

The Role of “Leads” in the Dynamic OLS Estimation of Cointegrating Regression Models

Kazuhiko Hayakawa* Eiji Kurozumi†

February 21, 2008

Abstract

In this paper, we consider the role of “leads” of the first difference of integrated variables in the dynamic OLS estimation of cointegrating regression models. Specifically, we investigate Stock and Watson’s (1993) claim that the role of leads is related to the concept of Granger causality by a Monte Carlo simulation. From the simulation results, we find that the dynamic OLS estimator without leads substantially outperforms that with leads and lags; we therefore recommend testing for Granger non-causality before estimating models.

JEL classification: C13; C22

Key Words: Cointegration; dynamic ordinary least squares estimator; Granger causality

*Hayakawa is a JSPS research fellow and acknowledges its financial support.

†Kurozumi’s research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology under Grants-in-Aid No. 17203016 and 18730142.

1 Introduction

Since the seminal work of Engle and Granger (1987), cointegrating regressions have become one of the standard tools in analyzing integrated ($I(1)$) variables. Although the ordinary least squares (OLS) estimator is consistent in the presence of a serial correlation in the error term and/or a correlation between the regressors and cointegration errors, it is well known that the OLS estimator contains the so-called second-order bias. In the literature, there are three typical estimators that deal with this problem: the fully modified OLS estimator proposed by Phillips and Hansen (1990), Park's (1992) canonical cointegrating regression estimator, and the dynamic OLS (DOLS) estimator of Phillips and Loretan (1991), Saikkonen (1991), and Stock and Watson (1993). These three estimators are known to be asymptotically equivalent and efficient. In this paper, we focus on the DOLS estimator among the three estimators and consider the role of “leads” of the first difference of the integrated variables in DOLS. Specifically, we focus on Stock and Watson's (1993, p.786) claim that leads are unnecessary if the cointegrating error does not Granger-cause the first-difference of $I(1)$ variables that appear in the right side. To the best of authors' knowledge, no studies have investigated the finite sample performance of the DOLS estimator without leads after Granger non-causality test. Thus, we investigate the case where leads are unnecessary for the DOLS method, and by using the Monte Carlo simulation, we demonstrate that in such a case, we can expect the improvement of the DOLS estimator in terms of the mean squared error (MSE) by excluding leads from the regressors.

2 Relation between Leads and Granger Causality

We consider a typical cointegrating regression model as follows:

$$\begin{aligned} y_t &= \alpha + \beta' x_t + u_{1t} = \theta' z_t + u_{1t} \\ \Delta x_t &= u_{2t} \end{aligned} \tag{1}$$

where $\theta = [\alpha, \beta']'$, $z_t = [1, x_t']'$ and x_t is an n -dimensional I(1) vector. As in Saikkonen (1991) we assume that $u_t = [u_{1t}, u_{2t}']'$ is a stationary process with

$$\sum_{j=-\infty}^{\infty} \|E[u_t u_{t-j}]\| < \infty, \quad \sum_{m_1, m_2, m_3=-\infty}^{\infty} |\kappa_{ijkl}(m_1, m_2, m_3)| < \infty$$

where $\kappa_{ijkl}(m_1, m_2, m_3)$ is the fourth order cumulant of u_t , and the spectral density matrix of u_t , denoted by $f_{uu}(\lambda)$, is bounded away from zero. Under this condition, u_{1t} is expressed as

$$u_{1t} = \sum_{j=-\infty}^{\infty} \Pi'_j u_{2t-j} + v_j \tag{2}$$

where

$$\Pi_j = \frac{1}{2\pi} \int_0^{2\pi} f_{12}(\lambda) f_{22}(\lambda)^{-1} \exp(ij\lambda) d\lambda$$

with $f_{12}(\lambda)$ and $f_{22}(\lambda)$ are the (1,2) and (2,2) blocks of $f_{uu}(\lambda)$. It is known that $\sum_{j=-\infty}^{\infty} \|\Pi_j\| < \infty$ and that v_t is a stationary process such that $E(u_{2s}v_t) = 0$ for all s and t . See also Brillinger (1981). By inserting (2) into (1), the model can be expressed as

$$y_t = \alpha + \beta' x_t + \sum_{j=-K}^K \Pi'_j \Delta x_{t-j} + \dot{v}_t \tag{3}$$

where $\dot{v}_t = v_t + \sum_{|j|>K} \Pi'_j u_{2t-j}$ and K is known as the lead-lag truncation parameter. Saikkonen (1991) showed that the OLS estimator of β based on (3) does not suffer from the second-order bias and is efficient in a certain class of distributions.

Let us consider the case where

$$\Pi_j = 0 \quad \text{for } \forall j < 0. \tag{4}$$

In this case, the model becomes

$$y_t = \alpha + \beta' x_t + \sum_{j=0}^K \Pi'_j \Delta x_{t-j} + \dot{v}_t \tag{5}$$

and then we do not have to include the leads of Δx_t as regressors. We therefore expect an improvement of the finite sample efficiency by estimating (5) because we do not have to include extra regressors. In this case, we note that condition (4) is related to the concept of Granger causality. According to Sims (1972) and Proposition 11.3 in Hamilton (1994), condition (4) holds if and only if u_{1t} does not Granger-cause

u_{2t} . In other words, it is possible to efficiently estimate the cointegrating regression model without any leads of the first difference of integrated variables if the past values of u_{1t} do not help to predict u_{2t} . Therefore, we recommend that the null of Granger non-causality be tested before estimating the cointegrating regression model.

Tests for Granger non-causality can be conducted by approximating the process of u_t by a finite-order vector autoregressive model: $u_t = \Psi_1 u_{t-1} + \Psi_2 u_{t-2} + \cdots + \Psi_p u_{t-p} + e_t$. Let $\hat{u}_t = [\hat{u}_{1t}, \hat{u}'_{2t}]'$, where $\hat{u}_{1t} = u_{1t} - (\hat{\theta} - \theta)' z_t$ is the regression residual from (1) with $\hat{\theta}$ being the OLS estimator of θ . We then estimate

$$\hat{u}_t = \Psi_1 \hat{u}_{t-1} + \Psi_2 \hat{u}_{t-2} + \cdots + \Psi_p \hat{u}_{t-p} + \hat{e}_t \quad (6)$$

and test the hypothesis that $H_0 : \Psi_{1,21} = \Psi_{2,21} = \cdots = \Psi_{p,21} = 0$ where $\Psi_{j,21}$ is the $(2, 1)$ block of Ψ_j and $\hat{e}_t = [\hat{e}_{1t}, \hat{e}'_{2t}]'$ with $\hat{e}_{1t} = e_{1t} - (I_{n+1} - \Psi_1 L - \cdots - \Psi_p L^p)(\hat{\theta} - \theta)' z_t$ and L being the lag operator. Although \hat{u}_{1t} includes an estimation error, its effect is asymptotically negligible. In fact, we can show that for $j = -p, \dots, p$,

$$\frac{1}{T} \sum_{1 \leq t-j, t \leq T} \hat{u}_{1t} \hat{u}_{1t-j} = \frac{1}{T} \sum_{1 \leq t-j, t \leq T} u_{1t} u_{1t-j} + O_p\left(\frac{1}{T}\right),$$

$$\frac{1}{T} \sum_{1 \leq t-j, t \leq T} \hat{u}_{1t} u'_{2t-j} = \frac{1}{T} \sum_{1 \leq t-j, t \leq T} u_{1t} u'_{2t-j} + O_p\left(\frac{1}{T}\right)$$

while for $j \geq 0$,

$$\frac{1}{\sqrt{T}} \sum_{1 \leq t-j, t \leq T} \hat{u}_{t-j} \hat{e}'_t = \frac{1}{\sqrt{T}} \sum_{1 \leq t-j, t \leq T} u_{t-j} e'_t + O_p\left(\frac{1}{\sqrt{T}}\right).$$

by the asymptotic technique explained in, for example, Chapters 17–19 in Hamilton (1994). If the evidence of Granger non-causality is observed by tests based on (6), we can expect the finite sample efficiency gain by excluding the leads of Δx_t from (3) and estimating (5).

We may also consider verifying condition (4) by investigating whether or not the regression error from (5) is serially uncorrelated. For this purpose, the portmanteau tests are available as explained in Lütkepohl (1993).

To demonstrate the case where the model can be expressed as (5), we consider the following case:

$$u_{1t} = \sum_{j=0}^{\infty} \phi_j \varepsilon_{1t-j} \quad \text{and} \quad u_{2t} = \varepsilon_{2t} \quad \text{where} \quad \sum_{j=0}^{\infty} j |\phi_j| < \infty$$

$$\text{and } \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \sim iid \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{pmatrix} \right]. \quad (7)$$

We then decompose ε_{1t} as

$$\varepsilon_{1t} = \varepsilon_{1\cdot 2t} + \tilde{\varepsilon}_{2t} \quad (8)$$

where $\varepsilon_{1\cdot 2t} = \varepsilon_{1t} - \sigma_{12}\Sigma_{22}^{-1}\varepsilon_{2t}$ and $\tilde{\varepsilon}_{2t} = \sigma_{12}\Sigma_{22}^{-1}\varepsilon_{2t}$. Note that $\varepsilon_{1\cdot 2t}$ is uncorrelated with all the leads and lags of ε_{2t} and $\tilde{\varepsilon}_{2t}$. Using this decomposition, u_{1t} can be expressed as

$$\begin{aligned} u_{1t} &= \sum_{j=0}^{\infty} \phi_j \varepsilon_{1\cdot 2t-j} + \sum_{j=0}^K \phi_j \tilde{\varepsilon}_{2t-j} + \sum_{j=K+1}^{\infty} \phi_j \tilde{\varepsilon}_{2t-j} \\ &= v_t + \sum_{j=0}^K \Pi'_j \Delta x_{t-j} + \sum_{j=K+1}^{\infty} \phi_j \tilde{\varepsilon}_{2t-j} \end{aligned} \quad (9)$$

where $\Pi'_j = \phi_j \sigma_{12} \Sigma_{22}^{-1}$ and $v_t = \sum_{j=0}^{\infty} \phi_j \varepsilon_{1\cdot 2t-j}$. Since $E(\varepsilon_{2s} \varepsilon_{1\cdot 2t}) = 0$ for all s and t , it is evident that v_t is uncorrelated with Δx_{t-j} for all j . The regression form in (5) is obtained by inserting (9) into (1).

3 Monte Carlo experiments

To investigate the finite sample performance of the dynamic OLS estimator without leads, we conduct a Monte Carlo experiment. We consider the following data generating process with one dimensional I(1) regressor:

$$y_t = 1 + x_t + u_{1t}, \quad \Delta x_t = u_{2t} \quad \text{with} \quad x_0 = 0,$$

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} u_{10} \\ u_{20} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \sim iidN \left(0, \begin{bmatrix} 1 & \sigma_{12} \\ \sigma_{12} & 1 \end{bmatrix} \right).$$

We set $T = 100$, $\sigma_{12} = \sigma_{21} = 0.4, 0.8$, and a_{11} and a_{22} are $0.2, 0.5, 0.8$. The computation was conducted by using the GAUSS matrix language, and the number of replications is 10,000 for all the cases with the first 50 observations discarded.

Note that we do not have to include leads when $a_{22} = 0$ whereas we do need leads when $a_{22} \neq 0$. The simulation results are summarized in Tables 1 and 2 (further simulation results are available from the author upon request).

For the choice of K , we use the general-to-specific method by Ng and Perron (1995) with 1% and 5% significant levels and information criteria, i.e., the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).

To test for Granger non-causality, we use the residuals obtained from the DOLS estimators with four types of K . With regard to the choice of p , we use the larger p selected by AIC and BIC, and conduct Wald test with 5% significant level. If the null of Granger non-causality H_0 is rejected we use leads and lags, and if H_0 is not rejected, we use only lags in the estimation.

Table 1 reports the case where we do not have to include leads in the regression model. From the table we observe that the dynamic OLS estimator after Granger non-causality test substantially outperforms that with leads and lags in all the cases. In particular, in terms of the MSE, the MSE of the DOLS estimator without leads are approximately half of that with leads and lags in many cases. We also observe that the bias of the estimator is much reduced when we use the information criteria for the selection of K .

Table 2 shows the results when $a_{22} \neq 0$; in this case we should include both leads and lags. We observe that the bias becomes larger in many cases when we do not include leads in the model. However, we also observe from Table 2 that the exclusion of leads results in the smaller MSE in some cases, even though both leads and lags should be included based on the asymptotic theory. The possible reason is that the degrees of freedom with both leads and lags are smaller than those with only lags and hence the finite sample performance with leads and lags becomes poor.

4 Conclusion

In this paper, we considered the role of leads of the first difference of the I(1) regressors in the dynamic OLS estimation. Based on a Monte Carlo simulation, we found that the dynamic OLS estimator without leads after Granger non-causality test substantially outperforms that with leads and lags when leads are, in fact, un-

necessary. We also found that even if both leads and lags are required in the model, the pretest of Granger non-causality results in a better finite sample performance in view of the MSE in some cases.

References

- [1] Brillinger, D. R. (1981). *Time Series Data Analysis and Theory*, Holden-Day, San Francisco.
- [2] Engle, R. F. and C. W. J. Granger (1987). “Co-integration and Error Correction: Representation, Estimation, and Testing,” *Econometrica*, 55, 251-276.
- [3] Hamilton, J. D. (1994). *Time Series Analysis*, Princeton, New Jersey.
- [4] Lütkepohl, H. (1993). *Introduction to Multiple Time Series Analysis*, 2nd. Springer-Verlag, New York.
- [5] Ng, S. and P. Perron (1995). “Unit Root Tests in ARMA Models with Data-dependent Methods for the Selection of the Truncation Lag,” *Journal of the American Statistical Society*, 90, 268-281.
- [6] Park, J. Y. (1992). “Canonical Cointegrating Regressions,” *Econometrica*, 60, 119-143.
- [7] 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.
- [8] Phillips, P. C. B. and M. Loretan (1991). “Estimating Long-run Economic Equilibria,” *Review of Economic Studies*, 58, 407-436.
- [9] Saikkonen, P. (1991). “Asymptotically Efficient Estimation of Cointegration Regressions,” *Econometric Theory*, 7, 1-21.
- [10] Sims, C. A. (1972). “Money, Income, and Causality,” *The American Economic Review*, 62, 540-552.
- [11] Stock, J. H. and M. W. Watson (1993). “A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems,” *Econometrica*, 61, 783-820.

Table 1: Simulation Results (Non-Causal Case)

a_{11}	a_{22}	σ_{12}	GS001		GS005		AIC		BIC		
			L&L	Lags	L&L	Lags	L&L	Lags	L&L	Lags	
0.2	0	0.4	BIAS	0.00294	0.00292	0.00096	0.00080	0.00245	0.00119	0.00430	0.00332
			STDEV	0.06261	0.04816	0.08115	0.05743	0.05452	0.04872	0.03911	0.03933
			MSE	0.00393	0.00233	0.00659	0.00330	0.00298	0.00238	0.00155	0.00156
0.5	0	0.4	BIAS	0.01008	0.01025	0.00266	0.00348	0.00648	0.00363	0.01584	0.01144
			STDEV	0.09891	0.07600	0.12518	0.08889	0.09918	0.08160	0.06377	0.06476
			MSE	0.00989	0.00588	0.01568	0.00791	0.00988	0.00667	0.00432	0.00433
0.8	0	0.4	BIAS	0.03775	0.04050	0.02043	0.02215	0.02525	0.01976	0.05711	0.04346
			STDEV	0.23113	0.17605	0.27694	0.20314	0.26759	0.20275	0.16099	0.16125
			MSE	0.05485	0.03263	0.07711	0.04175	0.07224	0.04150	0.02918	0.02789
0.2	0	0.8	BIAS	0.00374	0.00369	0.00126	0.00144	0.00209	0.00118	0.00611	0.00367
			STDEV	0.04145	0.03194	0.05211	0.03731	0.03808	0.03190	0.02730	0.02707
			MSE	0.00173	0.00103	0.00272	0.00139	0.00146	0.00102	0.00078	0.00075
0.5	0	0.8	BIAS	0.00891	0.00937	0.00353	0.00427	0.00578	0.00375	0.01456	0.00950
			STDEV	0.06694	0.05025	0.08134	0.05617	0.06914	0.05384	0.04746	0.04597
			MSE	0.00456	0.00261	0.00663	0.00317	0.00481	0.00291	0.00246	0.00220
0.8	0	0.8	BIAS	0.04592	0.04608	0.02964	0.02843	0.03128	0.02498	0.06808	0.04711
			STDEV	0.17302	0.13553	0.19058	0.13988	0.19115	0.14091	0.15098	0.13040
			MSE	0.03205	0.02049	0.03720	0.02038	0.03752	0.02048	0.02743	0.01922

Note: “GS001”, “GS005”, “AIC”, “BIC” denote the dynamic OLS estimator with K chosen by the general to specific approach with 1% and 5% significant levels, AIC, and BIC, respectively. “L&L” denotes the dynamic OLS estimator with the leads and lags, and “Lags” denotes the dynamic OLS estimator without leads after testing for Granger non-causality.

Table 2: Simulation Results (Causal Case)

a_{11}	a_{22}	σ_{12}	GS001		GS005		AIC		BIC		
			L&L	Lags	L&L	Lags	L&L	Lags	L&L	Lags	
0.2	0.2	0.4	BIAS	0.00028	-0.00035	-0.00045	-0.00032	0.00031	0.00000	-0.00030	-0.00032
			STDEV	0.05033	0.03857	0.06603	0.04694	0.04309	0.03821	0.03121	0.03137
			MSE	0.00253	0.00149	0.00436	0.00220	0.00186	0.00146	0.00097	0.00098
0.5	0.5	0.4	BIAS	-0.00029	-0.00041	-0.00023	-0.00009	-0.00004	-0.00009	-0.00033	-0.00023
			STDEV	0.05197	0.03909	0.06581	0.04722	0.04983	0.04165	0.03176	0.03208
			MSE	0.00270	0.00153	0.00433	0.00223	0.00248	0.00174	0.00101	0.00103
0.8	0.8	0.4	BIAS	0.00010	-0.00024	-0.00013	0.00004	-0.00041	-0.00026	-0.00011	-0.00015
			STDEV	0.05517	0.04239	0.06700	0.04911	0.06359	0.04884	0.03480	0.03503
			MSE	0.00304	0.00180	0.00449	0.00241	0.00404	0.00239	0.00121	0.00123
0.2	0.2	0.8	BIAS	0.00019	0.00022	0.00008	0.00027	0.00031	0.00025	0.00007	0.00008
			STDEV	0.03269	0.02496	0.04173	0.02950	0.02932	0.02501	0.02004	0.02031
			MSE	0.00107	0.00062	0.00174	0.00087	0.00086	0.00063	0.00040	0.00041
0.5	0.5	0.8	BIAS	0.00011	0.00011	-0.00009	0.00008	0.00011	0.00031	0.00010	0.00011
			STDEV	0.03359	0.02514	0.04266	0.02988	0.03392	0.02693	0.02031	0.02068
			MSE	0.00113	0.00063	0.00182	0.00089	0.00115	0.00073	0.00041	0.00043
0.8	0.8	0.8	BIAS	-0.00027	-0.00003	-0.00013	0.00008	-0.00003	0.00020	0.00024	0.00019
			STDEV	0.03714	0.02569	0.04441	0.03016	0.04230	0.03063	0.02334	0.02250
			MSE	0.00138	0.00066	0.00197	0.00091	0.00179	0.00094	0.00055	0.00051

Note: “GS001”, “GS005”, “AIC”, “BIC” denote the dynamic OLS estimator with K chosen by the general to specific approach with 1% and 5% significant levels, AIC, and BIC, respectively. “L&L” denotes the dynamic OLS estimator with the leads and lags, and “Lags” denotes the dynamic OLS estimator without leads after testing for Granger non-causality.