Long-Run Exclusion and the Determination of Cointegrating Rank: Monte Carlo Evidence

Takamitsu Kurita
Faculty of Economics
Fukuoka University

WP-2010-002
Long-Run Exclusion and the Determination of Cointegrating Rank: Monte Carlo Evidence

Takamitsu Kurita*
Faculty of Economics, Fukuoka University

January 12, 2010

Abstract

This note investigates long-run exclusion in a cointegrated vector autoregressive (VAR) model from the viewpoint of finite-sample statistical inference. Monte Carlo experiments show that, in various circumstances, a mis-specified partial VAR model, which is justified by the existence of a long-run excluded variable, can lead to better finite-sample inference for cointegrating rank than a fully-specified VAR model. Implications of long-run exclusion for econometric modelling are then considered based on the Monte Carlo study.

Key Words: Long-Run Exclusion, Cointegrating Rank, Cointegrated Vector Autoregressive Model, Monte Carlo Experiment.

JEL Classification Codes: C32, C52, C63.

1 Introduction

The objective of this note is to investigate long-run exclusion in a cointegrated vector autoregressive (VAR) model from the viewpoint of finite-sample statistical inference. Various Monte Carlo experiments are conducted for this purpose. The introductory section defines the concept of long-run exclusion using a cointegrated vector autoregressive (VAR) model and then describes what needs to be investigated in the Monte Carlo study.

Let us consider a $p$-dimensional cointegrated VAR($k$) model for $X_t$:

$$
\Delta X_t = \alpha \beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \varepsilon_t, \quad \text{for } t = 1, ..., T,
$$

(1)

where a sequence of innovations $\varepsilon_t$ has independent and identical normal $N(0, \Omega)$ distributions conditional on $X_{-k+1}, ..., X_0$, and the parameters are defined such that $\alpha, \beta \in \mathbb{R}^{p \times r}$ for $r < p$ and $\Gamma_i \in \mathbb{R}^{p \times p}$. The cointegrated VAR model is developed by Johansen (1988, 1996) for the purpose of modelling non-stationary time series data. The parameters $\alpha$ are called adjustment vectors, while $\beta$ are referred to as cointegrating vectors. The relationships $\beta' X_{t-1}$ are called cointegrating relationships, representing $r$-dimensional stationary linear combinations of non-stationary variables, and $r$ is referred to as cointegrating rank.

*Correspondence to: Faculty of Economics, Fukuoka University, Bunkei Center Building, 8-19-1 Nanakuma, Jonan-ku, Fukuoka, 814-0180, Japan. E-mail: tkurita@fukuoka-u.ac.jp

Let the process be decomposed as $X_t = (Y_t', Z_t')'$ for $Y_t \in \mathbb{R}^m$, $Z_t \in \mathbb{R}^{p-m}$ and $r \leq m$. The set of parameters and the error terms in (1) are also expressed as

$$
\alpha = \left( \begin{array}{c} \alpha_y \\ \alpha_z \end{array} \right), \quad \beta = \left( \begin{array}{c} \beta_y \\ \beta_z \end{array} \right), \quad \Gamma_i = \left( \begin{array}{cc} \Gamma_{yy,i} & \Gamma_{yz,i} \\ \Gamma_{zy,i} & \Gamma_{zz,i} \end{array} \right), \quad \varepsilon_t = \left( \begin{array}{c} \varepsilon_{yt} \\ \varepsilon_{zt} \end{array} \right), \quad \Omega = \left( \begin{array}{cc} \Omega_{yy} & \Omega_{yz} \\ \Omega_{zy} & \Omega_{zz} \end{array} \right).
$$

The long-run exclusion of $Z_{t-1}$ from $X_{t-1}$ is then defined as

$$
\beta' X_{t-1} = \left( \beta'_y, \beta'_z \right) \left( \begin{array}{c} Y_{t-1} \\ Z_{t-1} \end{array} \right) = \beta'_y Y_{t-1},
$$

that is, $\beta_z = 0$. The cointegrating relationships consist of $Y_{t-1}$ solely rather than any linear combinations of $Y_{t-1}$ and $Z_{t-1}$. It is natural to refer to this phenomenon as *long-run exclusion*, for the cointegrating relations are deemed to correspond to long-run economic relationships. See Juselius (2006, Ch.10) for testing long-run exclusion in the cointegrating VAR model.

Economic theory or insight may tell us in advance that some variables will be long-run excluded but rather may be relevant in terms of short-run influences and innovation correlations. In the context of the model formulated above, there are such economic variables as can be treated as $Z_t$ with $\beta_z = 0$ but probably $\Gamma_{yz,i} \neq 0$ for $i = 1, \ldots, k - 1$ and $\Omega_{yz} \neq 0$. A question of interest is then how informative such pre-knowledge of long-run exclusion is for statistical inference on the determination of the cointegrating rank.

More specifically, let us construct a partial model for $\Delta Y_t$ in such a way as

$$
\Delta Y_t = \alpha_y \beta'_y Y_{t-1} + \sum_{i=1}^{k-1} \Gamma_{yy,i} \Delta Y_{t-i} + u_{yt}, \quad (2)
$$

where $u_{yt}$ represents a stationary error process comprising omitted short-run dynamics and innovations. Model (2) is, obviously, mis-specified. However, in the presence of pre-knowledge of long-run exclusion, it is probably tempting for economists to estimate and analyse model (2) rather than model (1). This is because the full cointegrating relations $\beta'_y Y_{t-1}$ are embedded in (2), although the model is mis-specified, and (2) is a more tractable model than (1) as a result of the reduced dimension of the VAR system.

It is therefore important, from a practical viewpoint, to know how often we can get the right answer about the cointegrating rank $r$ by estimating the mis-specified model (2) only, as compared with estimating the full model (1). See Granger and Haldrup (1998) for similar issues in the context of separate cointegration, and see also Abadir, Hadri, and Tzavails (1996) and Juselius (2006, Ch.19) for problems associated with an increase in the dimension of the VAR model.

Econometric modelling in practice is often based on the analysis of a limited number of data. Thus, the testing issue above is of interest in finite-sample cases rather than in such cases as asymptotic theory can hold. It is known that the finite-sample properties of a cointegrating rank test is rather different from its limiting properties; see Cheung and Lai (1993) and Johansen (2002), *inter alia*. This fact motivates a simulation-based comparative study of the rank tests using both (1) and (2).

The organisation of this note is as follows. Section 2 reviews statistical inference for the cointegrating rank and considers how inference may be influenced by the structure of the
underlying stochastic trends. Section 3 then conducts Monte Carlo experiments; recursive rejection frequencies of the test statistic are calculated in order to examine differences in finite-sample performance of various versions of the rank test statistic. Section 4 discusses implications of the experiments’ results for econometric modelling. Finally, Section 5 provides the overall summary and conclusion. All the numerical analyses and graphics in this note use Ox (Doornik, 2007) and OxMetrics/PcGive (Doornik and Hendry, 2007).

2 Inference for the Cointegrating Rank

As demonstrated by Johansen (1996, Ch.6), the maximum likelihood analysis of cointegrating rank in the model (1) is based on reduced rank regression. The likelihood function for the full model (1) is maximized by regressing $X_t$ and $X_{t-1}$ on $X_{t-i}$ for $i = 1, \cdots, k-1$ to obtain residuals $R_{0t}$ and $R_{1t}$, respectively. The sample product moment matrix for the residuals is then defined as

$$
\left( \begin{array}{cc}
S_{00} & S_{01} \\
S_{10} & S_{11}
\end{array} \right) = \frac{1}{T} \sum_{t=1}^{T} \left( \begin{array}{c}
R_{0t} \\
R_{1t}
\end{array} \right) \left( \begin{array}{c}
R_{0t} \\
R_{1t}
\end{array} \right)^t.
$$

Based on the full model (1), the log-likelihood ratio (log LR) test statistic for the null hypothesis of at most $r$ cointegrating relations, $H(r)$, against $H(p)$ is given by

$$
\log LR \{ H(r) | H(p) \} = -T \sum_{i=r+1}^{p} \log(1 - \hat{\lambda}_i),
$$

where $1 \geq \hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_p \geq 0$ are solutions to the following generalised eigenvalue problem:

$$
det \left( \lambda S_{11} - S_{10} S_{00}^{-1} S_{01} \right) = 0.
$$

It is known that, as $T \to \infty$, (3) has the following asymptotic distribution:

$$
\log LR \{ H(r) | H(p) \} \overset{w}{\to} \text{tr} \left\{ \int_0^1 dB'_u \left( \int_0^1 B_u B'_u du \right)^{-1} \int_0^1 B_u dB'_u \right\},
$$

where $\overset{w}{\to}$ signifies weak convergence and $B_u$ is a $(p - r)$ dimensional standard Brownian motion for $u \in [0, 1]$. As demonstrated by Johansen (2002), inter alia, the finite-sample distribution of (3) is rather different from its asymptotic distribution given in (4). It is also known that the dimension of the VAR system also has influences on small-sample inference for cointegration. Thus, it is anticipated that, under various circumstances where $\beta_z = 0$ holds and the number of observations is small, a pseudo log LR test statistic for cointegrating rank constructed from the partial model (2) will be more likely to reach the right answer about $r$ than the exact log LR test given by (3).

In order to form a conjecture about the performance of the rank test based on the mis-specified model when $\beta_z = 0$, it is necessary to inspect $u_{y,t}$ in (2), in which $k = 2$ is chosen here for the sake of simplicity. Noting that $u_{y,t}$ is expressed as $u_{y,t} = \Gamma_{yz,t} \Delta Z_{t-1} + \varepsilon_{y,t}$, in which a marginal model for $\Delta Z_{t-1}$ is substituted, one finds

$$
\Delta Y_t = \alpha_y \beta'_y Y_{t-1} + \Gamma_{yy,t} \Delta Y_{t-1} + \Gamma_{yz,t} \Delta Z_{t-2} + \Delta Z_{t-2} + \varepsilon_{z,t-1} + \varepsilon_{y,t}.
$$
Thus, \( u_{y,t} \) is influenced by the cointegrating relations by way of \( \Delta Z_{t-1} \). Furthermore, this is re-expressed as

\[
\Delta Y_t = (\alpha_y + \Gamma_{yz,1}\alpha_z)\beta'_y Y_{t-1} + \Gamma_{yy,1}\Delta Y_{t-1} - \Gamma_{yz,1}\alpha_z\beta'_y \Delta Y_{t-1} + \Gamma_{yz,1}\left(\Gamma_{yz,2}\Delta Y_{t-2} + \Gamma_{zz,1}\Delta Z_{t-2} + \varepsilon_{z,t-1}\right) + \varepsilon_{y,t}.
\]

The expression above indicates that \( \Gamma_{yz,1}\alpha_z \) can have a significant effect on the performance of the rank test based on the mis-specified model. If \( \Gamma_{yz,1}\alpha_z \) acts as a factor weakening the adjustment mechanism \( \alpha_y \), thereby reducing impacts from \( \beta'_y Y_{t-1} \), the use of (2) will then lead to misleading inference on the cointegrating rank. We may therefore conjecture that, if \( (\alpha_y + \Gamma_{yz,1}\alpha_z) \) is close to a zero matrix or vector, the rank test based on the partial model will perform poorly in comparison with that based on the full model given by (1); otherwise the opposite will happen so we may rely on inference based on the partial model. The next section, using Monte Carlo experiments, investigates if this conjecture holds true or not.

## 3 Monte Carlo Experiments

This section conducts various Monte Carlo experiments to see the implications of long-run exclusion for finite-sample inference. The data generation process (DGP) for the experiments is given as follows:

\[
\begin{pmatrix}
\Delta y_{1,t} \\
\Delta y_{2,t} \\
\Delta z_t
\end{pmatrix} = \begin{pmatrix}
-0.1 & 1 & y_{1,t-1} \\
0.1 & -1 & y_{2,t-1} \\
-0.1 & 1 & \Delta y_{1,t-1}
\end{pmatrix}
\begin{pmatrix}
y_{1,t} \\
y_{2,t}
\end{pmatrix}
+ \begin{pmatrix}
0.2 & 0 & \zeta \\
0 & 0.1 & \eta \\
0 & 0 & 0.3
\end{pmatrix}
\begin{pmatrix}
\Delta y_{1,t-1} \\
\Delta y_{2,t-1} \\
\Delta z_{t-1}
\end{pmatrix}
+ \begin{pmatrix}
\varepsilon_{y1,t} \\
\varepsilon_{y2,t} \\
\varepsilon_{z,t}
\end{pmatrix},
\]

where

\[
\begin{pmatrix}
\varepsilon_{y1,t} \\
\varepsilon_{y2,t} \\
\varepsilon_{z,t}
\end{pmatrix} \sim \mathcal{N}
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
1 & 0.5 & 0.5 \\
0.5 & 1 & 0.5 \\
0.5 & 0.5 & 1
\end{pmatrix}.
\]

Note that \( z_t \) is long-run excluded from the cointegrating space in the DGP above, thus one can consider \( Y_t = (y_{1,t}, y_{2,t})' \) and \( Z_t = z_t \) in the context of the model discussed in Section 1. Thus, the DGP above indicates: \( \alpha_y = (-0.1, 0.1)' \), \( \Gamma_{yz,1} = (\zeta, \eta)' \) and \( \alpha_z = -0.1 \). A full model corresponding to (1) is formulated for \( (y_{1,t}, y_{2,t}, z_t)' \) with lag length 2, while a partial or mis-specified model corresponding to (2) is for \( (y_{1,t}, y_{2,t})' \) with the same lag length. Two sorts of rank test statistics are constructed accordingly, as discussed in the previous section. The experiments are categorised according to the following eight cases:

- **Case 1**: \( \zeta = 0.0 \) and \( \eta = 0.0 \) \( \Rightarrow \) \( \alpha_y + \Gamma_{yz,1}\alpha_z = (-0.10, 0.10)' \),
- **Case 2**: \( \zeta = 0.2 \) and \( \eta = 0.0 \) \( \Rightarrow \) \( \alpha_y + \Gamma_{yz,1}\alpha_z = (-0.12, 0.10)' \),
- **Case 3**: \( \zeta = 0.2 \) and \( \eta = 0.2 \) \( \Rightarrow \) \( \alpha_y + \Gamma_{yz,1}\alpha_z = (-0.12, 0.08)' \),
- **Case 4**: \( \zeta = 0.4 \) and \( \eta = 0.2 \) \( \Rightarrow \) \( \alpha_y + \Gamma_{yz,1}\alpha_z = (-0.14, 0.08)' \),
- **Case 5**: \( \zeta = 0.2 \) and \( \eta = -0.2 \) \( \Rightarrow \) \( \alpha_y + \Gamma_{yz,1}\alpha_z = (-0.12, 0.12)' \),
- **Case 6**: \( \zeta = 0.4 \) and \( \eta = -0.4 \) \( \Rightarrow \) \( \alpha_y + \Gamma_{yz,1}\alpha_z = (-0.14, 0.14)' \),
- **Case 7**: \( \zeta = -0.2 \) and \( \eta = 0.2 \) \( \Rightarrow \) \( \alpha_y + \Gamma_{yz,1}\alpha_z = (-0.08, 0.08)' \),
- **Case 8**: \( \zeta = -0.4 \) and \( \eta = 0.4 \) \( \Rightarrow \) \( \alpha_y + \Gamma_{yz,1}\alpha_z = (-0.06, 0.06)' \).
Case 1 corresponds to a benchmark case, where there is no impact from $\Delta z_{t-1}$ on $\Delta y_{1,t}$ and $\Delta y_{2,t}$. Effects of $\Delta z_{t-1}$ gradually increase from Case 2 to Case 4, thereby generating variations in $\alpha_y + \Gamma y_{z,1} \alpha_z$. In all of these four cases $\alpha_y + \Gamma y_{z,1} \alpha_z$ is far from a zero vector. Hence it is anticipated, from the argument in the previous section, that the performance of the rank test based on the partial model is better than that based on the full model. Turning to Cases 5 and 6, one finds that $\eta$ is negative, with the result that $\alpha_y + \Gamma y_{z,1} \alpha_z$ is amplified in absolute value in comparison with the previous four cases. It is therefore expected that the rank test based on the partial model performs much better than that based on the full model. In contrast, $\zeta$ holds a negative value in Cases 7 and 8, thereby reducing the volume of $\alpha_y + \Gamma y_{z,1} \alpha_z$. Thus, contrary to the previous six cases, the rank test based on the partial model will perform poorly in Cases 7 and 8.

The Monte Carlo study focuses on recursive rejection frequencies of the two rank test statistics: one is based on the partial model, while the other on the full model, as mentioned above. The nominal size or significance level is chosen for 5%, so that the calculated test statistics are compared with the corresponding 95% quantiles and it is recorded how often the test statistics reject the null hypotheses, $H_0 : r = 0$ and $H_0 : r \leq 1$. Since there is a single cointegrating relation or $r = 1$ in the DGP above, the rejection frequencies for $H_0 : r = 0$ should approach 100% as the sample size increases, while the rejection frequencies for $H_0 : r \leq 1$ should get closer to 5% with an increase in the sample size. That is, the former frequencies can be interpreted as measures of power in terms of $T$, while the latter frequencies correspond to measures of size distortions according to $T$. In the experiments, the sample size $T$ increases by 5 from 50 to 150, corresponding to finite-sample cases conceivable in applied macro analysis. The number of Monte Carlo replications is 10,000, and all the recursive rejection rates for the eight cases above are presented in Figures 1, 2, 3, and 4.

Figure 1: Recursive Rejection Frequencies: Cases 1 and 2
Figure 1 shows the results of the Monte Carlo experiments for Cases 1 and 2. Figure 1 (a) and (c) correspond to Cases 1 and 2 respectively, in which the null hypothesis is given by no cointegration or $H_0 : r = 0$, thus the rejection rates should rise up to the 100% level as the sample size increases. As expected, the rejection rates of the partial model approach the 100% level faster than that of the full model, according to Figure 1 (a) and (c). These results indicate that, in Cases 1 and 2, the rank tests based on the partial model tend to be more powerful than those based on the full model. Size properties of the rank tests can also be checked using Figure 1 (b) and (d), in which a set of thin dotted lines corresponds to a 95% confidence band for the 5% nominal level. In terms of size distortions, the test statistics based on the partial model perform a little better than those based on the full model, according to Figure 1 (b) and (d); the rejection rates in the case of the partial model come in the confidence band faster than those obtained from the full model. Thus, the experiment results for these two cases indicate the advantages of the partial model over the full model when the sample size is finite.

![Figure 1: Monte Carlo experiments for Cases 1 and 2](image)

Next, Figure 2 presents recursive rejection rates for Cases 3 and 4. In line with the previous figure, the null hypothesis for Figure 2 (a) and (c) is $H_0 : r = 0$, while the null hypothesis for Figure 2 (b) and (d) is $H_0 : r \leq 1$. Similar features are observed in Figure 2 as well, indicating that the use of the partial model can lead to better finite-sample inference in these cases.

Recursive rejection frequencies for Cases 5 and 6 are displayed in Figure 3. As discussed above, $\alpha_y + \Gamma_{yz,1}\alpha_z$ is amplified in these two cases, which may enhance the possibility that the partial model yields better finite-sample results than the full model. This conjecture turns out to be true, according to Figure 3. All the simulation results are in favour of the use of the partial model rather than the full model, when the number of observations is finite.

![Figure 2: Recursive Rejection Frequencies: Cases 3 and 4](image)
Figure 3: Recursive Rejection Frequencies: Cases 5 and 6

Figure 4: Recursive Rejection Frequencies: Cases 7 and 8
Finally, Figure 4 presents recursive rejection rates for Cases 7 and 8, in which the composite parameter $\alpha_y + \Gamma_{yz,1}\alpha_z$ is reduced, distinct from the previous cases investigated above. As expected from the fact that the composite parameter is close to a zero vector, the rank test based on the partial model performs very poorly in comparison with that based on the full model, in Figure 4 (a) and (c); the rejection rates of the former tests are far away from the 100% level even when $T = 150$. These results are in great contrast to those in the previous cases. With regard to the size properties of both tests, Figure 4 (b) and (d) show that the rank test based on the full model suffers from non-negligible size distortions, while that based on the partial model is a conservative test as its rejection rates converge to the nominal level from below rather slowly. Size control is usually seen as a fundamental requirement in conventional statistical inference; we may therefore conclude that the partial model leads to better size properties than the full model.

4 Implications for Modelling Non-Stationary Data

The Monte Carlo study above has provided much evidence in support of the view that the mis-specified partial model can be relied upon for the purpose of choosing the cointegrating rank; the partial model tends to perform better than the full model when $\alpha_y + \Gamma_{yz,1}\alpha_z$ is distinct from a zero vector, although mixed results are observed when $\alpha_y + \Gamma_{yz,1}\alpha_z$ is close to a zero vector. In practice, we often have no or little prior knowledge about the structure of the cointegrating vectors and short-run dynamics for the underlying data generating mechanism. Thus, some may feel that the advantages of the partial model over the full model, as demonstrated above, are difficult to justify in practical circumstances.

The determination of the cointegrating rank is, probably, one of the most difficult tasks in applied research using non-stationary data. Hence, we should make use of as much additional information as possible for this purpose, as discussed by Juselius (2006, Ch.8). The choice of the cointegrating rank relying on both the partial and full models, as demonstrated in the experiments above, should be regarded as a modelling strategy exploiting information available from the data. In fact, this modelling strategy is consistent with the modelling procedure of gradually increasing the information set, as suggested by Juselius (2006, Ch.19). In such a case as economic theory or insight suggest that some variables may be long-run excluded, it is desirable to analyse the data using both of the models and to perform a comparative study of the underlying cointegrating rank. If one obtains an identical result from the analysis of the two models, this can be seen as consolidated statistical evidence for the choice of the rank; otherwise one should proceed to detailed analysis of the full model, checking if the conjecture of long-run exclusion is correct and also investigating the structure of the underlying short-run dynamics. The joint analysis of the two models, thus, leads to more informative statistical investigation of the data than analysing either of the models solely.

5 Concluding Remarks

The objective of this note is investigating long-run exclusion in a cointegrated VAR model from the viewpoint of finite-sample statistical inference. Monte Carlo experiments has revealed that, in various circumstances, a mis-specified partial VAR model, which is justified by the existence of a long-run excluded variable, can lead to better finite-sample inference for cointegrating rank than a fully-specified VAR model. Implications of long-run exclu-
tion for econometric analysis are also considered based on the Monte Carlo study, leading to a feasible strategy for modelling non-stationary time series data.

References


