NOTE

A NOTE ON SPURIOUS COINTEGRATION

Philip Hans FRANSES

Department of Econometrics
Erasmus University Rotterdam
Tinbergen Institute

1 INTRODUCTION

Recently, two new test procedures for cointegration between trended time series have become available. In Johansen and Juselius (1990) a maximum likelihood method based on a nonstationary vector autoregression is developed and applied. In Boswijk (1989) a test procedure is derived from a conditional dynamic model. In Boswijk (1990) and Franses (1990) both approaches have been applied to empirical series, and a common conclusion is that they can yield approximately the same inference. A second result in Franses (1990) is that inference from both methods can be highly sensitive to residual autocorrelation in the initial model. In fact, when there is such correlation, one will be more inclined to reject the null hypothesis of no cointegration. In other words, spurious cointegration may occur in case of misspecified models. In this note this occurence will be highlighted in a small simulation experiment. Although there may be many more simulations to do, the general implication of this note is that thoroughly testing the initial model should be an essential prerequisite for testing for cointegration.

2 A SMALL EXPERIMENT

Consider the case where there are two time series, y_t and x_t, which might be cointegrated via a single cointegration relationship. For convenience, only the test equation in Boswijk (1989, eq. 16) is considered, although the experiment may naturally be extended to the maximum likelihood method. The test equation is

$$\Delta y_t = \alpha_0 + \alpha_1 T + \alpha_2 \Delta x_t + \lambda_1 y_{t-1} + \lambda_2 x_{t-1} + \sum_{i=1}^{p-1} (\delta_i \Delta y_{t-1} + \phi_i \Delta x_{t-1}) + \theta_t \quad (1)$$

where T denotes a trend variable, Δ is defined by $\Delta z_t \equiv z_t - z_{t-1}$. In case θ_t is white noise, a table with critical values is available (Boswijk 1989, table B3). The usual procedure in case of more complicated dynamics in θ_t is to allow p to capture these dynamics. The null hypothesis of no cointegration is given by the restriction $\lambda_1 = \lambda_2 = 0$, for in that case only differenced variables enter (1), and hence (1) can not be written in an error correction form. This null hypothesis is tested with a Wald test.

In the simulation experiment the expression in (1) and the Wald test are used to establish the presence of cointegration in case there are two data generating processes (DGPs), which are given by

(*i*) $\Delta y_t = 0.5 + 0.5\Delta x_t + 0.6\Delta y_{t-1} + \varepsilon_t$, $\quad \Delta x_t = \nu_t$

(*ii*) $\Delta y_t = 0.5 + \varepsilon_t - 0.6\varepsilon_{t-1}$, $\quad \Delta x_t = \nu_t$

where ε_t and ν_t are independently drawn from a standard normal distribution. Note that in case *(ii)* the series y_t and x_t follow distinct time series processes. Furthermore, the error processes are such that the p in (1) should probably be set equal to some value larger than 1. The DGPs are generated for 150 observations, of which only the last 100 are used to reduce initialization effects. For calculating the results for several values of p, the same generations of the standard normal error series are used. The Monte Carlo exercise is based on 100 replications, and all calculations are performed with TSP version 6.53 (1989). The empirical sizes of the test, i.e. the proportion of times the null hypotheses of no cointegration are incorrectly rejected, are displayed in the table below.

From this table it can be seen that in case of residual autocorrelation the occurence of spurious cointegration emerges. This applies certainly to case *(ii)*, although the y_t and x_t are generated independently. On the other hand, in case the value of p is such that it is reasonable to expect that the dynamics are modeled, one can observe that the empirical size approaches the nominal size. Furthermore, from the results for *(i)* it can be concluded that setting p at too large a value may also affect the empirical size.

Table 1 Empirical size of a cointegration test (100 replications)

		Test equation		
DGP	Nominal size	$p=1$	$p=2$	$p=5$
(i)	0.01	0.10	0.01	0.04
	0.05	0.17	0.07	0.13
	0.10	0.21	0.13	0.18
(ii)	0.01	0.88	0.40	0.10
	0.05	0.93	0.54	0.19
	0.10	0.95	0.68	0.25

3 CONCLUDING REMARKS

The tentative implication of the above small experiment is that neglected or too many dynamics in the conditional dynamic model can induce spurious cointegration. Part of these results may support those in Schwert (1989) where the notion of spurious integration is investigated. Of course, many more simulations in the present context can be done, such as those in which other data generating processes and numbers of observations other than 100 are considered. Furthermore, it might be interesting to see whether the above results carry over to the maximum likelihood procedure.

Anyhow, appropriately specifying the dynamics in the initial model is an important device. In case of indecisiveness with respect to model selection, it might be sensible to consider various choices of an initial model, as is done e.g., in Franses (1990). Finally, it seems worthwhile to investigate the power of the test procedures in case the DGP contains complex error dynamics.

REFERENCES

Boswijk, H.P. (1989), "Estimation and Testing for Cointegration with Trended Variables: A Comparison of a Static and a Dynamic Regression Procedure", Report AE 12/89, University of Amsterdam.

Boswijk, H.P. (1990), "On the Scope of Conditional Dynamic Modelling of Cointegrated Variables", *Tinbergen Institute Research Bulletin*, 2, 97-108.

Franses, P.H. (1990), "Selecting a Dynamic Model for Stock of Cars", Unpublished Manuscript, Econometric Institute, Rotterdam.

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

Schwert, G.W. (1989), "Tests for Unit Roots: A Monte Carlo Investigation", *Journal of Business & Economic Statistics*, 7, 147-159.