convenience. Similarly, by treating m_t as exogenous, equation (4) was interpreted as the monetarist equation of income. Sims' VAR approach objects to treating y_t as exogenous in (3) and m_t as exogenous in (4). Instead, it treats both variables as endogenous. To minimize the identification problem, we assume (3) and (4) are equilibrium relationships and introduce the controversial partial adjustment to derive the following short run specifications:

$$m_t = \mu_{11}\Big(a_{11} + a_{12}y_t + a_{13}T\Big) + (1 - \mu_{11})m_{t-1} + \xi_{1t} \quad (5)$$

$$y_t = \mu_{22}\Big(a_{21} + a_{22}m_t + a_{23}T\Big) + (1 - \mu_{22})y_{t-1} + \xi_{2t} \quad (6)$$

The corresponding reduced form equations can be expressed as:

$$m_t = \Pi_{11} + \Pi_{12}\ y_{t-1} + \Pi_{13}\ m_{t-1} + \Pi_{14}\ T + \upsilon_{1t} \quad (7)$$

$$y_t = \Pi_{21} + \Pi_{22}\ y_{t-1} + \Pi_{23}\ m_{t-1} + \Pi_{24}\ T + \upsilon_{2t} \quad (8)$$

Note that, in the reduced form equations, there are no contemporaneous variables like y_t in (7) or m_t in (8). In principle the 8 structural coefficients in (5) and (6) can be identified with the 8 reduced form coefficients in (7) and (8). However, this identification is possible with the controversial partial adjustment mechanism. Like GETS, VAR approach is also critical about the validity of the assumed adjustment mechanism because it may not be consistent with the underlying DGP. Therefore, VAR also suggests the use of a general dynamic specification such as the ARDL type to obtain the parsimonious specifications.

Since in the reduced form equations m and y are endogenous, equations (7) and (8) can be seen as the basic ARDL formulation of a simple VAR model. However, since the partial adjustment assumption is not appropriate, additional lagged variables are added to these two equations to improve their summary statistics and obtain white noise residuals. By using the tests for the optimal order of the VAR, the order of the VAR model can be determined, but it is not possible to identify the underlying structural coefficients from these simple VAR models and the Π_i are merely coefficients in the ARDL model. A simple example with the assumption that 3 lags are adequate could be as follows:

$$\begin{aligned} m_t =& a_{10} + a_{11}\ y_{t-1} + a_{12} y_{t-2} + a_{13} y_{t-3} \\ & b_{11}\ m_{t-1} + b_{12}\ m_{t-2} + b_{13}\ m_{t-3} + \gamma_1\ T + \upsilon_{1t} \quad (7A) \\ y_t =& a_{20} + a_{21}\ y_{t-1} + a_{22} y_{t-2} + a_{23} y_{t-3} \\ & b_{21}\ m_{t-1} + b_{22}\ m_{t-2} + b_{23}\ m_{t-3} + \gamma_2\ T + \upsilon_{1t} \quad (8A) \end{aligned}$$

Note that the number of parameters in this simple model, in addition to the 2 intercepts and 2 trend coefficients, are 12. Therefore, we need more than 16 observations to estimate this model. In general, the number of coefficients in a VAR model would be equal to $L \times N^2$ where L is the number of lags and N is the number of variables plus $2 \times N$ for the intercepts and trend. For this reason it is difficult to use VAR models to forecast macro variables in the developing countries with a limited number of annual observations for 30 or 40 years.

The main advantage of the VAR models is that they yield better forecasts (if the sample size is adequate) than the large-scale structural models based on the Cowles Commission methods. However, Bischoff et.al (2000) claim that judgementally adjusted large-scale structural models gave better forecasts than the VAR models of the US economic activity during 1981 to 1996. Furthermore, Evans (2003, p.446) says that VAR models have not passed the market test because no one offers forecasts based on them. Although, subsequent developments in building VAR models have aimed at developing structural VAR models, it does not seem to be possible, with the current state of the real business cycle theory, to deduce adequate identification restrictions in other than small VAR models.

3.3 Cointegration and VECM

This method developed by Johansen (1988) is undoubtedly the most widely used method in applied work. Models using this approach are known as vector error correction (VECM) or cointegrating var (CIVAR) models. VECM can be seen as scaled down VAR model in which the structural coefficients are identified.

The justification for the VECM approach is that identification and testing for the significance of the structural coefficients, underlying the theoretical relationships, is important. The simple VAR models do not identify structural coefficients nor do they take seriously the relevance of unit root tests. In GETS, although there is some awareness of the unit root characteristics of the variables, the crucial theoretical relationship, in the error correction part, is specified in the levels of the variables. The justification is that the underlying theory should be taken seriously. This does not mean that theory should be accepted. GETS treats economic theory as providing useful insight into what are the key explanatory variables. Furthermore, if the theory is well developed, it may also imply some restrictions on the magnitudes and signs of the coefficients of the explanatory variables. These implications can be tested with GETS in a simple manner, provided we agree that the underlying theory is satisfactory and the classification of variables into endogenous and exogenous is acceptable.

In contrast VECM, like VAR, treats all variables are endogenous, but limits the number of variables to those relevant for a particular theory. For example, if we are interested in estimating the demand for money and the theory implies that the log of real demand for money (m) depends on the log of real income (y) and the nominal rate of interest (r), these three variables form a system. According to VECM the assumption that m is the dependent variable and y and r are independent explanatory variables needs to be tested. Therefore, it starts with a general ARDL specification of a small system with these three variables. For simplicity we shall ignore r and write the two variable system of m and y, similar to equations (7) and (8), as follows:

$$m_t = \beta_{11} + \beta_{12}\ y_{t-1} + \beta_{13}\ m_{t-1} + \beta_{14}\ T + \upsilon_{1t} \tag{9}$$
$$y_t = \beta_{21} + \beta_{22}\ y_{t-1} + \beta_{23}\ m_{t-1} + \beta_{24}\ T + \upsilon_{2t} \tag{10}$$

It is necessary first to conduct unit root tests to verify that m and y are $I(1)$. If so, the above system can be easily transformed as follows. Subtract m_{t-1} from both sides of (9) and similarly subtract y_{t-1} from (10) to get:

$$\Delta m_t = \beta_{11} + \beta_{12} - \ y_{t-1} + (\beta_{13} - 1)\ m_{t-1} + \beta_{14}\ T + \upsilon_{1t} \tag{11}$$
$$\Delta y_t = \beta_{21} + (\beta_{22} - 1)\ y_{t-1} + \beta_{23}\ m_{t-1} + \beta_{24}\ T + \upsilon_{2t} \tag{12}$$

Note that Δm_t and Δy_t are I(0). If th levels of these variables move together, as the underlying theory implies, then they should be cointegrated. A linear combination of these levels is the cointegrating vector and it should be $I(0)$. Before the estimation procedure is considered, it is useful to understand an important point emphasized by VECM. It takes the view that the underlying theory of the demand for money makes an arbitrary assumption that y causes m. In the real world m and y change all the time and therefore causation could run in either direction and that needs to be tested. These causality tests are conducted in a simpler way as follows. The number of cointegrating vectors in a system, at the most, is one less than the number of variables in the system. Therefore, in our system of two variables of m and y, either there is one cointegrating vector or none. Suppose, there is one cointegrating vector. A test for causality is follows. Estimate the following system by using the lagged residuals from the cointegrating vector.

$$\Delta m_t = \lambda_1 \Big(m_{t-1} - \theta_0 - \theta_1 y_{t-1} - \theta_2 T \Big) + lagged\ changes\ of\ variables + e_{1t} \tag{13}$$

$$\Delta y_t = \lambda_2 \Big(m_{t-1} - \theta_0 - \theta_1 y_{t-1} - \theta_2 T \Big) + lagged\ changes\ of\ variables + e_{2t} \tag{14}$$

where λ_i measure the corresponding speed of adjustment in the variable with respect to their past deviation of the level of the variable from its equilibrium value. The levels term is the same as the ECM part in GETS. If λ_1 is significant in (13) and λ_2 is insignificant in (14), it can be concluded that m does not contribute to the explanation of the parameters of the equation for y. Therefore y can be treated as an exogenous variable (since m does not affect its value) and (13) can be estimated ignoring (14) to give the demand for money. This is known as the weak exogenity test. On the other hand if both λ_1 and λ_2 are significant, (13) and (14) should be estimated as a system by imposing the cross equation restrictions on their parameters. In this respect VECM becomes close to VAR and it is not known whether the estimated parameters are those of the demand for money or the monetarist income equation.[8] If both λ_1 and λ_2 are insignificant, there may not be any cointegrating vector and this means that the underlying theory (of the demand for money and/or the monetarist income equation) is inadequate.

The actual procedure for testing for unit roots and the number of the cointegrating vectors (CV) and estimation of

[8] In general many applied workers interpret the cointegrating vector as the demand for money without testing for causality. It is difficult to say that such interpretations are valid.

the parameters in CV is a standard procedure in many econometric softwares. We shall discuss this shortly. Suppose, it is found that there is a single CV for (13) and (14) and the exogenity test shows that y is weakly exogenous. How do we estimate the demand for money equation and test the theory of the demand for money? For illustrative purpose we shall use the single CV obtained by Rao and Singh (2005a) for Fiji (1972-2002) normalized on real money.

$$ln\left(\frac{M_t}{P_t}\right) = 1.133 ln\ Y_t - 0.037R_t \tag{15}$$

where $M = M1$, P = GDP deflator, Y = real GDP and R = a weighted average of the nominal rate of interest on time deposits. Weak exogenity tests showed that both Y and R can be treated as exogenous. Therefore (15) can be interpreted as the CV of the demand for money. It can be seen that the coefficient of $ln\ Y$ is close to unity and therefore it can be concluded that the income elasticity of the demand for money in Fiji is unity. The implied interest elasticity, at the mean interest rate of 6.97 is -0.286 is also plausible.

Without giving the full details of the estimation procedure of VECM models, the steps in estimation can be summarized as follows. Virtually all econometric softwares include these procedures but menu driven software packages like Microfit are easy to use.

1. Test for the stationarity of the relevant variables with the unit root tests and make sure, for example in the demand for money, that the relevant variables (real money, real income and nominal rate of interest) are I(1) and their first differences are I(0).
2. Search for the optimum lag length of the VAR to determine its order. Let us say that a second order VAR i.e., VAR(2) is the optimum.
3. Test for the number of the cointegrating vectors using 2 for the order of the VAR. In a model of 3 variables of

(15), at the most there will be 2 CVs.[9] Let us say that we found only 1 CV. It may look that this can be interpreted as the CV for demand for money. But be aware that it can be also interpreted as a monetarist version of the income equation or the rate of interest equation. Therefore, it is necessary to conduct the exogenity tests. Many software packages also give the relevant equations for conducting these exogenity tests. Let us say that we found that both income and rate of interest are weakly exogenous. Then the CV can be interpreted as that of the demand for money and normalized on real money as in equation (15).

4. The CV in (15) is only the long run relationship of the demand for money and we need a short run dynamic equation for the demand for money based on the ECM approach. For this purpose obtain the residuals from (15) and denote this as ECM. Now estimate the following short run equation, similar to equation (2) in GETS.

$$\begin{aligned}\Delta ln\Big(\frac{M_t}{P_t}\Big) =& a_0 + a_1 T + a_2\ \Delta ln\Big(\frac{M_{t-1}}{P_{t-1}}\Big) + a_3 \Delta ln\Big(\frac{M_{t-1}}{P_{t-1}}\Big) + \cdots \\ &+ b_1 \Delta ln\ Y_t + b_2 \Delta ln\ Y_{t-1} + b_3 \Delta ln\ Y_{t-2} + \cdots \\ &+ g_1 \Delta R_t + g_2 \Delta R_{t-1} + g_3 \Delta R_{t-2} + \cdots \\ &\lambda ECM_{t-1} + \epsilon_t \end{aligned} \tag{16}$$

5. Search for a parsimonious specification by using the variable deletion tests and by deleting the variables with insignificant coefficients to get the final short run dynamic equation. The sign of λ should be negative and its abso-

[9] If the number of CVs equals the number of variables, then the variables are stable in their levels i.e., I(0) and the equations can be estimated with the standard classical methods. If there is no cointegration vector, it implies that the underlying economic theory is inadequate and perhaps some other relevant variables are missing from the model.

lute value gives a rough idea of the speed of the adjustment process towards equilibrium.

Here is an example of an ultra parsimonious ECM equation for the demand for money in Fiji from Rao and Singh (2005a).

$$\begin{aligned}\Delta ln\Big(\frac{M_t}{P_t}\Big) =& - 3.047 - 0.002T - 1.114\ ECM_{t-1} \\ & (-10.94)^{*}(-0.706)^{*}\ \ (-11.05)^{*} \\ & + 1.114(\Delta lnY_t - \Delta lnY_{t-1}) - 0.820(\Delta lnY_{t-2} - \Delta lnY_{t-4}) \\ & \quad (11.05)^{*} \qquad\qquad\qquad (-3.59)^{*} \\ & - 0.039(\Delta R_t - \Delta R_{t-1}) + 0.279COUP - 0.114DEV2 \\ & \quad (6.43)^{*} \qquad\qquad\qquad (5.93)^{*} \qquad (1.93)^{**}\end{aligned} \tag{17}$$

Where the t-ratios are in the parentheses and * and ** denote, respectively, 5% and 10% significance. COUP is a dummy variable for the two political coups and DEV2 ia a dummy variable for the large devaluation of the domestic currency in 1998. Rao and Singh (2005a) also used GETS to estimate the same demand for money for Fiji and it is as follows:

$$\begin{aligned}\Delta ln\Big(\frac{M_t}{P_t}\Big) =& - 2.047\Big[- 1.109\ ln\Big(\frac{M_{t-1}}{P_{t-1}}\Big) \\ & (-5.34)^{*} \quad (-5.21)^{*} \\ & + 1.109lnY_{t-1} - 0.031R_{t-1}\Big] + 1.742\Delta Y_t \\ & \quad (5.21)^{*} \qquad (-2.53)^{*} \qquad (4.71)^{*} \\ & 0.082\Delta Y_{t-4} - 0.045\Delta R_t + 0.247COUP \\ & \quad (1.69)^{**} \qquad (-2.68)^{*} \qquad (4.07)^{*}\end{aligned} \tag{18}$$

Note that the expression in the square brackets is the error correction part comparable to the CV in (15). It can be seen that there are no significant differences in the income and interest rate elasticities given by GETS and VECM. The coefficients of the COUP dummy variable are also similar. Although in some versions of this GETS equation the devaluation dummy was significant, it is dropped in (18) because it was insignificant. The dynamic structure of these two estimates are different and it is difficult to say which is better. The summary statistics of both the VECM and GETS equations (not reported) are similar although the SEE of (17) at 0.055 is less than 0.083 for (18), but in some other applications GETS performed better than VECM. In the estimates of the consumption function for Fiji, the SEE of the GETS equation was 0.016 compared to 0.031 of the VECM estimate.

4. WHICH ONE TO USE?

It is obvious from section 2 that this is a hard question to answer given the hard-line methodological positions. Pragmatism suggests that when both the thesis and the anti-thesis are equally justified, perhaps it is wise to select the synthesis. If these alternative approaches are seen, using the three-fold classification of Smith (2000), they are alternative techniques to summarize data. When properly used, there may not be significant differences in their summaries. Consider an observation by Granger (1997) on models based on alternative techniques:

> " It is impossible to decide between models on purely intellectual grounds, one individual can only say why he or she may have a preference. Given the choice between an apple, an orange, or a banana, one person could say that the public will not buy a fruit that is not yellow or another that he prefers fruit that is spherical but the only way to decide which consumers really prefer is to make all three fruit available and see which is purchased. The equivalence with models is to make several available and see which is selected and used in practice. This choice is made easier if a sequence of outputs, such as forecasts, are made and can be evaluated. The parallel concept of

> the price of the fruit in my example is the human capital required to use a modeling technique, if a method is very complex and needs a great deal of experience before it can be used with confidence it will be viewed as expensive compared to a simple or pre-packaged, automated procedure. There is therefore little controversy if we agree to try several modeling methods and let competition decide between them. " (p.176).

The methodological message of Granger is clear. On the basis of his observations we may say that it is better to use all the three (or all the six) alternative methods to estimate and test theoretical relationships. If they all yield a similar summary of the facts, that would enhance our confidence. It is only when there are conflicting summaries, it is necessary to select a more reliable technique. Although VECM looks complicated, it is not difficult because most econometric softwares have the necessary routines to reduce computational problems. Nevertheless, VECM is more demanding than GETS, especially for those with a modest background in the estimation theory. Furthermore, VECM is becoming increasingly complex with the developments in the unit root tests based on endogenous structural breaks and if in the seasonally unadjusted quarterly and monthly data there are significant seasonal patterns. Although there are seasonal unit root tests and cointegration methods, these are computationally demanding. Therefore, in many applications when seasonally adjusted data are not available, such data are seasonally adjusted, for example, with the U.S. Census Bureau seasonal adjustment programme X12.[10] In a limited number of comparisons between the estimates with GETS and VECM, we have obtained very similar results on the demand for money reported in (17) and (18), consumption and gasoline price adjustment equations for Fiji. In the gasoline price adjust-

[10] Eviews has a few seasonal adjustment options including X12.

ment equations with US monthly data, GETS gave better results than a variant of VECM based on the Granger two-step method; see Rao and Singh (2005a, 2005b and 2005c) and Rao and Rao (2005a and 2005b).[11] While a limited number of empirical studies are inadequate for generalizations, we hope that other investigators would report their experiences with these alternative techniques. In the meantime we may draw a tentative conclusion that given its simplicity, perhaps GETS is likely to be widely used by many applied economists not only because it is simpler to use but its results do not seem to conflict with VECM and perhaps also with those based on the alternative techniques, not examined in this paper.

5. CONCLUSIONS

In this paper we have examined a few methodological difficulties in choosing between alternative techniques to estimate time series models. From a limited number of empirical works in which both GETS and VECM are used there seem to be no major conflicts in their conclusions. Furthermore, as Granger has observed, no matter how complicated and sophisticated is a technique, it is unlikely that we can ever find the true DGP underlying in our models. Therefore, for pragmatic reasons, a simpler method like GETS that adequately summaries the data and captures the DGP is likely to be widely used by the applied economists. Nevertheless, more empirical evidence, based on these alternative approaches, is necessary to draw a few stylized conclusions on their relative merits and computational ease. It is hoped that these methodological issues would interest other researchers.

Finally, as Harvey (1997) has observed, so far economists using cointegration and error correction models did not discover any new results to contradict theories. We interpret

[11] The estimated coefficients in the CVs in these relationships and other summary statistics are very similar, although there are differences in the structures of the short run dynamic adjustments.

this observation as highlighting the need to proceeding further and develop appropriate multi-variable models to explain the dynamics of more than a limited number of variables and their use for decision making by the policy makers. For this purpose, perhaps GETS, because of its simplicity, is more appropriate than the VECM techniques.

References

Alagoskoufis, G. and Smith, J. J., (1991) "On error correction models: specification, interpretation, estimation", *Journal of Economic Surveys*, pp.97-132.

Bahmani-Oskooe, M. and Rehman, H., (2005) "Stability of the money demand function in Asian developing countries", *Applied Economics*, pp.773-792.

Bischoff, C. W., Belay, N. and Kang, I., (2000) "Bayesian VAR forecasts fail to live up to their promise", *Business Economics*, pp.19-29.

Engel, R. F. and Granger, C. W. J., (1987) "Cointegration and error correction representation, estimation and testing", *Econometrica*, pp.251-276.

Evans, M. K., *Practical Business Forecasting*, Oxford: Blackwell Publishers.

Eviews 5 (2004) Quantitative Micro Software, Irvine CA.

Fair, R. C., (2004) *Estimating How the Macroeconomy Works*, available at http://fairmodel.econ.yale.edu/

Granger, C.W.J., (1997) "On modeling the long run in applied economics", *Economic Journal*, pp.169-177.

Harvey, A., (1997) "Trends, cycles and autoregressions", *Economic Journal*, pp.192-201.

Hendry, D. F., (1987) "Econometrics methodology: a personal perspective", in T. F. Bewley (ed.) *Advances in Econometrics Fifth World Congress*, Vol.2. Cambridge: Cambridge University Press.

Hendry, D. F. Pagan, A. R. and Saragan, D. J., (1984) "Dynamic Specification", in *Handbook of Econometrics Vol.2*, (eds.) Grilliches, Z. and Intriligator, M.D., Amsterdam:

North-Holland, pp.1023-1100.

Hendry, D. F. and Krolzig, H. M., (2001) *Automatic Econometric Model Selection* , Timberlake Consultants Press.

——— (2005) "The properties of automatic GETS modeling", *Economic Journal*, pp.C32-C61.

Hoover, K. D. and Perez (1999) "Data mining reconsidered: encompassing and the general to specific approach to specification search", *Econometrics Journal*, pp.167-91.

Johansen, S., (1988) "Statistical analysis of cointegrating vectors", *Journal of Economics Dynamics and Control*, pp.231-254.

Owen, P.D., (2002) "General-to-Specific modeling using PcGets", *Economics Discussion Papers No.0213*, University of Otago.

Pesaran, M. H. and Smith, R. J., (1991) "The interaction between theory and observations in economics", *the Economic and Social Review*, pp.1-23.

Pesaran, M. H. and Shin, Y., (1995) "An autoregressive distributed lag modeling approach to cointegration analysis", in Storm, S. , Holly, A. and Diamond, P. (eds.) *Centennial Volume of Rangar Frisch*, Econometric Society Monograph, Cambridge: Cambridge University Press.

Pesaran, M. H. and Pesaran, B., (1997) *Working with Microfit 4.0*, (Oxford: Oxford University Press).

Pesaran, H.M. Shin, Y. and Smith, R. J., (2001) "Bounds testing approaches to the analysis of level relationships", *Journal of Applied Econometrics*, pp.289-236.

Phillips, P.C.B. and Hansen, B. E., (1990) "Statistical inference in instrumental variables regression with I(1) processes", *Review of Economic Studies*, pp. 99-125.

Rao, B.B. and Rao, G., (2005a) "Is the gasoline price adjustment equation for Fiji asymmetric?" *Working Paper No.*

2005/10, Department of Economics, the University of the South Pacific.

———- (2005b) " Further evidence on asymmetric US gasoline price responses", *Working Paper No. 2005/11*, Department of Economics, the University of the South Pacific.

Rao, B.B. and Singh, R., (2005a) "A cointegration and error correction approach to the demand for money in Fiji: 1971-2002", forthcoming, *the Pacific Economic Bulletin*.

————- (2005b) "A consumption function for Fiji", forthcoming, *ICAR Journal of Applied Economics*.

———— (2005c) "Demand for money in Fiji with Pc Gets", forthcoming, *Applied Economics Letters*.

Sims, C., (1980) " Macroeconomics and reality", *Econometrica*, pp.1-48.

Smith, R. P., (2000) "Unit Roots and all that: the impact of time-series methods on macroeconomics", in Backhouse, R. E. and Salanti, A. (eds.) *Macroeconomics and the Real World*, (Oxford: Oxford University Press).