市場經濟研究 제41집 3호 (2012년) pp. 1～14

IV Coefficient Tests for a Unit Root

Kyung So Im*, Junsoo Lee**

We examine the property of the LM coefficient tests using instrumental variables (IV) estimation. As in Im and Lee (2011), we utilize stationary instrumental variables but consider coefficient tests rather than t-tests. Under the null hypothesis, the proposed coefficient statisics converge to the standard normal distribution. This result follws since the score vector in the derivation of the LM statistic from the likelihood function converges to a constant. This makes a contrast from the DF version tests whose distributions are nonstandard.

Key Words: Unit Root Tests, Stationary Instrumental Variables, Standard Normal Distributions

JEL Classification: C12, C15, C22

* Kyung So Im, Senior Economist, Division of Insurance and Research, Federal Deposit Insurance Corporation (FDIC), 550 17th Street, NW, Washington, DC, USA (e-mail: kim@fdic.gov);

** Junsoo Lee (Corresponding author), Professor, Department of Economics, Finance and Legal Studies, University of Alabama, Box 870224, AL 35487 USA (email:jlee@cba.ua.edu, phone: 205-348-8978, fax: 205-348-0590)

Ⅰ. Introduction

The limiting distributions of usual unit root tests are typically nonstandard. The distributions depend on various functionals of Brownian motions reflecting different deterministic components of the time series as well as the detrending methods. Since the specific expressions of the nonstandard asymptotic distributions vary over different models, different sets of critical values must be provided for every test. However, it is possible to have unit root tests with a standard normal distribution if we deviate from the traditional approach. The work of So and Shin (1999) is enlightening in this regard. They use a sign function as an instrumental variable (IV) and their test statistics using IV estimation have the standard normal distribution.

Im and Lee (2011) suggest using stationary instrumental variables and their IV tests have this same feature. That is, they use $w_{t-1} = y_{t-1} - y_{t-1-m}$ as an instrument for y_{t-1} in the regression of $\triangle y_t$ on y_{t-1}. When m is a fixed finite number, $y_{t-1} - y_{t-1-m}$ is a stationary process and we can see that the sample moment $\sqrt{T}\sum_{t=1}^{T}(y_{t-1} - y_{t-1-m})\triangle y_t$converges to a normal distribution. Therefore, the asymptotic distributions of the corresponding unit root tests will be standard normal. This result continues to hold with different deterministic terms or detrending methods in the underlying model. Im and Lee (2011) use the LM detrending method since the LM tests are shown to be more powerful than the DF tests; see Vougas (2003).

These IV unit root tests are based on the t-statistic in the testing regression. That is, the standard normal result can be obtained with the usual t test, but the normality result does not hold in general for the coefficient tests which use the coefficient estimate $\hat{\beta}$ of y_{t-1}. In this paper we show that the coefficient tests based on the LM detrending method will also have the normal distribution. Although the distribution is not standard normal, the modified coefficient tests can have the standard normal distribution when using a proper normalization procedure. This result is due to a unique feature of the LM based tests using IV estimation. While the LM coefficient tests use a score vector involving a possibly nonstationary term y_{t-1}under the null, it converges to a constant. However, when the DF type detrending method is used, the corresponding coefficient tests will not have the normal distribution. Actually, no other existing unit root tests utilizing the coefficient tests will have this property.

Ⅱ. Stationary IV Tests Based on the LM detrending method

Suppose we have data y_t for $t = 0, 1, 2, ..., T$, generated as

$$y_t = d_t + x_t \tag{1}$$

Here, we let $d_t = z_t^{'}\gamma$be the deterministic component of y_t for which we can consider general models with various types of deterministic functions. x_t is the stochastic component of the series following an autoregressive process

$$x_t = \varnothing x_{t-1} + \varepsilon_{t,} \tag{2}$$

where ε_t is the innovation term and is assumed to have zero mean and satisfy the following assumption.

Assumption 1 $\{\varepsilon_t\}$ is a martingale difference process satisfying.

$E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, ...) = 0$, and $E(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, ...) = \sigma^2$, for $t = 1, 2, ...,$ with $O < \sigma^2 < \infty$.

We assume that the initial value is finite such that $y_o = O_p(1)$ Combining (1) and (2), we have

$$(1 - \varnothing L)y_t = (1 - \varnothing L)z_t^{'}\gamma + \varepsilon_t, \tag{3}$$

and the testing regression model is

$$\triangle y_t = \beta y_{t-1} + (1 - \varnothing) z_t^{'} \gamma + \varnothing \triangle z_t^{'} \gamma + \varepsilon_t, \tag{4}$$

where $\beta = \varnothing - 1$. Interest centers on testing the null hypothesis $\beta = 0$ against the alternative hypothesis $\beta < 0$. When $z_t = (1, t)'$, we can consider a model with a non zero mean and linear trend. We can also have models with structural breaks by introducing a dummy variable when a break occurs between $t = T_B$ and $T_B + 1$

$$D_t = \begin{cases} 0 \text{ if } t \leq T_B \\ 1 \text{ if } t > T_B \end{cases} \tag{5}$$

Then, we can let $z_t = (1, t, D_t)'$ for a model with a level shift, or $z_t = (1, t, D_t, tD_t)'$ for a model with a trend-shift. Note that the term $\triangle z_t$ drops out from the regression when z_t is a trend function t, but remains in z_t when z_t contains dummy variables. When a time series contains a non zero mean or other deterministic terms, including a linear trend, we need to control their effects. This is the detrending procedure. One popular method has been to adopt a regression in level to estimate these coefficients to detrend the series. This is the detrending procedure adopted in the DF type tests. An alternative detrending method is available. One may estimate the coefficients of deterministic terms from the regression using differenced data and then use these coefficients to detrend the data. This detrending method was suggested by Bhargava (1986) and Schmidt and Phillips (1992, SP hereafter). We call this the LM detrending method and call the resulting tests as LM tests, since the restriction is first imposed under the null. Vougas (2003) and Lee and Strazicich (2003) discuss the advantages of the LM based tests.

The LM test follows three steps. Step (1): We estimate the parameters of the deterministic term by imposing the restriction under the null hypothesis $(\beta = \varnothing - 1 = 0)$ using differenced data rather than estimating in levels. Thus, we consider the following regression and obtain $\hat{\gamma}$.

$$\triangle y_t = \triangle z_t' \gamma + \varepsilon_t, \; t = 1, 2, \ldots, T. \tag{6}$$

Step (2): We remove the deterministic trend based on the estimated parameter values from the differenced data in the first step.

$$\tilde{y}_t = y_t - z_t' \hat{\gamma}, \;\; t = 1, 2 \ldots, \; T. \tag{7}$$

Step (3): We conduct unit root tests using the detrended data obtained in step (2).

$$\triangle y_t = \beta \tilde{y}_{t-1} + \triangle z_t' \delta + error. \tag{8}$$

Note that one can possibly use $\triangle \tilde{y}_t$ as a dependent variable in the above regression. In that case, we can drop $\triangle z_t' \delta$ to have another LM test as done in Schmidt and Lee (1991). We follow the first two steps to remove the deterministic component of the data. However, in step (3), we use instrumental variable estimation rather than ordinary least squares estimation. To do so, we define the instrumental variable as

$$\tilde{\omega}_t = \tilde{y}_{t-1} - \tilde{y}_{t-m-1}. \tag{9}$$

Then, we consider IV estimation using $\tilde{w}_t$ for $\tilde{y}_{t-1}$ to estimate equation (8). For example, when we consider the model with a linear trend, we have

$$\hat{\beta}_{LM-IV} = \frac{\sum_{t=1}^{T} \tilde{\omega}_t \triangle \tilde{y}_{t-1} - T^{-1} \sum_{t=1}^{T} \tilde{\omega}_t \sum_{t=1}^{T} \triangle \tilde{y}_{t-1}}{\sum_{t=1}^{T} \tilde{\omega}_t \tilde{y}_{t-1} - T^{-1} \sum_{t=1}^{T} \tilde{\omega}_t \sum_{t=1}^{T} \tilde{y}_{t-1}} \tag{10}$$

and

$$t_{LM-IV} = \frac{\sum_{t=1}^{T} \tilde{\omega}_t \triangle \tilde{y}_{t-1} - T^{-1} \sum_{t=1}^{T} \tilde{\omega}_t \sum_{t=1}^{T} \triangle \tilde{y}_{t-1}}{\tilde{\sigma} \sqrt{\sum_{t=1}^{T} \tilde{\omega}_t^2 - T^{-1} \left(\sum_{t=1}^{T} \tilde{\omega}_t \right)^2}} \tag{11}$$

where $\tilde{\sigma}^2$ is an estimator of σ^2. It is obtained by

$$\tilde{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} \left(\triangle y_t - \hat{\beta}_{LM-IV} \tilde{y}_{t-1} - \triangle z_t' \tilde{\delta} \right)^2$$

For the models with structural breaks and others, the same expressions in the above can be used after controlling the effects of additional regressors.

Theorem 1: Under Assumption 1 and the null hypothesis,

$$\sqrt{T} \hat{\beta}_{LM-IV} \xrightarrow{d} \frac{2}{\sqrt{m}} W(1). \tag{12}$$

In the Appendix, we prove the above theorem for the cases where $z_t = (1, t)'$ and $z_t = (1, t, D_t, tD_t)'$. Im and Lee (2011) examine the t_{LM-IV} test. But, our focus in this paper is on the distribution of $\hat{\beta}_{LM-IV}$. To begin with, we can easily show from the expressions for the t-statistic in (11) that the distribution of the t-statistic does not depend on parameters in the deterministic terms. The second terms in each of the numerator and the denominator in (11) are asymptotically degenerate and the first term in the numerator of (11) is expressed as the stationary moment conditions. However, at the first sight, it does not seem that the coefficient tests will have the normal distribution. This is so, since unlike the t-statistics in (11), the denominator of the coefficient estimator $\hat{\beta}_{LM-IV}$ in (10) has the term $\tilde{y}_{t-1}$. That is, the first term in the denominator of (10) involves a moment condition with a non stationary term $\tilde{y}_{t-1}$ and one may think that this could lead to a nonstandard distribution. It is interesting to see how the coefficient test $\sqrt{T}\hat{\beta}_{LM-IV}$ has the normal distribution as well. We find that

$$T^{-1}\sum_{t=1}^{T}\tilde{\omega}_t\tilde{y}_{t-1}\xrightarrow{p}-0.5\sigma^2 \tag{13}$$

The above term is a score vector in the derivation of the LM statistic from the likelihood function. Intuitively speaking, since this term converges to a constant term and other non degenerate terms are stationary moments, the resulting distribution of $\hat{\beta}_{LM-IV}$ is normal. We show details of the proof in the Appendix for two general models. As such, one can consider a normalized test and utilize the normal approximation by standardizing the estimator. Owing to the standard normal result, it is obvious that the asymptotic distribution is free of any nuisance parameters. We note that the Dickey Fuller version coefficient tests do not have this feature since their score vector does not converge to a constant term. The DF version coefficient tests have nonstandard distributions. Note that the above standard normality result holds asymptotically for a finite value of $m \ll T$. However, the finite sample performance of the tests will depend on the choice of m. Usual model selection procedues such as minimizing the sum of squared residuals can be applied in finite samples, as done for t_{LM-IV} in Im and Lee (2011). Also, note that the IV tests can be more useful for the models with structural changes. A preliminary examination reveals that the LM-IV tests are less sensitive to the nuisance parameters regarding structural breaks than the IV tests of So and Shin (1999). However, the IV tests of So and Shin are designed to capture non-linearity; thus, they are more robust to various types of nonlinearity and asymmetric distributions.

Ⅲ. Concluding Remarks

The work of So and Shin (1999) and other extensions of IV unit root tests are based on t-statistics in the testing regression. They have the standard normal distributions. In this paper, we have considered the LM coefficient tests using stationary instrumental variables. Unlike other types of unit root tests, we show that the LM coefficient tests have the normal distribution. The underlying reason is that the score vector in the LM procedure converges to a constant term that simplifies the required asymptotics.

(접수일: 2012. 05. 06. / 수정일: 2012. 07. 17. / 게재확정일: 2012. 07. 17.)

References

1. Bhargava, A., (1986), “On the theory of testing for unit roots in observed time series,” *Review of Economic Studies* 52(3), 369-384.

2. Dickey, D.A, and W.A. Fuller (1979), “Distribution of the Estimators for Autoregressive Time Series with a Unit Root,” *Journal of the American Statistical Association* 74, 427-431.

3. Hamilton, J., (1994), *Time Series Analysis*, Princeton University Press.

4. Im, K., and J. Lee (2011), “Testing for Unit Roots Using Weighted Moment Conditions,” Working Paper, University of Alabama.

5. Lee, J., and M.C. Strazicich (2003), “Minimum LM Unit Root Test with Two Structural Breaks,” *Review Economics and Statistics* 85(4), 1082-1089.

6. So, B.S., and D.W. Shin (1999), “Cauchy Estimators for Autoregressive Processes with Applications to Unit Root Tests and Confidence Intervals,” *Econometric Theory* 15(2), 165-176.

7. Schmidt, P., and J. Lee (1991), “A Modification of the Schmidt-Phillips Unit Root Test,” *Economics Letters* 36(3), 285-289.

8. Schmidt, P., and P.C.B. Phillips (1992), “LM Tests for a Unit Root in the Presence of Deterministic Trends, “*Oxford Bulletin of Economics and Statistics* 54(3), 257-287.

9. Vougas, D.V., (2003), “Reconsidering LM Unit Root Testing,” *Journal of Applied Statistics* 30(1), 727-741.

10. White, H. (2001), *Asymptotic Theory for Econometricians*, Academic Press.

Appendix

Lemma 1: Let

$$\widetilde{A}_T = \sum_{t=1}^{T} \widetilde{\omega}_t \widetilde{y}_{t-1} - T^{-1} \sum_{t=1}^{T} \widetilde{\omega}_t \sum_{t=1}^{T} \widetilde{y}_{t-1}$$

$$\widetilde{B}_T = \sum_{t=1}^{T} \widetilde{\omega}_t \triangle y_t - T^{-1} \sum_{t=1}^{T} \widetilde{\omega}_t \sum_{t=1}^{T} \triangle \widetilde{y}_t \tag{A1}$$

where $\widetilde{y}_{t-1}$ and $\widetilde{\omega}_t$ are defined in (7) and (8), respectively. Under the null hypothesis and Assumption 1, we have

$$T^{-1} \widetilde{A}_T \xrightarrow{p} \frac{1}{2} m\sigma^2, \tag{A2}$$

and

$$T^{-1/2} \widetilde{B}_T \xrightarrow{d} \sqrt{m}\,\sigma\, W(1).$$

Proof:

Distribution of $T^{-1}\widetilde{A}_T$

The transformed data, which are obtained after the effect of the deterministic components is removed, do not depend on γ Therefore, we do not lose generality by assuming that $\gamma = 0$ We then have, under the null hypothesis

$$T^{-1}\widetilde{A}_T = T^{-1} \sum_{t=1}^{T} (\xi_t - \triangle^m z'_{t-1}\hat{\gamma})(S_{t-1} - z'_{t-1}\hat{\gamma}) - T^{-2} \sum_{t=1}^{T} (\xi_t - \triangle^m z'_{t-1}\hat{\gamma}) \sum_{t=1}^{T} (S_{t-1} - z'_{t-1}\hat{\gamma}) \tag{A3}$$

where $\triangle^m z_t = z_t - z_{t-m}$, $S_t = \sum_{j=1}^{t} \varepsilon_j$, and $\xi_t = \varepsilon_{t-1} + \ldots + \varepsilon_{t-m}$. We evaluate this term for each of two models.

(i) Model with a trend:

We have $z_t = t$, and $\triangle z_t = 1$. Then, we obtain $\hat{\gamma} = \frac{1}{T} \Sigma_{t=1}^{T} \triangle y_t = \frac{1}{T} \Sigma_{t=1}^{t} \varepsilon_t = \bar{\varepsilon}$.

$$\sqrt{T}\bar{\varepsilon} \xrightarrow{d} \sigma W(1) \tag{A4}$$

Note that $\triangle^{m} z_t = m$ Equation (A3) becomes

$$\begin{aligned} T^{-1}\widetilde{A}_T = {} & T^{-1}\sum_{t=1}^{T}(\xi_t - m\bar{\varepsilon})(S_{t-1} - t\bar{\varepsilon}) \\ & - T^{-2}\sum_{t=1}^{T}(\xi_t - m\bar{\varepsilon})\sum_{t=1}^{T}(S_{t-1} - t\bar{\varepsilon}) \end{aligned} \tag{A5}$$

For the first term of (A5), we have

$$\begin{aligned} & T^{-1}\sum_{t=1}^{T}\xi_t S_{t-1} \xrightarrow{d} \frac{1}{2}m\sigma^2\left[W(1)^2 + 1\right] \\ & T^{-1}\sum_{t=1}^{T}\xi_t t\bar{\varepsilon} \xrightarrow{d} m\sigma^2\left[W(1) - \int_0^1 W(r)dr\right]W(1) \\ & T^{-1}\sum_{t=1}^{T} m\bar{\varepsilon}S_{t-1} \xrightarrow{d} m\sigma^2\int_0^1 W(r)dr\, W(1) \\ & T^{-1}\sum_{t=1}^{T} m\bar{\varepsilon}^2 t \xrightarrow{d} \frac{1}{2}m\sigma^2 W(1)^2 \end{aligned} \tag{A6}$$

Combining the terms in (A6), we obtain

$$T^{-1}\sum_{t=1}^{T}(\xi_t - m\hat{\gamma})(S_{t-1} - t\hat{\gamma}) \xrightarrow{p} \frac{1}{2}m\sigma^2 \tag{A7}$$

The second term of (A5) is ignorable asymptotically since

$$T^{-1/2}\sum_{t=1}^{T}(\xi_t - m\bar{\varepsilon}) \xrightarrow{d} m\sigma W(1) - m\sigma W(1) = 0 \tag{A8}$$

(ii) Model with a trend shift:

We have $z_t = [t, D_t, tD_t]$ and $\triangle z_t = [1, \triangle D_t, \triangle(tD_t)]$ Noting that $\triangle D_t = 1$ at $t = T_B + 1$, and $\triangle D_t = 0$ for all other $t's$, numerically identical estimators of γ_1 and γ_3 can be obtained from the regression of $\triangle y_t$ on $\triangle z_t$ amounts to the regression of $\triangle y_t$ on $[1, \triangle(tD_t)]$ omitting $T_B + 1's$ row in the regression. We therefore have: $\hat{\gamma}_1 = \overline{\triangle y_1}$, $\hat{\gamma}_3 = \overline{\triangle y_2} - \overline{\triangle y_1}$, where $\overline{\triangle y_1} = \frac{1}{T_B}\Sigma_{t=1}^{T_B}\triangle y_t = (y_{T_B} - y_0)/T_B$ and $\overline{\triangle y_2} = \frac{1}{T-(T_B+1)}\Sigma_{t=T_B+2}^{T}\triangle y_t = (y_T - y_{T_B+1})/[T-(T_B+1)]$.
Using the fact that the (T_B+1)-th residual is zero when $\triangle D_t$ is included in the regression, we obtain $\hat{\gamma}_2 = \triangle y_{T_B+1} - \hat{\gamma}_1 - (T_B+1)\hat{\gamma}_3$. Using these estimates, we can obtain the detrended series: $\tilde{y}_t = y_t - \hat{\gamma}_1 t - \hat{\gamma}_2 D_t - \hat{\gamma}_3 tD_t$, which will be reduced to a simple expression:

$$\tilde{y}_t = \begin{cases} y_t - t\overline{\triangle y_1}, \ \text{for } t = 0, 1, \ldots, T_B \\ y_0 + (y_t - y_{T_B+1}) - [t-(T_B+1)]\overline{\triangle y_2}, \ \text{for } t = T_B+1, \ldots, T \end{cases} \tag{A9}$$

To show this we make use of the the identity $T_B\overline{\triangle y_1} = y_B - y_0$. Note in particular $\tilde{y}_0 = \tilde{y}_{T_B} = \tilde{y}_{T_B+1} = \tilde{y}_T = y_0$ in the transformed data. Under the null hypothesis, we have

$$\tilde{y}_t = \begin{cases} S_t - t\bar{\varepsilon}_1, \ \text{for } t = 0, 1, \ldots, T_B \\ S_0 - (S_t - S_{T_B+1}) + [t-(T_B+1)]\bar{\varepsilon}_2, \ \text{for } t = T_B+1, \ldots, T \end{cases}$$

where $\bar{\varepsilon}_1 = \frac{1}{T_B}\Sigma_{t=1}^{T_B}\varepsilon_t$, and $\bar{\varepsilon}_2 = \frac{1}{T-(T_B+1)}\Sigma_{t=T_B+2}^{T}\varepsilon_t$. For simplicity assume $y_0 = 0$, so that $\tilde{y}_0 = \tilde{y}_{T_B} = \tilde{y}_{T_B+1} = \tilde{y}_T = 0$. Actually, this holds when we obtain the detrended series by subtracting the initial observation effect with $\tilde{y}_t = (y_t - z_t'\hat{\gamma}) - (y_1 - z_1'\hat{\gamma})$. It will be constructive to split the data into two periods. Let i and j index the first and the second period, respectively, and let S_{1t} and S_{2t} be the partial the partial sum processes for each period starting from zero. Also $\xi_{1i} - m\bar{\varepsilon}_1$ and $\xi_{2j} - m\bar{\varepsilon}_2$ be the instruments for period 1 and 2. We then have

$$\begin{aligned} T^{-1}\tilde{A}_T = {} & T^{-1}\left[\begin{array}{l}\sum_{i=1}^{T_B}(\xi_{1i} - m\bar{\varepsilon}_1)(S_{1,i-1} - i\bar{\varepsilon}_1) + \\ \sum_{j=1}^{T-(T_B+1)}(\xi_{2j} - m\bar{\varepsilon}_2)(S_{2,j-1} - j\bar{\varepsilon}_2)\end{array}\right] \\ & - T^{-1/2}\sum_{t=1}^{T}[(\xi_{1i} - m\bar{\varepsilon}_1) + (\xi_{2i} - m\bar{\varepsilon}_2)] \times \\ & T^{-3/2}\left[\sum_{i=1}^{T_B}(S_{1,i-1} - i\bar{\varepsilon}_1) + \sum_{j=1}^{T-(T_B+1)}(S_{1,j-1} - j\bar{\varepsilon}_2)\right] \end{aligned} \quad \text{(A10)}$$

For the first term of (A10), we use the result in (A7) to obtain

$$T^{-1}\left[\begin{array}{l}\sum_{i=1}^{T_B}(\xi_{1i} - m\bar{\varepsilon}_1)(S_{1,i-1} - i\bar{\varepsilon}_1) + \\ \sum_{j=1}^{T-(T_B+1)}(\xi_{1j} - m\bar{\varepsilon}_2)(S_{1,j-1} - j\bar{\varepsilon}_2)\end{array}\right]$$

$$\xrightarrow{p} \lambda\frac{1}{2}m\sigma^2 + (1-\lambda)\frac{1}{2}m\sigma^2 = \frac{1}{2}m\sigma^2 \quad \text{(A11)}$$

The second term is negligible asymptotically. To see this apply the result in (A8):

$$T^{-1/2}\sum_{t=1}^{T}[(\xi_{1i}-m\bar{\varepsilon}_1)+(\xi_{2i}-m\bar{\varepsilon}_2)]=o_p(1). \qquad \text{(A12)}$$

Distribution of $T^{-1/2}\widetilde{B}_T$

Under the null hypothesis,

$$T^{-1/2}\widetilde{B}_T = T^{-1/2}\sum_{t=1}^{T}\xi_t\varepsilon_t - T^{-3/2}\sum_{t=1}^{T}\xi_t\sum_{t=1}^{T}\varepsilon_t \xrightarrow{d} \sqrt{m}\,\sigma^2 W(1) \qquad \text{(A13)}$$

This result follows since for the first term, we have $T^{-1/2}\sum_{t=1}^{T}\xi_t\varepsilon_t \xrightarrow{d} \sqrt{m}\,\sigma^2 W(1)$, and the second term is $O_p(T^{-1/2})$

Proof of Theorem 1:

Since we have

$$\sqrt{T}\hat{\beta}_{LM-IV}=\frac{T^{-1/2}\widetilde{B}_T}{T^{-1}\widetilde{A}_T}\xrightarrow{d}\frac{\sqrt{m}\,\sigma^2 W(1)}{(1/2)m\sigma^2}=\frac{2}{\sqrt{m}}W(1)$$

where $\widetilde{A}_T$ and $\widetilde{B}_T$ are defined in (A1), we apply the result for $T^{-1}\widetilde{A}_T$ in (A12) and the result in (A13) for $T^{-1/2}\widetilde{B}_T$. We note that the term $T^{-1}\widetilde{A}_T$ is given as a score vector in the derivation of the LM statistic; see the Appendix of Schmidt and Phillips (1992). The proof for $\sqrt{T}\hat{\beta}_{LM-IV}$ is complete by applying the continuous mapping theorem.