How useful is yet another data-driven bandwidth in long-run variance estimation?:
A simulation study on cointegrating regressions

Masayuki Hirukawa *

Faculty of Economics, Setsunan University, 17-8 Ikeda Nakamachi, Neyagawa, Osaka, 572-8508, Japan

ARTICLE INFO

Article history:
Received 4 November 2009
Received in revised form 12 January 2011
Accepted 2 February 2011
Available online 12 February 2011

JEL classification:
C13
C14
C22

Keywords:
Bandwidth
Cointegration
Kernel
Long-run variance
Simulation

ABSTRACT

This paper investigates how bandwidth choice rules in long-run variance estimation affect
finite-sample performance of efficient estimators for cointegrating regression models. Monte Carlo results indicate that
Hirukawa’s (2010) bandwidth choice rule contributes bias reduction in the estimators.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

In time series econometrics, estimating the long-run variance (“LRV”) matrix of a random vector process is essential for empirical research on estimation (e.g. generalized method of moments) and testing (e.g. standard error calculation and unit-root testing) problems. This paper focuses on a standard, kernel smoothing approach to LRV estimation, and investigates finite-sample performance of the bandwidth choice rule newly proposed by Hirukawa (2010) via Monte Carlo simulations, in comparison with two most popular ones by Andrews (1991) and Newey and West (1994). The simulation study is conducted in the context of cointegrating regressions. While there is rich literature in simulation studies on cointegrating regressions, this paper contributes the literature in the following two respects. First, while Andrews’ (1991) and Newey and West’s (1994) bandwidth choice rules (“A rule” and “NW rule”) are frequently used in simulation studies, very little is known about finite-sample performance of Hirukawa’s (2010) solve-the-equation plug-in bandwidth choice rule (“SP rule”). To the best of our knowledge, there are only two simulation results on the SP rule, both of which are available in Hirukawa (2010). Neither is on cointegrating regressions. Second, the SP rule is expected to contribute the bias correction method by Kurozumi and Hayakawa (2009) for the fully modified least squares (“FMLS”; Phillips and Hansen, 1990) and canonical cointegration regression (“CCR”; Park, 1992) estimators of cointegrating vectors. When the I(1) regressors are endogenous and/or the regression errors are serially correlated in cointegrating regressions, the ordinary least squares (“OLS”) estimators of cointegrating vectors suffer so-called the “second-order bias”. Although FMLS and CCR are proposed as devices for correcting the second-order bias nonparametrically by means of LRV estimators, these methods work poorly in the presence of strong serial dependence in regression errors. Then, Kurozumi and Hayakawa (2009) develop a further bias reduction method for FMLS and CCR when regression errors obey an AR(1) model with the AR coefficient moderately close to unity. However, their simulation results indicate that average lengths of bandwidths from the A and NW rules are too long and too short for the purpose of bias reduction. Apparently, there is a need for a bandwidth choice rule that tends to yield intermediate lengths. The SP rule is expected to be a remedy, as suggested in the next section.

2. Three Bandwidth Formulae in LRV Estimation

To illustrate the difference in three bandwidth choice rules, consider a problem of estimating the LRV of a zero-mean scalar process $h_t$, where the LRV is defined as $\omega = \sum_{j=-\infty}^{\infty} \gamma(j) = \sum_{j=-\infty}^{\infty} E(h_t h_{t-j})$. Given $T$ observations $\{h_t\}_{t=1}^{T}$, a kernel $k(\cdot)$ and a bandwidth $M$, the kernel estimator of $\omega$ is given by a weighted sum of sample autocovariances
\[ \hat{\omega} = \sum_{j=1}^{t-1} k(j/M) \hat{y}(j), \text{ where } \hat{y}(j) = T^{-1} \sum_{s=\min(T,j+1)}^{t-1} b_{h(j)} \hat{e}(s). \]

Each of the three rules is an estimator of the minimizer of the asymptotic mean squared error (“AMSE”) of \( \hat{\omega} \). The minimizer (= AMSE-optimal bandwidth) is given by

\[
M^* = \left( \frac{\hat{q}^2}{\sum_{j=1}^{t-1} k(j/M)^2} \right)^{-1/2} T^{-1/2} \sum_{j=1}^{t-1} k(j/M)^2, \\

b^*(M^*) = \left( \frac{\mathbf{M}^{(0)}(b^*(M^*))}{\sum_{j=1}^{t-1} k(j/M)^2} \right)^{1/2} T^{-1/2} \sum_{j=1}^{t-1} k(j/M)^2, \\
\text{ where } q(z) \text{ is the characteristic exponent of } k(z) \text{ that satisfies } k_j = \lim_{m \to 0} \left( 1 - \frac{k(m)}{k(0)} \right)^{1/m}. \]

Because of the endogeneity and the serially correlated regression error in Eq. (1), \( \hat{\theta}_{\text{OLS}} \) (= OLS estimator of \( \theta \)) is consistent but inefficient due to the second-order bias. FMLS and CCR are developed as efficient estimation methods that eliminate the second-order bias. Let \( \Omega = \Sigma + \Gamma + \Gamma' \) and \( \Lambda = \Sigma + \Gamma \). The normalized bandwidth (AMSE) is given by

\[
\hat{\theta}_{\text{FMLS}} = \left( \sum_{t=1}^{T} \mathbf{z}^2 \right)^{-1} \left( \sum_{t=1}^{T} \mathbf{z} \mathbf{y}_t - \mathbf{T} \mathbf{y}_t \right), \\
\hat{\theta}_{\text{CCR}} = \left( \sum_{t=1}^{T} \mathbf{z}^2 \mathbf{z}_t \right)^{-1} \left( \sum_{t=1}^{T} \mathbf{z}^2 \mathbf{y}_t \right), \]

where \( \mathbf{y}_t = \mathbf{y}_t - \hat{\Omega}_{11} \hat{\Omega}_{12} \hat{\Delta} \) and \( \hat{\Sigma} = \left( \hat{\Sigma}_{11} \left( \hat{\Omega}_{11} \right)^{-1} \right) \). CCR employs the transformed data

\[
x_t^* = x_t - \hat{\Sigma}^{-1} \hat{\Delta} \mathbf{z}_t \quad \text{and} \quad \hat{\Sigma} = \left( \hat{\Sigma}_{11} \left( \hat{\Omega}_{11} \right)^{-1} \right) \]

where \( \hat{\Sigma} \) is a consistent estimator of \( \Sigma \). Defining \( z_t^* = (1, x_t^* \gamma) \) yields the CCR estimator of \( \theta \) as

\[
\hat{\theta}_{\text{CCR}} = \left( \sum_{t=1}^{T} x_t^* x_t^* \right)^{-1} \left( \sum_{t=1}^{T} x_t^* y_t \right). \]

However, FMLS and CCR work poorly when the regression error \( u_t \) exhibits strong serial dependence. Kurozumi and Hayakawa (2009) consider further bias reduction in FMLS and CCR for the N local-to-unify system such that \( \rho \) is modeled as \( \rho = 1 - N/C, \) where \( N \) satisfies \( N \to \infty \) but \( N = O(T) \). In addition, if \( M/N \to d_M \) for a bandwidth of the LRV estimators \( M \) and if another bandwidth \( M_c \) satisfies \( M_c = o(N) \), \( c \) can be consistently estimated by

\[
\hat{c} = \left( N/2 \right) \left( \hat{\Omega}_{1111} / \hat{\Omega}_{111} \right), \]

where \( \hat{\Omega}_{1111} = T^{-1} \sum_{t=1}^{T} \mathbf{z}_t \mathbf{z}_t \) and \( \hat{\Omega}_{111} = \sum_{j=1}^{M_c} \left( \frac{1}{M_c} \right) k \left( j/M \right) \mathbf{z}_t^2 \mathbf{z}_t \).

The bias-corrected FMLS and CCR (“FMLS-BC” and “CCR-BC”) estimators of \( \theta \) can be obtained by replacing \( \left( \hat{\Omega}_{11}, \hat{\Delta} \right) \) in \( \hat{\theta}_{\text{FMLS}} \) and \( \hat{\theta}_{\text{CCR}} \) with \( \left( \hat{\Omega}_{1111}, \hat{\Delta} \right) = \left( \hat{\Omega}_{1111} / \hat{r} \hat{\Lambda}_{1111} / \hat{r} \right) \), where \( \hat{r} = \hat{d}_M \int_0^r k(r) \exp \left( -\hat{c} d_M r \right) dr \) to yield a bias-corrected FMLS-BC and CCR-BC. Kurozumi and Hayakawa (2009) suggest that extra tuning parameters are set equal to \( d_M = 1, \) \( N = M, \) and \( M_c = M_c^{2/3} \).

In this simulation study, \( \theta_1 \) is estimated by OLS, FMLS, CCR, FMLS-BC, and CCR-BC for each data set. Finite-sample performance of each estimator is evaluated by its bias and mean squared error (‘MSE’). The Parzen kernel is employed for all LRV estimators, and three bandwidths are computed via the A, NW, and SP rules. The lag length for the truncated estimator of the normalized curvature in the NW rule is set equal to \( \lfloor T/100 \rfloor^{4/23} \) for each combination of parameter values, 10,000 data sets of sample sizes \( T = 100 \) or \( T = 300 \) are simulated.

### 3. Monte Carlo Simulations

This paper adopts the experimental design of Kurozumi and Hayakawa (2009). Let the data \( \{ (y_t, x_{t1}) \}_{t=1}^{T} \) be generated by

\[
y_t = \theta_0 + \theta_1 x_t + u_{t2} \quad \text{with} \quad u_{t2} \sim \text{N}(0, \sigma^2), \]

where the error term \( u_t = (u_{t1}, u_{t2}) \) obeys

\[
u_{t1} = \rho u_{t-1} + \epsilon_{t1}, \quad u_{t2} = \epsilon_{t2}. \]

True parameter values of \( \theta = (\theta_0, \theta_1)' \), \( \rho \), and \( \sigma^2 \), are \( \theta_0 = \theta_1 = 1, \rho \in [0.7, 0.8, 0.85, 0.9, 0.95], \) and \( \sigma^2 \in [0.4, 0.8] \). For each combination of parameter values, 10,000 data sets of sample sizes \( T = 100 \) or \( T = 300 \) are simulated.

### 4. Conclusion

Table 1 reports biases and MSEs of the five estimators of \( \theta_1 \). Because the results for \( \sigma^2 = 0.8 \) are qualitatively similar, only those
Results on MSEs are mixed. While four estimators appear to be negligible.

In the presence of moderate to high persistence in regression errors, bandwidths are relatively small, and thus their influences over the estimation results are often most biased.

The results are summarized as follows:

- FMLS and CCR are less biased than OLS, and FMLS-BC and CCR-BC are less biased than FMLS and CCR. However, the advantage in bias of FMLS-BC and CCR-BC over FMLS and CCR appears to be diminishing as \( \rho \) increases. In addition, FMLS and FMLS-BC tend to be less biased than CCR and CCR-BC.

- FMLS-BC and CCR-BC, as \( \rho \) gets closer to unity, the SP rule is more likely to give the smallest bias among three bandwidth choice rules.

For each of FMLS and CCR, when \( \rho \) is close to unity, the SP rule tends to yield the smallest bias among three bandwidth choice rules. Nonetheless, estimators based on the A rule are often most biased.

### References


