MONASH

No. 16/92

MONASH
UNIVERSITY

AUSTRALIA

DIFFERENT ESTIMATORS OF COINTEGRATING VECTORS

AND THEIR IMPACT ON SHORT RUN DYNAMICS

Baiding Hu and Brett Inder

Working Paper No. 16/92

December 1992

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 7326 0366 8

DIFFERENT ESTIMATORS OF COINTEGRATING VECTORS AND THEIR IMPACT ON SHORT RUN DYNAMICS

Baiding Hu and Brett Inder

Working Paper No. 16/92

December 1992

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS COMMERCE & MANAGEMENT

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

Different Estimators of Cointegrating Vectors and their Impact on Short Run Dynamics

by

Baiding Hu
La Trobe University

and

Brett Inder
Monash University

Address all correspondence to:

Baiding Hu
Department of Econometrics
La Trobe University
Bundoora Victoria 3083
Australia

ABSTRACT

We use a Monte Carlo study to compare the precision of estimates of the parameters of an Error Correction Model when different estimators of the long run relationship are employed. We also compare forecasting performance.

October 1992

JEL Classification Number: C22

1. INTRODUCTION

In the context of modelling economic time series with unit roots, the Error Correction Model (ECM) representation suggests a natural strategy for identifying and estimating the long run and short run characteristics of a relationship. Engle and Granger (1987) introduced the two-step procedure, the first step of which enables estimation of the cointegrating vector by a straightforward ordinary least squares (OLS) regression involving the levels of the variables. All dynamics can be ignored and endogeneity of any of the variables has no effect asymptotically due to the unit root in the data. In the second step of the procedure, these estimated long run parameters are used in an ECM to obtain estimates of the short run dynamics.

Many authors have expressed concern at the use of OLS in levels to estimate the cointegrating vector. There is a clear loss of efficiency through omitting dynamics, and inference based on standard test statistics is in general not valid. Consequently, a number of alternative methods for estimating cointegrating vectors have been proposed (see Phillips and Loretan, 1991, for one recent survey). These represent different approaches to allowing for the dynamics in the system. Comparative studies also exist which seek to determine which of these alternatives is the most appropriate (again see Phillips and Loretan, Phillips, 1988, Inder, 1993, among others). The criteria for evaluating seem to be related to the degree of precision or bias in the estimators of the cointegrating vector, as well as validity of inference. The conclusion reached in most studies is that it matters which estimator is used; those which allow for dynamics in the data generating process and possible endogeneity of the explanatory variables seem to give much more reliable estimates and inference.

As various new estimators of the cointegrating vector have been proposed, the question of how the short run parameters are to be estimated has generally been overlooked. Presumably one could follow the approach of Engle and Granger (1987): find estimates of the cointegrating vector, and then substitute these into the ECM to estimate the short run dynamics.

The question we wish to address in this note is: *does a different choice of estimator for the cointegrating vector have much effect on estimation of the short run dynamics in the ECM?* Our basis for comparison is the root mean square error (RMSE) of estimates of the parameters of the ECM, and mean square error (MSE) of forecasts up to four periods ahead. It could be expected that different estimates of the cointegrating vector may have little effect on short run estimates, as the asymptotic properties of the various estimators are the same. Differences may, however show up in finite samples. We investigate this possibility by way of a simple Monte Carlo (MC) study involving two different estimators of the co-integrating vector.

2. MODEL AND MONTE CARLO DESIGN

Consider the following model for y_t

$$\alpha(L)y_t = \mu + \beta'(L)x_t + u_t\,, \quad (t=1,.....,T) \tag{1}$$

where y_t is a scalar, x_t is a k x 1 vector of explanatory variables, u_t is a stationary error term, and $\alpha(L)$ and $\beta(L)$ are p^{th} and q^{th} order polynomials in the lag operator.

We assume that x_t is generated by the process

$$x_t = x_{t-1} + \delta + v_t\,, \tag{2}$$

where v_t is stationary. By (2) we are implying that each regressor is integrated of order 1, or I(1), and (1) implies that y_t and x_t are cointegrated.

The cointegrating vector implied by (1) is (-1, λ'), where $\lambda = \beta(1)/\alpha(1)$. An ECM representation of (1) is given by

$$\delta_1(L)\Delta y_t = \mu + \delta_2'(L)\,\Delta x_t - \alpha(1)[y_{t-p} - \lambda' x_{t-p}] \tag{3}$$

where $\delta_1(L) = \dfrac{\alpha(L) - \alpha(1)L^p}{1-L}$, and $\delta_2(L) = \dfrac{\beta(L) - \beta(1)L^p}{1-L}$

(see Inder, 1993). Having obtained estimates of λ, the remaining parameters of (3) can be estimated by OLS. It is these estimates we wish to evaluate. We will also investigate the forecasting performance of the model with different estimates of λ.

Engle and Granger (1987) show that λ can be estimated consistently by an OLS regression of y_t on x_t; we will describe this as the OLS estimator. The alternative estimator considered in this paper involves estimating (1) by OLS, and then estimating λ by $\hat{\beta}(1)/\hat{\alpha}(1)$; that is, all possible dynamics is included in the model. This we will describe as the unrestricted ECM estimator [see Inder (1993), Banerjee et al (1986) and Stock (1987)].

A more comprehensive study would need to include other contenders, including Phillips and Hansen's (1990) fully modified OLS, Saikkonen's (1991) asymptotically efficient estimator, and others. This study could be seen as a pilot investigation of the question; the results below based on just OLS and unrestricted ECM estimators will give us a clue as to the effect of improved efficiency in the long run estimation on ECM estimation.

The details of the design of the MC study are given below. There are three regressors with the disturbances generated by

$$u_t = \rho_{11} u_{t-1} + \varepsilon_t ,$$

$$v_t = \rho_2 v_{t-1} + e_t ,$$

where $\rho_2 = \text{diag}(\rho_{21}, \rho_{22}, \rho_{23})$, and ε_t and e_t are independent and identically distributed standard normal variables. This implies that x_t is exogenous in (1), making the unrestricted ECM estimator asymptotically valid (see Phillips, 1988).

The parameters of the data generating process (DGP) are given by:
T=50 and 200, k=3, p=q=4, μ=0.2, $(\rho_{11}, \rho_{21}, \rho_{22}, \rho_{23})$ = (0.5, 0.4, 0.3 0.2), δ' = (0.191, 1.759, 1.42), β_0' = (0.58, 0.45, -0.78). Models were then generated with no dynamics $(\beta_1' = \beta_2' = \beta_3' = \beta_4', \alpha' = 0)$, dynamics in terms of either lagged y (α' = (0.2, 0.15, 0.1, 0.07) or lagged x (β_1' = (0.55, 0.3, -0.68), β_2' = (0.5, 0.2, -0.6), β_3' = (0.45, 0.2, -0.5), β_4' = (0.4, 0.14, -0.32)), and then with dynamics including both lags of y and of x. Initial x and y values were selected to give realistic looking series;

specifically,

$y_0 = y_{-1} = y_{-2} = y_{-3} = 3, \quad x_0' = (10, 7, 9), \quad x_{-1}' = (10, 6, 5), \; x_{-2}' = (10, 8, 10),$

$x_{-3}' = (10, 10, 9).$

All results are based on 5000 replications for each experiment.

3. RESULTS AND DISCUSSION

3.1 Precision of Estimates

The overall precision is measured by RMSE, which is obtained by averaging the individual RMSE's for each estimated short run parameter. Table 1 contains all the RMSE's for the different DGPs and estimators that are employed in this study. The following observations may be made about the results:

(i) When all the dynamics are included in the DGP, the Unrestricted ECM (UECM) exhibits noticeable superiority over OLS. In the small sample case (T=50), the latter's RMSE is 0.39, more than twice as big as that of the former (0.17). Even in the large sample case where T=200, the superiority of the UECM over OLS is still preserved as the RMSEs for the OLS and UECM are 0.28 and 0.21 respectively.

(ii) The overall precision appears better again for the UECM than for the OLS if only partial dynamics are included in the DGP $(\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ or $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0)$, i.e. the UECM model is partially overspecified. The table presents to us a picture that as long as there are dynamics in the DGP, the UECM will yield better short run parameter estimates. Specifically, in small samples (T=50) when all lagged explanatory variables are excluded $(\beta_1 = \beta_2 = \beta_3 = 0)$, the UECM gives an RMSE of 0.17 as compared to 0.21

given by the OLS. As sample size increases to 200, the former has an RMSE of 0.10 while the latter 0.12. When all lagged dependent variables are excluded from (9) $(\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0)$, the RMSEs of the UECM and the OLS are 0.17 versus 0.35 in the small sample size (T=50) and 0.18 versus 0.25 in the large sample size (T=200).

(iii) In the case where no dynamics are present in the DGP, the UECM model is completely overspecified. From the table we see that only under these circumstances does OLS defeat the UECM. The former has a RMSE of 0.17 versus 0.19 for the latter in small samples and 0.1 versus 0.11 in large samples. Thus, it appears that the overspecification problem matters less in large samples than in small samples. The price paid in overspecifying the dynamics seems to be very small though, compared to the hazards of underspecifying.

3.2 Forecasting performance of the ECM

It may be reasonable to use the ECM to forecast economic variables. Given the two different sets of long run and short run parameter estimates, we are able to compare the estimators' forecasting performance. Table 2 contains the results on the precision of the forecasts of the estimated ECM. We observe that, similar to the situation above, exclusion of dynamics does incur a substantial penalty. When all dynamics are present and T=50, the forecasting error of OLS exceeds that of the UECM, and the magnitude of the difference between the two estimators' forecasting error increases as the forecasting period becomes greater. This holds for all cases where dynamics are present. The difference becomes smaller as the dynamics are gradually removed from the DGP. Not surprisingly, OLS outperforms the UECM when dynamics are absent. Regardless of the presence or absence of dynamics, the difference between the forecasting errors of the two estimators tends to be negligible as sample size increases. This indicates that forecasting precision may be insensitive to the dynamics in the DGP so long as the sample size is large.

4. Conclusions

This paper has been motivated by the fact that the estimation of long run parameters has received a great deal of attention and the estimation of dynamics, in contrast, seems to be largely ignored. As such, this paper investigates the effects of the choice of estimator of the long run parameters on that of the short run parameters. It turns out that an estimator that gives more precise long run parameter estimates also gives more precise short run parameter estimates (the ECM parameters). Our Monte Carlo evidence has revealed that as far as estimation of short run parameters is concerned it is unwise to discard possible dynamics. When samples are small, disregarding possible dynamics in the first-step can only lead to much poorer precision of long run parameter estimates as well as that of short run parameters in the second-step. This situation can be alleviated somewhat when samples are large.

Since I(1) variables are well modelled by an ECM, the task of forecasting I(1) variables can be performed within the ECM framework. Our Monte Carlo results have brought some ideas as to the precision of forecasts of an ECM when its long run part is estimated by the two different estimators. The evidence has shown that the forecasting precision of an ECM is sensitive to the choice of estimator of the long run parameters. The estimator that gives good short run parameter estimates often brings about more reliable forecasts.

Table 1

RMSEs of the short-run parameters estimates

DGP		Estimators in 1st-step	
		OLS	Unrestricted ECM
Dynamics in both x and y	T=50	0.390	0.170
	T=200	0.280	0.210
Dynamics in x only	T=50	0.350	0.170
	T=200	0.250	0.180
Dynamics in y only	T=50	0.210	0.170
	T=200	0.120	0.100
No Dynamics	T=50	0.170	0.190
	T=200	0.100	0.110

Table 2

Precision of Forecasts from the ECM

DGP	Estimators in 1st-step		OLS	Unrestricted ECM
Dynamics in both x and y	Forecasting MSE T=50	1st-period	14.391	9.763
		2nd-period	15.571	10.631
		3rd-period	17.498	10.834
		4th-period	19.814	11.506
	Forecasting MSE T=200	1st-period	7.835	7.496
		2nd-period	8.168	7.688
		3rd-period	8.223	7.601
		4th-period	8.326	7.572
Dynamics in x	Forecasting MSE T=50	1st-period	13.278	9.702
		2nd-period	14.452	10.413
		3rd-period	15.905	10.519
		4th-period	17.804	11.154
	Forecasting MSE T=200	1st-period	7.534	7.194
		2nd-period	7.848	7.420
		3rd-period	7.819	7.362
		4th-period	7.940	7.398
Dynamics in y	Forecasting MSE T=50	1st-period	11.032	9.664
		2nd-period	11.798	10.265
		3rd-period	12.783	10.534
		4th-period	13.801	11.058
	Forecasting MSE T=200	1st-period	7.076	6.934
		2nd-period	7.379	7.200
		3rd-period	7.199	7.026
		4th-period	7.335	7.094
No Dynamics	Forecasting MSE T=50	1st-period	9.391	9.684
		2nd-period	9.724	10.198
		3rd-period	9.855	10.408
		4th-period	10.181	10.944
	Forecasting MSE T=200	1st-period	6.932	6.943
		2nd-period	7.201	7.321
		3rd-period	6.966	6.989
		4th-period	7.074	7.053

References

Banerjee, A., Dolalo, 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.

Bewley, R.A., 1979, The direct estimation of the equilibrium response in a linear dynamic model, Economics Letters 3, 357-361.

Engle, R.F. and C.W.J. Granger, 1987, Co-integration and error correction: representation, estimation and testing, Econometrica 55, 251-276.

Granger, C.W.J. and A.A. Weiss, 1983, Time series analysis of error-correcting models, in Samuel Karlin, Takeshi Amemiya and Leo A. Goodman. eds. Studies in econometrics, time series, and multivariate statistics, Academic Press, New York.

Inder, B. , 1993, Estimating long run relationships in Economics: A comparison of different approaches, forthcoming, Journal of Econometrics.

Johansen, S., 1988, Statistical analysis of Cointegrating vectors, Journal of Economic Dynamics and Control 12, 231-54.

Nelson, C.R. and C. Plosser, 1982, Trends and random walks in macroeconomic time series; Some evidence and implications, Journal of Monetary Economics 10, 139-62.

Phillips, P.C.B. and B.E. Hansen, 1990, Statistical inference in instrumental variables regression with I(1) processes, Review of Economic Studies 57, 99-125.

Phillips, P.C.B. and M. Loretan, 1991, Estimating long run equilibria, Review of Economic Studies 58, 407-436.

Saikkonen, P. , 1991, Asymptotically efficient estimation of cointegration regressions, Econometric Theory 7, 1-21.

Stock, J. ,1987, Asymptotic properties of least square estimators of cointegrating vectors, Econometrica 55, 381-86.