Cointegration Tests Using Instrumental Variables Estimation and the Demand for Money in England

Kyung So Im
University of Central Florida

Junsoo Lee*
University of Alabama

Walter Enders
University of Alabama

May 15, 2005

Abstract

In this paper, we propose new cointegration tests based on *stationary* instrumental variables estimation in a single equation model. An important property of our tests is that the asymptotic distribution is standard normal. As such, our IV cointegration tests provide a convenient solution to the nuisance parameter dependency problem found in existing cointegration tests. In addition, the asymptotic distribution of the IV tests does not depend on the dimension of the regressors or differing deterministic terms. We apply the new IV tests to estimate the demand for money equation for England.

JEL Classification: C12, C15, C22
Key Words: Cointegration Test, Stationary Instrumental Variables, Standard Normal Distribution

*Corresponding author: Department of Economics, Finance and Legal Studies, Box 870224, University of Alabama, Tuscaloosa, AL 35487, e-mail: jlee@cba.ua.edu, phone: 205-348-8978, fax: 205-348-0590.

1 Introduction

The pioneering work by Engle and Granger (1987, EG hereafter) addresses the issue of testing whether non-stationary time series are cointegrated and provides relevant asymptotic theory. It is well known that the limiting distribution of the EG cointegration test is non-standard, following the functional of vector Brownian motions. The distribution depends on the dimension of the regressors as well as the documented deterministic terms. Thus, one needs to use different critical values for different specifications of the model. Various extensions of cointegration tests have been suggested and the distributions of these existing cointegration tests are also nonstandard.

In this paper, we propose new cointegration tests based on instrument variables (IV) estimation in a single equation framework. We utilize *stationary* instrumental variables. One important feature of our stationary IV cointegration tests is that their asymptotic distributions are standard normal. As a result, the distributions of the IV tests are free of any nuisance parameter. Additionally, this result implies that the limiting distributions of the IV tests do not depend on the dimension of the regressors or differing deterministic terms. This is a convenient result. Regardless of various model specifications, one can use the same critical value of -1.645 at the 5% significance level. This feature cannot be found in existing cointegration tests based on OLS estimation. More importantly, the asymptotic normality result for the IV cointegration tests implies that we can provide a solution to the problem of nuisance parameter dependency when testing for cointegration. Im and Lee (2005) have shown that the IV approach can be utilized for unit root tests and the resulting tests are free of the nuisance parameters that exist in the models with structural changes. The nature of the nuisance parameter problem is quite different in cointegration tests, and the present paper seeks to resolve the following problems of exisiting cointegration tests.

First, the IV cointegration provides a solution to the nuisance parameter problem of cointegration test statistics based on the error correction model (ECM). Many authors have noted that the asymptotic distribution of the ECM based cointegration test statistic depends on a nuisance parameter when it is a mixture of two quite different distributions, viz, the DF type non-standard distribution and the standard normal distribution. The nuisance parameter describes the relative importance of each of these distributions; see Boswijk (1994) and Zivot (2000), among others. The nuisance parameter is unknown and can take any unknown value. This dependency problem has often hinders researchers from applying the ECM cointegration test during the past decade. However, when the IV estimation is applied to the ECM the distribution of the resulting t-statistic is standard normal.

Second, by utilizing IV estimation in the EG type procedure, the IV cointegration test also provides a solution to the common factor restriction problem imposed in the EG test. One well known problem is that the EG procedure imposes a strict restriction which may not hold. Although the long-run coefficient in the cointegration relationship usually differs from the short-run coefficient obtained from a regression using differenced series, the long-run and short-run coefficients are assumed to be equivalent in the EG procedure. Kremers, Ericsson and Dolado (1992) refer to this restriction as a common factor restriction (CFR). Another serious problem of the EG test is that the test loses power when the signal-noise ratio increases. One way to resolve this problem is to add additional terms using the differenced regressors, but in this case the problem renders to that of the ECM based test. Our IV cointegration tests are free of this type of problem as we do not impose a common factor restriction. In addition, interestingly, although the OLS based EG test loses power under the alternative when the signal-noise ratio increases, the EG type IV test gains power when the ratio increases. The reason for increased power is because the added regressors enter as stationary covariates in our testing regressions. Using stationary covariates has been advocated for unit root tests by Hansen (1995), who showed that the power of unit root tests can be improved. A similar advantage can be utilized in the IV cointegration tests. The difference is that the distribution of the IV cointegration tests still remain as standard normal and is free of the nuisance parameter problem, although unit root tests using covariates induce a nuisance parameter problem. We can use additional stationary covariates as regressors or additional instruments when testing for cointegration. In either case, the resulting cointegration tests will not entail nuisance parameters.

Finally, the IV tests can be adopted in an autoregressive distributed lag (ADL) model. The ADL test requires weak exogeneity of the regressors; see Banerjee, Dolado and Mestre (1998). Under this condition, the ADL model yields efficient estimates of the parameters in the model. The ADL test is free of the CFR and its null distribution does not depend on the signal-noise ratio. However, the ADL test depends on nuisance parameters that enter the deterministic components of the data and the suggested IV tests are free of this complication. For instance, when the cointegration model involves structural changes, the corresponding tests that control for their effects will usually depend on nuisance parameters indicating the location of breaks. This is very cumbersome since the tests already depend on the number of regressors and other deterministic components. In contrast, the asymptotic distribution of the IV cointegration tests remain as standard normal, regardless of structural changes, differing deterministic terms or the number of regressors.

The remainder of the paper is organized as follows. In Section 2, we illustrate the major testing models and examine their relationships. In Section 3, we consider the stationary IV cointegration tests in the error correction model and the EG procedure. We provide relevant asymptotic results. In Section 4, we examine the finite sample performance of the tests. Section 5 provides concluding remarks.

2 Single Equation Cointegration Models

As Ericsson and MacKinnon (2002) explain, there have been three main approaches to testing for cointegration: a system based full information maximum likelihood estimation of a vector error correction model as developed by Johansen (1989), a two step procedure based on a single-equation regression suggested by Engle and Granger (1987), and a single-equation conditional error correction model suggested by Banerjee *et al.* (1986). Ericsson and MacKinnon discuss the advantages and disadvantages of each of these approaches. Perhaps, the Johansen procedure has been the most popular in applied works since it provides the most efficient estimates of the parameters in the model. Ericsson and MacKinnon point out that Johansen's test requires a well-specified full information system and that there are cases in which testing for cointegration can be performed properly in a single equation model under certain conditions. Zivot (2000) discusses various examples for which economic theory can imply a single cointegrating vector and explains why it is reasonable to test for cointegration in single equation models in such cases. In this paper, we also deal with single equation cointegration models. To begin with, we consider a VAR(p) model

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

$$\Pi(L)x_t = \varepsilon_t, \tag{2}$$

where y_t, $t = 1, 2, ..., T$, is a $n \times 1$ vector of $I(1)$ process, d_t denotes deterministic terms, and x_t is the stochastic component following an autoregressive process with $\Pi(L) = I - \sum_{i=1}^{p-1} \Pi_i L^i$ and $\varepsilon_t \sim iid,\ N(0, \Sigma)$. The normality assumption of the error term is made for convenience, but this assumption is not required for the asymptotic results. One can consider the deterministic term with $d_t = c_1$ for the model with a constant, or $d_t = c_1 + c_2 t$ for a model with a trend, but one may also allow for additional polynomial trends and dummy variables to capture seasonality or structural change. It is convenient to write (1) and (2) as a vector error correction model

$$\Delta y_t = c_{11}^* + c_2^* t + \delta\alpha' y_{t-1} + \Gamma(L)\Delta y_{t-1} + \varepsilon_t. \tag{3}$$

Here, we let $\Pi = -\Pi(1)$ and express Π in a cointegrated system as

$$\Pi = \delta\alpha' = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} (1, -\beta'),$$

where δ_1 is a scalar and δ_2 and β are $(n-1)\times 1$ vectors. We can express the conditional equation of Δy_{1t} given Δy_{2t} and other past values of y_t, and the corresponding marginal equation of Δy_{2t} as follows

$$\Delta y_{1t} = (d_{11}+d_{12}t)+\delta_{1.2}(y_{1,t-1}-\beta' y_{2,t-1})+\phi'\Delta y_{2t}+C_{11}\Delta y_{1,t-1}+C_{12}'\Delta y_{2,t-1}+\varepsilon_{1.2t}, \tag{4}$$

$$\Delta y_{2t} = \delta_2(y_{t-1} - \beta' y_{2,t-1}) + \Gamma_{21}\Delta y_{1,t-1} + \Gamma_{22}'\Delta y_{2,t-1} + \varepsilon_{2t}, \tag{5}$$

where $\varepsilon_{1.2t} = \varepsilon_{1t} - \phi'\varepsilon_{2t}$ such that $E(\varepsilon_{1.2t}\varepsilon_{2t}) = 0$. Many authors suggest that if y_{2t} is weakly exogenous for the parameters δ_1 and β , then these parameters can be efficiently estimated in the conditional error correction model (4). The weak exogeneity of y_{2t} implies that $\delta_2 = 0$, so that y_{2t} is not error-correcting. In this case, we obtain that $\delta_{1.2} = \delta_1$; see Harbo *et al.* (1998). Under this assumption, the parameters of interest can be efficiently estimated in the conditional error correction model (4) without having to refer to (5) together. Then, we can rewrite (4) as

$$\Delta y_{1t} = (d_{11} + d_{12}t) + \delta_1 z_{t-1} + \phi'\Delta y_{2t} + C_{11}\Delta y_{1,t-1} + C_{12}'\Delta y_{2,t-1} + v_t, \tag{6}$$

with $z_{t-1} = y_{1,t-1} - \beta' y_{2,t-1}$ and one can test the null hypothesis of no cointegration against the alternative hypothesis of cointegration

$$H_0 : \delta_1 = 0, \quad \text{against} \quad H_1 : \delta_1 < 0.$$

To test this hypothesis, we use the usual t-statistic. This is the ECM based cointegration test for which the cointegrating vector β is assumed to be prespecified; see Banerjee *et al.* (1986), Kremers *et al.* (1992), and Zivot (2000), among others. On the other hand, the conditional ECM can be reparameterized as the conditional autoregressive distributed lag (ADL) model

$$\Delta y_{1t} = (d_{11}+d_{12}t)+\delta_1 y_{1,t-1}+\gamma' y_{2,t-1}+\phi'\Delta y_{2t}+C_{11}\Delta y_{1,t-1}+C_{12}'\Delta y_{2,t-1}+v_t. \tag{7}$$

We can relate the above equation to (6) with $\delta_1\beta' = \gamma'$. It is clear that $\delta_1 = 0$ implies $\gamma = 0$ as well, but δ_1 is unaffected by an arbitrary value of γ, say, $\gamma^* = \delta_1\beta^{*\prime}$ where β^* is an arbitrary long-run coefficient. Thus, the null of no cointegration can be tested on $\delta_1 = 0$ against $H_1 : \delta_1 < 0$ in (7). The resulting t-statistic from (7) is another version of the ECM version cointegration, but it is obtained from the unrestricted ADL model. Banerjee *et al.* (1998) adopt this test and provide critical values for the models with a constant and trend with the number of regressors up to 5.

The Engle and Granger cointegration test is based on the two step procedure. In the first step, the OLS estimate of β, say $\hat{\beta}$, is obtained in the regression of y_{1t} on y_{2t}. Then, the EG cointegration test is based on the t-statistic on $\delta_1 = 0$ in the regression,

$$(\Delta y_{1t} - \hat{\beta}'\Delta y_{2t}) = \delta_1(y_{1,t-1} - \hat{\beta}' y_{2,t-1}) + C(L)(\Delta y_{1t} - \hat{\beta}'\Delta y_{2t}) + u_t. \tag{8}$$

Obviously, the same value of $\hat{\beta}$ is used in both sides of the equation. This implies that the short-run coefficient in the regression of Δy_{1t} on Δy_{2t} is assumed to be equivalent to the long-run coefficient in the regression of y_{1t} on y_{2t}. To see this in more detail, we suppress the deterministic terms and the lagged terms of Δy_{1t} and Δy_{2t} in (6) and relate it to (8) as

$$\Delta y_{1t} - \beta'\Delta y_{2t} = \delta_1(y_{1,t-1} - \beta' y_{2,t-1}) + (\phi' - \beta')\Delta y_{2t} + e_t. \tag{9}$$

We rewrite this as

$$\Delta z_t = \delta_1 z_{t-1} + e_t^*, \tag{10}$$

where

$$e_t^* = (\phi' - \beta')\Delta y_{2t} + e_t. \tag{11}$$

It is clear that the restriction $\phi = \beta$ is imposed in the EG procedure. This restriction is referred to as the common factor restriction which may or may not hold. When the common factor restriction does not hold, the EG test can lose power; see Kremers *et al.* (1992). Note that the EG test does not require the weak exogeneity assumption, which is necessary for the ECM based test. In addition, we note that the asymptotic distributions of both

ECM and EG type cointegration tests depend on the dimension of integrated regressors and differing deterministic terms. The ECM and ADL type tests are free of the CFR problem, but they depend on the dimension of regressors and deterministic terms.

In estimating δ_1 in each of (6) and (8), OLS estimation has been adopted in the literature. The limiting distribution of the resulting t-statistic based on OLS estimation is already provided; see Kremers *et al.* (1992) and Boswijk (1994) for the ECM cointegration test, Banerjee *et al.* (1998) for the ADL cointegration, and Engle and Granger (1987) and Phillips and Ouliaris (1990) for the EG type cointegration test. As noted previously, it is known that the asymptotic distribution of the t-statistic based on the ECM is a mixture of the DF type non-standard distribution and a standard normal distribution. Furthermore, it depends on the nuisance parameter which describes the relative importance of each of these distributions; see Boswijk (1994) and Zivot (2000). The nuisance parameter is unknown, and this dependency has been a source of the problem that hinders from using the ECM cointegration test. The asymptotic distribution of the EG cointegration test is also non-standard, having a Dickey-Fuller type distribution. The EG test is free of the nuisance parameter problem, but it loses power when the signal-noise ratio increases. In the next section, we demonstrate how the IV cointegration tests provide solutions to these problems of the OLS based cointegration tests.

3 Stationary IV Cointegration Tests

In this paper, we consider instrumental variables (IV) estimation of δ_1 in each of (6), (7) and (8). Specifically, we define the instrumental variable, w_t, differently for each model, as

$$\begin{aligned} w_t &= z_{t-1} - z_{t-m-1}, \quad \text{for } z_{t-1} \text{ in (6)} \\ w_t &= (w_{1t},\ w_{2t}'), \quad \text{for } (y_{1,t-1},\ y_{2,t-1}') \text{ in (7)} \\ w_t &= \hat{z}_{t-1} - \hat{z}_{t-m-1}, \quad \text{for } \hat{z}_{t-1} \text{ in (8)}, \end{aligned} \tag{12}$$

where $m \ll T$ is a finite positive integer, $w_{1t} = y_{1,t-1} - y_{1,t-m-1}$, and $w_{2t}' = y_{2,t-1}' - y_{2,t-m-1}'$. Here, $\hat{z}_{t-1} = y_{1,t-1} - \hat{\beta} y_{2,t-1}$ for which β is estimated from the regression of y_{1t} on y_{2t} as

$$y_{1t} = (d_{11} + d_{12}t) + \beta' y_{2t} + error. \tag{13}$$

In a more general case where the error terms are serially correlated, the instrumental variable needs to be adjusted by subtracting more lags. For instance, w_t is modified as $w_t = z_{t-1} - z_{t-1-p-m}$ for the ECM model. For each of the regressions in (6), (7), and (13), a constant term (d_{11}), or both the constant and a trend function $(d_{11} + d_{12}t)$ can be included. The t-statistic on $\delta_1 = 0$ using the IV estimation with the corresponding instrumental variables w_t is our suggested test statistic for cointegration in each of three models

$$t_i = \frac{\hat{\delta}_{1,IV-i}}{s(\hat{\delta}_{1,IV-i})}, \qquad i = \textit{ECM, ADL, and EG.} \tag{14}$$

Then, we refer to each of the corresponding t-statistics as t_{ECM}, t_{ADL}, and t_{EG}, respectively. The distribution of these statistics is shown to follow a standard normal distribution. To see this in more detail, we suppress the deterministic terms for simplicity and rewrite each of the testing regressions. For example, in considering the ECM test, we express the conditional ECM model in (6) as

$$\Delta y_{1t} = \delta_1 z_{t-1} + \pi' q_t + v_t, \tag{15}$$

where $q_t = (\Delta y_{1,t-1}, .., \Delta y_{1,t-p+1}, \Delta y_{2t}', \Delta y_{2,t-1}', .., \Delta y_{2,t-p+1}')'$ and $\pi = (c_{11,1}, .., c_{11,p-1}, \phi', c_{21,1}', .., c_{21,p-1}')'$. In practice, when the cointegrating vector is unknown, the term z_{t-1} in (15) is not feasible. In that case, we replace z_{t-1} in (15) with $\hat{z}_{t-1}$ and construct the IV as $w_t = \hat{z}_{t-1} - \hat{z}_{t-m-1}$, for $\hat{z}_{t-1}$. Then it can be shown that

$$t_{ECM} = \frac{\sum_{t=1}^{T} w_t \Delta y_{1t} - \sum_{t=1}^{T} w_t q_t' \left(\sum_{t=1}^{T} q_t q_t' \right)^{-1} \sum_{t=1}^{T} q_t \Delta y_{1t}}{\hat{\sigma} \sqrt{\sum_{t=1}^{T} w_t^2 - \sum_{t=1}^{T} w_t q_t' \left(\sum_{t=1}^{T} q_t q_t' \right)^{-1} \sum_{t=1}^{T} q_t w_t}}, \tag{16}$$

where

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} \left(\Delta y_{1t} - \hat{\delta}_{1,ECM} \hat{z}_{t-1} - \hat{\pi}' q_t \right)^2.$$

The asymptotic distribution of t_{ECM} is given as follows.

Theorem 1 *Suppose that y_t is generated by the DGP (**??**) with $\delta_1 = 0$, and the t-statistic in (14) is obtained by (16). Then as $T \to \infty$,*

$$t_{ECM} \xrightarrow{d} N(0,1). \tag{17}$$

Proof. See the Appendix. ■

Since the asymptotic distribution of t_{ECM} is standard normal, it is clear that the distribution does not depend on any nuisance parameters. In addition, the asymptotic distribution of the IV statistics is unaffected by the dimension of the regressors (y_{2t}) or the deterministic terms. It is also standard normal when a constant or trend function is allowed in the testing regression. This finding can be extended when dummy variables accounting for structural change are added or when a polynomial time trend is included. The distribution remains as standard normal. Intuitively speaking, this is so because the second term in each of the numerator and denominator in (16) disappear asymptotically. They remain as $O_p(T^{-1/2})$ or $O_p(T^{-1})$ when additional deterministic terms or stationary terms are included in q_t. In addition, asymptotic normality of the test statistic holds as augmented terms are added. If potentially non-stationary terms are added, we need to instrument them to obtain the normality result. This is the case with the ADL based test. The normality result follows since the remaining first term in each of the numerator and denominator is expressed by stationary terms. That is, w_t in (12) is stationary by construction under both the null of no cointegration and the alternative hypothesis. The same normality results are expected to hold for t_{ADL} and t_{EG} owing to the same reasoning.

Corollary *Suppose that y_t is generated by the DGP (**??**) with $\delta_1 = 0$, and the t-statistics t_{ADL} and t_{EG} are obtained as in (14). Then as $T \to \infty$*

$$t_{ADL} \xrightarrow{d} N(0,1) \text{ and } t_{EG} \xrightarrow{d} N(0,1). \tag{18}$$

The property of the IV cointegration test makes a sharp contrast with the OLS based test statistic whose distribution is a mixture of the DF type non-standard distribution and a standard normal distribution, implying that the OLS based statistic depends on a nuisance parameter indicating the relative contributions of the two different distributions. It is well know that the OLS based statistic also depends on the deterministic terms and dimension of stochastic terms describing the non-standard distribution.

We note that the coefficient estimates, $\hat{\delta}_{1,i}$, $i = ECM$, ADL or EG, do not follow a normal distribution. Their distribution is a mixture of a non-standard and standard normal distribution. It has similar properties to those of the OLS based statistics shown in Zivot (2000) and Hansen (1995). One difference is that $\hat{\delta}_{1,i}$ is a$\sqrt{T}$-consistent estimator, instead of a T-consistent estimator. Note that the instrument w_t is asymptotically uncorrelated with the instrumented variable under the null hypothesis of no cointegration. However, under the alternative hypothesis, their correlation coefficient is $1 - \delta_1^m$, which essentially implies consistency of the test under the alternative.

These are asymptotic results in large samples and may not hold in finite samples. When additional deterministic terms are added or the dimension of the regressors increases, a slower convergence rate can be observed in finite samples. In this case, the issue of choosing proper lags and valid instrumental variables arises. However, this outcome occurs mostly in small samples−as the sample size increases, the bias term disappears. It is important to note that our result holds asymptotically with proper values of m, which increases as T increases. In finite samples, the terms that disappear asymptotically remain as bias terms. As such, although the asymptotic result is rather straightforward for a finite value of m, the size property of the test in finite samples depends on the selected value of m. At the same time, a moderately big value of m is necessary for obtaining desirable power properties. No immediate theoretical guidance is readily available in selecting the optimal value of m. Some practical guidance on the choice of m can be found for different values of T from the simulation results in Table 1.

For the EG type IV test, we suggest augmenting Δy_{2t} in (8). As discussed in Kremers *et al.* (1992), we lose potentially valuable information from Δy_{2t} by omitting it in the EG testing regression. Adding Δy_{2t} amounts to not imposing the common factor restriction. As we see in (11), by adding Δy_{2t} in the testing regression, we do not necessarily impose the restriction that $\phi = \beta$. Thus, we have

$$\Delta \hat{z}_t = \delta_1 \hat{z}_{t-1} + \phi' \Delta y_{2t} + u_t, \tag{19}$$

where $\hat{z}_t$ is the residual from the regression in (13). We refer to the resulting t-statistic for $\delta_1 = 0$ as t_{EG}^{+}. Note that one cannot add Δy_{2t} in the usual EG test based on OLS estimation. When Δy_{2t} is added, the distribution of the resulting test depends on the nuisance parameter, which is essentially the same problem of the OLS based ECM test.

4 Simulations

In this section, we investigate small sample properties of the IV cointegration tests through Monte Carlo simulations. We use different values of $m =$

$1, .., 9$. We consider four IV statistics described in the previous section. They are t_{ECM}, t_{ADL}, t_{EG}, and t_{EG}^{+}. In addition, for comparison, we report the simulation results using the OLS based tests. They are denoted as t_{ECM-O}, t_{ADL-O}, and t_{EG-O}, where the subscript "o" was added to signify the use of OLS estimation. We simulated new critical values for the OLS based tests for each of different values of k, T and different models, and used them to compute the size and power of the tests. The simulated critical values are provided in the footnote of corresponding tables. On the other hand, for the IV tests, we use the asymptotic one-sided asymptotic critical value -1.645 of the standard normal distribution at the 5% significance level for all cases. We adopt the following data generating process (DGP) as in Kremers *et al.* (1992)

$$\begin{aligned} \Delta y_{1t} &= \phi' \Delta y_{2t} + \delta_1 (y_{1t} - \beta' y_{2t}) + v_t, \\ \Delta y_{2t} &= u_t \\ \begin{pmatrix} v_t \\ u_t \end{pmatrix} &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix} \right). \end{aligned} \tag{20}$$

All simulation results are based on 20,000 replications. We denote k as the number of integrated regressors, viz the row dimension of y_{2t}. When $\delta_1 = 0$, the DGP implies no cointegration. When $\delta_1 < 0$, the DGP implies cointegration. We start with the case of $k = 1$, in which we examine the effect of using different signal-noise ratios on the size and power. We set $\sigma_v = 1$ and $\beta = 1$. We denote $s = \sigma_u/\sigma_v$ and define the signal-noise ratio as $q = -(\phi - 1)s$. Then, we examine the cases with $(\phi, s) = (1.0, 1)$, $(0.5, 6)$, and $(0.5, 16)$ such that $q = 0$, 3, and 8. Following this, we examine whether the tests are sensitive to the number of integrated regressors with $k = 3$. When $k = 3$, we use the same set of values of (ϕ, s) for all integrated regressors. We examine two different models; one is the drift model with $d_t = c_1$ (the results are reported in Tables 1 - 4) and the other is the trend model with $d_t = c_1 + c_2 t$ (the results are reported in Tables 5 - 8). The instrumental variables are given in (12). A special attention will be given to robustness of the standard normal result of the IV tests under different model specifications and different values of q. For each model, we report the results with $T = 100$ and 300. We also examined the cases of $T = 500$ or $1,000$, but these results are not much different from the results of $T = 300$, except for more desirable results of increased power or such.

In Table 1, we report the size of various tests under the null of no cointegration with $k = 1$ and 3, when we include a constant term in d_t. The OLS based tests should have the correct size as customized critical values are used. It appears that all IV tests show more or less the correct size, which lies within a $\pm 1\%$ error from the nominal 5% size, when a proper value of m is selected. One exception is that no values of m give the correct size for t_{EG}. Also, it is apparent that the sizes of all IV tests are not changed much with $k = 3$ when compared to the case with $k = 1$. We observe that almost no changes are observed for t_{ECM}. This result indicates that the IV tests are invariant to k, the number of endogenous integrated variables in the model. More importantly, it is shown that the IV test t_{ECM} is invariant to the signal-noise ratio (q), although the corresponding OLS based test t_{ECM-O} critically depends on q. Thus, the IV test resolves the nuisance parameter dependency problem of the OLS based ECM test t_{ECM-O}. All other IV tests also seem insensitive to different values of q under the null.

For the OLS based tests, we can use customized critical values for a given model when they have non-standard distributions. For the IV tests, we wish to use only one asymptotic critical value, -1.645, for all different cases with different model specifications. To this end, we wish to select a correct value of m that gives a correct size under the null. However, no clear guidance is readily available regarding the selection of m. Theoretically speaking, any finite value of m will lead to the asymptotic standard normal result, but this is not necessarily the case in finite samples. The best values of m that lead to a correct size under the null vary over different models in finite samples. Our simulation results may provide guidance in this regard. As such, to have a better sense of this matter, we underlined the values of m that give the closest size to the nominal 5% size for each case.

In Figure 1, as an example, we illustrate the pdf of the empirical distribution of t_{ECM} using different values of q, k and m under the null hypothesis when $T = 100$. These empirical distributions are based on the kernel estimation using a Gaussian kernel function. The solid curve depicts the pdf of the standard normal distribution. We can observe that the pdfs of the IV statistics are close to the pdf of the standard normal distribution, regardless of the values of q. On the contrary, we observe in Figure 2 that pdfs of the t-statistics of the OLS based ECM test (t_{ECM-O}) vary significantly over differing values of q. As q increases, the pdf moves to the right away from the pdf with $q = 0$ (solid curve on the left). Thus, they show serious size distortion when $q > 0$.

In Table 2, we examine the size adjusted power of the tests when $\delta_1 = -0.1$. We observe that the IV ECM test t_{ECM} performs well. The t_{ECM} test is more powerful than other IV tests and is fairly comparable to the OLS based ADL tests. For instance, the power of t_{ECM} is 0.257 when $m = 7$, while the power of t_{ADL-O} is 0.244. The OLS ECM test t_{ECM-O} appears more powerful than any others but t_{ECM-O} is not a valid test as examined in Table 1. Thus, we exclude this test from further discussion. In general, the ECM and ADL based tests are more powerful than the EG tests. Our interest lies in examining the effect of q. The power of the EG tests (t_{EG} and t_{EG-O}) decreases as q increases. Under the null, these tests do not depend on the signal-noise ratio q, but they depend on q under the alternative. A similar result for the OLS based EG test was discussed in Kremers *et al.* (1992). The source of this problem is that these tests omit the term Δy_{2t} from the testing regression. However, the modified EG test t^{+}_{EG} is not subject to this problem when Δy_{2t} is added as in (19). The power of the t^{+}_{EG} test increases as the signal-noise ratio increases. As noted in the previous section, adding Δy_{2t} in the EG procedure amounts to relaxing the common factor restriction (CFR). In the OLS framework, we cannot add Δy_{2t} since it leads to the nuisance parameter problem and invalid tests. But, in the IV framework the null distribution remains as standard normal, implying absence of the nuisance parameter dependency problem, and the power increases under the alternative. The power of the ECM and ADL tests increases as q increases. This phenomenon is observed for both IV and OLS based tests (t_{ECM}, t_{ADL}, and t_{ADL-O}). Thus, in addition to the t^{+}_{EG} test, both t_{ECM} and t_{ADL} also solve the problem of losing power in the usual EG tests when the signal-noise ratio increases.

In Figure 3, we plot the pdf of the empirical distribution of the t-statistics for the IV test t_{ECM} by varying the values of the signal-noise ratio (q) under the alternative hypothesis when $T = 100$ and $k = 1$ or 3. We wish to illustrate graphically the effect of using different values of q on the power of the test. It is clear that the pdfs of the IV statistic of t_{ECM} shifts leftward away from the pdf of the normal distribution, thereby gaining power as q increases. The underlying logic for increased power is similar in nature to the finding of Hansen (1995) who showed that power of the usual unit root tests increases by adding stationary covariates. In our framework, Δy_{2t} works the same as stationary covariates, and the gain in power is bigger when the variance of Δy_{2t} gets bigger so that the signal-noise ratio increases.

This effect is enhanced as the dimension of Δy_{2t} (k) increases with more regressors. Then, the power increases further as k increases. The far-left truncated curve shows the case when $k = 3$. This phenomenon is the opposite direction of the OLS based tests whose power decreases as k increases. In general, power noramlly decreases as k increases since additional parameters need to be estimated. In the IV tests, the effect of increasing power with additional covariates is usually bigger than the effect of decreasing power. If not, the test loses power. This occurs in the t^+_{EG} test; it loses power as k increases. However, it is encouraging that the power of t_{ECM} is not reduced as k increases. Indeed, in some cases, the power of the t_{ECM} test increases; see the case in Table 2 when $q = 3$ as k increases from 1 to 3. We note that the IV version of the ADL test (t_{ADL}) is less powerful than the OLS version of the ADL test (t_{ADL-O}). This result is expected since the OLS based tests are usually more powerful. However, the t_{ECM} test is fairly comparable to t_{ADL-O}.

In Tables 3 and 4, we replicate the same simulations as in Tables 1 and 2, but only with a large sample size of $T = 300$. The properties of the tests are similar to those with $T = 100$, except that the power of the test increases as the sample size increases. It is clear that a higher value of m is needed for the IV tests to have a correct size under the null when the sample size increases. We underlined the values of m that give the size closest to the nominal 5% size for each case.

In Tables 5 - 8, we report the results with the trend model. Note that the same asymptotic critical value of -1.645 is used in all cases for the IV tests. The outcome remains similar to that in the previous simulation results with only a constant and the basic results on the properties of the tests remain unchanged. We do not observe any significant size distortion or significant loss of power by adding the trend function. The power of all tests decreases somewhat, when compared with the drift model, but this result is expected when we deal with more general models. Again, we conclude that the standard normal result still holds in the trend model.

Overall, the IV tests are reasonably robust to different model specifications. This is an expected outcome due to the fact that the IV cointegration tests do not depend critically on the usual deterministic terms. As noted, our tests do not critically depend on the number of integrated regressors. We expect that similar results will follow when structural breaks are allowed. Most important, the IV tests are invariant to the nuisance parameters that make the OLS based tests invalid. The IV cointegration tests do not depend on the signal-noise ratio under the null, and their power increases as the signal-noise ratio increases under the alternative.

5 Empirical Example: The Demand for Narrow Money in the U.K.

We now demonstrate an empirical example of using the IV cointegration tests. To avoid the appearance of selecting an arbitrary set of variables, we illustrate our IV cointegration test using the data set complied by Hendry and Ericsson (1991) for their study of the demand for narrow money in the U.K during the 1964:Q3 ~1989:Q2 period. This data is widely available, has been studied intensively, and has interesting time-series properties.[1] For our purposes, the data is particularly appealing because it has been estimated and tested using the Johansen FIML procedure as well as the single equation-based procedures. For example, Hendry (1995) used the Johansen methodology to obtain the following estimates of possible cointegrating relationships among the variables of the money demand function

$$\begin{bmatrix} 1.00 & -1.00 & -7.34 & 7.65 & -0.0005 \\ -0.06 & 1.00 & -3.38 & 0.86 & -0.0059 \\ -0.29 & 0.69 & 1.00 & -0.63 & -0.0025 \\ 0.03 & -1.58 & 1.10 & 1.00 & 0.0097 \end{bmatrix} \begin{bmatrix} (m-p)_t \\ i_t \\ \Delta p_t \\ R_t \\ t \end{bmatrix}, \qquad (21)$$

where m_t = the log of nominal narrow money, i_t = the log of real total final expenditure (TFE) at 1985 prices, p_t = the log of the TFE deflator, R_t = the difference between the three-month local authority interest rate and a learning-adjusted retail sight-deposit interest rate, and t = time index.

The λ_{max} and λ_{trace} statistics indicate exactly two cointegrating vectors among the four variables. A long-run money demand relationship can be identified by imposing a unitary income elasticity of demand, equal coefficients for Δp_t and R_t, and a zero time trend on the most significant cointegrating relationship. The key feature of the resulting sub-system is that i_t, Δp_t,and R_t are weakly exogenous for $(m-p)_t$; as such, the money demand function can be estimated in a single-equation framework. For example, Ericsson and Mackinnon (2002) use identical data to estimate the following error-correction model using OLS[2]

[1]The data we use is available at http://www.nuff.ox.ac.uk/users/hendry/.

[2]Notice that this result is a reparameterized version of Ericsson and Mackinnon's (2002) equation (30). We are able to reproduce their results to two significant decimal places.

$$\Delta(m-p)_t = -0.088(m-p)_{t-1} - 0.696\Delta p_{t-1} - 0.611R_t \quad (22)$$
$$+0.095i_{t-1} - 0.174\Delta(m-p-i)_{t-1} - 0.0498.$$

The test for cointegration can be conducted by comparing the t-statistic on the coefficient $(m-p)_{t-1}$ to the appropriate critical value tabulated by Ericsson and Mackinnon (2002). Since there are four variables in the cointegrating relationship and there is only an intercept term in the regression equation, it is appropriate to use the $\kappa_c(4)$ statistic. The asymptotic critical value for $\kappa_c(4)$ at the 1% level is -4.35. As such, it is possible to reject the null hypothesis of no cointegration at conventional levels.

A critical issue in any cointegration analysis is the selection of the proper set of deterministic regressors. Although there is little economic reason to include a linear or a quadratic trend in the money demand function, Ericsson and MacKinnon (2002) do report the effects of including such deterministic regressors. In the presence of the quadratic trend, the estimated coefficient on $(m-p)_{t-1}$ remains at -0.088 but the t-statistic falls to -3.36. With a quadratic trend and four variables in the cointegrating relationship, the appropriate critical values are given by the table for $\kappa_{ctt}(4)$. The asymptotic critical value for $\kappa_{ctt}(4)$ at the 5% and 1% levels are -4.52 and -5.18, respectively. As such, they are not able to reject the null hypothesis of no cointegration at conventional levels.

Notice that the critical values for the cointegration test depend on the number of non-stationarity variables in the model. Within the sample period under consideration, the inflation rate (Δp_t) acts as a unit root process; the sample value of τ_μ in a standard Dickey-Fuller test is -2.53 whereas the critical value at the 10% level is -2.58. However, if U.K. inflation were actually $I(0)$, it would be necessary to test for cointegration using the $\kappa_c(4)$ or $\kappa_{ctt}(4)$ statistics. A similar problem would result if the interest rate and inflation were cointegrated. No such problem exists for our IV cointegration test since the distribution is normal regardless of the deterministic regressors and the number of $I(1)$ variables included in the estimating equation, in which case we can instrument the $I(1)$ variables.

We begin by re-estimating the model in (22) using $w_t = (m-p)_{t-1} - (m-p)_{t-9}$ as an instrument for $(m-p)_{t-1}$ and obtain the IV estimates as

$$\Delta(m-p)_t = -0.103(m-p)_{t-1} - 0.684\Delta p_{t-1} - 0.671R_t + 0.100i_{t-1} - 0.1914\Delta(m-p-i)_{t-1} - 0.053. \quad (23)$$

Notice that the point estimates of the coefficients are all quite similar to that of (22). The t-statistics, not shown, for $(m-p)_{t-1}$ rises from -7.85 to -5.77. Since the distribution of the t-statistic of the IV estimate is normally distributed, we can reject the null of no cointegration at conventional levels.

When we follow Ericsson and MacKinnon (2002) and include a quadratic time trend as instruments and regressors, the t-statistic for $(m-p)_{t-1}$ rises to -2.18 with a *prob*-value of 0.030, from a standard normal distribution. Thus, even if the quadratic trend is included, we are still able to reject the null hypothesis of no cointegration. In addition, the IV estimator allows us to include the quadratic trend as an instrument but not as a regressor. When we use this method we obtain very strong evidence of cointegration. It is also important to note that the speed-of-adjustment coefficient on the error correction term is almost exactly the same as that in (22). Consider

$$\Delta(m-p)_t = -0.090(m-p)_{t-1} - 0.700\Delta p_{t-1} - 0.619R_t + 0.095i_{t-1} - 0.176\Delta(m-p-i)_{t-1} - 0.029. \quad (24)$$

Using quarterly data, it seems natural to use an instrument in the form $w_t = (m-p)_{t-1} - (m-p)_{t-n}$ where n is a multiple of 4. Nevertheless, to provide an idea of the sensitivity of the results to the choice of n, we estimated an equation in the form of (22) using values of n ranging from 4 to 16. The resulting t-statistic for the error-correction term are given as

n	4	5	6	7	8	9	10	11	12
t	−2.57	−3.38	−4.98	−5.77	−6.60	−6.10	−7.91	−7.23	−6.82

Notice that the null hypothesis of no cointegration can be rejected for all values of n. Thus, our results seem pretty robust.

These results hinge on the assumption that all regressors are stationary since we do not utilize other regressors in constructing instruments. It is possible that they are non-stationary, in which case we instrument them as well. This becomes the ADL based estimation in our example. We find almost identical qualitative results when we estimate the model as an ADL.

Specifically, when we use $(m-p)_{t-1}-(m-p)_{t-9}$, $\Delta p_t - \Delta p_{t-8}$, $R_t - R_{t-8}$ and $i_t - i_{t-9}$ as instruments for $(m-p)_{t-1}$, Δp_t, R_t and i_{t-1}, we obtain

$$\begin{aligned}\Delta(m-p)_t &= -0.083(m-p)_{t-1} - 0.700\Delta p_{t-1} - 0.618 R_t \\ &\quad +0.078 i_{t-1} - 0.158\Delta(m-p-i)_{t-1} + 0.083.\end{aligned} \tag{25}$$

Given that the t-statistic, reported as -5.29, for the error correction term is normally distributed, we can reject the null hypothesis of no cointegration at any conventional significance level. The findings are quite robust to using other values of n. The t-statistics for the error correcting term for other values of n are

n	4	5	6	7	8	9	10	11	12
t	-6.34	-5.90	-5.62	-5.90	-4.15	-5.05	-3.89	-3.52	-5.93.

6 Concluding Remarks

In this paper we have proposed new cointegration tests using stationary instrumental variables. Unlike the usual cointegration tests using OLS estimates, the asymptotic distributions of the IV cointegration tests are standard normal. As such, the tests are free of nuisance parameters. This is an important feature of the IV cointegration tests. Also, the distribution does not depend on the deterministic terms or the dimension of the regressors. Furthermore, by adding Δy_{2t}, the EG type tests are free of the problem of imposing incorrect common factor restrictions. Therefore, our IV tests provide solutions to the limitation of each of the existing cointegration tests.

While we have examined standard cases of single equation cointegration models, the idea can be extended to other important cases where nuisance parameter dependency may be an issue. An immediate extension of our tests may be to allow for structural change. As in IV unit root tests of Im and Lee (2005), the resulting cointegration tests are expected to be free of the nuisance parameters indicating the location of breaks. Also, given that IV tests are free of nuisance parameters, it may be reasonable to extend our cointegration tests to the panel framework. While using stationary covariates has been advocated by Hansen (1995), who showed that the power of unit root tests can be improved, it also appears possible to improve the power of cointegration tests by utilizing stationary covariates. In our case, we can use stationary covariates as regressors or additional instruments, but in either case the tests will not entail nuisance parameters.

References

[1] Banerjee, A., J.J. Dolado, D. Hendry, and G.W. Smith (1986), "Exploring Equilibrium Relationships in Econometrics Through Static Models: Some Monte Carlo Evidence," *Oxford Bulletin of Economics and Statistics*, 48, 3, 253-277.

[2] Benerjee, A., J.J. Dolado, and R. Mestre (1998), "Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework," *Journal of Time Series Analysis*, 19, 3, 267-283.

[3] Boswijk, H. P. (1994), "Testing for an Unstable Root in Conditional and Structural Error Correction Models," *Journal of Econometrics*, 63, 37-60.

[4] Enders, W. and P. Siklos (2001), "Cointegration and Threshold Adjustment." *Journal of Business and Economic Statistics* 19, 166 – 76.

[5] Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251-276.

[6] Ericsson, N.R., and J.G. MacKinnon (2002), "Distributions of Error Correction Tests for Cointegration," *Econometrics Journal*, 5, 285-318.

[7] Hansen, B. (1995), "Rethinking the univariate approach to unit root tests: How to use covariates to increase power," *Econometric Theory*, 11, 1148-1171.

[8] Harbo, I., S. Johansen, B. Nielsen, and A. Behbek (1988), "Asymptotic Inference on Cointegrating Rank in Partial Systems," *Journal of Business and Economic Statistics*, 16, 4, 388-399.

[9] Hendry, D.F. (1995), "Econometrics and business cycle empirics," *Economic Journal*, 105, 1622-1636.

[10] Hendry, D. F. and Ericsson, Neil R. (1991). "Modeling the demand for narrow money in the United Kingdom and the United States," *European Economic Review*, 35 (4), 833-881.

[11] Im, K., and J. Lee (2005), "Testing for Unit Roots with Stationary Instrumental Variables," manuscript.

[12] Johansen, S. (1989), "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, 231-254.

[13] Kremers, J.J.M., N.R. Ericsson and J.J. Dolado (1992), "The Power of Cointegration Tests," *Oxford Bulletin of Economics and Statistics*, 54, 3, 325-348.

[14] Phillips, P.C.B., and S. Ouliaris (1990), "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica, 58, 165-193.*

[15] Phillips, P.C.B, J. Y. Park and Y. Chang (2004), "Nonlinear instrumental variable estimation of an autoregression," *Journal of Econometrics*, 118, 219-246.

[16] Saikkonen, P. (1991), "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory*, 7, 1-21.

[17] 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, 165-176.

[18] Zivot, E. (2000), "The Power of Single Equation Tests for Cointegration When the Cointegrating Vector is Prespecified," *Econometric Theory*, 16, 407-439

7 APPENDIX

Lemma 2 *Assume that* $\{\varepsilon_t\}_{t=1}^{\infty}$ *is an iid process with mean zero, variance* σ^2*, and finite fourth moment. Define a partial sum process* $S_{[rT]} = \Sigma_{j=1}^{[rT]}\varepsilon_j$*, with* $r \in [0,1]$ *and* $\xi_t = \varepsilon_{t-1} + ... + \varepsilon_{t-m}$*, where* m *is a finite positive integer. Then,*

$$T^{-1}\sum_{t=1}^{T} S_{t-1}\varepsilon_t \xrightarrow{d} \frac{1}{2}\sigma^2\left[W(1)^2 - 1\right] \qquad (A1)$$

$$T^{-1/2}\sum_{t=1}^{T} \xi_t\varepsilon_t \xrightarrow{d} \sqrt{m}\sigma^2 W(1). \qquad (A2)$$

$$T^{-1}\sum_{t=1}^{T} \xi_t^2 \xrightarrow{p} m\sigma^2. \qquad (A3)$$

Proof. *(A1) is a standard result. (A3) can be easily seen since* $\xi_t = \varepsilon_{t-1} + ... + \varepsilon_{t-m}$*. For (A2), note that* $\xi_t S_{t-1} = \xi_t(S_{t-1} - S_{t-1-m} + S_{t-1-m})$ $= \xi_t(\xi_t + S_{t-1-m})$*. We have* $T^{-1}\Sigma_{t=1}^{T}\xi_t^2 \xrightarrow{p} m\sigma^2$ *by the strong law of large numbers and* $T^{-1}\Sigma_{t=1}^{T}\xi_t S_{t-1-m} \xrightarrow{d} \frac{1}{2}m\sigma^2\left[W(1)^2 - 1\right]$ *from (A1). Then, (A3) follows since* $\{\xi_t\varepsilon_t\}_1^{\infty}$ *is a martingale difference series with variance* $m\sigma^4$*.* ■

Proof of Theorem 1

Let

$$B_T = \sum_{t=1}^{T} w_t\Delta y_{1t} - \sum_{t=1}^{T} w_t q_t'\left(\sum_{t=1}^{T} q_t q_t'\right)^{-1}\sum_{t=1}^{T} q_t\Delta y_{1t} \qquad (A4)$$

$$C_T = \sum_{t=1}^{T} w_t^2 - \sum_{t=1}^{T} w_t q_t'\left(\sum_{t=1}^{T} q_t q_t'\right)^{-1}\sum_{t=1}^{T} q_t w_t. \qquad (A5)$$

We first examine the distribution of $\frac{1}{\sqrt{T}}B_T$. Rewrite (A4) as

$$\frac{1}{\sqrt{T}}B_T = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} w_t\Delta y_{1t} - \frac{1}{\sqrt{T}}\sum_{t=1}^{T} w_t q_t'\left(\frac{1}{T}\sum_{t=1}^{T} q_t q_t'\right)^{-1}\frac{1}{T}\sum_{t=1}^{T} q_t\Delta y_{1t}. \qquad (A6)$$

Looking at the first term on the right hand side of (A6), we first note that w_t can be written as $w_t = z_{t-1} - z_{t-2} + z_{t-2} - .. + z_{t-m} - z_{t-m-1} = \Delta z_{t-1} + .. +$

Δz_{t-m}. Then, from (10) and (11), we have $w_t = e^*_{t-1} + .. + e^*_{t-m}$. Also, for the null distribution, we can express $\Delta y_{1t} = e_t$ without loss of generality since the null distribution is unaffected by the parameters in the term q_t. Note that we can have $\Sigma_{t=1}^T (e^*_{t-1} + .. + e^*_{t-m})e_t = \Sigma_{t=1}^T (e_{t-1} + .. + e_{t-m})e_t$ when past values of Δy_{1t} and Δy_{1t} are uncorrelated with e_t. Therefore, under the null hypothesis, we have

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t \Delta y_{1t} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} (e_{t-1} + .. + \ e_{t-m}) e_t \to \sqrt{m} a(1)^{-1} \sigma^2 W(1). \quad \text{(A7)}$$

This result follows from (A2). We now show that the second term on the right side of (A6) is $O_p(T^{-1/2})$. The first product term $\frac{1}{\sqrt{T}} \Sigma_{t=1}^T w_t q_t' = O_p(1)$, follows a similar result as in (A7). It is clear that $\left(\frac{1}{T} \Sigma_{t=1}^T q_t q_t'\right)^{-1} = O_p(1)$ since q_t contains only $I(0)$ processes. It is also clear that $\frac{1}{T} \Sigma_{t=1}^T q_t \Delta y_{1t} = O_p(T^{-1/2})$, since $\frac{1}{\sqrt{T}} \Sigma_{t=1}^T q_t \Delta y_{1t} = O_p(1)$. Then, the second term on the right hand side of (A6) is $O_p(T^{-1/2})$. When q_t contains deterministic terms including a time trend, the normalization factor of T can be changed accordingly, but the order of convergence of these terms will not be affected. Thus, the second term on the right hand side of (A6) is still $O_p(T^{-1/2})$ when additional deterministic terms are included. Therefore,

$$\frac{1}{\sqrt{T}} B_T \to \sqrt{m} a(1)^{-1} \sigma^2 W(1). \quad \text{(A8)}$$

We next examine the distribution of $\frac{1}{T} C_T$. We rewrite (A5) as

$$\frac{1}{T} C_T = \frac{1}{T} \sum_{t=1}^{T} w_t^2 - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t q_t' \left(\frac{1}{T} \sum_{t=1}^{T} q_t q_t' \right)^{-1} \frac{1}{T^{1.5}} \sum_{t=1}^{T} q_t w_t. \quad \text{(A9)}$$

Again, we can show that the second term on the right hand side of (A9) is degenerate. It is clear that under the null hypothesis

$$\frac{1}{T} \sum_{t=1}^{T} w_t^2 \to m a(1)^{-2} \sigma^2, \quad \text{(A10)}$$

following (A3). Then, since $\frac{1}{\sqrt{T}} \Sigma_{t=1}^T w_t q_t' = O_p(1)$, and $\frac{1}{T} \Sigma_{t=1}^T q_t q_t' = O_p(1)$, the second term on the right hand side of (A9) is $O_p(T^{-1})$. This result holds when additional deterministic terms are added to q_t. Thus,

$$\frac{1}{T} C_T \to m a(1)^{-2} \sigma^2. \quad \text{(A11)}$$

Therefore, combining the results in (A8) and (A11), we can show that

$$t_{IV-ECM} = \frac{\frac{1}{\sqrt{T}} B_T}{\hat{\sigma}\sqrt{\frac{1}{T} C_T}} \xrightarrow{d} \frac{\sqrt{m} a(1)^{-1} \sigma^2 W(1)}{\sigma \sqrt{a(1)^{-2} m \sigma^2}} = W(1) \sim N(0,1). \qquad \text{(A12)}$$