WHAT ARE WE LEARNING ABOUT THE LONG-RUN?

Clive W.J. Granger

Abstract

An attempt is made to link together earlier definitions of the long-run found in micro and macro economics with recent developments in econometrics, specifically cointegration. It is suggested that the links are not strong and that most of the previous work in econometric theory has been unnecessarily over-precise. Unit root processes can be replaced by processes that approximate them without loss of interpretation. The possibility of embedding cointegration theory into a very general nonlinear theory is suggested. An example uses a nonlinear relationship between UK short and long run interest rate proposed by Frank Paish.

Key words:
The long-run in microeconomics, the long-run in macroeconomics, cointegration, approximating unit roots, cointegration in nonlinear models.

*Department of Economics, University of California, San Diego. Presented as the Frank Paish Lecture, Royal Economic Society Meeting, London, May, 1992. Work supported partially by NSF Grant SES 903037. The final version of the paper was written whilst I was Foundation BBV fellow at Carlos III University, Madrid.
1. Introduction

The most famous truism in economics is the statement by Lord Keynes, that "in the long run we are all dead." It does not imply that the long run is unimportant, after all institutions can exist for a long time, the Royal Economic Society being an example, and most of us are altruistic enough to be concerned about the economic well-being of our children and grandchildren. What Keynes was actually emphasizing was that the study of the short run is also important, as his statement continues. "Economists set themselves too easy, too useless a task if in tempestuous seasons they can only tell us that when the storm is long past, the ocean is flat again." I doubt if that is the kind of forecast made by current economists.

Until recently there has been a rather strange division of labour amongst academic economists with economic theorists generally being concerned with equilibrium, which in the macroeconomy is associated with the long run, and being relatively little concerned with disequilibrium, the short run. On the other hand econometricians have largely concentrated on short run dynamics, and even simultaneity, and have given less attention to long-run relationships. This is illustrated by that usual omission of the entry "equilibrium" in the indexes of econometrics texts, including the three volume Handbook of Econometrics (Grilliches and Intrilligator (1984)). If macrotheories are about equilibria and econometric techniques are not, it becomes difficult for these theorems to be tested on actual data.
A major reason for interest in the long run is to study the impact of changes in the economy, such as new policies, tastes and technologies, the effects of which may take some time to be fully understood. A university may decide to try to improve the quality of its graduates by imposing stricter requirements on its entrants but the effects of the new rule will take three or more years before they can be observed, for example. Other reasons for delayed impacts are the familiar price stickiness, and labour and supplier contracts.

There seems to be no clear definition of what precisely is the long-run; there really being a continuum from contemporaneous effects through the short and middle run out to the long run, with no precise division from one region to another. However, it may be possible to distinguish the long run by its possession of a property. For example, Mankiw (1991) is a recent macro text says "most macroeconomists believe that the crucial difference between the short run and the long run is the behaviour of prices. In the long run, prices are flexible and can therefore respond to changes in supply or demand. In the short run, however, many prices are "struck" at some predetermined level" (page 215). He also describes "short run issues" as "year to year economic fluctuations," thus attaching a time-span to the short run. I find it difficult to believe that macroeconomists would generally agree on anything and doubt if they would all find prices as the main determining variable. What is interesting is that time series econometricians find prices to be amongst the smoothest of series, often designating them I(2), meaning that they should be differenced twice to achieve stationarity.

A related but different definition is associated with Marshall and is found in microeconomic texts in chapters discussing cost and production functions. Production of a company is believed to depend on several distinct factors some of which are fairly easy to change (hours worked, say) but others can be changed quickly with difficulty or at high cost. The "long run" occurs when all factors can be altered, leading to production levels which minimize the long run cost function. The prices of factors which arise from
This solution are said to be from a long run equilibrium. Clearly, this definition of the long run will produce different time periods for the different industries and sectors that make up the economy. The long run for the whole economy will correspond to that of the slowest industry (shipbuilding rather than electronics) or sector (government rather than services, perhaps).

These definitions of the long run, and others discussed by Panico and Petri (1987) are imprecise and are of little direct use by econometricians or by others attempting to interpret models. A frequently re-occurring theme is that in the long run the economy can be considered to be in equilibrium. There are, of course, many types of equilibrium, and in microeconomics these can be reached rapidly but this seems to be less true in macroeconomics. Recent developments in econometrics have attempted to include an equilibrium concept, whether or not this has been successful will be discussed below.

2. The Smooth, Dominant Component

Figure 1 shows what may be considered to be a typical plot of a monthly macroeconomic variable for the period of roughly 1950 to 1990. The actual variable shown is the logarithm of U.S. commercial and industrial loans outstanding in 1982 dollars. There is seen to be a distinction upward movement, which can be called a trend, with possibly a change in slope in the early seventies, giving a broken trend. The data also contains long, smooth swings, some of which can be linked with the business cycle. Superimposed on everything are various small oscillations. What is clear is that this variable contains a dominant component — in terms of its contribution to the total variance — which is generally smooth but is probably not deterministic. Being smooth, this component will have a high correlation between terms that are not too distant and is naturally associated
with the long run. There are many statistical models that produce such smooth components, including trends with breaks of the form

\[ T_t = \begin{cases} \alpha t^\theta & t < t_0 \\ \beta t^\phi & t \geq t_0 \end{cases} \]

where \( 0 < \theta, \phi < 1 \), say, and long cycles with possibly stochastic phases. An important class of relevant models are the ARIMA\((p, 1, q)\) discussed in Box and Jenkins (1970), where a series \( X_t \) is generated by

\[ a_p(B)(1-B)X_t = b_q(B)\epsilon_t \quad (1) \]

where \( \epsilon_t \) is a white noise, \( B \) is the backward operator, so that \( B^kX_t = X_{t-k} \) and \( a_p(B), b_q(B) \) are polynomials in \( B \) of orders \( p \) and \( q \) respectively, such that \( a_p(1) \) and \( b_q(1) \) are not zero. As the autoregressive (i.e. left hand) component of (1) is zero if \( B=1 \), \( X_t \) is called a unit root process. These processes have long swings, increasing variances and tend towards no particular value. Other processes have similar properties, for example

\[ X_t = \alpha X_{t-1} + \epsilon_t \quad (2) \]

where \( \alpha \) is a stochastic process generated independently of \( X_t \) and having \( \mathbb{E}[\alpha_t] = 1 \). Such an \( X_t \) may be called a stochastic with root process. There are also many series that are nearly unit root processes in various ways of measuring this concept. These processes together have similar properties, are smooth compared to the usual stationary series and will here be lumped together under the title "Generic Unit Root" or GUR processes. Given the amount (i.e. length in years) of data available in the macro economy it is almost certainly impossible to distinguish between members of this general class and I believe that
attempts to find methods which can discriminate in practice are of limited usefulness. Asymptotically, of course, the models may behave quite differently and for very long term forecasting deciding on a particular model may seem worthwhile. However, for such forecasts, one has to make the assumption that the future generating mechanism is the same as that of the past. This is likely to be a very poor assumption. The thing we know about the long run is that we should expect structural breaks, occasional large ones and probably many smaller ones. The economies of Eastern Europe make a good, topical example. There are many tests for unit root processes, some of which are discussed in Engle and Granger (1991) and are often called I(1) tests. The typical test has as it's null hypothesis the presence of a unit root and most GUR series will fail to reject this hypothesis. However, I will assume that the change of a GUR process will typically be rejected by a unit root test.

GUR series are found frequently in macroeconomics. For example, Stock and Watson (1990) studied 163 individual U.S. series. Excluding three growth rates (i.e. changes, in logs) and fifteen interest rate spreads because of "cointegration" to be explained below, the find that 131 of the remaining 148 variables are not rejected by a unit root test, that is over 88% of the remaining series. It is perhaps worth recording that the 17 non-GUR series included several unemployment series, some wage rates and housing starts. Apparent GUR variables included production, inventories, new and unfilled orders, money stock, and interest rates. Further evidence of the presence of smooth, dominant components in many national log real GNP per capita series is provided by Koop (1991). The presence of a GUR in many series is a clear empirical fact which cannot be ignored by economic theory, although the theory could be helpful in limiting the classes of models considered to explain this fact.

Let $X_t$ some time series and suppose that a future value $X_{t+h}$ is to conditioned on some information set $I_t$ available now and including current and previous values of $X_t$. Considering just conditional means, define
to be the best, least squares forecast of $X_{t+h}$ based on $I_t$. The variable $X_t$ can be called "short-memory is mean" if $f_{t,h}$ trends to a constant (the unconditional mean) as $h$ becomes large for every $t$, so that the information in $I_t$ does not help make long-run forecasts in mean. If this is not true, so that $f_{t,h}(I_t)$ remains a function of $I_t$ for $h$ large and most $t$, the variable can be considered "long-memory in mean." (I am ignoring various technicalities such as limit cycles and deterministic processes). Unit root processes are long-memory in mean for all $h$ and GUR are so for at least many $h$ values. It is easy to extend this definition to long and short memory in variance and in distribution. A series with a smooth, dominant component will have both a long-memory in mean and a short-memory in mean component and these are closely related to the familiar "permanent" and "transitory" decomposition. If a series $X_t$ is written as a weighted sum of previous shocks and a constant

$$X_t = m + C_0 \varepsilon_t + C_1 \varepsilon_{t-1} + \ldots + C_j \varepsilon_{t-j} + \ldots$$  

where it is assumed that $\varepsilon_t$ is a series with zero mean and zero autocorrelation, $\rho_k = \text{cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$, $k > 0$, then consider the sequence of weights $C_j$, $j = 0, 1, \ldots$. If $C_j \to c$, some non-zero constant, $X_t$ can be rewritten as

$$X_t = c \sum_j \varepsilon_{t-j} + \sum_j d_j \varepsilon_{t-j}$$

where $d_j = C_j - c \to 0$ as $j$ becomes large. The first component is a unit root process and is
usually equated with the permanent part of the series and the second part will be stationary, provided $\sum d_j$ is finite, and is equated with the transitory part. It should be noted that the shocks are not permanent or transitory as they are identical in the two components here. Extensions to the multivariate case is possible and potentially useful for interpretation, although the usual identification problems arise and the shocks can enter the representation nonlinearly, which leads to rich interpretation possibilities but difficult specification and estimation questions.

The exact, asymptotic properties of (3) should not be taken too seriously for practical interpretation of the properties of actual economic variables. The difference between $c_j$ tending to a non-zero constant or tending very slowly to zero is a very slight one given the realities of our data. A word such as "permanent" should not be taken literally. I would like to refer to the pleasantly and usefully vague definition of permanent income, "The average income expected to be received over a period of years" in the glossary of Milton Friedman and Walter Heller (1969).

3. Attractors

There are a number of situations in economics in which when one variable is plotted against another — producing what is sometimes called the phase diagram — the variables appear to behave as though the economy prefers to be near some region, which might be called the attractor, than elsewhere. An obvious example is if one plots the price of an agricultural commodity, say tomatoes, in the North of the country verses the price of the same commodity in the South. If the two prices differ sufficiently, the profit motive will encourage entrepreneurs to buy in the North (if that is the cheaper region) and transport to the South and sell there. This activity will increase demand and thus prices in the North and increase supply and decrease prices in the South and the prices will be driven towards equality by market pressures. In this case, the attractor will be the 45° line. It will not be
capturing as if prices are equal a non-symmetric shock can produce prices that are not
equal but the mechanism will be inclined to move then towards each other again. The
movement towards the attractor will not be instantaneous, because of organizational
arrangements that have to be made, and the strength of attraction will be great when the
prices differ greatly but will be virtually nonexistent when the prices are nearly equal. The
attractor will then be a band around the 45° line rather than a sharp line.

If a single series is plotted against time, its mean is an attractor if it is
short-memory in mean, but it has no attractor if it is a generalized unit root process (at all
if it is unit root without drift, effectively otherwise unless it has a trend when infinity may
be thought of as an attractor).

Suppose that a pair of series $X_t, Y_t$ are both unit root processor and that $X - AY$ is
an attractor, it follows that $Z_t = X_t - AY_t$ is short-memory in mean and that the two
series have the "common factor" representation

$$X_t = AP_t + TX_t$$
$$Y_t = P_t + TY_t$$

where $P_t$ is the "permanent" long-memory factor, $TX_t, TY_t$ are "transitory"
short-memory components. When $X_t, Y_t$ are unit root processes with such a linear
attractor they are called "cointegration." There now is a great deal of literature on such
processes, some of which is included in Engle and Granger (1991). It is well known that
such series must appear, at least, to have been generated by an "error correction model."

$$\Delta X_t = \gamma_1 \Delta X_{t-1} + \text{lags } \Delta X_t, \Delta Y_t + e_{1t}$$
$$\Delta Y_t = \gamma_2 \Delta X_{t-1} + \text{lags } \Delta X_t, \Delta Y_t + e_{2t}$$

(6)

with at least one of $\gamma_1, \gamma_2 \neq 0$. It is possible to estimate the permanent component directly
as \( P_t = c_1 X_t + c_2 Y_t \)

where \( c_1 \gamma_1 + c_2 \gamma_2 = 0, \ c_1 + c_2 = 1 \)

Applying this linear transformation to (6) implies that \( \Delta P_t \) is not Granger caused by \( z_t \) in the long run, as discussed by Gonzalo and Granger (1992). It seems that economists should be interested in determining and interpreting this common factor as it is the driving force that determines many economic relationships. Early experience with its use suggest that some surprising results can be found. For example, Yoon (1992) using U.S. monthly data for the period 1952.1 to 1990.1 for the real variables log GNP (minus government expenditure), log consumption and log investment found each pair to be cointegrated and with consumption to be a close approximate to the common permanent factor. This means that if one wants to analyze the causes of the long run movements of the system of three variables it is sufficient to just use log consumption in this analysis, as this contains all of the long-run information in the system.

In the two variable case, if the long memory component \( P_t \) is defined as above, then each variable is a weighted sum of \( P_t \) and the short memory component \( z_t \). Because of this, or directly from (5), it has been suggested that the attractor can be viewed as an equilibrium from two viewpoints:

(i) if the system \( (X_t, Y_t) \) is afflicted by no further shocks after time \( t_0 \), then the process \( X_t, Y_t \) will tend towards some point on the attractor as \( t \) becomes large and \( t > t_0 \) or

(ii) if \( \hat{X}_{t+h} \) denotes the (least-squares) forecast of \( X_{t+h} \) made at time \( t \), then the pair of forecasts \( (\hat{X}_{t+h}, \hat{Y}_{t+h}) \) will tend towards the attractor as the horizon \( h \) becomes large.

The first case is a thought experiment of little practical relevance, a continuing flow of unexpected shocks to the economy appears to be inevitable. The second case points out that forecasts from a cointegrated system "hang together" in ways that other forecasts may
not and do correspond to a certain type of equilibrium in mean. However, once the confidence intervals around the forecasts are added, the picture is less clear. Once the forecasts are on the attractor, the confidence intervals will consist of a sequence of overlapping sausage or balloon shapes, with approximately constant width in the dimension orthogonal to the attractor but increasing roughly proportionally to $h^{1/2}$ along the attractor, because $P_t$ is a GUR process. Thus, the sequence of forecasts as $h$ increases will reach an equilibrium in mean but certainly not in variance and thus not in distribution. Whether or not this is acceptable for the attractor to be equated with an equilibrium I leave to others to discuss.

Generalizations to many variables is quite straightforward and various forms of nonlinearity can be considered. The attractor need not be a straight line but can be a curve such as

$$z_t = g_1(X_t) - g_2(Y_t)$$

provided that at least one of the functions involved is unbounded and $z_t$ is short-memory in mean. Estimation questions have been discussed in Granger and Hallman (1991) although further experience and evaluation is required. An alternative way to introduce non-linearity in the system is through the error-correction model, so that in (6) $\gamma_j z_{t-1}$ is replaced by $\gamma_j(\tilde{z}_{t-1})$, $j = 1, 2$ for some functions $\gamma(\cdot)$ such that $\gamma(0) = 0$. It is generally true that a function of a process that is short-memory in mean and has a constant distribution around this mean, will also be short-memory in mean. It follows that error-correction models need not have the error-correction term involving $z_{t-1}$ coming in linearly. As always, a richer interpretation is now possible. If $\gamma(x)$ is a non-symmetric function, it says that the strength of attraction may be different on one side of the attractor than the other, as is plausible in some economic situations, as discussed in Granger and Lee (1989) for example.
These developments can be embedded into a single, more general framework using Langevin equations from physics. If \( \mathbf{X}_t \) is now a vector of variables, suppose that the dynamic generating mechanism involves a single lag (so that technically, we are in a state-space formulation) and is

\[
\Delta \mathbf{X}_{t-1} = -\mathbf{P}'(\mathbf{X}_t) + \xi_{t-1}
\]

where the dash indicates that the derivative is taken with respect to the components of the vector and \( \xi_{t-1} \) is a zero mean independent series. To consider the interpretation of (7), suppose that \( \mathbf{X}_t \) is just a single series, \( P(0) = 0 \) and \( P(x) \) is monotonically non-decreasing as \( x \) moves away from 0, so that \( P(x) \gtrless 0 \), \( P'(x) \gtrless 0 \). The function \( P(x) \) is called the "potential" at \( x \) and in physical terms may correspond to potential energy. \( x = 0 \) is the attractor, being the value at which \( P(x) \) takes its minimum. For a given \( \mathbf{X}_t \), the expected size of a change in the series is \( -P'(\mathbf{X}_t) \) which is the step towards the attractor if there is no shock. With shocks \( \xi_{t-1} \), the actual step could be away from the attractor but would average \( -P'(\mathbf{X}_t) \) if starting at \( \mathbf{X}_t \). The analogy is with a small particle on a bowl-shaped surface but in a viscous fluid, so that it does not immediately move to the lowest point on the bowl, and then hit by random shocks. In economics the viscous fluid might be replaced by concepts such as sticky prices and potential energy by potential profits from a marketing situation or potential benefits in other cases. Figure 2 shows a bivariate function which is non-symmetric and with a linear attractor. Clearly a potential function can be chosen with a nonlinear attractor along its base and with any shaped sides. Several such functions can be superimposed, if the condition that the potential function is non-decreasing is not required, leading to several attractors possibly crossing, each with their own potential wells and maybe separated by regions that can be thought of as repellers. All of the concepts in the old deterministic stability literature, such as stable and unstable equilibria, can be included in what is now a stochastic and nonlinear...
framework. It seems that the scope for econometric modelling of economic theories is substantial although virtually all of the practical questions in this area have yet to be tackled.

4. Example Explaining Long−Term Interest Rates?

As this lecture is given in honour of Professor Frank Paish it is appropriate to illustrate some of the points made using his work. In his book *Long−term and Short−term Interest Rates in the U.K.* (1966) he asks "what is the major determinant of the long−term interest rate?" His answer was money/GNP which was then decomposed as cash/GNP + bank deposits/GNP. The first of these terms was thought to be roughly constant and so the final term was taken to be the main determinant. Figure 3 shows the Yield on 2½% consols (Y) on the vertical axis and the ratio of bank deposits to national income (%) (X) on the horizontal axis using annual U.K. figures for 1921−65. As is seen, two separate curve, were required to give a good fit to the data. The curves are

(i) \[ Y_t = 0.254 + 1.73X_t + e_t \]  \hspace{1cm} (16.2)  
\[ R^2 = 0.70, \text{ Durbin/Watson } = 0.77 \]

for the years 1921−33, 1947−65, and

(ii) \[ Y = 0.25 + 2.06X_t^{1/2} + e_t \] for the period 1934−46

One t−value is shown although it is of little relevance given the low Durbin−Watson statistic.

I would like to thank Gawon Yoon for the empirical work of this section.
The second period includes the war years, which are certainly a structural break but Paish does not comment why 1934 should be the year in which the models switch. It does not seem to have been an exceptional year economically although it is when a period of negative inflation ended. For the complete series, both $Y$ and $X$ are not rejected by an $I(1)$ test but if the model in (i) is fitted to the whole period, the residuals are also not rejected by such a test. However, taking the residuals from each of the models for their appropriate years does give a series which is rejected by an $I(1)$ test and does appear to be short-memory in mean. Thus, for the period 1921–65, yields and bank deposits over national income seem to be non-linearly and regime-switching cointegrated.

The model was extended to quarterly data for the period 1963–86, giving now a reasonable sample size of 90, with the variables now defined as:

$$Y_t = \text{yield on 20-year government bond}$$
$$X_t = \frac{\text{demand deposits (other than government)}}{\text{GNP}}$$

The relationship found was

$$Y = -3.42 + 15.23 X_{t-1} + \epsilon_t$$

$$R^2 = 0.72, \text{Durbin–Watson} = 0.82$$
$$\text{Dickey–Fuller} = 4.63 \text{ (95% level is } -3.4)$$
$$\text{Augmented Dickey–Fuller} = -2.66$$

The values taken by the Durbin–Watson and Dickey–Fuller test statistics suggest that the residuals are short-memory in mean, being rejected by these $I(1)$ tests. However, if demand deposits are replaced by demand plus time deposits, this result no longer holds. The tests suggest that $Y, X$ and $X_{t-1}$ and all GUR. Thus, $Y$ and $X_{t-1}$ appear to be
cointegrated and Paish's conjecture still appears reasonable, although only a contemporaneous relationship has been fitted. More interpretation is possible from the error-correction model

\[
\Delta(Y_t) = 0.008 e_{t-1} - 0.25 \Delta(Y_{t-1}) + 0.016 \Delta Y_{t-1} + \epsilon_{t-1}
\]

\[t=1.98, \quad (2.3)\]

\[R^2 = 0.25, \quad DW = 1.99\]

\[
\Delta Y_t = -0.13 e_{t-1} + 2.65 \Delta(Y_{t-1}) + 0.03 \Delta Y_{t-1} + \epsilon_{2t}
\]

\[t=1.45, \quad (1.63)\]

\[R^2 = 0.13, \quad DW = 1.98\]

Interpreting the t-values just shown suggests that \( Y_t \) obeys a random walk, with none of the explanatory variables being significant, but change in GNP/bank deposits is explained by lagged error-correction and changes in yield terms. Thus, evidence of causality from yields to GNP/bank deposits has been found, somewhat related to a Baumol–Tobin money demand equation but that there is no causality in the other direction. The conjecture made by Paish in the 1966 book is not seen to be supported by modern methods of dynamic modelling although an interesting non-linearity is discovered.

5. Conclusions

In this lecture, I have tried to survey briefly and critically a number of methods of considering the long-run that are currently in use by econometricians. I have deliberately not been too detailed and careful with the definitions used because I feel that much of the present literature has a feel of pseudo-precision that detracts from our understanding of the actual economy. What are we learning about the long run is what are the difficulties with our techniques and in which directions pay-off are likely to be greatest. The long run is inevitably a difficult area for research, new information accumulates very slowly but
there seems to be plenty of new approaches which are promising, par-
icularly using non-linearity.
References


Log Commercial and Industrial Loans Outstanding from Jan 1953 to Dec 1990 in 1982 Dollars (dol.)
FIG. 2 Nonsymmetric potential function
Fig. 3 U.K. bank deposits and long-term interest rates, 1921–55