Cointegration Tests Using Instrumental Variables with an Example of the U.K. Demand for Money

Walter Enders Kyung So Im Junsoo Lee*

May 20, 2008

Abstract

In this paper, we propose new cointegration tests based on *stationary* instrumental variables in a single equation model as well as in a system of equations. An important property of our tests is that the asymptotic distribution is standard normal or chi-square. As such, the asymptotic distribution of the IV tests does not depend on the number of the regressors, differing deterministic terms, structural changes, and even statinary covariates. Thus, our IV cointegration tests have operating advantages in the presence of nuisance parameters. Moreover, we show that including stationary covariates increases considerably the power of the tests without affecting size. We illustrate the use of the tests by examining the demand for money in the U.K.

JEL Classification: C12, C15, C22

Key Words: Cointegration Test, Stationary Instrumental Variables, Standard Normal Distribution

*Walter Enders, Professor and Bidgood Chair of Economics and Finance, Department of Economics, Finance and Legal Studies, University of Alabama, Tuscaloosa, AL 35487 (email: wenders@cba.ua.edu); Kyung So Im, Economist, Division of Insurance and Research, Federal Deposit Insurance Corporation (FDIC), 550 17th Street, NW, Washington, DC (e-mail: Kim@FDIC.gov); and Junsoo Lee (Corresponding author), Professor, Department of Economics, Finance and Legal Studies, University of Alabama, Box 870224, AL 35487 (email: jlee@cba.ua.edu, phone: 205-348-8978, fax: 205-348-0590).

1 Introduction

The commonly used cointegration tests have nonstandard distributions that depend on a number of nuisance parameters. The appropriate critical values depend on the number of $I(1)$ variables (k), and the form of the deterministic regressors. In many cases, it is possible to tabulate critical values for various values of T, k and different combinations of the deterministic regressors that can appear in the DGP. Nevertheless, there are circumstances in which it is cumbersome to provide tabulated critical values. For example, if dummy variables are used to capture the effects of structural breaks, the critical values will depend on the location of the break dates. It is simply difficult to provide critical values of all possible combinations of break locations for different values of k and differing deterministic terms. This issue is even more problematic when stationary $I(0)$ covariates are present in the cointegration regression. It is well-known that the presence of $I(0)$ variables in a cointegration test induces a nuisance parameter problem as the appropriate critical values depend on the nature of the stationary covariates. In such cases, a number of applied authors exclude the $I(0)$ variables from the model; for example, see Wohar and Rapach (2002). But, excluding them is not necessarily the best option. Given that the critical values for these applications are so idiosyncratic, it is possible to bootstrap the appropriate confidence intervals. However, as detailed in Harris and Judge (1998) bootstrapped critical values for cointegration tests have poor size properties. Their point is that the *a priori* notion that the bootstrap should perform better than an asymptotic approximation is not generally true when dealing with nonergodic processes.

In this paper, we propose a set of new cointegration tests based on *stationary* instrumental variable (IV) estimation. We first develop tests in the single equation framework. Then, we extend the tests to a system of equations. An important feature of the new tests is that their asymptotic distributions are standard normal or chi-square throughout. As such, the limiting distributions of the IV tests do not depend on the number of the regressors, the nature of deterministic terms, or the presence of $I(0)$ variables. This is a useful result since it alleviates much of the need to obtain new critical values for all different cases. For instance, dummy variables and stationary covariates can be included in the model without the need to simulate new critical values or use a bootstrap method. Regardless of different model specifications, one can use the same asymptotic critical value following the standard normal or chi-square distributions rather than non-standard distributions. More importantly, we demonstrate that adding stationary covariates increases considerably the power of the IV cointegration tests without affecting the size properties under the null. The asymptotic normality result holds for residual-based cointegration tests, such as the Engle and Granger (1987, EG hereafter) test, but also holds for tests based on error-correction models (ECMs) and autoregressive distributed lag (ADL) models. ECM and ADL models can be quite powerful.

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 ECM and ADL framework and in the EG procedure. We consider both single equation tests and tests in a system of equations. Asymptotic properties are provided. In Section 4, we examine the finite sample performance of the tests and show that they have good size and power properties. In Section 5 we illustrate the appropriate use of the test by examining the well-known example of U.K. money demand. Section 6 provides concluding remarks.

2 Cointegration Models

As Ericsson and MacKinnon (2002) explain, there are three main approaches to testing for cointegration: the full information maximum likelihood estimation of a vector error correction model as developed by Johansen (1989), the two step procedure based on regression residuals suggested by Engle and Granger (1987), and single-equation ECM and ADL models of the type 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 work since it provides the most efficient set of parameter estimates. Nevertheless, there are circumstances in which testing for cointegration can be performed properly in a single equation model. As discussed in Zivot (2000), the ECM and ADL models have seen a resurgence. Unlike a system-based VAR approach, the single-equation models often provide a very parsimonious representation of the variable of interest. Moreover, the single equation models are easily adaptable to the particular functional form suggested by economic theory and they have a flexible dynamic representation. To better understand relationships among the three approaches, consider the 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 $k \times 1$ 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. Although it is customary to use $d_t = c_1$ for a model with a constant, or $d_t = c_1 + c_2 t$ for a model with a trend, it is also possible to include polynomial trends and dummy variables to capture seasonality or structural change. If the variables are cointegrated (1) and (2) can be written as the vector error correction model

$$\Delta y_t = c_1^* + c_2^* t + \delta \alpha' y_{t-1} + \Gamma(L) \Delta y_{t-1} + \varepsilon_t. \tag{3}$$

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 $(k-1) \times 1$ vectors. For our purposes, it is convenient to express the individual equations of (3) using orthogonalized errors.

It should be clear that the conditional equation of Δy_{1t}, given Δy_{2t} and $\Gamma(L)\Delta y_t$, and the corresponding marginal equation of Δy_{2t} can be written as

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

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

where $v_t = \varepsilon_{1t} - \phi' \varepsilon_{2t}$ such that $E(v_t \varepsilon_{2t}) = 0$. It is well known 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

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

This is the ECM based cointegration test for which the cointegrating vector β is assumed to be pre-specified; 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 $H_0 : \delta_1 = 0$ against $H_1 : \delta_1 < 0$ in (7). The resulting t-statistic from (7) is another version of the ECM 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. In light of imposing the absence of any level relationship between y_{1t} and y_{2t}, Boswijk (1994) considers the joint hypothesis $H_0 : \delta_1 = \gamma = 0$ in (7) and suggests an alternative ADL test; see also Pesaran *et al.* (2001).

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. Kremers, Ericsson and Dolado (1992) refer to this restriction as a common factor restriction (CFR). When the CFR 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. Note that the ADL tests also depend on the dimension of regressors and deterministic terms.

In estimating δ_1 in each of (6), (7), 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) for the ECM cointegration test, Banerjee *et al.* (1998) and Boswijk (1994) for the ADL cointegration test, 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 the standard normal distribution. Furthermore, it depends on the nuisance parameter which describes the relative importance of each of these distributions. The nuisance parameter is unknown, and this dependency has been a reason that hinders from using the ECM cointegration test. Recently, Zivot (2000) suggests estimating the nuisance parameter in a nonparametric way by using a long-run variance estimate. However, the ECM test still depends on stationary covariates and deterministic terms. 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 the 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

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

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 practice, when the cointegrating vector is unknown, the term z_{t-1} in (6) is not feasible. In that case, we replace z_{t-1} in (6) with $\hat{z}_{t-1}$ and construct the IV as $w_t = \hat{z}_{t-1} - \hat{z}_{t-m-1}$, for $\hat{z}_{t-1}$. 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, and a similar adjustment is necessary for other tests. 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. However, the asymptotic critical values of the IV cointegration tests are not changed. The t-statistic on $\delta_1 = 0$ using the IV estimation with the corresponding instrumental variables is our suggested statistic testing for cointegration in each of three models

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

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 $q_t = (1, t, \Delta y_{1,t-1}, .., \Delta y_{1,t-p+1}, \Delta y'_{2t}, \Delta y'_{2,t-1}, .., \Delta y'_{2,t-p+1})'$ and $\pi = (d_{11}, d_{12}, c_{11,1}, .., c_{11,p-1}, \phi', c'_{21,1}, .., c'_{21,p-1})'$. 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 our suggested IV cointegration tests is given as follows.

Theorem 1 *Suppose that* y_t *is generated as a vector time series as in (1), and the t-statistic in (14) is obtained by (16). Then under the null of no cointegration, as* $T \to \infty$,

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

Proof: See the Appendix.

Since the asymptotic distribution of each test is standard normal, it is clear that the distribution does not depend on any nuisance parameters. As such, 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 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 involving q_t in each of the numerator and denominator in (16) disappear asymptotically. The second term on the numerator is $O_p(T^{-1/2})$ and the second term in the denominator is $O_p(T^{-1})$. Note that additional deterministic terms and 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 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. When the null hypothesis is given as the joint hypothesis $H_0 : \delta_1 = \gamma = 0$ in (7), we can consider the Wald type test as in Boswijk (1994). Since each t-statistic for δ_1 or γ_i, for $i = 1, .., k-1$, has a standard normal distribution, the resulting Wald statistic will have a chi-square distribution with degree of freedom equal to the number of restrictions. We call this Wald test as $ADL2^*$, and will examine this test further along with the system based IV tests.

The property of the IV cointegration test makes a sharp contrast with the OLS based test. For example, the distribution of the OLS based ECM test is a mixture of the DF type non-standard distribution and a standard normal distribution, implying that it depends on a nuisance parameter indicating the relative contributions of the two different distributions. It is well known 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; see Im and Lee (2007) for further analysis of the distribution of the coefficient estimates. However, the corresponding t-statistic for δ_1 has the standard normal distribution. One interesting question to examine is the effect of the contemporaneous term Δy_{2t} in the conditional models on the properties of the single equation based IV tests. For example, we express equation (15) as $\Delta y_{1t} = \delta_1 z_{t-1} + \phi' \Delta y_{2t} + v_t$, while we suppress augmented terms for simplicity. In the OLS based ECM tests, the distribution of the t-statistic for δ_1 is a mixture of a non-standard and standard normal distribution. The reason is that the term Δy_{2t} acts like stationary covariates. Hansen (1995) has already shown that adding stationary covariates to a unit root testing regression will result in a mixture of these two distributions. This same phenomenon occurs in the OLS based cointegration tests as demonstrated by Kremers *et al.* (1992) and Zivot (2000). However, our suggested IV cointegration tests are free of this difficulty. Instead, the stationary term Δy_{2t} serves to improve the power of the IV ECM cointegration tests under the alternative without affecting the distribution under the null. We will examine this issue more carefully in the next section.

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 they may not necessarily hold well 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, some asymptotically negligible terms can remain as bias terms. Thus, 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. We suggest selecting the value of m that minimizes the sum of squared residuals in the testing regression. In addition, the simulation results in the next section may provided some limited guidance on the choice of m; one may choose the optimal value of m which balances size and power properties for each of different values of T.

For the EG type IV test, we modify (8) and suggest augmenting Δy_{2t}. 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

$$\Delta \hat{z}_t = \delta_1 \hat{z}_{t-1} + \phi' \Delta y_{2t} + u_t, \tag{18}$$

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 situation of the OLS based ECM test. Although Δy_{2t} affects the null distribution of the OLS based ECM tests, we will show that the presence of the term Δy_{2t} in the conditional model makes the OLS based EG type test lose power under the alternative. On the other hand, our IV based EG test using (18) avoids this problem as we we will see in the next section.

We next consider the IV cointegration tests in a system of equations. We rewrite the vector error correction model in (3) as

$$\Delta y_{1t} = (d_{11} + d_{12}t) + \delta_1 z_{t-1} + C_{11}(L)\Delta y_{1,t-1} + C_{12}'(L)\Delta y_{2,t-1} + e_{1t} \tag{19}$$

$$\Delta y_{2t} = (d_{21} + d_{22}t) + \delta_2 z_{t-1} + C_{21}(L)\Delta y_{1,t-1} + C_{22}'(L)\Delta y_{2,t-1} + e_{2t}, \tag{20}$$

or ADL version models as follows.

$$\Delta y_{1t} = (d_{11} + d_{12}t) + \delta_{11} y_{1,t-1} + \delta_{12}' y_{2,t-1} + C_{11}(L)\Delta y_{1,t-1} + C_{12}'(L)\Delta y_{2,t-1} + e_{1t} \tag{21}$$

$$\Delta y_{2t} = (d_{21} + d_{22}t) + \delta_{21} y_{1,t-1} + \delta_{22}' y_{2,t-1} + C_{21}(L)\Delta y_{1,t-1} + C_{22}'(L)\Delta y_{2,t-1} + e_{2t}. \tag{22}$$

Note that the contemporaneous terms Δy_{1t} or Δy_{2t} are omitted in the above system of equations. Here, we let $x = (x_{m+1}, .., x_T)'$ denote the main regressor whose coefficient we test on. That is, $x_t = z_{t-1}$ for the ECM tests, and $x_t = y_{1,t-1}$ or $x_t = (y_{1,t-1},\ y_{2,t-1})$ for the ADL tests, and w_t is the corresponding instrument for x_t. Then, we let q_t denote the remaining regressors in each equation in the system. $\widetilde{w}$ and $\widetilde{x}$ as the residuals from the regression of w_t or x_t on q_t. For example, we define $q_t = (1,\ t,\ \Delta y_{1,t-1},\ ..,\ \Delta y_{1,t-p+1},\ \Delta y_{2t}',\ \Delta y_{2,t-1}',\ ..,\ \Delta y_{2,t-p+1}')'$ for the ECM tests. Then, we have $\widetilde{w} = M_q w$ and $\widetilde{x} = M_q x$ with $M_q = I_{T-m} - q(q'q)^{-1}q'$, where $q = (q_{m+1}, .., q_T)'$. We can use the traditional 3SLS estimator.

$$\widehat{\theta} = (\widehat{X}'\widehat{\Omega}^{-1}\widehat{X})^{-1}\widehat{X}'\widehat{\Omega}^{-1}\Delta y \tag{23}$$

where Δy is a vector that stacks Δy_{1t} and $\Delta y_{2t}'$, and we let $\widehat{X} = P_w X$ with $X = I_p \otimes \widetilde{x}$, $P_w = W(W'W)^{-1}W'$, $W = I_p \otimes \widetilde{w}$; and $\widehat{\Omega} = \widehat{\Sigma} \otimes I_T$, which is obtained by using the error variance estimate $\widehat{\Sigma}$ in the system of equations of (19) and (20), or (21) and (22).

The null of no cointegration can be expressed as $H_0 : \delta_1 = \delta_2 = 0$ in (19) and (20) for the system IV-ECM test. We can use the usual Wald statistic for this hypothesis. We refer to the resulting test as W_{ECM}. Similarly, the null of no cointegration is given as $H_0 : \delta_{11} = \delta_{21} = 0$ in (21) and (22) for the system IV-ADL test, which we denote as W_{ADL}. Alternatively, following Boswijk (1994), we can consider the restriction on all level data under the null, $H_0 : \delta_{11} = \delta_{21} = 0$ and $\delta_{12} = \delta_{22} = 0$. The resulting Wald statistic is referred to as W_{ADL2}. We note that the 3SLS estimates can be obtained equivalently from the 3SLS GMM estimator or the IV estimates in each equation independently from other equations, given that common regressors (x_t, q_t) are used in each equation. However, the Wald statistic should incorporate information on the correlation structure of the error variance Σ. We consider the following Wald statistic

$$W_{ECM},\ W_{ADL},\ W_{ADL2} = vec(R\widehat{\theta})'[R(\widehat{\Sigma}^{-1} \otimes \widetilde{x}'\widetilde{w}(\widetilde{w}'\widetilde{w})^{-1}\widetilde{w}'\widetilde{x})R']^{-1}vec(R\widehat{\theta}) \tag{24}$$

where R is a selection matrix that chooses the relevant parameters under the null in the corresponding equations.

Corollary 2 *Suppose that y_t is generated as a vector time series as in (1), and the Wald-statistic is obtained by (24). Then under the null of no cointegration, as $T \to \infty$,*

$$W_{ECM},\ W_{ADL},\ W_{ADL2} \longrightarrow \chi_{d.f.} \tag{25}$$

where $d.f.$ is the number of restrictions under the null.

Proof: See the Appendix.

Another important question to examine is how our system based cointegration tests will be affected by including additional stationary covariates, say s_t, to the list of remaining regressors q_t in the above system of equations. Owing to the same logic that any stationary term in q_t, or the added term Δy_{2t}, does not affect the null distribution, we expect that adding s_t will not affect the null distribution of our IV tests in the system equation framework. We will examine this issue in the next section.

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 begin by examining single equation based tests. 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. All simulation results are based on 20,000 replications.

We wish to examine the size and power properties of our IV tests. One focus of our simulations is to examine the effect of including Δy_{2t}. We adopt the following data generating process (DGP) as in Kremers *et al.* (1992)

$$\begin{aligned} \Delta y_{1t} &= \phi' \Delta y_{2t} + \delta_1 (y_{1,t-1} - \beta' y_{2,t-1}) + v_t, \qquad (26) \\ \Delta y_{2t} &= u_t \\ \begin{pmatrix} v_t \\ u_t \end{pmatrix} &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix} \right). \end{aligned}$$

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 $\beta = 1$. Letting $\sigma_u^2 = var(u_t) = var(\Delta y_{2t})$, we examine the effect of including Δy_{2t} on the size and power properties by using the signal-noise ratio $q = -(\phi - 1)s$, where $s = \sigma_u / \sigma_v$. Thus, the effect of Δy_{2t} is captured by two components: the coefficient of Δy_{2t} (ϕ) and the relative variance of Δy_{2t} (s). Then, we examine the cases with (ϕ, s) = (1.0, 1), (0.5, 6), and (0.5, 16) such that $q = 0$, 3, and 8. Special attention will be given to the robustness of the standard normal result under different values of q. Following this, we also examine whether the tests are sensitive to different model specifications and the number (k) of integrated regressors . 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 1A, 1B, 2A, and 2B) and the other is the trend model with $d_t = c_1 + c_2 t$ (the results are reported in Tables 3A, 3B, 4A, and 4B). The instrumental variables are given in (12). 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.

4.1 Performance of the Single Equation Based Tests

In Table 1A, we report the size of various tests under the null of no cointegration with $k = 1$ and 3, when we include only 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 integrated regressors 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 it. 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 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 empirical applications, one can select m as the value that minimizes the sum of squared residuals or the determinant of the estimated error variances in the testing equations.

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 1B, we examine the size adjusted power of the tests for the model with drift 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} critically depends on the nuisance parameter of the signal-noise ratio as examined in Table 1A. Thus, we exclude this test from further discussion. In general, the ECM and ADL based IV tests are more powerful than the EG tests. As noted previously, a focus of our interest is to examine the effect of q under the alternative. The power of the EG tests (t_{EG} and t_{EG-O}) decreases as snr increases. Under the null, these tests do not depend on 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 EG testing regression. The modified EG test t_{EG}^{+} is not subject to this problem when Δy_{2t} is added as in (18). The power of the t_{EG}^{+} test increases as snr 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, adding Δy_{2t} induces the nuisance parameter problem. But, in the IV framework the null distribution remains as standard normal, implying absence of the nuisance parameter dependency problem, and the power increases when we add Δy_{2t}. The power of the ECM and ADL tests also 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 explained by Hansen (1995) who showed that power in the usual unit root tests increases when adding stationary covariates. In our framework, Δy_{2t} works the same as when including stationary covariates, and the gain in power is bigger when the variance of Δy_{2t} increases, or its coefficient increases, so that the signal-noise ratio increases.

This effect is enhanced as the dimension (k) of Δy_{2t} 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 normally 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 1B 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 2A and 2B, we replicate the same simulations as in Tables 1A and 1B, 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 3A, 3B, 4A and 4B, we report the results with the trend model. Note that the same asymptotic critical value of -1.645 is used again 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. Similar results will follow when structural breaks are allowed. Most important, the IV tests are invariant to the nuisance parameters. Moreover, our IV cointegration tests have the desired feature that they do not depend on the signal-noise ratio under the null, and their power increases as the signal-noise ratio increases under the alternative.

4.2 Comparison with Other Tests

Next, in order to illustrate the importance of using stationary covariates, we compare our tests to the standard EG and Johansen (1989) cointegration tests. Consider the following DGP for the variables y_{1t}, y_{2t} and y_{3t}

$$\begin{aligned} \Delta y_{1t} &= -\delta(y_{1,t-1} - y_{2,t-1}) + \beta y_{3t} + \varepsilon_{1t} \\ y_{2t} &= y_{2,t-1} + \varepsilon_{2t} \\ y_{3t} &= \rho y_{3,t-1} + \varepsilon_{3t} \end{aligned} \tag{27}$$

where $\varepsilon_{1t}, \varepsilon_{2t}$, and ε_{3t} are iid, $N(0, I)$ random variables (so that all are unit-variance and mutually uncorrelated), and the parameters are such that $|\rho| < 1$, $\beta \geq 0$, and $\delta \geq 0$. Notice that y_{2t} is a unit-root process and that y_{3t} is stationary. For $\delta > 0$, y_{1t} is cointegrated with y_{2t} and for $\beta > 0$, y_{3t} acts as a stationary covariate. The nature of our simulation experiment is to suppose that a researcher is trying to determine whether y_{1t} is cointegrated with y_{2t}. The researcher correctly surmises that y_{1t} and y_{2t} are unit-root processes and uses the OLS-based EG and Johansen (1989) tests ignoring the fact that y_{3t} is a stationary covariate. As such, the EG test (t_{EG-O}) is conducted by regressing y_{1t} on y_{2t} and performing the usual residual based test using an equation in the form of (10). The Johansen (1989) test is conducted by determining the rank of Π in in the system

$$\Delta y_t = \Pi y_{t-1} + \varepsilon_t \tag{28}$$

where $y_t = [y_{1t}, y_{2t}]'$ and $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}]'$. We parameterize the model using $\rho = (0.9, 0.95), \beta = (0, 1.0, 5.0)$, and $\delta = (0.0, 0.1, 0.5)$.

As such, there are two important cases:

1. $\beta = 0, \delta = 0:$ This is the situation in which there is no cointegration so that y_{1t}, y_{2t}, and y_{3t} are all independently distributed random variables. The Johansen test should indicate that $rank(\Pi) = 0$ and the EG should not reject the null hypothesis of no cointegration.

2. $\delta > 0$: In this situation y_{1t} and y_{2t} are cointegrated. The y_{1t} series is independent of y_{3t} if $\beta = 0$ but not if $\beta > 0$. Nevertheless, the Johansen test should indicate that $rank(\Pi) = 1$ and the EG should reject the null hypothesis of no cointegration.

We also use our three tests to check for cointegration. For the IV EG (t_{EG}) and IV ECM tests (t_{ECM}) , we regress y_{1t} on y_{2t} and save the residuals as $\widehat{z}_{t-1}$. We then estimate an equation in the form of (6) for the IV ECM test and **(9)** for the IV EG test using $w_t = \widehat{z}_{t-1} - \widehat{z}_{t-5}$ as a stationary instrument for $\widehat{z}_{t-1}$ in each case. For the IV ADL test (t_{ADL}) , we estimate an equation in the form of (7) and use $w_{1t} = y_{1,t-1} - y_{1,t-5}$ and $w_{2t} = y_{2,t-1} - y_{2,t-5}$ as instruments for $y_{1,t-1}$ and $y_{2,t-1}$, respectively. Since the tests are invariant to the number of $I(1)$ and $I(0)$ variables, we use critical values from a normal distribution.

The top portion of Table 5 reports the results of the various tests using 2000 simulated series with 100 observations for each of the parameter sets. Columns 4 and 6, respectively, report the relative frequency of instances where the $\lambda_{trace}(1)$ and $\lambda_{\max}(1)$ tests rejected the null of no cointegration.[1] Respectively, columns 8 through 11, report the relative frequency of instances in which the EG test (t_{EG_O}) , our IV ECM test (t_{ECM}), our IV ADL test (t_{ADL}), and our IV EG test (t_{EG}) rejected the null hypothesis of no cointegration. The entries in the table refer to the number of instances in which the null hypothesis was rejected using 5% critical values for each test.

First consider the results for the parameter values $\beta = \delta = 0$. Notice that the empirical size of the $\lambda_{trace}(1)$ test is about double the nominal size of 5% in that the null hypothesis of no cointegration was rejected in 211 (10.6%) of the 2000 replications . The more specific $\lambda_{\max}(1)$ performs better in that the null hypothesis of no cointegration was rejected in 141 (7.1%) of the 2000 replications. The empirical size of the EG and the IV version of the EG test are both close to 5%, while the IV ECM and IV ADL tests are about as oversized as the $\lambda_{trace}(1)$ test. There are several other important points to note about the results:

1. For $\beta = 0$ and $\delta > 0$, the power of the $\lambda_{trace}(1)$ and $\lambda_{\max}(1)$ tests always exceeds that of the EG test and our three IV tests. As expected, a properly specified full-information

[1] Our notation is such that if $rank(\Pi) = r$, $\lambda_{trace}(r)$ is the Johansen (1989) test with the null hypothesis $rank(\Pi) = r - 1$, and the alternative of $rank(\Pi) \geq r$. Similarly, $\lambda_{\max}(r)$ is the test with the null hypothesis $rank(\Pi) = r - 1$, and the alternative of $rank(\Pi) = r$. Note that we do not report $\lambda_{trace}(2)$ and $\lambda_{\max}(2)$ in the top half of Table 9 since the researcher is assumed to know that both variables are $I(1)$.

maximum likelihood approach works better than the two-step EG approach or the single-equation models.

2. For $\delta = 0$, the size properties of the Johansen test deteriorate as the importance of the y_{3t} in the DGP increases. For $\beta = 1$ and $\beta = 5$, the null of no cointegration is rejected by the $\lambda_{trace}(1)$ and $\lambda_{\max}(1)$ tests in more than half of the Monte Carlo trials. For example, when $\delta = 0, \rho = 0.90$, and $\beta = 1.0$, the null hypothesis of no cointegration is rejected in 1192 (59.6%) of the Monte Carlo trials. The addition of the stationary covariate y_{3t} into the DGP leads the Johansen test to incorrectly conclude that a linear combination of y_{1t} and y_{2t} is stationary. In contrast, when $\delta = 0$, and $\beta > 0$, the EG and all three IV tests are all undersized in that the null hypothesis of no cointegration is correctly accepted in almost all instances.

3. For $\delta > 0$ and $\beta > 0$,the IV ECM and the IV ADL tests have better power than the Johansen test and the EG test. The EG test usually has better power than the IV EG test. Although the presence of the stationary covariate diminishes the power of the Johansen test, it increases the power of the IV ECM and the IV ADL tests. For example, for $\delta = 0.1$ and $\rho = 0.9$, the $\lambda_{trace}(1)$ test finds cointegration in 521 trials (26.1%) whereas the IV ECM and ADL tests find cointegration in 1013 (50.7%) and 1723 (86.2%) of the trials, respectively.

For our purposes, the key point is that our IV tests (particularly the ECM and ADL versions of the tests) can have better size and power than the Johansen test. In the absence of any cointegration, the presence of a stationary covariate in the DGP, but not in the testing equations, creates size problems for the Johansen test that are not present in our tests. When there is a cointegrating relationship, the presence of a stationary covariate in the DGP, but not in the testing equations, creates a loss of power for the Johansen test. We are also encouraged by the fact that our tests work about as well as the Johansen test when $\beta = \delta = 0$ and better than the Johansen test when $\delta = 0$ and $\beta > 0$. The inclusion of the stationary covariate, even when $\beta = 0$, does not create any serious size problems for our tests relative to the Johansen test.

We now consider a related experiment such that the researcher incorrectly believes that y_{3t} is an $I(1)$ variable that belongs in the cointegrating relationship. The EG test is conducted by regressing y_{1t} on y_{2t} and y_{3t} and performing the test using (10). The Johansen test is conducted by determining the rank of Π in (28) where $y_t = [y_{1t}, y_{2t}, y_{3t}]'$ and $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}]'$. Due to the extra variable, the important cases to consider are

1. $\beta = 0$, $\delta = 0$: This is the situation in which there is no cointegration so that y_{1t}, y_{2t} and y_{3t} are all independently distributed random variables. Since y_{3t} is stationary, the Johansen test should find $rank(\Pi) = 1$ due to the trivial cointegrating vector indicating that y_{3t} ~ $I(0)$.The EG should not reject the null hypothesis of no cointegration.

2. $\delta > 0$: In this situation y_{1t} and y_{2t} are cointegrated and y_{1t} is independent of y_{3t} if $\beta = 0$. The Johansen test should indicate that $rank(\Pi) = 2$ (since y_{1t} and y_{2t} are cointegrated and y_{3t} is stationary). The EG should reject the null hypothesis of no cointegration.

For comparability, our three IV tests are also conducted assuming that the researcher believes that y_{3t} is $I(1)$. For the IV EG and IV ECM tests, we regress y_{1t} on y_{2t} and y_{3t} and save the residuals as $\widehat{z}_{t-1}$. We then estimate an equation in the form of (6) for the IV ECM test and (18) for the IV EG test using $w_t = \widehat{z}_{t-1} - \widehat{z}_{t-5}$ as a stationary instrument for $\widehat{z}_{t-1}$ in each case. For the IV ADL test, we estimate an equation in the form of (7) and use $w_{1t} = y_{1,t-1} - y_{1,t-5}$, $w_{2t} = y_{2,t-1} - y_{2,t-5}$ and $w_{3t} = y_{3,t-1} - y_{3,t-5}$ as instruments for $y_{1,t-1}, y_{2,t-1}$ and $y_{3,t-1}$, respectively. All of the results are contained in the lower portion of Table 5.

For the parameter values $\beta = \delta = 0$, the $\lambda_{trace}(1)$ test is most likely to find a cointegration relationship followed by the IV ECM test. For example, for $\rho = 0.9$, the $\lambda_{trace}(1)$ test properly indicates $rank(\Pi) = 1$ in 324 instances (16.2%) while the IV ECM test is correct in only 209 instances (10.5%). The other key results are

1. For $\beta = 0$ and $\delta > 0$, the $\lambda_{trace}(1)$ and $\lambda_{\max}(1)$ test do quite well in indicating the presence of a cointegrating vector. However, the $\lambda_{trace}(2)$ and $\lambda_{\max}(2)$ tests are only indicate the presence of a second cointegrating vector moderately well. For example, when $\delta = 0.1$ and $\rho = 0.90$, the $\lambda_{trace}(2)$ test finds $rank(\Pi) = 1$ in only 362 (18.1%) of the Monte Carlo trials. For this same set of parameter values, the IV ECM and IV ADL tests find cointegration in 596 (29.8%) and 582 trials (29.1%), respectively.

2. For $\delta = 0$ and $\beta > 0$, the $\lambda_{\max}(1)$ and $\lambda_{trace}(1)$ tests correctly reject the null hypothesis of no cointegration in all 2000 Monte Carlo trials. The EG, IV EG, and IV ADL tests find little evidence of cointegration while the IV ECM test finds a cointegrating relationship in slightly more than 5% of the trials.

3. For $\delta > 0$ and $\beta > 0$, the Johansen test should indicate two cointegrating vectors. However, an examination of the table shows that the various versions of the Johansen test generally indicate the presence of only one cointegrating vector. The EG test performs very poorly for $\delta = 0.1$ but works well for $\delta = 0.5$. While the presence of the stationary covariate diminishes the power of the Johansen test, it increases the power of the IV ECM and the IV ADL tests. The three IV tests all have better power than the $\lambda_{trace}(2)$ and $\lambda_{\max}(2)$ tests.

The importance of these results is that price of improperly estimating the order of integration of y_{3t} is small in our IV tests. Although the Johansen (1982) test almost always rejects the null of no cointegration, it has trouble finding the two cointegrating relationships when y_{3t} is stationary.

4.3 System Based IV Tests and the Effects of Stationary Covariates

To examine the performance of the system based IV cointegration tests (W_{ECM}, W_{ADL}, and W_{ADL2}), we perform Monte Carlo experiments using four different data generating process (DGP). DGP1 is the same as the DGP in (27), which adds a stationary covariate variable y_{3t} with $\rho < 1$. This experiment is designed to examine the effect of including stationary covariates on the size and power of the system based IV cointegration tests. As in the single equation tests, we wish to confirm whether the system based IV tests also remain robust to adding y_{3t} under the null, while their power increases as the signal-noise ratio increases. In the next three DGPs, we examine the size properties of the system based IV cointegration tests under different model specifications. In DGP2, we allow for non-zero autocovariances in the error variance with $\Sigma = \begin{bmatrix} 1.0 & -0.3 \\ -0.3 & 1.0 \end{bmatrix}$ in (3). DGP3 and DGP4 adopt the same DGP considered in Li and Lee (2006), which are based on the following equations

$$\begin{aligned} \Delta y_{1t} &= \phi_1 \Delta y_{1,t-1} + \phi_2 \Delta y_{2,t-1} + \varepsilon_{1t} \\ \Delta y_{2t} &= \phi_3 \Delta y_{1,t-1} + \phi_4 \Delta y_{2,t-1} + \varepsilon_{2t}. \end{aligned} \tag{29}$$

DGP3 employs the parameter values $(\phi_1, \phi_2, \phi_3, \phi_4)$ = (-0.2, 0.0, -0.1, -0.2), while DGP4 uses the values $(\phi_1, \phi_2, \phi_3, \phi_4)$ = (-0.2, -0.1, -0.1, -0.2). These experiments are designed to examine whether the IV cointegration tests are robust to non-zero long-run correlations in a long-run covariance matrix; that is, $\rho^{+2} \equiv \Omega_{11}^{-1}\Omega_{12}\Omega_{21}\Omega_{22}^{-1} \neq 0$, where the partial sum process of $(\varepsilon_{1t}, \varepsilon_{2t})'$ converges to a vector Brownian process with the long-run variance matrix having $\Omega_{ij}, i, j = 1, 2$, as its components.

In Table 6, we report the size and power of the system based IV cointegration tests when the values of β and ρ vary. It is clear that when $\delta = 0.0$, the sizes of all three IV tests do not change much for different values of β and ρ, and the results with $m = 3$ or 5 are quite similar. (To spare space, we do not report the results using other values of m.) Thus, we can say that all three system based IV tests are fairly invariant to the nuisance parameter under the null. As in the single equation based tests, we observe that the power of all three tests increases as the signal-noise ratio increases (i.e., as either β or ρ increases). When the sample size increases, the power increases accordingly as expected. Lastly, we examine the sensitivity of the system based IV cointegration tests in different model specifications. In Table 7, we document the size properties under the different DGPs as described above. The simulation results show that all three tests are mostly robust to the non-zero error structure (DGP2) and long-run endogeneity (DGP3 and DGP4). The size of the tests are close to their nominal sizes in all cases. We also considered different lag structures and different values of m under the same simulation designs, but these results are self explanatory. Overall, the properties of the system based tests are quite similar to those of the single equation based tests.

5 The Demand for Narrow Money in the U.K.

We now illustrate the appropriate use of our IV test using an extended example from the literature. To avoid the appearance of selecting an arbitrary set of variables, we use 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.[2] 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

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

demand function

$$\begin{bmatrix} 1.00 & -1.00 & -7.34 & 7.65 & -0.0005 \\ -0.06 & 1.00 & -3.38 & 0.86 & -0.0059 \\ -0.29 & 0.69 & 1.00 & -0.63 & -0.0025 \\ 0.03 & -1.58 & 1.10 & 1.00 & 0.0097 \end{bmatrix} \begin{bmatrix} (m-p)_t \\ i_t \\ \Delta p_t \\ R_t \\ t \end{bmatrix}, \qquad (30)$$

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[3]

$$\begin{aligned} \Delta(m-p)_t \quad = \quad & -0.088(m-p)_{t-1} - 0.696\Delta p_{t-1} - 0.611R_t \\ & +0.095i_{t-1} - 0.174\Delta(m-p-i)_{t-1} - 0.0498. \end{aligned} \qquad (31)$$

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

[3]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.

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, it is unclear as to whether the inflation rate (Δp_t) acts as a unit root process since 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 $\kappa_c(3)$ or $\kappa_{ctt}(3)$ 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 (31) 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

$$\begin{aligned} \Delta(m-p)_t &= -0.103(m-p)_{t-1} - 0.684\Delta p_{t-1} - 0.671R_t \\ &\quad +0.100i_{t-1} - 0.1914\Delta(m-p-i)_{t-1} - 0.053. \end{aligned} \tag{32}$$

Notice that the point estimates of the coefficients are all quite similar to that of (31). 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 (31). Consider

$$\begin{aligned}\Delta(m-p)_t &= -0.090(m-p)_{t-1} - 0.700\Delta p_{t-1} - 0.619R_t \\ &\quad +0.095i_{t-1} - 0.176\Delta(m-p-i)_{t-1} - 0.029.\end{aligned} \tag{33}$$

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 (31) 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 quite 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

$$\begin{aligned}\Delta(m-p)_t &= -0.083(m-p)_{t-1} - 0.700\Delta p_{t-1} - 0.618R_t \\ &\quad +0.078i_{t-1} - 0.158\Delta(m-p-i)_{t-1} + 0.083.\end{aligned} \tag{34}$$

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 propose a new set of cointegration tests based on stationary instrumental variables. Unlike the usual cointegration tests, the asymptotic distributions of the IV cointegration tests are free of nuisance parameters and have standard normal or chi-square distributions. This is a key feature of the IV cointegration tests. The appropriate critical values do not depend on the deterministic terms, the number of $I(1)$ variables, or the presence of stationary covariates. This is important because simulating different critical values corresponding to unknown nuisance parameter values seems cumbersome or impossible in some cases, and bootstrapping may not work well using potentially nonergodic variables. Moreover, there is no need to bootstrap critical values even when stationary covariates are included. Our simulations illustrate this point by showing that the IV tests can have better size and power properties than the Johansen (1989) test when the researcher is unsure about the behavior of stationary covariates.

While we have examined standard cases of single equation cointegration models, the same idea can be usefully 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. Im and Lee (2006) have shown that the IV approach can be utilized for unit root tests with structural changes and the tests are free of nuisance parameters. Similarly, IV cointegration tests allowing for structural changes are also expected to be free of the nuisance parameters indicating the location of breaks. Based on our findings it also appears possible to improve the power of cointegration tests by additional stationary covariates or more instruments. In either case, the resulting cointegration tests will not entail nuisance parameters. Moreover, it seems reasonable to extend our cointegration tests to the panel framework. Finally, it appears promising to consider threshold cointegration tests using IV estimation as an extension of Enders and Siklos (2001). In this case, we expect that the statistic testing for asymmetric nonlinearity have standard distributions regardless of whether the data are cointegrated or not. This topic is pursued elsewhere.

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] Banerjee, 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] Harris R.I.D.1; Judge G, (1998), "Small Sample Testing for Cointegration Using the Bootstrap Approach," *Economics Letters*, 58, 31-37.

[12] Im, K., and J. Lee (2007), "The asymptotic Distribution of the Coefficient Estimates in IV Unit Root and Cointegration Tests Using Stationary Instrumental Variables," manuscript.

[13] Im, K., and J. Lee (2006), "Testing for Unit Roots with Stationary Instrumental Variables," manuscript.

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

[15] 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.

[16] Li, J. and J. Lee (2006), "Single-equation ADL Tests for Threshold Cointegration," manuscript.

[17] Ng, S. and P. Perron (1995), "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association* 90, 269-281.

[18] Pesaran, M.H., Y. Shin and R.J. Smith (2001), Bounds testing approaches to the analysis of level relationships, *Journal of Applied Econometrics* 16, 289–326.

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

[20] Wohar, M and D. Rapach (2002), "Testing the Monetary Model of Exchange Rate Determination: New Evidence from a Century of Data," *Journal of International Economics*, 58, 359-385.

[21] 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 1. Assume that $\{\varepsilon_t\}_{t=1}^{\infty}$, is an *iid* process with mean zero, variance σ^2, and $E|\varepsilon_t|^{2k} < \infty$, for some $k \geqslant 2$. 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,

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

$$T^{-1}\sum_{t=1}^{T} \xi_t^2 \xrightarrow{p} m\sigma^2 \quad \text{(A2)}$$

$$T^{-1/2}\sum_{t=1}^{T} \xi_t\varepsilon_t \xrightarrow{d} \sqrt{m}\sigma^2 W(1) \quad \text{(A3)}$$

where $W(1) = Z$.

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

Lemma 2. Suppose that a vector of a time series process $\varepsilon_t = (\varepsilon_{it}, .., \varepsilon_{pt})'$, $t = 1, .., \infty$, is a stationary *iid* $(0, \Sigma)$ sequence of random variables, where $\varepsilon_{it}, i = 1, .., p$, satisfies the condition, $E|\varepsilon_{it}|^{2k} < \infty$, for some $k \geqslant 2$. Similarly, we let $\xi_t = (\xi_{it}, .., \xi_{pt})'$, with $\xi_{it} = \varepsilon_{i,t-1} + ...+ \varepsilon_{i,t-m}$. Then,

$$T^{-1}\sum_{t=1}^{T} \xi_t\xi_t' \xrightarrow{p} m\Sigma \quad \text{(A4)}$$

$$T^{-1/2}\left[\sum_{t=1}^{T}\xi_{1t}\varepsilon_{1t}, .., \sum_{t=1}^{T}\xi_{pt}\varepsilon_{pt}\right]' \xrightarrow{d} \sqrt{m}\Sigma W_p \quad \text{(A5)}$$

where W_p is a p-variate normal random variable.

Proof: (A4) is a multivariate extension of (A2) when ξ_t is a vector, $p \succeq 1$. For this, we use $E(\varepsilon_{it}\varepsilon_{is}) = \Sigma_{ij}$ if $t = s$ and $E(\varepsilon_{it}\varepsilon_{is}) = 0$ if $t \neq s$, for $i, j = 1, .., p$. (A5) is extended from (A3) where a vector $\xi_t = (\xi_{it}, .., \xi_{pt})'$ is used. The result can be shown by noting

$$\begin{aligned} E\left[T^{-1/2}\sum_{t=1}^{T}[\xi_{1t}\varepsilon_{1t}, .., \xi_{pt}\varepsilon_{pt}]'\right] &= \mathbf{0}_{p\times 1} \\ var\left[T^{-1/2}\sum_{t=1}^{T}[\xi_{1t}\varepsilon_{1t}, .., \xi_{pt}\varepsilon_{pt}]'\right] &= m^2\Sigma\Sigma'. \end{aligned}$$

Proof of Theorem 1

We first prove the standard normality result of the t-statistic of the single equation based IV cointegration tests. Let

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

$$C_T = \sum_{t=1}^{T} w_t^2 - \sum_{t=1}^{T} w_t q_t' \left(\sum_{t=1}^{T} q_t q_t' \right)^{-1} \sum_{t=1}^{T} q_t w_t. \tag{A7}$$

We first examine the distribution of $\frac{1}{\sqrt{T}} B_T$. We rewrite (A6) as

$$\frac{1}{\sqrt{T}} B_T = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t \Delta y_{1t} - \frac{1}{\sqrt{T}} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t q_t' \left(\frac{1}{T} \sum_{t=1}^{T} q_t q_t' \right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} q_t \Delta y_{1t}. \tag{A8}$$

We note that the first term on the right hand side of (A8) is expressed as the product of stationary terms. First, we can follow Zivot (2000) and Hansen (1995), and express the conditional error correction model (6) as

$$a(L)\Delta \alpha' y_t = \delta_1 \alpha' y_{t-1} + b(L)' \Delta y_{2t} + v_t = \delta_1 \alpha' y_{t-1} + e_t,$$

where $a(L) = 1 - C_{11}(L)L$, $e_t = b(L)'\Delta y_{2t} + v_t$, and $b(L) = (\phi' - \beta') + [C_{12}(L) + C_{11}(L)\beta]L$. Then, for the ECM test, under the null hypothesis ($\delta_1 = 0$), we have $\Delta y_{1t} = a(L)^{-1}(\pi_2' q_{2t} + v_t) = \psi(L)\Delta y_{2t} + v_t^* \equiv e_t^*$. We can also show from the above equation that $\Delta z_t = a(L)^{-1}(b(L)\Delta y_{2t} + v_t) = \psi^*(L)\Delta y_{2t} + v_t^*$ under the null hypothesis . Note that $\psi(L)$ is not distinguishable from $\psi^*(L)$. Then, letting $a(L) = a(1) + a^*(L)(1-L)$, we can have $\Delta z_t \approx e_t^*$ under the null. Then, we can have $\Sigma_{t=1}^{T}(e_{t-1}^* + .. + e_{t-m}^*)e_t = \Sigma_{t=1}^{T}(e_{t-1} + .. + e_{t-m})e_t$ as past values of Δy_{1t} and Δy_{2t} are uncorrelated with e_t. Similarly, for the ADL test, we have $w_t = (\Delta y_{1,t-1} + .. + \Delta y_{1,t-m}, \Delta y_{2,t-1} + .. + \Delta y_{2,t-m}) = (e_{t-1}^* + .. + e_{t-m}^*, e_{2,t-1} + .. + e_{2,t-m})$. Moreover, for the EG test, we have $w_t = e_{t-1}^* + .. + e_{t-m}^*$ from (10) and (11), since the null distribution is unaffected by the parameters in the term q_t. Therefore, utilizing the results in (A2), we have under the null hypothesis for each of the IV ECM, ADL and EG tests

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t \Delta y_{1t} \to \sqrt{m} a(1)^{-1} \sigma^2 W(1). \tag{A9}$$

We now show that the second term on the right side of (A8) 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 (A9). 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 $\frac{1}{\sqrt{T}} \Sigma_{t=1}^{T} q_t \Delta y_{1t} = O_p(1)$. Then, the second term on the right hand side of (A8) 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 (A8) is still $O_p(T^{-1/2})$ when additional deterministic terms are included. Therefore, we have

$$\frac{1}{\sqrt{T}} B_T \rightarrow \sqrt{m} a(1)^{-1} \sigma^2 W(1). \tag{A10}$$

We next examine the distribution of $\frac{1}{T} C_T$. We rewrite (A7) as

$$\frac{1}{T} C_T = \frac{1}{T} \sum_{t=1}^{T} w_t^2 - \frac{1}{T} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t q_t' \left(\frac{1}{T} \sum_{t=1}^{T} q_t q_t' \right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} q_t w_t. \tag{A11}$$

Again, we can show that the second term on the right hand side of (A11) is degenerate. It is clear that following (A3) we have under the null hypothesis

$$\frac{1}{T} \sum_{t=1}^{T} w_t^2 \rightarrow m a(1)^{-2} \sigma^2. \tag{A12}$$

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 (A11) is $O_p(T^{-1})$. This result holds when additional deterministic terms are added to q_t. Thus, we have

$$\frac{1}{T} C_T \rightarrow m a(1)^{-2} \sigma^2. \tag{A13}$$

Therefore, combining the results in (A10) and (A13), we can show that

$$t_{ECM},\ t_{ADL},\ t_{EG} = \frac{\frac{1}{\sqrt{T}} B_T}{\hat{\sigma} \sqrt{\frac{1}{T} C_T}} \xrightarrow{d} \frac{\sqrt{m} a(1)^{-1} \sigma^2 W(1)}{\sigma \sqrt{a(1)^{-2} m \sigma^2}} = W(1) \sim N(0,1). \tag{A14}$$

To consider a general case with a system of p equations, we let $\eta = (\eta_1,\ ..,\ \eta_p)'$ with $\eta_i = T^{-1/2} \Sigma_{t=1}^{T} \lambda' V_i^{-1/2} \xi_{it} \varepsilon_{it}$, and $V = var(T^{-1/2} \xi_t' \varepsilon_t)$, where λ is a vector satisfying $\lambda' \lambda = 1$. Following White (1999, Chapter 5), we have

$$E(\lambda' V^{-1/2} \xi_t \varepsilon_t | F_{t-1}) = 0 \tag{A15}$$

$$T^{-1} \sum_{t=1}^{T} [\lambda' V_i^{-1/2} \xi_{it}' \varepsilon_{it} \varepsilon_{it}' \xi_{it} V_i^{-1/2} \lambda - \lambda' V_i^{-1/2} V_i V_i^{-1/2} \lambda] \rightarrow 0 \tag{A16}$$

where $\{F_{t-1},\ \xi_t\}$ is an adapted stochastic sequence. Therefore, we have $\eta_i \,\tilde{}\, N(0,1)$ by the martingale central limit theorem, when we utilize the result $V_i = m\sigma_i^4$ as given in (A3) of Lemma

1. For the system based Wald statistics and the Wald statistics for the joint restrictions in a single equation, we use $V = (\sqrt{m}\Sigma)(\sqrt{m}\Sigma)'$ and $\eta = (\eta_1, .., \eta_p)'$ to show that the same result in (A16) holds in the multivariate normal distribution. Specifically, for the system based Wald tests, we use a properly defined selection matrix R and $\widetilde{w}$ for each of the Wald type tests, as they are defined to describe the relevant hypothesis using (7). Then, we rewrite (7) as

$$\begin{aligned} & W_{ECM},\ W_{ADL},\ W_{ADL2} \\ = \ & vec[(\widetilde{w}'\widetilde{x})^{-1}\widetilde{w}'u^*]'[\widehat{\Sigma}^{-1} \otimes \widetilde{x}'\widetilde{w}(\widetilde{w}'\widetilde{w})^{-1}\widetilde{w}'\widetilde{x}]^{-1}vec[(\widetilde{w}'\widetilde{x})^{-1}\widetilde{w}'u^*] \\ = \ & tr[(\widetilde{w}'u^*\widehat{\Sigma}^{-1/2})'(\widetilde{w}'\widetilde{w})^{-1}(\widetilde{w}'u^*\widehat{\Sigma}^{-1/2})] \end{aligned} \tag{A17}$$

where $u^* = (u_1, .., u_p)$. Then, in light of (A4) and (A5), it is clear that each of our Wald statistics has a chi-square distribution with degree of freedom equal to the number of restrictions. For the $ADL2^*$ test in a single equation on the joint restriction, $\delta_1 = \gamma = 0$ in (7), we can rewrite (24) as

$$ADL2^* = \widehat{\theta_1}'[\widehat{\Sigma}^{-1} \otimes \widetilde{x}'\widetilde{w}(\widetilde{w}'\widetilde{w})^{-1}\widetilde{w}'\widetilde{x}]^{-1}\widehat{\theta_1} = tr[(\widetilde{w}'u_1\widehat{\Sigma}^{-1/2})'(\widetilde{w}'\widetilde{w})^{-1}(\widetilde{w}'u_1\widehat{\Sigma}^{-1/2})] \tag{A18}$$

where $\widehat{\theta_1} = (\widetilde{w}'\widetilde{x})^{-1}\widetilde{w}'\Delta y_1$ and $\theta_1 = (\delta_1,\ \gamma')'$ is the coefficient of $(y_{1,t-1},\ y'_{2,t-1})'$. This is a special case of (A17), and it is clear that $ADL2^*$ has a chi-square distribution.

Table 1A Size of the Single Equation Tests Under the Null (δ = 0, T = 100; Drift Model)

		k = 1							k = 3						
q	m	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}
0	1	0.049	0.050	0.051	0.014	0.016	0.037	0.030	0.050	0.049	0.050	0.014	0.018	0.106	0.065
	2				0.030	0.035	0.065	0.049				0.031	0.042	0.161	0.093
	3				0.042	0.049	0.083	0.064				0.044	0.066	0.204	0.118
	4				0.053	0.060	0.096	0.072				0.055	0.088	0.238	0.135
	5				0.060	0.074	0.106	0.080				0.062	0.108	0.269	0.149
	6				0.069	0.084	0.118	0.088				0.071	0.126	0.291	0.164
	7				0.074	0.096	0.124	0.092				0.075	0.145	0.316	0.175
	8				0.081	0.102	0.131	0.096				0.081	0.158	0.336	0.185
	9				0.088	0.112	0.135	0.100				0.088	0.173	0.355	0.193
3	1	0.010	0.050	0.051	0.007	0.016	0.037	0.030	0.005	0.049	0.050	0.005	0.018	0.106	0.065
	2				0.019	0.035	0.065	0.049				0.016	0.042	0.161	0.093
	3				0.028	0.049	0.083	0.064				0.024	0.066	0.204	0.118
	4				0.034	0.060	0.096	0.072				0.029	0.088	0.238	0.135
	5				0.037	0.074	0.106	0.080				0.033	0.108	0.269	0.149
	6				0.040	0.084	0.118	0.088				0.037	0.126	0.291	0.164
	7				0.042	0.096	0.124	0.092				0.040	0.145	0.316	0.175
	8				0.045	0.102	0.131	0.096				0.040	0.158	0.336	0.185
	9				0.047	0.112	0.135	0.100				0.042	0.173	0.355	0.193
8	1	0.005	0.050	0.051	0.006	0.016	0.037	0.030	0.003	0.049	0.050	0.004	0.018	0.106	0.065
	2				0.015	0.035	0.065	0.049				0.014	0.042	0.161	0.093
	3				0.024	0.049	0.083	0.064				0.022	0.066	0.204	0.118
	4				0.028	0.060	0.096	0.072				0.027	0.088	0.238	0.135
	5				0.030	0.074	0.106	0.080				0.029	0.108	0.269	0.149
	6				0.033	0.084	0.118	0.088				0.032	0.126	0.291	0.164
	7				0.036	0.096	0.124	0.092				0.036	0.145	0.316	0.175
	8				0.039	0.102	0.131	0.096				0.035	0.158	0.336	0.185
	9				0.040	0.112	0.135	0.100				0.037	0.173	0.355	0.193

Note: For the OLS based tests, the 5% critical values were simulated. For T = 100, they are -2.881, -3.247 and -3.363 for $t_{ECM\text{-}O}$, $t_{ADL\text{-}O}$ and $t_{EG\text{-}O}$ tests, when k = 1. The corresponding critical values are -2.890, -3.807 and -4.175, when k = 3. For the IV tests, -1.645 was used for all cases.

Table 1B Size-adjusted Power Under the Alternative (δ = -0.1, T = 100; Drift Model)

		k = 1							k = 3						
q	m	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}
0	1	0.319	0.244	0.220	0.205	0.159	0.156	0.167	0.305	0.182	0.140	0.196	0.110	0.098	0.121
	2				0.215	0.170	0.159	0.180				0.216	0.128	0.100	0.132
	3				0.228	0.184	0.171	0.193				0.218	0.136	0.105	0.146
	4				0.236	0.198	0.176	0.205				0.226	0.140	0.108	0.149
	5				0.250	0.199	0.178	0.216				0.238	0.142	0.109	0.163
	6				0.252	0.194	0.186	0.215				0.243	0.139	0.115	0.163
	7				0.257	0.202	0.197	0.226				0.247	0.138	0.118	0.167
	8				0.257	0.204	0.200	0.238				0.251	0.139	0.118	0.169
	9				0.255	0.205	0.204	0.238				0.250	0.144	0.120	0.174
3	1	1.000	0.954	0.066	0.657	0.449	0.092	0.261	1.000	0.962	0.036	0.899	0.505	0.050	0.159
	2				0.791	0.551	0.097	0.331				0.974	0.622	0.050	0.218
	3				0.869	0.617	0.099	0.385				0.992	0.672	0.052	0.260
	4				0.915	0.661	0.102	0.427				0.996	0.703	0.053	0.281
	5				0.940	0.687	0.100	0.455				0.997	0.713	0.050	0.307
	6				0.953	0.699	0.101	0.467				0.997	0.720	0.049	0.316
	7				0.964	0.712	0.101	0.485				0.998	0.727	0.049	0.326
	8				0.966	0.720	0.099	0.507				0.998	0.727	0.048	0.336
	9				0.970	0.724	0.099	0.513				0.997	0.727	0.049	0.343
8	1	1.000	1.000	0.051	0.988	0.808	0.082	0.469	1.000	1.000	0.033	0.999	0.820	0.047	0.226
	2				0.999	0.871	0.086	0.571				1.000	0.869	0.047	0.307
	3				1.000	0.898	0.087	0.608				1.000	0.886	0.049	0.355
	4				1.000	0.912	0.088	0.636				1.000	0.893	0.050	0.381
	5				1.000	0.915	0.087	0.650				1.000	0.897	0.047	0.402
	6				1.000	0.918	0.088	0.659				1.000	0.897	0.046	0.415
	7				0.999	0.923	0.090	0.671				1.000	0.899	0.046	0.426
	8				1.000	0.925	0.085	0.681				1.000	0.900	0.045	0.438
	9				0.999	0.925	0.086	0.688				1.000	0.896	0.046	0.445

Table 2A Size of the Single Equation Tests Under the Null ($\delta = 0$, $T = 300$; Drift Model)

		$k = 1$							$k = 3$						
q	m	$t_{\text{ECM-O}}$	$t_{\text{ADL-O}}$	$t_{\text{EG-O}}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}	$t_{\text{ECM-O}}$	$t_{\text{ADL-O}}$	$t_{\text{EG-O}}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}
0	1	0.052	0.050	0.047	0.008	0.009	0.022	0.019	0.053	0.046	0.048	0.009	0.010	0.060	0.042
	2				0.021	0.022	0.041	0.033				0.022	0.021	0.094	0.061
	3				0.028	0.031	0.052	0.044				0.030	0.033	0.117	0.075
	4				0.036	0.039	0.060	0.050				0.036	0.044	0.130	0.085
	5				0.042	0.044	0.067	0.056				0.044	0.054	0.145	0.091
	6				0.048	0.051	0.074	0.060				0.049	0.064	0.160	0.099
	7				0.051	0.055	0.077	0.062				0.054	0.076	0.173	0.106
	8				0.055	0.063	0.084	0.066				0.059	0.081	0.187	0.114
	9				0.062	0.071	0.087	0.069				0.060	0.087	0.200	0.122
3	1	0.006	0.050	0.047	0.005	0.009	0.022	0.019	0.006	0.046	0.048	0.004	0.010	0.060	0.042
	2				0.016	0.022	0.041	0.033				0.013	0.021	0.094	0.061
	3				0.022	0.031	0.052	0.044				0.019	0.033	0.117	0.075
	4				0.028	0.039	0.060	0.050				0.024	0.044	0.130	0.085
	5				0.031	0.044	0.067	0.056				0.026	0.054	0.145	0.091
	6				0.033	0.051	0.074	0.060				0.030	0.064	0.160	0.099
	7				0.035	0.055	0.077	0.062				0.032	0.076	0.173	0.106
	8				0.037	0.063	0.084	0.066				0.034	0.081	0.187	0.114
	9				0.037	0.071	0.087	0.069				0.035	0.087	0.200	0.122
8	1	0.002	0.050	0.047	0.004	0.009	0.022	0.019	0.004	0.046	0.048	0.003	0.010	0.060	0.042
	2				0.014	0.022	0.041	0.033				0.012	0.021	0.094	0.061
	3				0.020	0.031	0.052	0.044				0.018	0.033	0.117	0.075
	4				0.025	0.039	0.060	0.050				0.023	0.044	0.130	0.085
	5				0.027	0.044	0.067	0.056				0.024	0.054	0.145	0.091
	6				0.028	0.051	0.074	0.060				0.029	0.064	0.160	0.099
	7				0.032	0.055	0.077	0.062				0.032	0.076	0.173	0.106
	8				0.033	0.063	0.084	0.066				0.031	0.081	0.187	0.114
	9				0.032	0.071	0.087	0.069				0.032	0.087	0.200	0.122

Note: For the OLS based tests, the 5% critical values were simulated. For $T = 300$, they are are -2.872, -3.237 and -3.355 for $t_{\text{ECM-O}}$, $t_{\text{ADL-O}}$ and $t_{\text{EG-O}}$ tests, when $k = 1$. The corresponding critical values are -2.858, -3.772 and -4.107, when $k = 3$. For the IV tests, -1.645 was used for all cases.

Table 2B Size-adjusted Power Under the Alternative (δ = -0.1, T = 300; Drift Model)

		$k = 1$							$k = 3$						
q	m	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}
0	1	0.996	0.974	0.968	0.369	0.281	0.312	0.321	0.995	0.892	0.809	0.368	0.221	0.218	0.249
	2				0.439	0.363	0.367	0.395				0.437	0.297	0.265	0.316
	3				0.507	0.434	0.434	0.457				0.506	0.348	0.306	0.370
	4				0.559	0.485	0.486	0.520				0.543	0.396	0.348	0.420
	5				0.608	0.537	0.525	0.569				0.590	0.427	0.387	0.468
	6				0.642	0.569	0.575	0.609				0.626	0.463	0.402	0.500
	7				0.680	0.605	0.617	0.657				0.659	0.485	0.429	0.537
	8				0.706	0.616	0.644	0.684				0.684	0.499	0.467	0.554
	9				0.723	0.636	0.661	0.699				0.715	0.510	0.458	0.578
3	1	1.000	1.000	0.867	0.941	0.794	0.281	0.806	1.000	1.000	0.460	0.999	0.839	0.182	0.843
	2				0.993	0.885	0.344	0.928				1.000	0.892	0.214	0.926
	3				0.999	0.923	0.393	0.964				1.000	0.913	0.235	0.951
	4				1.000	0.940	0.436	0.980				1.000	0.923	0.275	0.968
	5				1.000	0.948	0.471	0.988				1.000	0.929	0.305	0.974
	6				1.000	0.952	0.515	0.992				1.000	0.932	0.312	0.979
	7				1.000	0.955	0.545	0.994				1.000	0.935	0.319	0.984
	8				1.000	0.957	0.568	0.995				1.000	0.934	0.345	0.985
	9				1.000	0.959	0.593	0.997				1.000	0.937	0.341	0.987
8	1	1.000	1.000	0.849	1.000	0.949	0.279	0.985	1.000	1.000	0.443	1.000	0.932	0.178	0.945
	2				1.000	0.964	0.342	0.993				1.000	0.948	0.208	0.967
	3				1.000	0.971	0.390	0.996				1.000	0.956	0.232	0.978
	4				1.000	0.974	0.430	0.997				1.000	0.962	0.270	0.983
	5				1.000	0.977	0.467	0.998				1.000	0.963	0.303	0.985
	6				1.000	0.978	0.512	0.999				1.000	0.964	0.309	0.988
	7				1.000	0.979	0.538	0.999				1.000	0.968	0.317	0.991
	8				1.000	0.978	0.563	0.999				1.000	0.966	0.344	0.993
	9				1.000	0.980	0.583	1.000				1.000	0.966	0.340	0.993

Table 3A Size of the Single Equation Tests Under the Null ($\delta = 0$, $T = 100$; Trend Model)

		$k = 1$							$k = 3$						
q	m	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}
0	1	0.052	0.050	0.049	0.043	0.044	0.068	0.056	0.050	0.049	0.049	0.042	0.044	0.142	0.092
	2				0.070	0.076	0.100	0.079				0.071	0.094	0.210	0.128
	3				0.092	0.107	0.127	0.099				0.095	0.138	0.267	0.161
	4				0.110	0.131	0.145	0.114				0.112	0.179	0.315	0.183
	5				0.128	0.152	0.161	0.128				0.130	0.217	0.354	0.207
	6				0.142	0.173	0.179	0.138				0.143	0.249	0.384	0.228
	7				0.156	0.195	0.191	0.149				0.156	0.279	0.414	0.246
	8				0.171	0.216	0.204	0.156				0.170	0.306	0.442	0.268
	9				0.185	0.234	0.215	0.166				0.184	0.333	0.463	0.283
3	1	0.003	0.050	0.049	0.019	0.044	0.068	0.056	0.002	0.049	0.049	0.014	0.044	0.142	0.092
	2				0.036	0.076	0.100	0.079				0.030	0.094	0.210	0.128
	3				0.048	0.107	0.127	0.099				0.041	0.138	0.267	0.161
	4				0.057	0.131	0.145	0.114				0.048	0.179	0.315	0.183
	5				0.060	0.152	0.161	0.128				0.050	0.217	0.354	0.207
	6				0.062	0.173	0.179	0.138				0.056	0.249	0.384	0.228
	7				0.065	0.195	0.191	0.149				0.059	0.279	0.414	0.246
	8				0.071	0.216	0.204	0.156				0.059	0.306	0.442	0.268
	9				0.074	0.234	0.215	0.166				0.058	0.333	0.463	0.283
8	1	0.001	0.050	0.049	0.014	0.044	0.068	0.056	0.001	0.049	0.049	0.012	0.044	0.142	0.092
	2				0.030	0.076	0.100	0.079				0.027	0.094	0.210	0.128
	3				0.038	0.107	0.127	0.099				0.036	0.138	0.267	0.161
	4				0.045	0.131	0.145	0.114				0.041	0.179	0.315	0.183
	5				0.047	0.152	0.161	0.128				0.044	0.217	0.354	0.207
	6				0.047	0.173	0.179	0.138				0.047	0.249	0.384	0.228
	7				0.049	0.195	0.191	0.149				0.050	0.279	0.414	0.246
	8				0.052	0.216	0.204	0.156				0.049	0.306	0.442	0.268
	9				0.057	0.234	0.215	0.166				0.048	0.333	0.463	0.283

Note: For the OLS based tests, the 5% critical values were simulated. For $T = 100$, they are -3.453, -3.727 and -3.838 for $t_{ECM\text{-}O}$, $t_{ADL\text{-}O}$ and $t_{EG\text{-}O}$ tests, when $k = 1$. The corresponding critical values are -3.458, -4.184 and -4.530, when $k = 3$. For the IV tests, -1.645 was used for all cases.

Table 3B Size-adjusted Power Under the Alternative (δ = -0.1, T = 100; Trend Model)

		k = 1							k = 3						
q	m	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}
0	1	0.175	0.158	0.158	0.134	0.118	0.113	0.122	0.176	0.133	0.115	0.131	0.097	0.082	0.100
	2				0.137	0.128	0.117	0.132				0.135	0.110	0.088	0.114
	3				0.150	0.134	0.129	0.145				0.145	0.108	0.089	0.120
	4				0.152	0.138	0.137	0.155				0.146	0.114	0.094	0.128
	5				0.156	0.138	0.138	0.159				0.151	0.118	0.098	0.129
	6				0.154	0.136	0.139	0.157				0.147	0.115	0.097	0.133
	7				0.151	0.134	0.142	0.163				0.150	0.110	0.101	0.135
	8				0.154	0.139	0.147	0.166				0.151	0.115	0.106	0.139
	9				0.149	0.131	0.142	0.164				0.154	0.109	0.104	0.140
3	1	0.999	0.759	0.014	0.585	0.302	0.036	0.081	1.000	0.838	0.014	0.867	0.416	0.027	0.044
	2				0.735	0.390	0.038	0.098				0.964	0.532	0.027	0.058
	3				0.826	0.440	0.038	0.110				0.988	0.580	0.027	0.063
	4				0.876	0.480	0.038	0.124				0.993	0.616	0.026	0.070
	5				0.910	0.503	0.035	0.128				0.996	0.636	0.025	0.069
	6				0.928	0.517	0.033	0.126				0.997	0.633	0.023	0.071
	7				0.942	0.520	0.030	0.129				0.997	0.636	0.022	0.074
	8				0.947	0.534	0.029	0.133				0.998	0.638	0.022	0.075
	9				0.951	0.530	0.026	0.134				0.998	0.625	0.021	0.072
8	1	1.000	0.997	0.006	0.984	0.728	0.025	0.097	1.000	0.999	0.012	0.999	0.842	0.024	0.038
	2				0.998	0.829	0.025	0.134				1.000	0.898	0.024	0.064
	3				0.999	0.871	0.022	0.158				1.000	0.914	0.022	0.081
	4				0.999	0.891	0.024	0.178				1.000	0.921	0.023	0.088
	5				1.000	0.899	0.022	0.186				1.000	0.921	0.022	0.089
	6				0.999	0.904	0.020	0.189				1.000	0.921	0.020	0.093
	7				0.999	0.907	0.020	0.195				1.000	0.921	0.021	0.096
	8				1.000	0.910	0.018	0.203				1.000	0.920	0.019	0.095
	9				1.000	0.909	0.016	0.204				1.000	0.916	0.018	0.097

Table 4A Size of the Single Equation Tests Under the Null (δ = 0, T = 300; Trend Model)

		k = 1							k = 3						
q	m	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}	$t_{ECM\text{-}O}$	$t_{ADL\text{-}O}$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}
0	1	0.049	0.047	0.048	0.026	0.025	0.040	0.036	0.049	0.048	0.045	0.027	0.024	0.080	0.058
	2				0.048	0.049	0.066	0.055				0.047	0.049	0.119	0.079
	3				0.059	0.063	0.078	0.063				0.061	0.069	0.145	0.092
	4				0.070	0.073	0.089	0.072				0.071	0.085	0.166	0.105
	5				0.079	0.089	0.098	0.081				0.078	0.101	0.186	0.116
	6				0.085	0.099	0.107	0.090				0.086	0.120	0.205	0.130
	7				0.092	0.108	0.117	0.091				0.093	0.136	0.226	0.140
	8				0.097	0.116	0.122	0.096				0.098	0.147	0.240	0.147
	9				0.106	0.126	0.127	0.101				0.101	0.161	0.254	0.153
3	1	0.002	0.047	0.048	0.014	0.025	0.040	0.036	0.002	0.048	0.045	0.012	0.024	0.080	0.058
	2				0.029	0.049	0.066	0.055				0.026	0.049	0.119	0.079
	3				0.038	0.063	0.078	0.063				0.035	0.069	0.145	0.092
	4				0.043	0.073	0.089	0.072				0.037	0.085	0.166	0.105
	5				0.046	0.089	0.098	0.081				0.043	0.101	0.186	0.116
	6				0.050	0.099	0.107	0.090				0.046	0.120	0.205	0.130
	7				0.054	0.108	0.117	0.091				0.048	0.136	0.226	0.140
	8				0.056	0.116	0.122	0.096				0.048	0.147	0.240	0.147
	9				0.054	0.126	0.127	0.101				0.049	0.161	0.254	0.153
8	1	0.001	0.047	0.048	0.011	0.025	0.040	0.036	0.001	0.048	0.045	0.009	0.024	0.080	0.058
	2				0.025	0.049	0.066	0.055				0.024	0.049	0.119	0.079
	3				0.035	0.063	0.078	0.063				0.031	0.069	0.145	0.092
	4				0.040	0.073	0.089	0.072				0.034	0.085	0.166	0.105
	5				0.040	0.089	0.098	0.081				0.038	0.101	0.186	0.116
	6				0.042	0.099	0.107	0.090				0.041	0.120	0.205	0.130
	7				0.045	0.108	0.117	0.091				0.044	0.136	0.226	0.140
	8				0.046	0.116	0.122	0.096				0.044	0.147	0.240	0.147
	9				0.047	0.126	0.127	0.101				0.042	0.161	0.254	0.153

Note: For the OLS based tests, the 5% critical values were simulated. For T = 300, they are are -3.414, -3.708 and -3.794 for $t_{ECM\text{-}O}$, $t_{ADL\text{-}O}$ and $t_{EG\text{-}O}$ tests, when k = 1. The corresponding critical values are -3.412, -4.153 and -4.448, when k = 3. For the IV tests, -1.645 was used for all cases.

Table 4B Size-adjusted Power Under the Alternative (δ = -0.1, T = 300; Trend Model)

		$k = 1$							$k = 3$						
q	m	$t_{\text{ECM-O}}$	$t_{\text{ADL-O}}$	$t_{\text{EG-O}}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}	$t_{\text{ECM-O}}$	$t_{\text{ADL-O}}$	$t_{\text{EG-O}}$	t_{ECM}	t_{ADL}	t_{EG}	t^{+}_{EG}
0	1	0.949	0.892	0.882	0.283	0.245	0.252	0.263	0.944	0.766	0.711	0.279	0.214	0.188	0.218
	2				0.342	0.314	0.304	0.331				0.349	0.281	0.233	0.279
	3				0.412	0.378	0.366	0.392				0.400	0.332	0.267	0.347
	4				0.454	0.426	0.406	0.452				0.443	0.378	0.305	0.385
	5				0.489	0.460	0.446	0.495				0.479	0.399	0.337	0.413
	6				0.526	0.496	0.491	0.522				0.510	0.432	0.349	0.450
	7				0.576	0.529	0.523	0.562				0.546	0.454	0.366	0.470
	8				0.589	0.539	0.550	0.588				0.570	0.467	0.405	0.493
	9				0.604	0.564	0.570	0.610				0.596	0.481	0.411	0.507
3	1	1.000	1.000	0.496	0.919	0.774	0.194	0.622	1.000	1.000	0.226	0.998	0.902	0.133	0.669
	2				0.989	0.888	0.235	0.776				1.000	0.951	0.152	0.802
	3				0.999	0.937	0.280	0.860				1.000	0.963	0.164	0.860
	4				1.000	0.958	0.306	0.903				1.000	0.968	0.188	0.888
	5				1.000	0.967	0.337	0.928				1.000	0.971	0.209	0.909
	6				1.000	0.972	0.366	0.943				1.000	0.974	0.207	0.923
	7				1.000	0.976	0.375	0.952				1.000	0.975	0.207	0.936
	8				1.000	0.977	0.393	0.960				1.000	0.976	0.230	0.944
	9				1.000	0.979	0.415	0.968				1.000	0.978	0.225	0.950
8	1	1.000	1.000	0.426	1.000	0.974	0.187	0.904	1.000	1.000	0.206	1.000	0.976	0.128	0.827
	2				1.000	0.986	0.227	0.948				1.000	0.980	0.148	0.887
	3				1.000	0.990	0.270	0.962				1.000	0.984	0.160	0.916
	4				1.000	0.991	0.292	0.972				1.000	0.985	0.184	0.929
	5				1.000	0.991	0.323	0.978				1.000	0.986	0.203	0.942
	6				1.000	0.992	0.345	0.982				1.000	0.986	0.199	0.954
	7				1.000	0.992	0.359	0.986				1.000	0.988	0.203	0.960
	8				1.000	0.992	0.372	0.988				1.000	0.988	0.223	0.966
	9				1.000	0.992	0.393	0.991				1.000	0.989	0.219	0.972

Table 5. Comparison of the Various Cointegration Tests

CASE 1: Only y_t and x_{1t} are used in the test

β	δ	ρ	$\lambda_{trace}(1)$	$\lambda_{trace}(2)$	$\lambda_{max}(1)$	$\lambda_{max}(2)$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}
0	0.0	0.90	**0.106**	na	0.071	na	0.069	0.098	0.080	0.069
		0.95	0.106	na	0.071	na	0.069	0.099	0.088	0.069
	0.1	0.90	0.538	na	0.405	na	0.172	0.309	0.286	0.221
		0.95	0.538	na	0.405	na	0.172	0.307	0.282	0.221
	0.5	0.90	1.000	na	1.000	na	0.995	1.000	0.953	0.974
		0.95	1.000	na	1.000	na	0.995	1.000	0.961	0.974
1.0	0.0	0.90	0.596	na	0.579	na	0.018	0.030	0.027	0.012
		0.95	0.715	na	0.699	na	0.015	0.028	0.026	0.015
	0.1	0.90	**0.251**	na	0.245	na	0.025	0.405	0.595	0.004
		0.95	0.367	na	0.366	na	0.016	0.494	0.764	0.002
	0.5	0.90	0.162	na	0.093	na	0.150	0.442	0.630	0.096
		0.95	0.143	na	0.099	na	0.071	0.310	0.613	0.052
5.0	0.0	0.90	0.617	na	0.599	na	0.017	0.028	0.022	0.013
		0.95	0.729	na	0.712	na	0.016	0.024	0.023	0.015
	0.1	0.90	**0.261**	na	0.262	na	0.036	**0.507**	**0.862**	0.020
		0.95	0.385	na	0.389	na	0.026	0.544	0.922	0.010
	0.5	0.90	0.075	na	0.048	na	0.194	0.430	0.815	0.149
		0.95	0.096	na	0.071	na	0.093	0.367	0.831	0.083

CASE 2: All three variables are used in the test

β	δ	ρ	$\lambda_{trace}(1)$	$\lambda_{trace}(2)$	$\lambda_{max}(1)$	$\lambda_{max}(2)$	$t_{EG\text{-}O}$	t_{ECM}	t_{ADL}	t_{EG}
0	0.0	0.90	**0.162**	0.029	0.102	0.011	0.059	**0.105**	0.072	0.084
		0.95	0.112	0.018	0.071	0.007	0.068	0.107	0.076	0.086
	0.1	0.90	0.459	**0.181**	0.239	0.056	0.125	**0.298**	**0.291**	0.240
		0.95	0.370	0.114	0.219	0.031	0.132	0.296	0.282	0.245
	0.5	0.90	1.000	0.301	1.000	0.181	0.975	1.000	0.959	0.972
		0.95	1.000	0.183	1.000	0.113	0.978	1.000	0.959	0.972
1.0	0.0	0.90	1.000	0.173	1.000	0.112	0.014	0.038	0.023	0.016
		0.95	1.000	0.216	1.000	0.130	0.013	0.055	0.022	0.028
	0.1	0.90	1.000	0.209	1.000	0.119	0.011	0.285	0.957	0.269
		0.95	1.000	0.192	1.000	0.117	0.015	0.372	0.957	0.328
	0.5	0.90	1.000	0.265	1.000	0.160	0.970	1.000	0.968	0.997
		0.95	1.000	0.175	1.000	0.103	0.969	1.000	0.973	0.998
5.0	0.0	0.90	1.000	0.175	1.000	0.112	0.016	0.041	0.023	0.028
		0.95	1.000	0.218	1.000	0.129	0.015	0.057	0.022	0.030
	0.1	0.90	1.000	0.209	1.000	0.123	0.009	0.290	0.990	0.272
		0.95	1.000	0.191	1.000	0.119	0.013	0.387	0.988	0.340
	0.5	0.90	1.000	0.255	1.000	0.149	0.979	1.000	0.993	1.000
		0.95	1.000	0.175	1.000	0.099	0.969	1.000	0.988	1.000

Table 6. Size and Power of the System Based IV Cointegration Tests
In the Presence of Stationary Covariates

T	β	δ	ρ	$m = 3$			$m = 5$		
				W_{ECM}	W_{ADL}	W_{ADL2}	W_{ECM}	W_{ADL}	W_{ADL2}
100	0	0.0	0.90	0.055	0.067	0.062	0.065	0.077	0.072
			0.95	0.051	0.056	0.056	0.071	0.082	0.073
		0.1	0.90	0.162	0.162	0.199	0.238	0.255	0.306
			0.95	0.150	0.163	0.200	0.255	0.261	0.314
	1.0	0.0	0.90	0.051	0.056	0.056	0.055	0.064	0.053
			0.95	0.052	0.052	0.047	0.052	0.065	0.055
		0.1	0.90	0.458	0.615	0.660	0.366	0.576	0.614
			0.95	0.726	0.819	0.847	0.607	0.754	0.782
	5.0	0.0	0.90	0.053	0.050	0.055	0.058	0.066	0.051
			0.95	0.051	0.053	0.058	0.043	0.060	0.052
		0.1	0.90	0.932	0.971	0.979	0.844	0.963	0.968
			0.95	0.983	0.993	0.995	0.935	0.983	0.985
300	0	0.0	0.90	0.054	0.056	0.052	0.057	0.058	0.057
			0.95	0.053	0.055	0.054	0.054	0.059	0.060
		0.1	0.90	0.297	0.362	0.460	0.482	0.552	0.652
			0.95	0.297	0.357	0.444	0.491	0.563	0.666
	1.0	0.0	0.90	0.052	0.057	0.050	0.052	0.054	0.049
			0.95	0.048	0.046	0.046	0.055	0.056	0.053
		0.1	0.90	0.883	0.964	0.972	0.664	0.854	0.878
			0.95	0.991	0.998	0.999	0.956	0.983	0.987
	5.0	0.0	0.90	0.052	0.056	0.055	0.052	0.053	0.056
			0.95	0.046	0.049	0.045	0.050	0.054	0.053
		0.1	0.90	0.999	0.999	1.000	0.969	0.992	0.994
			0.95	1.000	1.000	1.000	0.998	1.000	1.000

Table 7. Size of the System Based IV Cointegration Tests
Under Different Model Specifications

T	DGP	$m = 3$			$m = 5$		
		W_{ECM}	W_{ADL}	W_{ADL2}	W_{ECM}	W_{ADL}	W_{ADL2}
100	2	0.054	0.055	0.054	0.058	0.063	0.055
	3	0.055	0.059	0.055	0.057	0.057	0.057
	4	0.058	0.057	0.055	0.054	0.056	0.058
300	2	0.049	0.055	0.055	0.049	0.046	0.051
	3	0.046	0.044	0.051	0.059	0.056	0.057
	4	0.045	0.043	0.047	0.058	0.054	0.051

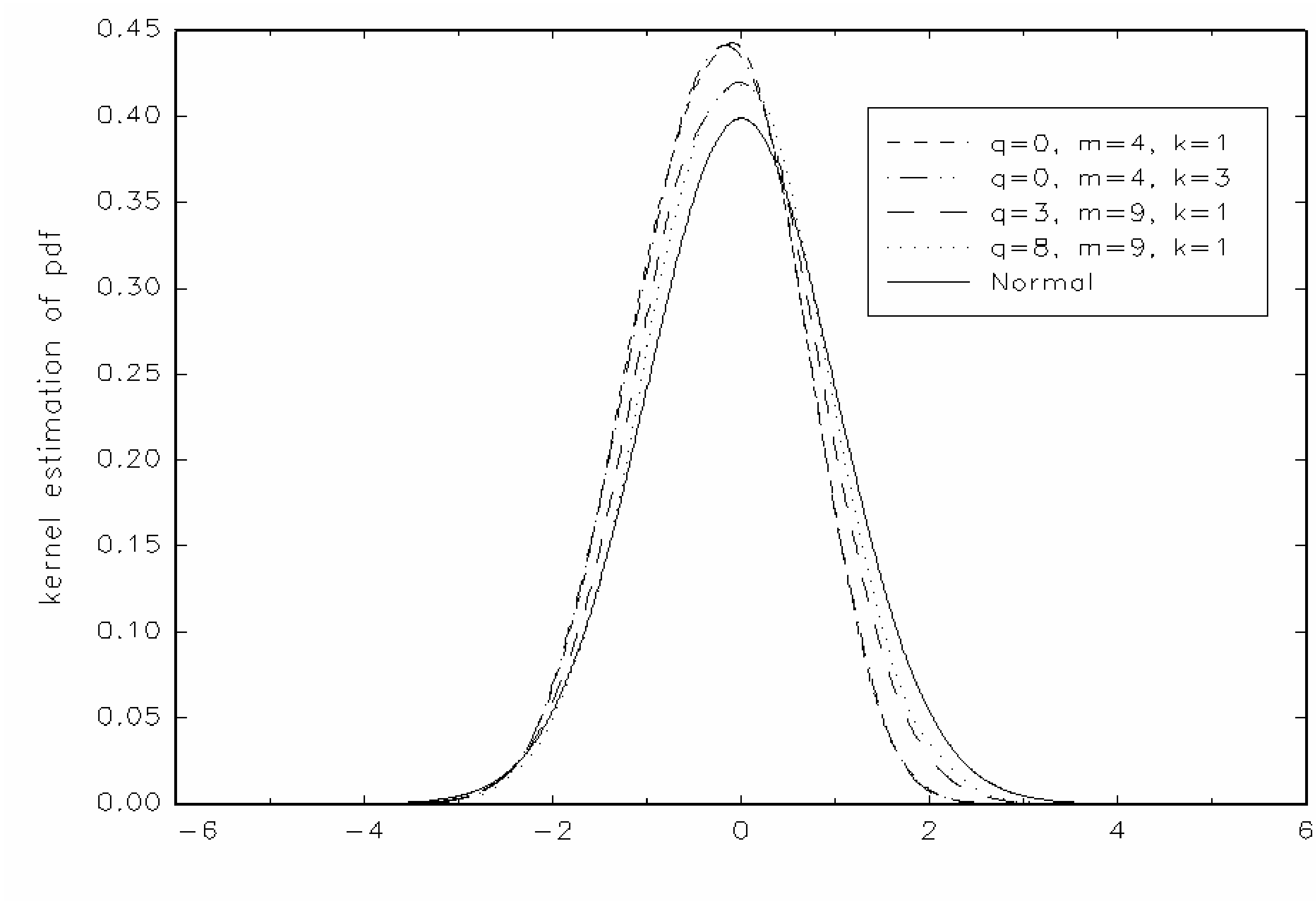

Figure 1. The Empirical Distribution of the IV test (t_{ECM}) Under the Null
(δ_1 = 0.0, T = 100, drift model)

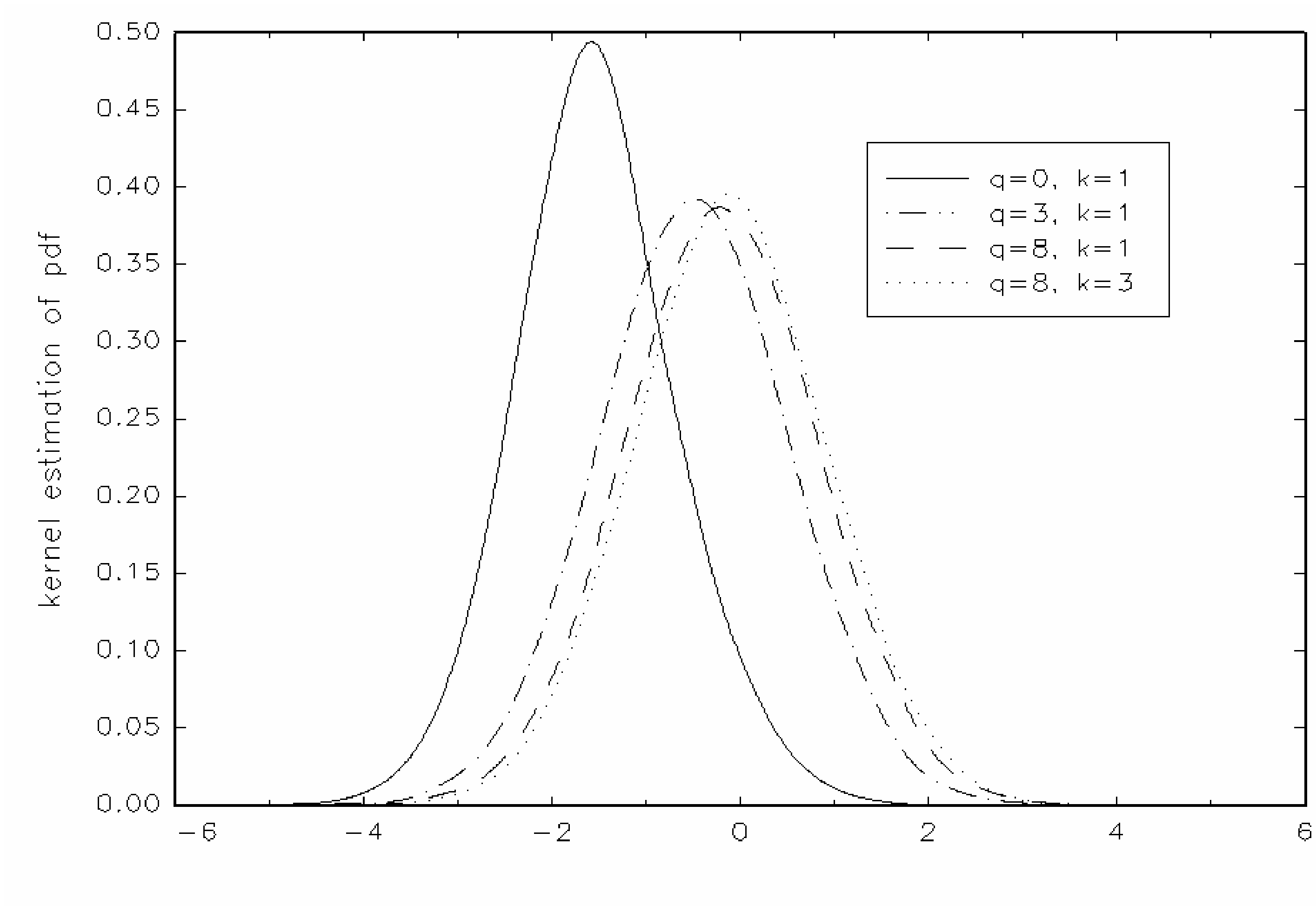

Figure 2. The Empirical Distribution of the OLS test ($t_{\text{ECM-O}}$) Under the Null
(δ_1 = 0.0, T = 100, drift model)

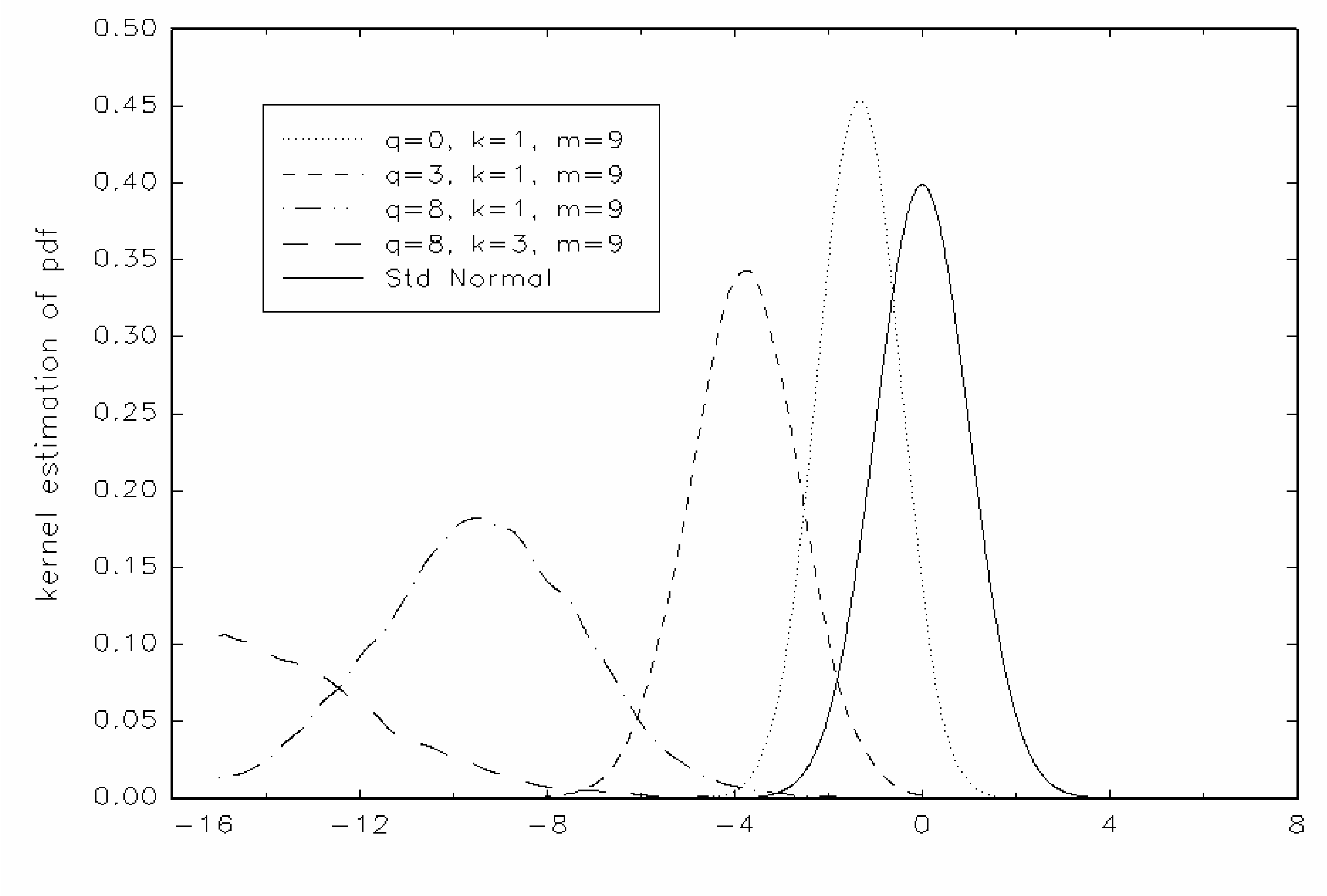

Figure 3. The Empirical Distribution of the IV test (t_{ECM}) Under the Alternative
(δ_1 = -0.1, T = 100, drift model)