On the scope of conditional dynamic modelling of cointegrated variables

Boswijk, H.P.

Published in:
Tinbergen Institute Research Bulletin

ON THE SCOPE OF CONDITIONAL DYNAMIC MODELLING OF COINTEGRATED VARIABLES

by

H. Peter BOSWIJK

Department of Actuarial Science and Econometrics
University of Amsterdam

1. INTRODUCTION

The theory of cointegration was developed by Engle and Granger (1987). The notion behind this theory is that economic time series that fluctuate widely individually with increasing variance, may be tied together by equilibrium relationships. Thus, although the original series are non-stationary, there may exist stationary functions (e.g. linear combinations) of those processes, representing deviations from equilibria.

One of the most striking results of the above cointegration theory is that inference on the parameters of the equilibrium relationships (the long-run multipliers) can be conducted consistently without having to specify the dynamic properties of these relationships. This led Engle and Granger to recommend a two-step procedure. The first step consists of the estimation of the long-run multipliers in a static regression equation (the cointegrating regression) and testing for cointegration. The dynamics is specified only in the second step, where an error-correction model is formulated and estimated using the residuals from the first step as equilibrium errors.

Recently this procedure has become very popular in empirical work. This may be explained by the ease of application and interpretation of the first step. Indeed, the notion of cointegration and the two-step procedure seem to be virtually conflated. In this paper however, a case is made for the specification of a dynamic model prior to the analysis of the long-run multipliers. The procedure we shall propose can be regarded as an intermediate case between the two-step procedure and the maximum likelihood

Tinbergen Institute Research Bulletin, Vol. 2 (1990), No. 2, pp. 97-108

analysis of the full vector autoregressive system, as proposed by Johansen (1988, 1989). We shall compare the merits of this conditional dynamic regression procedure with those of its two competitors, and discuss conditions for its applicability.

The plan of the paper is as follows. In section 2 we introduce some properties and representations of cointegrated systems. In the third section the dynamic regression procedure is presented. The underlying assumptions are discussed in section 4, where the proposed procedure is also compared with the two-step procedure and the full information maximum likelihood procedure. In section 5 an empirical case is presented where the dynamic procedure seems to be superior to the two-step procedure. In the final section some concluding remarks are made.

2. REPRESENTATIONS OF COINTEGRATED SYSTEMS

We consider an $n\times 1$ vector time series z_t, all components of which are integrated of order 1 ($I(1)$). This means that, although z is non-stationary, its first difference Δz_t is (second order) stationary, where $\Delta=(1-L)$ and L is the lag operator, such that $L^k z_t = z_{t-k}$ ($k = 1,2,\ldots$). The most significant properties of $I(1)$ series are that their variance increases with time, and that they contain a random walk or stochastic trend.

An $I(1)$ vector is said to be cointegrated if there exist r stationary linear combinations u_{it} ($i = 1,\ldots,r$, $0 < r <$ n) with:

$$u_{it} = \alpha_i' z_t \tag{2.1}$$

The cointegrating vectors α_i are not linearly dependent, so that the $n\times r$ matrix $\alpha = (\alpha_1,\ldots,\alpha_r)$ has rank r. Although the components of z_t contain a stochastic trend, this is not the case for the linear combinations u_{it}. Hence the components of z_t must have common trends that cancel out in (2.1). Note that the cointegrating matrix α is not unique; if $\alpha^* = \alpha Q$ with Q non-singular, then the columns of α^* also constitute cointegrating vectors, because $Q'u_t$ is stationary if u_t is. The occurrence of cointegration may be explained by the existence of economic equilibrium relationships ($\alpha_i'z = 0$). Because the economy prevents some components of z_t to drift too far from their equilibrium values, the deviations from equilibrium u_{it} are stationary.

This interpretation is in accordance with the error correction representation:

$$\Delta z_t = \mu + \gamma\alpha' z_{t-1} + \sum_{j=1}^{p} A_j \Delta z_{t-j} + \eta_t + \sum_{j=1}^{q} \delta_j \eta_{t-j} \tag{2.2}$$

or

$$A(L)\Delta z_t = \mu + \gamma\alpha' z_{t-1} + \delta(L)\eta_t \tag{2.3}$$

where μ is an n-vector of intercepts, α and γ are $n\times r$ matrices, $A(L)$ is an $n\times n$ matrix lag polynomial of order p, $\delta(L)$ is a scalar lag polynomial of order q, and η_t is an n dimensional white noise vector with covariance matrix Ω. From this representation we can see that the change in z_t follows a dynamic process but also depends on the vector of equilibrium errors from the previous period, $u_{t-1} = \alpha' z_{t-1}$. Under a stability condition on γ, (2.2) ensures that $\alpha' z_t$ is stationary. This is the consequence of the Granger Representation Theorem (see Engle and Granger (1987)), which states that cointegration implies (and is implied by) an error correction mechanism.

We assume that r dependent variables are associated with the r equilibrium relationships, analogous to the simultaneous equations model. This assumption is essentially an identification device. It enables a unique representation of the cointegrating matrix α to be identified from a conditional model for y_t given x_t. Hence we partition $z_t' = (y_t', x_t')$, where y_t is an $r\times 1$ vector of dependent variables and x_t is $(n-r)\times 1$ and contains exogenous variables. This classification is not meant to imply any causality from x to y, although such considerations will influence the choice of y_t. If each component of y_t enters at least one of the r cointegrating relationships, we can normalize α as

$$\alpha' = (I_r : -\Theta') \tag{2.4}$$

so that the equations $u_t = \alpha' z_t$ can be rearranged as

$$y_t = \Theta' x_t + u_t \tag{2.5}$$

The components of Θ are called long-run multipliers.

The aim of the conditional dynamic representation is to model explicitly the dynamics of the relationship between x_t and y_t, which is implicitly embodied in u_t in (2.5). We partition the parameters of (2.3) in correspondence with the partitioning of z_t, so that $\mu' = (\mu_1',\mu_2')$, $\gamma' = (\gamma_1',\gamma_2')$, $\eta_t' = (\eta_{1t}',\eta_{2t}')$ and

$$\mathrm{A}(L) = \begin{bmatrix} \mathrm{A}_{11}(L) & \mathrm{A}_{12}(L) \\ \mathrm{A}_{21}(L) & \mathrm{A}_{22}(L) \end{bmatrix}, \qquad \Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \tag{2.6}$$

In order to obtain a conditional dynamic model, we pre-multiply (2.3) by an nxn matrix P:

$$P = \begin{bmatrix} I_r & -\Omega_{12}\Omega_{22}^{-1} \\ 0 & I_{n-r} \end{bmatrix} \tag{2.7}$$

This results in

$$\Gamma(L)\Delta y_t = \kappa + \lambda(y_{t-1} - \Theta' x_{t-1}) + \mathrm{B}(L)\Delta x_t + \delta(L)\varepsilon_t \tag{2.8}$$

$$\mathrm{A}_{22}(L)\Delta x_t = \mu_2 + \gamma_2(y_{t-1} - \Theta' x_{t-1}) + \mathrm{A}_{21}(L)\Delta y_t + \delta(L)\eta_{2t} \tag{2.9}$$

with

$$\Gamma(L) = \mathrm{A}_{11}(L) - \Omega_{12}\Omega_{22}^{-1}\mathrm{A}_{21}(L), \qquad \kappa = \mu_1 - \Omega_{12}\Omega_{22}^{-1}\mu_2,$$

$$\mathrm{B}(L) = \mathrm{A}_{21}(L) - \Omega_{12}\Omega_{22}^{-1}\mathrm{A}_{22}(L), \qquad \lambda = \gamma_1 - \Omega_{12}\Omega_{22}^{-1}\gamma_2.$$

Note that $\Gamma(0) = I_r$, $\mathrm{B}(0) = \Omega_{12}\Omega_{22}^{-1}$, $\mathrm{A}_{22}(0) = I_{n-r}$ and $\mathrm{A}_{21}(0) = 0$. Furthermore, it is easily seen that $P\Omega P'$ is block-diagonal, so that $\mathrm{E}[\varepsilon_t\eta_{2s}'] = 0$, $\forall\ t,s$. Therefore (2.8) may be interpreted as a conditional model for Δy_t given Δx_t and the past history of y_t and x_t. This interpretation is justified if η_t is a Gaussian innovation process, which we shall assume in the sequel (see Boswijk (1989) for a more elaborate derivation of the conditional model). Note that Δx_t is predetermined in (2.8), because only lagged y_t's enter (2.9) and η_{2t} is uncorrelated with ε_s, $\forall\ s,t$.

3. A DYNAMIC REGRESSION PROCEDURE

The aim of the procedure proposed in this section is inference on the long-run relationship(s) between the components of z_t. We distinguish three stages :

1. Testing for cointegration,
2. Estimation of the cointegrating vector(s),
3. Hypothesis testing on the cointegrating vector(s).

We shall first briefly outline these stages for the Engle and Granger procedure and for the Johansen analysis; next we shall discuss them for the conditional dynamic model.

The first step of the two-step procedure is to estimate the long-run multipliers by OLS in the separate equations of (2.5); serial correlation and correlation between u_t and x_t is neglected at this stage, because it does not influence the consistency of the estimators. Next, the residuals of the regressions are used for a test on cointegration. Engle and Granger do not discuss hypothesis tests on the cointegrating vectors. These are analyzed in Park and Phillips (1988), where it is shown that the dynamic misspecification of (2.5) induces quite some problems in the construction of test statistics, which can only be solved by rather complicated correction methods.

In the analysis of Johansen the cointegrating matrix α is estimated from the system of equations (2.2), under the assumptions that $\delta(L) = 1$ and η_t i.i.d. $N(0,\Omega)$. The number of cointegrating relationships r is established via a sequence of likelihood ratio tests, which have a non-standard distribution, tabulated in Johansen (1989). Finally, hypotheses on both α and γ can be tested by likelihood ratio tests, which have the usual χ^2 distribution under the null hypothesis.

Central to the procedure proposed in this paper is the conditional dynamic model (2.8). If $r = 1$, it reduces to

$$
\begin{aligned}
\gamma(L)\Delta y_t &= \kappa + \lambda(y_{t-1} - \theta' x_{t-1}) + \beta(L)'\Delta x_t + \delta(L)\varepsilon_t \\
& \kappa + \lambda y_{t-1} + \pi' x_{t-1} + \beta(L)'\Delta x_t + \delta(L)\varepsilon_t \qquad (3.1)
\end{aligned}
$$

where $\pi = -\lambda\theta$ is an $(n-1)$-vector, $\gamma(L)$ is a p-th order lag polynomial and $\beta(L)$ is an $(n-1)\times 1$ vector lag polynomial of the same order, and κ, λ and ε_t are scalars. Note that this is a reparametrized ARMAX model. Here we shall concentrate on this case, but in the next section we shall discuss the possible generalizations to multivariate models for the case where $r>1$.

3.1 Testing for cointegration

The Granger Representation Theorem tells us that cointegration is equivalent with an error correction model. Therefore a test on the significance of the error correction term is in fact a test for the null hypothesis that z_t is not cointegrated against the alternative of cointegration. Such a test can be performed in at least two ways.

The first one is a Wald test on the n restrictions $\lambda = 0$ and $\pi = 0$ in (3.1). This test was proposed by Boswijk (1989) for the case that $\delta(L) = 1$, i.e. in the context of an ARX or AD (autoregressive-distributed lag) model. In that paper the asymptotic distribution of this test under the null hypothesis is derived and tabulated. Because the series are non-stationary, this asymptotic distribution differs from the χ^2_n distribution and has larger critical values.

The second one was proposed by Dolado *et al.* (1989). They consider a t-test on the significance of the error correction term, where θ is replaced by a hypothesized value. The conceptual disadvantage of this procedure is that it combines stages 1 and 3. However, it may have better power properties than the Wald test.

3.2 Estimation of the cointegrating vector

Because of the normalization restriction, the estimation of the cointegrating vector of (3.1) amounts to the estimation of the vector of long-run multipliers θ. If $\delta(L) = 1$, an obvious estimator of θ is the indirect least squares estimator

$$\tilde{\theta} = -\hat{\pi}/\hat{\lambda}, \tag{3.2}$$

with $\hat{\pi}$ and $\hat{\lambda}$ the OLS estimators of (3.1). This estimator is consistent for θ. If $\delta(L) \neq 1$, we could replace the OLS estimators by estimators that take into account the moving average structure of the disturbances. The statistical properties of this generalization have not yet been investigated.

3.3 Hypothesis testing on the cointegrating vector

We consider the hypothesis that θ satisfies a linear restriction :

$$H_0 \, : \, R\theta \, = \, q. \tag{3.3}$$

We can construct a Wald test for this restriction using an asymptotic covariance matrix derived from a first order Taylor series expansion of (3.2). However, better small sample behaviour can be expected from a Wald test on a linearized version of H_0, viz.

$$H_0' \, : \, R\pi \, + \, q\lambda \, = \, 0 \tag{3.4}$$

These two tests are shown to be asymptotically equivalent under the null and a sequence of local alternatives in Boswijk (1990). Under the null they are asymptotically distributed as χ^2 only if x_t is weakly exogenous for θ in (3.1) (see Dolado *et al.* (1989) and the next section).

4. DISCUSSION

The most essential assumption for the conditional dynamic approach is that the normalization (2.4) of α is valid. This entails that each component of y_t is part of at least one cointegrating relationship, and that x_t is not cointegrated. If this assumption is not valid, α cannot be identified from (2.8). The question whether or not x_t is cointegrated can be tested with the above mentioned tests. If it is not, we can proceed to the next question, which is whether y_t is cointegrated with x_t.

If $r > 1$, problems can arise because we may have modelled too many or too few long-run relationships. In the first case α will not be identified, whereas in the latter case the model will be unstable and Θ will not be defined. Therefore, we need a test to assess the number of cointegrating relationships. A generalization of the Wald test mentioned in the previous section is a test for the significance of λ and $\lambda\Theta$ in (2.8). This test has a null hypothesis of no cointegration and an alternative of r cointegrating vectors. Therefore it is not appropriate for this problem, because it can not test intermediate cases. Once r is correctly specified, the estimation and testing methods from the previous section may be generalized to multivariate models.

Another condition which influences the performance of the proposed procedure is weak exogeneity of x_t. Let Z_t denote the collection of all observations on z_s up to period t, including all relevant initial values. The density of the vector z_t conditional upon its past can be factorized as:

$$D(z_t|Z_{t-1},\psi) = D_1(y_t|x_t,Z_{t-1},\psi_1)\ D_2(x_t|Z_{t-1},\psi_2), \tag{4.1}$$

where ψ are the parameters of the data generating process of z_t, and ψ_1 and ψ_2 are the parameters of the conditional density of y_t and the marginal density of x_t, respectively. x_t is said to be weakly exogenous for the parameters of interest θ, if θ is defined solely from ψ_1 and is not related to ψ_2. If we assume that (2.2) represents the data generating process of z_t, then D_1 and D_2 are characterized by (2.8) and (2.9). Because the parameters of interest are the components of θ, the vector of long-run multipliers, x_t is weakly exogenous if and only if $\gamma_2 = 0$, i.e. if x_t is not error-correcting. This is a property of the system (2.2), so it can be tested in the full information maximum likelihood framework.

The assumption of weak exogeneity is not essential to the first two stages of the conditional dynamic procedure. It is however necessary for the asymptotic distribution of the hypothesis tests on θ to be χ^2. If it does not hold, then this distribution depends on a nuisance parameter (see Park and Phillips (1988)). The assumption of weak exogeneity is also beneficial to the performance of the dynamic procedure because it means that efficient inference on θ is possible in this setting.

Comparing the conditional dynamic approach to the two-step procedure, there seems to be little rationale for the latter. The methods are not computationally more involved, except when moving average effects are present (if $\delta(L) \neq 1$). We can expect the indirect least squares estimator $\tilde{\theta}$ to be less biased and more efficient than the static regression estimator, which has only asymptotic justification. For some Monte Carlo evidence of this, see Banerjee *et al.* (1986). Moreover, the construction of hypothesis tests is considerably easier in a dynamic framework.

The choice between our procedure and the full information maximum likelihood analysis is less clear-cut. The latter is definitely more flexible with respect to the number of cointegrating vectors and the normalization of α, and does not require any exogeneity assumptions. Still there are a number of cases where the conditional model is preferable.

First, for macroeconomic time series the number of observations is usually quite small, which may hinder a sound dynamic specification of a vector autoregressive system; an extra lag implies n^2 additional parameters to be estimated. Second, a conditional model may exhibit more structural stability and less sensitivity to regime changes than a vector autoregression. Finally, if x_t is weakly exogenous then the conditional dynamic procedure is efficient, so that there is no need to estimate the full system.

5. AN APPLICATION : THE DEMAND FOR MONEY IN DENMARK

In this section we shall apply the techniques considered in section 3 to a model for the demand for money in Denmark. Quarterly data (1974/I - 1987/III) on the log of real broad (M2) money m, the log of real income y, the bond interest rate i^b and the deposit rate i^d are from Johansen (1989), where the maximum likelihood techniques are applied. We shall briefly survey his results.

The outcome of the likelihood ratio tests on cointegration is that there is one long-run relationship between m_t, y_t, i^b_t and i^d_t, i.e. $r = 1$. With the coefficient of m normalized to 1, the estimated long-run equation is,

$$m = 1.03\ y - 5.21\ i^b + 4.22\ i^d + 6.06 \tag{5.1}$$

The restrictions that the long-run multiplier of income is equal to one ($\theta_1 = 1$) and that the parameters of i^b and i^d are equal except for opposite signs ($\theta_2 = -\theta_3$) cannot be rejected. The restricted long-run equation looks like

$$m - y = -5.88\ (i^b - i^d) + 6.21 \tag{5.2}$$

Note that this is a relationship between minus the log of the velocity of money ($m - y$) and the cost of holding money ($i^b - i^d$), because m includes interest-bearing money. The estimates of the error correction coefficients γ (see (2.2)) are found to be $\hat{\gamma} = (-0.177,\ 0.095,\ 0.023,\ 0.032)$. In respect of the discussion in the previous section, it is of interest to test whether γ can be restricted to $(\gamma_1, 0, 0, 0)'$, because in that case $(y_t, i^b, i^d)'$ is weakly exogenous for the long-run parameters. The outcome of this test depends on whether or not θ is restricted; on the whole the validity of this restriction seems doubtful.

Using the same data we estimate an equation like (3.1), with $p = 1$, $q = 0$ and three quarterly dummies added to the regressors (as in Johansen (1989)). As we want to model the demand for money, we choose m_t as the dependent variable and y_t, i^b and i^d as exogenous variables. This equation appears to be misspecified; the Lagrange Multiplier test for first order serial correlation yields a statistic of $F(1,37) = 4.64$, and the $\chi^2(2)$ test statistic on normality of the disturbances equals 110.83. After some vain experimentation with the lag structure, both problems were "solved" by the inclusion of a dummy variable for 1984/IV; in that period a large outlier appears in the original residuals (for an explanation of this outlier, see Juselius (1989)).

The Wald test statistic on cointegration (see section 3.1) is equal to 40.3, which is highly significant at a size of one per cent, so that the null hypothesis of no cointegration is convincingly rejected. Our estimated long-run relationship is

$$m = 0.96\ y - 4.59\ i^b + 2.59 i^d + 6.52 \tag{5.3}$$

Note that the estimators are reasonably close to the maximum likelihood estimators, although the estimated effect of the deposit rate is smaller. We use the second version of the Wald test mentioned in section 3.3 for the hypotheses $(\theta_1 = 1)$ and $(\theta_2 = -\theta_3)$. The first hypothesis results in an F-type statistic of $F(1,37) = 0.09$. Note that we can only use critical values from the F distribution if $(y_t, i^b, i^d)'$ is weakly exogenous. However, if this is not the case critical values will be larger, so that the null hypothesis $(\theta_1 = 0)$ can in no way be rejected. The test statistic for the second hypothesis equals $F(1,38) = 8.10$. Although we can not make a clear decision because the F-distribution may not be applicable, the value of the test statistic at least casts some doubt on this hypothesis. Estimation with the restriction $(\theta_1 = 1)$ imposed yields

$$m - y = -4.51\ i^b + 2.55\ i^d + 6.29, \tag{5.4}$$

whereas the estimated error correction coefficient is $\hat{\lambda} = -0.28$.

Finally, we estimate the long-run parameters in a static regression equation:

$$m = 1.29\ y - 2.63\ i^b + 0.64\ i^d + 4.40. \tag{5.5}$$

Note that this equation differs considerably from both (5.1) and (5.3). Although we have not attempted to construct test statistics for the hypotheses, it seems likely that they will be rejected. Next, we perform an Augmented Dickey-Fuller test with four lags, since fewer lags induce serial correlation (possibly due to seasonal effects). The test statistic equals -3.84; comparing this with table 2b from Phillips and Ouliaris (1988) we conclude that the hypothesis of *no* cointegration is not rejected by this procedure at a five per cent significance level and is only scarcely rejected at a ten per cent significance level.

Summarizing the empirical results of this section we may conclude that the dynamic approach performs quite satisfactorily compared to the maximum likelihood procedure. As far as testing for cointegration and estimation of the cointegrating vector is concerned, the results of the two procedures are very similar, even though the conditioning variables may not be weakly exogenous. The two-step procedure fails to identify a cointegrating relationship; if we ignore this and interpret the static regression estimates as long-run multiplier estimates, they differ quite sharply from the maximum likelihood estimates.

6. CONCLUDING REMARKS

In this paper we have proposed a procedure to conduct inference on cointegrating relationships in conditional dynamic models. We have argued that if more than one long-run relationship is suspected, the scope of this procedure may be limited and the maximum likelihood analysis of the full system will be more flexible. If this is not the case the proposed framework may provide a good alternative, especially if only a small sample is available. In a simple empirical application we have seen a case in which both procedures to a large extent yield the same conclusions, whereas the static regression procedure seems to break down.

ACKNOWLEDGEMENTS

Helpful comments from Philip Hans Franses and Jan Kiviet are gratefully acknowledged.

REFERENCES

Banerjee, A., Dolado, J.J., Hendry D.F. and G.W. Smith (1986), "Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence", *Oxford Bulletin of Economics and Statistics*, 48, 253-277.

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), "Hypothesis Testing on Long-Run Multipliers in Dynamic Linear Regression Models with Integrated Processes", unpublished paper, University of Amsterdam.

Dolado, J.J., Ericsson N.R. and J.J. Kremers (1989), "Inference in Conditional Dynamic Models with Integrated Variables", paper presented at the ESEM 1989, Munich.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55, 251-276.

Johansen, S. (1988), "Statistical Analysis of Cointegrating Vectors", *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1989), *Likelihood Based Inference on Cointegration. Theory and Applications*, Centro Universitario di Econometria, Bologna.

Juselius, K. (1989), "Stationary Equilibrium Error Processes in the Danish Money Market. An Application of ML Cointegration", paper presented at the ESEM 1989, Munich.

Park, J.Y. and P.C.B. Phillips (1988), "Statistical Inference in Regressions with Integrated Processes: Part 1", *Econometric Theory*, 4, 468-497.

Phillips, P.C.B. and S. Ouliaris (1988), "Asymptotic Properties of Residual Based Tests for Cointegration", Cowles Foundation Discussion Paper No. 847-R, Yale University, forthcoming in *Econometrica*.