



Munich Personal RePEc Archive

# **A Perspective on Unit Root and Cointegration in Applied Macroeconomics**

Weshah Razzak

Department of Labour

2003

Online at <http://mpra.ub.uni-muenchen.de/1970/>  
MPRA Paper No. 1970, posted 2. March 2007

**A Perspective on  
Unit Root and Cointegration in Applied Macroeconomics**

**W A Razzak  
Labour Market Dynamics  
Department of Labour  
56 The Terrace  
Wellington, New Zealand  
2007**

**Abstract**

I discuss econometric issues of high relevance to economists in central banks whose job is to interpret the permanency of shocks and provide policy advice to policymakers. Trend, unit root, and persistence are difficult to interpret. There are numerous econometric tests, which vary in their power and usefulness. I provide a set of strategies on dealing with macro time series.

**JEL Classifications:** C13, C22, E1, F4

**Keywords:** Unit root, trend, persistence, cointegration

**Forthcoming in  
The International Journal of Applied Econometrics and Quantitative  
Studies, issue 1, 2007**

## 1. Introduction

Economists in policy institutions and particularly in central banks have one major task: assess the permanency of the shocks hitting the economy. Policies that aim at neutralising the effects of particular permanent shocks, are pretty much functions of the success of the central bank's economists in sifting through the information and providing a clear picture about the economy to the decision maker. Economists' job is to understand, at least approximately, the Data Generating Process (DGP). One major complication in this exercise is to determine whether the aggregation is stochastic or deterministic. In doing so, unit roots and cointegration time series analyses are pretty much at the heart of the matter. Box and Jenkins (1970) proposed a strategy to studying time series, which consists of: (1) pre-analysis of the data, (2) modelling the specification of the data, and (3) model evaluation. I will discuss all of the three strategy points and discuss strategies for empirical research in this area.

All test statistics are derived under restrictive assumptions such as, a known DGP, normality, asymptotic theory etc. Most of the assumptions are not met in reality and, therefore, researchers have to make judgements and interpret the results of their tests. This is not always straightforward, and it is not all science. Although most statistical tests seem very rigorous, there is a lot of subjectivity when interpreting the results.

The note is organised as follows: Next I discuss the issue of trend, which is really the driving motive of this note. In section 3, I briefly talk about popular unit root tests and discuss some empirical issues and problems, and provide strategies for interpreting results and making decisions about the nature of the data. Section 4 is about cointegration. Section 5 is reserved for some modelling techniques to estimate integrated time series.

## 2. Trend, unit root and persistence

Almost all macroeconomic time series display trends when plotted against time. Just saying that a time series has a deterministic trend component  $f(t)$  is probably insufficient because we have to say something about  $f$ ; for example, one cannot accept that  $\cos(t)$  is a trend.

Granger (1995) talks about "extended memory" where the conditional expected mean over a forecast horizon  $s$  is:

$$1 \quad f_{t,s} = E(x_{t+s} | \Omega_t)$$

The information set  $\Omega$  includes current and past values of the time series. If  $f$  does not tend to a constant as  $s$  becomes large the time series is said to have "extended memory." The random walk process is a good example of such a time series. Extended memory processes can also be defined in terms of the impact of current shocks on the process. Persistence is determined by the size of the autoregressive coefficient and by the variance of the shock in this simple AR process,

$y_t = \rho y_{t-1} + \varepsilon_t$  such that the root of the characteristic equation is close to unity. Therefore, a unit root process is a good example of a persistent time series.

The nature of the trend has implications regarding the “persistence”, which is the nature of smoothness in the data that is related to the long-run component of time series. Phillips (2003) is an excellent review of the issues related to this subject. Section 5 in his article is under the title: “No One Understands Trends” says it all. Indeed, trend is rather the most interesting and challenging aspect of the history of time series regression analysis.

Pre-testing the data aims at assessing the nature of the trend; this then determines how to model the data. Trend can be deterministic or stochastic. These are two different DGP, and have different economic and policy implications. It is not possible for macroeconomists who conduct applied research to ignore “persistence”. Nelson and Plosser (1982) raised the question about the nature of trend. The nature of trend, e.g., liner or stochastic, determines the method of statistical analysis, but it is very difficult to tell the difference between different types of trends. All trends are significant in time series regression analysis. Stochastic trends make things even more difficult because they are synonymous with unit roots, whose regressions have non-standard limiting distributions.

When the trend is assumed to be deterministic, economists regress the time series on a constant term, and a linear trend. The residuals of that regression are the de-trended time series. Presumably, this de-trended time series’ mean, variance and covariance are not functions of time. Thus, this time series is called trend-stationary.

Nelson and Plosser (1982) argued that the trend in most macroeconomic time series is not deterministic, but rather stochastic; i.e., slowly varying. Removal of trend from such time series can be accomplished by first differencing, which is a popular filter (though there are many different ways to remove trend from the data). The resulting time series is called differenced-stationary.

The difference between trend and differenced-stationary time series is that the trend-stationary time series tends to return to a fixed deterministic trend function or the time series would fluctuate around a fixed trend function. The differenced-stationary time series, however, has no tendency to return to a fixed trend function. It simply grows at a rate  $\beta$  from its current position.

Differencing might render the data stationary, most of the time. A stationary time series will be  $I(0)$ , but not all  $I(0)$  time series are stationary. Some AR (1) model’s can be stationary but not  $I(0)$ . Also, not all non-stationary times series are  $I(1)$ .

The inability to determine the nature of the trend results in misspecification with all common consequences such as inconsistency of the coefficients. If

$\lim_{N \rightarrow \infty} |\hat{\beta} - \beta| < \delta = 1$  the consistency means that the probability limit of  $\hat{\beta}$  is  $\beta$ .

Determining the nature of the trend turns out to be a very serious and difficult task.

Nelson and Plosser also argued that unit root implies that the source of business cycle fluctuations in GDP are REAL shocks not MONETARY ones. The finding that GDP has a stochastic trend is widely accepted now. The implication about the source of shocks created a stir; see McCallum (1986) and West (1988) for example. The argument that macroeconomic time series display stochastic trends and fluctuate randomly without a tendency to come back to a fixed trend also threatened the concept of equilibrium, which economists like very much.

### 3. Unit Root: formal testing

A conclusion that has been reached by many macroeconomists (see Stock (1991), Cochrane (1991), Rudebusch (1993), Christiano and Eichenbaum (1990) is that it is rather difficult to settle the issue of unit root using the existing data. A unit root has a lot more serious implications than a root slightly less or greater than one. But it is widely accepted that formal statistical tests cannot tell the difference between a root that is one and another that is 0.98. Therefore, macroeconomists' job is to assess the "persistence" of the data. Non-rejection of the null hypothesis of the unit root, at best, can be interpreted as evidence of "persistence."

Dickey and Fuller and Said and Dickey in a series of influential papers (1979, 1981, 1984 etc) provided the statistical theory and formal tests for unit root. If

$y_t = \rho y_{t-1} + u_t$ , where  $u_t$  is iid  $N(0, \sigma^2)$  and the initial observation  $y_0$  is zero, the OLS

estimate of  $\rho$  is  $\hat{\rho} = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_{t-1}^2}$ . If the true value  $|\rho| < 1$ , then

$\sqrt{T}(\hat{\rho} - \rho) \rightarrow N(0, 1 - \rho^2)$ . However, if  $\rho = 1$  then  $\sqrt{T}(\hat{\rho} - \rho)$  has a variance equal to zero thus,  $\sqrt{T}(\hat{\rho} - 1) \xrightarrow{p} 0$ . It turned out to be the case that we have to have  $T\hat{\rho}$  than the usual OLS  $\sqrt{T}\hat{\rho}$ , which means that the unit root coefficient converges at a faster rate  $T$  than the coefficient in the usual stationary process OLS regression. The limiting distribution is not standard and is derived in Dickey and Fuller.

The test in its simple format and without going into the theory itself (the Augmented Dickey-Fuller test is due to Said and Dickey (1984)) uses an OLS regression of the first difference of the time series itself (or its log) on a constant term, perhaps a time trend, lag of the level of the time series and lag polynomial of the differenced time series (lagged dependent variables) to soak up the serial correlation. The regression's  $t$  statistic on the coefficient of the lagged level (normally labelled  $\rho$ ) is not standard. Its limiting distribution is derived, and the critical values are different from what you find in the back of your textbook. Because we run the regression in first differences, the null hypothesis is that  $\rho = 0$ . Rejection of the null hypothesis that this coefficient is zero is taken as evidence that the level of the time series (or the log level) has a unit root. If the hypothesis is not rejected, the time series is persistent and the trend is not deterministic.

Note that neither the constant term, nor the trend coefficients in the level regression are standard. They also have different critical values from those for  $\rho$ . Also dummy variables included in such regressions will have non-standard distributions.

There are many tests for unit root, which I am not going to list, but there are a few well known and popular tests such as the Phillips-Perron due to Phillips (1987) and Phillips and Perron (1988), the Elliott-Rothenberg-Stock (1996) and Elliott (1999). Sims (1988) provided a Bayesian technique to test for unit root. What is interesting in all the empirical Bayesian analysis is that the prior for  $\rho$  was much smaller than 1! It is not surprising that this kind of analysis almost always rejects the unit root hypothesis.

The Phillips-Perron test deals with the case where the residuals in the regression follow an AR process and perhaps heteroskedastic. The usual outcome of such an OLS regression is that the estimated coefficient is inconsistent. They provide a method of modifying the statistics to deal with the inconsistency of the coefficients. Elliott et al are nothing but the GLS estimator of the ADF OLS type regressions. In essence, they deal with the same problem Phillips and Perron dealt with.

The ADF, the Phillips-Perron, Elliott et al, etc. have different powers against stationary alternatives. There are many studies of power, but Fuller et al (1994) shows that an unconditional ML estimator is most powerful test for unit root. It also solves the small sample problem usually found in these tests. Dickey and Gonzalez introduced that test several years ago, but I am not sure if the paper is published. However, if the test, any test, rejects the null hypothesis (here is a unit root) then the power argument is irrelevant. If the hypothesis is rejected easily, then we only need to use one test.

Think of the following hypothesis: the windshield of your car is unbreakable. Say that you have two tests: one is powerful, where you hit the windshield with a sledgehammer; and the other is weak, where you hit the windshield with a plastic hammer. Say that you hit the windshield with the plastic hammer and you break it. You reject the hypothesis that your car windshield is unbreakable. Does it matter for the result whether you used a plastic hammer, a mallet or sledgehammer? The answer is definitely no. Thus, the power of the test is not an issue when the test rejects the null. The power of the test becomes an issue when the hypothesis is not rejected, and when this occurs often. If the test does not reject the null hypothesis of unit root, then we have to convince the readers and ourselves that there is no power problem.

The nature of shocks also changes over time. If demand shocks become persistent, then it might be difficult to tell the difference between “real shocks” i.e., shocks to the aggregate supply and monetary shocks. That is why it is difficult for many to accept Nelson and Plosser’s conclusion.

The other problem in testing for unit root is the determination of the model including the specification of lag structure in the regression. This is a specification problem, which affects the result. Said and Dickey (1984) tested whether additional lags are jointly significant using  $F$  tests. Others used “general-to-specific” method based on  $t$  statistic on the coefficient associated with last lag in the estimated auto-regression. The procedure selects a number of lags  $k^*$  from a more general structure of length

$k$  such that the coefficient on the last lag is significant, and that the last coefficient in an auto-regression of order greater than  $k^*$  is insignificant up to a maximum order  $k$  max .

Applied economists also use other different criteria, e.g., Akaike:  $-2(l/T) + 2(k/T)$ , Schwartz:  $-2(l/T) + k \log(T)/T$ , their modified versions, and general to specific etc., and each one would result in a particular optimal number of lags, and with it the test statistic, of the lagged level variable in the OLS regression that I described earlier, changes. In some cases it is significant and in others it is not.

Almost all tests for unit root suffer from finite sample bias problem (chronic) because the critical values are computed under asymptotic assumptions.

Many researchers report values for  $\rho$  and  $\tau_\rho$  that are also positive, and conclude that the time series have unit roots. This is quite a common misinterpretation of the results because a positive  $\rho$  means the time series has an “explosive root” not a unit root. Most of the time, the finding of a positive  $\rho$  is a result of misspecification of the model and incorrect number of lags. It is rather difficult to think of many macroeconomic time series that have explosive roots.

There is another way for assessing the trend. It is quite possible that time series have both deterministic and stochastic trends with different weights. An unobserved component approach (Kalman filtering) was used to test this hypothesis. The time series can have a cyclical component that might be an AR process and a stochastic component that is a RW and the two are independent (are they really?). Remember that the sum of a trend-stationary process and a differenced-stationary process is itself a differenced-stationary process. Therefore, the evidence in favour of the unit root is strong, yet it is not possible given the data we have to reach a reliable conclusion on this matter.

If the time series has a break due to changes in regimes or policy then most of the available unit root tests that fit straight lines through the data will miss the break, whether it is in the mean or the slope (change in the growth rate) or both, and fails to reject the unit root hypothesis. Perron (in a series of well-cited papers) developed techniques to test for unit root under the assumptions that the time series has breaks either in the mean, the slope or both. Perron also derived the tests for endogenous breaks.

Where do we go from here? Let me summarise “a” methodology to deal with the unit root when conducting research. The task is to make a case for or the lack of persistence because this would determine what to do next; i.e., modelling, predictions...etc.

Plot the data and show they have (or do not have) trend. I do not discuss seasonal patterns, but there are tests for seasonal unit roots, where there is a unit root at the seasonal peak. The data might have a deterministic seasonal pattern.

If the data trend, I would start with the ADF test.

What model do we fit: a model without a constant term and without a trend, or with a constant term, but no trend, or with a constant term and a trend? It seems that most macro data would require both a constant (for the mean) and a trend so I would only fit a model with a constant and a time trend.

To determine the lag structure I would experiment with all possible techniques. I would examine the autocorrelation functions, and use all the information criteria that are readily available in software can be used, and the general-to-specific approach and the Said-Dickey approach can also be used. At the end one can hope to make a reasonable judgement on the lag structure. The results of the unit root test are sensitive to the lag structure and one should be careful to not commit a misspecification.

If the ADF test rejects the null hypothesis then I would stop and conclude that my data have persistence. If it does not reject the null hypothesis, I recommend using all available alternative tests for unit root. Every possible method to determine the lag structure should also be implemented. The idea is to build a consensus about the nature of the trend that emerges from these different tests. Sometimes, one test suggests that the time series has a unit root and the other tests says no it does not. In a situation like this one, we might want to use another test with a different null hypothesis, e.g., KPSS or a Bayesian test (we cannot compare powers because the ADF and the KPSS have different null hypotheses). But, the result might help us draw a conclusion about the nature of the persistence.

Choosing the sample span is important. The ADF test statistic tabulates critical values for a sample of 25 observations. I think it would be a stretch to test for unit root in time series that are shorter than 25 observations.

I would always try to avoid using data that have structural breaks. We might have ex-post knowledge about changes in regimes. Perron provides many tests. However, implementing them in short data is problematic because these tests normally run a regression (different models) over a certain number of observations (15% from the beginning of the sample) then run rolling regressions one observation at the time and look at maximum  $t$  or  $R^2$  to identify breaks. If these tests are not possible to use because of the lack of degrees-of-freedom, we probably have to use ex-post information about the breaks, and search in a narrower interval for them rather than searching over the entire sample.

There seems to be confusion about structural changes. I do not think we are talking about policy regime changes a la the Lucas Critique here. Perron and other tests do not test for such a break in the data, but rather a break resulting from a shock. Ex-post information, therefore, are essential in this regard to identify the nature of the break because it will help us identify the break.

#### **4. Cointegration**

Now we move from testing and modelling a uni-variate time series to bi- and multi-variate time series. Cointegration was first understood as “spurious regression”, which occurs when a pair (or more) of seemingly independent time series, with properties such as RW or unit roots, appear to be related. Typically, we find high  $R^2$



and low  $DW$  statistics in an OLS regression, Granger and Newbold (1974) and in Phillips (1986). One way around this problem is differencing the  $I(1)$  data, which renders them stationary  $I(0)$ , and where the statistics have standard limiting distributions. The second solution is “cointegration.”

If the data ( $y$  and  $x$ ) have unit roots, then the OLS estimator in the regression of  $y$  on  $x$  (where  $x$  is assumed to be exogenous) is inconsistent depending on the DGP of the data. But, regardless of whether the OLS estimator is inconsistent or consistent, its sampling distribution does not have finite moments. It is possible to make inference, but it is inappropriate.

If the first difference does not render the data  $I(0)$  one can second difference the data. An OLS regression of the second differenced data induces serial correlation in the residuals, but its distribution is  $I(0)$ . The OLS estimator will be unbiased and consistent, but it will be inefficient and its sampling variance will not be consistently estimated in general (proof).

The number of differencing is also related to the number of roots in level time series. There is a relatively new literature that discusses double unit root  $I(2)$  series. Johansen is probably the best reference for this literature. There are two more issues to think about: First, a time series that has one unit root and another that has a double unit root can still be cointegrated, where the resulting linear combination is  $I(1)$ . This is the “variance reduction” property that cointegration possesses. Second, what does  $I(2)$  mean in economics? Granger argues that if the second difference of a time series is the white noise of  $\varepsilon_t$  then

$$2 \quad y_t = \varepsilon_t + 2\varepsilon_{t-1} + 3\varepsilon_{t-2} + 4\varepsilon_{t-3} + \dots + k\varepsilon_{t-k},$$

which seems to imply that a shock at time  $t - k$  has a much stronger impact on the time series than a shock at time  $t - 1$ . Whether that makes sense or not is questionable.

It turns out that over-differencing is not a serious issue for the OLS estimator, Plosser and Schwert (1977). The estimator would be unbiased and consistent. But, inference is possible when the serial correlation in the residuals is recognised and treated properly. Also, under-differencing would not result in serious estimation problems provided that we treat the serial correlation. Dickey and Pantula (2002) provide testing procedures for higher order differencing of time series.

Two or more integrated time series of any order can be cointegrated if there exists a linear combination of the two that is of a lower order of integration, e.g.,  $I(1) \rightarrow I(0)$  or  $I(2) \rightarrow I(1)$ . If this is true, the OLS estimator of the regression in the levels is consistent.

Linear stationary relationship is often referred to as “the long run” or “equilibrium” but it is not really equilibrium in an economic sense. Think about cointegration as a drunken man and his dog. Both walk in a random manner, but they are always together. If you can locate one of them you can locate the other one easily, as they do

not wander away from each other for a long time. Also, it is rather difficult to conclude that there is a long-run equilibrium when the equilibrium error has a variance that grows with time. That is the cointegrating vector can be  $I(0)$  in the mean but  $I(1)$  in the variance.

I think it is appropriate to interpret the stationary linear combination between two or more integrated time series to imply that there is a stable long-run relationship. If, indeed, two or more time series each of which is  $I(1)$  are cointegrated then there exists a stationary representation that is called the error-correction representation, due to Granger.

Testing the null hypothesis of “no cointegration” is similar to testing for unit root. Granger (1983) and Engle-Granger (1987) discuss the classic case for a bi-variate system, where they introduce six different statistics to test the residuals of the OLS regression of the levels’ data for unit root. Among those statistics they suggest using the ADF. The ADF statistic to test for unit root in the residuals has a different distribution than the ADF itself. Testing the residuals for unit root using the Phillips-Perron test would yield the Phillips-Ouliaris (1990) test for the null hypothesis that there is no cointegration.

Just as difficult to determine whether a time series has a unit root it is difficult to determine whether two  $I(1)$  time series are cointegrated. There are problems with the (weak) power of the tests, small sample problems and structural breaks.

There is a strand of literature that deals with time series that are neither  $I(1)$  nor  $I(0)$ . Time series is fractionally integrated if  $d$  in the  $ARIMA(p, d, q)$  is neither 1 nor zero. These fractionally integrated time series can among themselves have fractional cointegration (See Richard Baillie and Peter C. B. Phillips for references). When  $d > 0.5$  the time series is non-stationary, but it does not have a unit root, and when  $d < 0.5$ , the time series is stationary.

Valid inference about cointegration requires long span of data. Would a sample of five, ten or fifteen years (annual data) be appropriate to test for cointegration (no cointegration) in the data? My feeling is no because the span is perhaps short. Many studies tried to assess the PPP hypothesis for example, and found that it only holds in samples of 100 or more years. Also, one cannot use 60 or 120 monthly observations and think the problem is solved.

The answer is no because testing for cointegration in 120 monthly observations is not different for 10 annual data points as Hakkio and Rush (1991) important papers shows.

Let  $x_t = x_{t-1} + e_t$ , where  $e_t$  is normal *iid* and  $y_t = x_t + v_t$ , where  $v_t$  is also normal and *iid*,  $v_t = \rho v_{t-1} + u_t$ . If  $\rho < 1$ ,  $x$  and  $y$  are cointegrated. If  $\rho = 1$ ,  $x$  and  $y$  are not cointegrated. The choice of the frequency of the data for the analysis of cointegration will depend on the size of  $\rho$ . It is found that if  $\rho$  is high, e.g., 0.90, then it won’t help to use monthly data.

So what do we do? Testing the significance of the error correction term in the error correction regression is informative in determining cointegration. Thus, if the data  $x$  and  $y$  have unit roots (or at least if the hypothesis of unit root is not rejected), run a regression in levels, obtain the residuals, and run a second OLS regression in differences with the lagged residuals. Provided that the residuals of the error correction equation are *iid*, we can test whether the coefficient of the lagged residuals is significantly different from zero using the  $t$  statistic (look for a high value); if it is not zero then  $x$  and  $y$  are most probably to be cointegrated. I would also check whether the  $DW$  statistic is low and  $R^2$  is high in the levels' regression.

Just like in the unit root case, I would also recommend applying several different tests, as many as you can, and try to reach a conclusion by examining all tests' results. Do not depend on one test. The use of different tests to examine the data is not new. The literature includes numerous studies. For example, see Gonzalo (1994) and Guisan (2004) for comparisons of tests and methodologies.

All tests for unit root and cointegration are sensitive to the lag length. Remember that we don't know the true model or the DGP. These tests are liable to misspecifications. Different information criteria give different suggestions about the lag structure, and hence change the results. Make sure that you apply as many techniques as you can when determining the lag structure. It is often a good strategy to report them all and provide a clear assessment of the results.

When we have more than two time series, the bivariate Engle-Granger type test might not be appropriate. The Johansen-Juselius test is popular and used widely to test for cointegration. Now these tests are found in all software and researchers use them easily.

In the case of one unit root, this test is equal to the square of the ADF. So in the case where we have two variables and one cointegrating vector we would expect that the ADF test when applied to the cointegrating relationship to find it stationary, i.e., rejecting the null hypothesis that it is  $I(1)$ . But, empirically, this might not be true. This gives you an idea about how difficult it is to pin down these unit roots and cointegrating relationships.

One of the experiments I have run uses the following VAR to generate the stationary data:

$$\Delta y_t = \theta_1 \Delta y_{t-1} + \dots + \theta_{k-1} \Delta y_{t-k+1} - \alpha \beta' y_{t-k} + e_t$$

using a sequence of *iid*  $P$  – dimensional Gaussian random variables with mean zero and a variance – covariance matrix  $\Lambda$ ,  $y_t$  is a vector of two time series  $y_1$  and  $y_2$  each has an exact unit root and the matrices  $\alpha$  and  $\beta$  are of dimension  $p \times r$ .

A random sample of size 213 observations corresponding to monthly data is generated. The first 50 observations were thrown away to avoid starting problems. The innovation variances were set equal to 0.003, 0.005 throughout the experiment. I experimented with the values of  $\theta$ . We impose the cointegration on the system. The

idea is to compute the cointegration relationship  $Z_t = \beta'y_t$ , which we know is stationary by construction and then test it for unit root using the ADF. The same number of lags in the Johansen system is used in the ADF and everything else is also fixed. The only difference between the theory and this experiment is that we estimated  $\beta$  from the data. The Monte-Carlo experiment is replicated 10000 times.

For each value of  $\theta$ , I computed the Johansen statistics and their distribution, then I computed the ADF statistic conditional on the Johansen statistics that are rejecting the null hypothesis that there is no cointegrated relationship. The ADF fails to reject the null hypothesis that  $Z_t = \beta'y_t$  is  $I(1)$  on average. Also, the percentage of the ADF that rejects the null hypothesis:  $Z_t$  is  $I(1)$ , increases with the autocorrelation, i.e., the size and the sign of  $\theta$ . The difference between the two tests increases when the root is slightly more than or less than one, say 1.0003 instead of 1.

The Johansen-Juselius method also suffers from a finite sample size problem. The critical values are obtained under the theoretical assumption of normality. Therefore, one must correct these  $CV$  for small samples. Cheung and Lai (1993) is a nice way to deal with this problem.  $CR(\infty)$  is the Johansen critical value.  $CR(T)$  is the finite sample critical value. If  $CR(\infty) = CR(T)$  then there is no bias. The bias happens to be a function of  $T/(T - nk)$ , where  $T$  is the effective number of observations,  $n$  is the number of variables in the system, and  $k$  is the number of lags. The critical value is  $CR(T) = CR(\infty) * T/(T - nk)$ .

A lot has been said about the estimated  $\alpha$  and  $\beta$  in the Johansen model:

$$4 \quad \Delta y_t = \alpha\beta'x_{t-1} + e_t$$

Some interpret  $\beta$  as long-run estimated elasticities, for example, the marginal propensity to consume out of permanent income in the income-consumption regression. However, researchers have noticed that the estimates of  $\beta$  are usually very large in magnitudes, i.e. larger than what theory predicts. Wickens (1996) shows that the cointegrating vectors cannot have structural interpretation unless some economic structure is imposed on the Johansen system. However, it is quite possible that econometric models perform much worse when they are constrained. This result stems from attempts to identify the parameters by relying on information that are not in the sample, but in theory.

Also, people often interpret  $\alpha$  as a “speed of adjustment” I think this is misleading, and might be totally wrong.

Given the stationary process:

$$5 \quad y_t = \rho y_{t-1} + e_t \text{ where } e_t \sim iid N(0, \sigma^2)$$

The characteristic polynomial is:

$$6 \quad A(z) = 1 - \rho z$$

With the root  $\rho^{-1}$ .

If  $|\rho| < 1$  then  $\rho^t$  is the exponential decrease of the initial value and therefore, economists interpret  $\ln \rho$  or  $1 - \rho$  as the “speed of adjustment.”

In the cointegration simple case above, the stationary process is:

$$7 \quad \beta'x_t = \sum_{i=0}^{t-1} (I + \beta'\alpha)^i \beta'e_{t-i} + (I + \beta'\alpha)^t \beta'x_0$$

The speed of adjustment is the eigenvalues of the matrix  $I + \beta'\alpha$ .

The problem becomes more complex when the system includes lags. Therefore, I would not interpret  $\alpha$  as a speed of adjustment.

Before I finish with the Johansen test I would refer you to the CATS manual about the Johansen test. It actually acknowledges that the decision that  $P$  time series are cointegrated is pretty subjective and depends on many factors, where the researcher has to make a call at the end of the day.

The following strategy is recommended: (1) If the variables involved are time series and we are almost convinced that they are persistent (the hypothesis that they have unit roots cannot be rejected in many tests and many different models), we only need to know how are they related economically to determine whether they are likely to be cointegrated or not. For example, if we are testing money, output and the price level, where each has a unit root, it would not be very uncommon to assume that these three variables are related to each other in the long run in a stable relationship. (2) If the data are not long, and here I mean less than 20 years, then it is probably difficult to test for cointegration or even make sense of it. In the case of small sample, but a clear economic theory, I would prefer to impose the cointegration on the system. (3) The reason is that I impose the cointegration assumption and not the unit root on the data is because I think that we have economic theory that helps us in the case of cointegration, but I am not sure of any economic theory about the nature of the trend. (4) If we are interested in the  $\beta$ 's, the long – run coefficients, I would not recommend using the Johansen method to uncover them. This leads me to discuss the next issue of directly estimating the  $\beta$ 's, which has some clear benefits.

## 5. Phillips et al

Phillips and Hansen (1991) and Phillips and Loretan (1991) are two methods worth understanding.

Assume that we have two vectors of unit root time series that are also cointegrated. The system is given by:

$$8 \quad y_t = \alpha + \beta'x_t + v_t$$

and,

$$9 \quad x_t = x_{t-1} + u_t$$

Where  $v_t$  and  $u_t$  are vectors of disturbances with classical assumptions.

Assume the errors follow a general MA representation:

$$10 \quad \begin{bmatrix} v_t \\ u_t \end{bmatrix} = \Gamma(L)\varepsilon_t = \begin{bmatrix} \Gamma_{11}L & \Gamma_{12}L' \\ \Gamma_{21}L & \Gamma_{22}L \end{bmatrix} \varepsilon_t$$

$$E\varepsilon_t = 0$$

$$E\varepsilon_t\varepsilon_s' = V\mathcal{V},$$

All we are saying is that the disturbances are serially correlated and that is what you expect when you regress two unit root variables on each other.

$$11 \quad V = \begin{bmatrix} \sigma_v & V_{12}' \\ V_{21} & V_{22} \end{bmatrix}$$

How do estimate  $\alpha$  and  $\beta$ ? We know OLS is inappropriate because the test statistics are not standard.

Recall that the Engle-Granger method and ECM are appropriate only if  $V_{12} = 0$ .

The Phillips et al estimator, let:

$$12 \quad v_t = \sum_{i=-p}^p \gamma_i u_{t+i} + \omega_t$$

Where  $u_t = \Delta x_t$  from above.

Substitute back in  $y_t$ , we get the Phillips–Hansen:

$$13 \quad y_t = \alpha + \beta'x_t + \sum_{i=-p}^p \gamma_i \Delta x_{t+i} + \omega_t$$

The above gets rid of all nuisance parameters,  $\Gamma_{ij}$  and  $V_{ij}$  and allows valid testing using standard statistics. It is called the Fully Modified OLS. To compute the statistics, we need to estimate the variance of the residuals  $\omega_t$ . The modified  $t$  or  $Z$  times the modifier, which is  $s_\omega / \lambda$ , where  $s$  is the standard error of the residuals and  $\lambda$  is scalar  $= \sigma\psi(1)$  and  $\omega_t = \psi(L)\eta_t$ . There are two ways to model and estimate  $\omega_t = \psi(L)\eta_t$ . First, we can fit a MA process such as:  $\Omega(L)\omega_t = \eta_t$ , which is an OLS regression

$$14 \quad \omega_t = a_1\omega_{t-1} + \dots + a_k\omega_{t-k} + e_t$$

and

$$15 \quad \lambda = \frac{\hat{\sigma}}{1 - a_1 - \dots - a_k}, \text{ and}$$

$$16 \quad \sqrt{\hat{\sigma}^2} = \hat{\sigma} = \frac{1}{T - k} \sum_{t=k+1}^T \hat{e}_t^2.$$

The other way to estimate  $\lambda$  is to use the Newey-West, which also fits a general MA model of order  $k$  :

$$17 \quad \hat{\lambda}^2 = \hat{b}_0 + 2 \sum_{i=1}^k \left(1 - \frac{i}{k+1}\right) \hat{b}_i, \text{ where}$$

$$18 \quad \hat{b}_i = \frac{1}{T} \sum_{t=i+1}^T \hat{\omega}_t \hat{\omega}_{t-i}$$

The Phillips-Loretan is the Two-Sided Non-Linear Dynamic Least Squares. This estimator is given by:

$$19 \quad y_t = \alpha + \beta x_t + \sum_{i=-p}^p \gamma_i \Delta x_{t+i} + \rho(y_{t-1} - \alpha - \beta x_{t-1}) + \varepsilon_t$$

If one error correction term is not sufficient we have to add another. These regressions are non-linear in the parameters and should be estimated with restrictions above. They are most efficient, and equivalent to a Maximum Likelihood Estimator of a System of Equations. They resolve the endogeneity problem we face all the time (single-equation bias), and they also provide asymptotically valid inferences.

Over-fitting is a problem, thus we have to make sure that we have just enough lags and leads to get rid of the serial correlation. Check the residuals for serial correlation.

I prefer the Kolmogorov-Smirnov test, which is a non-parametric test that takes the residuals above and checks how far they are away from the density function of a white noise process.

The tests statistics in this regression are asymptotically normal.

What happens if the RHS variables are cointegrated themselves? Phillips (in an unpublished manuscript) shows that the dynamic becomes simpler (I have a paper on this, let me know if you would like a copy).

Why this method is recommended? Here are my reasons: (1) Macroeconomic/economic theory focuses on the point elasticities of the long run relationships. For example, what is the marginal propensity to consume out of income? What is the income elasticity with respect to the exchange rate? What is the semi-interest rate elasticity with respect to real money balances? And so on so forth. These are questions that most central banks investigate. This method provides us with such estimates. (2) We have now accepted the fact that most these macro variables are integrated time series, and therefore, OLS is inappropriate. (3) We rely on economic theory about equilibrium in the sense that the variables like income and consumption, or money and interest rate do not wander away from each other in the long run. (4) Phillip's information criteria is actually useful in this case because it allocates the cointegration vector along with the optimal number of lags required. (5) The method is equivalent to estimating a system of equations using ML, thus it is not

a subject to endogeneity problem. (6) The fact that the long – run coefficients are estimated jointly with the short-run dynamics is appealing since the two are not separated, and that is where the efficiency of the estimator arises. Barnhart, McNown and Wallace (1999) compared the Phillips-Loretan methods to many other tests to examine the forward exchange rate unbiasedness hypothesis.

## **6. A final word**

To close, I have to say a few words about models' evaluation. The fact remains that we will not be able to find out what the DGP is precisely. We also are not able to precisely pin down the nature of the trend. Yet, we are still required to make judgments about “persistence” and the permanency of the shocks and make recommendations to policymakers. That said, we would be better off examining variety of models instead of one single model. Granger (2000) recommends Thick Modelling as a strategy. This strategy stems from portfolio theory, where investors diversify their investments to minimise risk. Similarly, policy makers should minimise risk of policy errors arising from reliance on wrong macroeconomic model. Razzak (2002) is one example of how to model and forecast inflation.

(1) Several different models should be fit to the data; (2) stable models (parameters that do not change over time when policy changes) are to be picked out; (3) outputs such as the forecasts can be evaluated. But, the average of the forecasts over different stable models can have superior properties to a single model's forecast.



**References:**

Barnhart, S. W., R. McNown and M. S. Wallace, "Non – Informative Tests of the Unbiased Forward Exchange Rate," *Journal of Financial and Quantitative Analysis*, Vol. 34 No.2, June (1999).

Box G E P, and G M Jenkins, Time Series Analysis, Forecasting and Control, Holden Day, San Francisco, (1970).

Cheung, Y W and K S Lai, "Finite-sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration," *Oxford Bulletin of Economics and Statistics* 55, 3 (1993), 313-28.

Christiano, L. and M Eichenbaum, "Unit Roots in Real GNP: Do We Know, and Do We Care?" *Carnegie-Rochester Conference Series on Public Policy*, Vol. 32 (Spring 1990), 7-62.

Cochrane, J, "Comments," *NBER Macroeconomics Annual 1991*,

Dickey, D A and W Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of The American Statistical Association* (June 1979), 423-31.

Dickey, D A and W Fuller, "The Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica* 49 (July 1981), 1057-72.

Dickey D A and S Pantula, "Determining the Order of Differencing in Autoregressive Processes," *Journal of Business and Economic Statistics*, (January 2002) 20(1), 18-24

Elliott, G., "Efficient Tests for a Unit Root when the Initial Observation is Drawn from its Unconditional Distribution," *International Economic Review*, Vol. 140, No. 3, (1999), 767-783.

Engle, R and C W J Granger, "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, vol. 55 (March 1987), 251-76.

Fuller, W, G Farias and S Pantula, "A Comparison of Unit-Root Test Criteria," *Journal of Business and Economic Statistics* (October 1994) 12(4), 449-5.

Granger, C W J, "Cointegrated Variables and Error Correcting Models," WP 83-13, San Diego: University of California, 1983.

Granger, C W J., "Non Linear Relationships Between Extended Memory Series," *Econometrica* 63 (1995), 265-280.

Granger, C W J and P Newbold, "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2 (1974), 111-20.

Gonzalo, J., "Five Alternative Methods of Estimating Long-Run Equilibrium Relationships," *Journal of Econometrics*, Vol.60 issue 1-2, (1994), 203-233.

Guisan, M. Carmen, "A Comparison of Causality Tests Applied to Bilateral Relationship Between Consumption and GDP in the USA and Mexico," *International Journal of Applied Econometrics and Quantitative Studies*, Vol. 1, Issue 1, (2004).

Hakkio, C and M Rush, "Cointegration: How Short is the Long Run?" *Journal of International Money and Finance* (1991), 571-81.

Johansen, S, "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control* 12 (1988), 231-54.

Johansen, S and K Juselius, "The Full Information Maximum Likelihood Procedure of Inference on Cointegration with Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics* 52 (May 1990), 169-210.

McCallum, B T, "On 'Real' and 'Sticky-Price' Theories of the Business Cycle," *Journal of Money, Credit and Banking* Vol. 18 (November 1986), 397-414.

McCallum, B T, "Unit Roots in Macroeconomic Time series: Some Critical Issues," *Federal Reserve Bank of Richmond Economic Quarterly*, Vol. 79/2 (Spring 1993), 13-43.

Nelson, C R and C Plosser, "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, Vol. 10 (September 1982), 139-62.

Plosser, C and W G Schwert, "Estimation of Noninvertible Moving Average Process: The Case of Overdifferencing," *Journal of Econometrics*, Vol. 6 (August 1977), 199-224.

Rudebusch, G, "The Uncertain Unit Root in Real GNP," *American Economic Review*, Vol. 83 (March 1993), 264-72.

Perron, P, "Testing for Unit Root in a Time Series with a Changing Mean," *Journal of Business and Economics Statistics*, 8 (1990), 153-62.

Perron, P (1989), "The great crash, the oil price shock and the unit root hypothesis," *Econometrica*, 57, 1361-1402.

Perron, P, "Trends and random Walks in Macroeconomic Time Series," *Journal of Economic Dynamics and Control* 12 (1988), 297-332.

Perron, P,” Further Evidence on Breaking Trend Functions in Macroeconomic Variables,” *Journal of Econometrics* 80, (1997).

Phillips, P C B, “Laws and Limits of Econometrics,” *Economic Journal*, Vol. 113, No. 486, (March 2003), c26-c52.

Phillips, P C B and B Hansen, “Statistical Inference in Instrumental Variables Regression with I (1) Processes,” *Review of Economic Studies* 57, (1990), 99-125.

Phillips, P C B and M. Loretan, “Estimating Long-Run Equilibria”, *The Review of Economic Studies* 58(3) (1991), 407-436.

Phillips, P C B and S Ouliaris, “Asymptotic Properties of Residual Based Tests for Cointegration.” *Econometrica* 58 (1990), 165–193.

Phillips, P C B, “Time Series Regressions with a Unit Root”, *Econometrica* 55 (1987), 277-301.

Phillips, P C B, “Understanding Spurious Regressions in Econometrics,” *Journal of Econometrics*, Vol. 33, (1986), 311-40.

Razzak, W, “Monetary policy and forecasting inflation with and without the output gap,” *Reserve Bank of New Zealand Discussion Paper* 2002/03.

Sims, C A, “Bayesian Skepticism on Unit Root Econometrics,” *Journal of Economic Dynamics and Control*, Vol. 12, (June/September 1988), 463-74.

Said S and D A Dickey, “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order,” *Biometrika* 71 (1984), 599-608.

Stock, J, “Confidence Intervals for the Largest Autoregressive Root in U.S. Macroeconomic Time Series,” *Journal of Monetary Economics*, Vol. 28, (December 1991), 435-59.

West, K, “On the Interpretation of the Near Random Walk Behaviour in GNP,” *American Economic Review*, Vol. 78 (march 1988), 202-11.