Econometrics Journal (2002), volume **5**, pp. 285–318.

Distributions of error correction tests for cointegration

NEIL R. ERICSSON[†] AND JAMES G. MACKINNON[‡]

[†]*Stop 20, Division of International Finance, Federal Reserve Board, 2000 C Street, NW, Washington, DC 20551, USA*
E-mail: ericsson@frb.gov

[‡]*Department of Economics, Queen's University, Kingston, Ontario, Canada K7L 3N6*
E-mail: jgm@qed.econ.queensu.ca;
Homepage: www.econ.queensu.ca/faculty/mackinnon

Received: January 2000

Summary This paper provides densities and finite sample critical values for the single-equation error correction statistic for testing cointegration. Graphs and response surfaces summarize extensive Monte Carlo simulations and highlight simple dependencies of the statistic's quantiles on the number of variables in the error correction model, the choice of deterministic components, and the sample size. The response surfaces provide a convenient way for calculating finite sample critical values at standard levels; and a computer program, freely available over the Internet, can be used to calculate both critical values and p-values. Two empirical applications illustrate these tools.

Keywords: *Cointegration, Critical value, Distribution function, Error correction, Monte Carlo, Response surface.*

1. INTRODUCTION

Three general approaches are widely used for testing whether or not non-stationary economic time series are cointegrated: single-equation static regressions, due to Engle and Granger (1987); vector autoregressions, as formulated by Johansen (1988, 1995); and single-equation conditional error correction models, initially proposed by Phillips (1954) and further developed by Sargan (1964). While all three have their advantages and disadvantages, testing for cointegration with any of these approaches requires non-standard critical values, which are usually calculated by Monte Carlo simulation. Engle and Granger (1987) tabulate a limited set of critical values for their procedure. MacKinnon (1991) derives a more extensive set with finite sample corrections based on response surfaces, and MacKinnon (1996) provides a computer program to calculate critical values for Engle and Granger's test at any desired level. Johansen (1988), Johansen and Juselius (1990), and Osterwald-Lenum (1992) include critical values for the Johansen procedure under typical assumptions about deterministic terms and the number of stochastic variables. Johansen (1995), Doornik (1998), and MacKinnon *et al.* (1999) provide more accurate estimates of these critical values, with the last of these papers also providing computer programs to calculate critical values and p-values.

By contrast, critical values for the single-equation error correction procedure are scant, perhaps because error correction models substantially predate the literature on cointegration. Banerjee *et al.* (1993) tabulate critical values for an error correction model with two variables at three sample sizes; and Banerjee *et al.* (1998) list critical values for models with two through six variables at five sample sizes. Harbo *et al.* (1998), MacKinnon *et al.* (1999), and Pesaran *et al.* (2000) list asymptotic critical values for a related but distinct procedure for single- and multiple-equation error correction models.

The current paper addresses this dearth by providing an extensive set of cointegration critical values for the single-equation error correction model. These critical values include finite sample adjustments similar to those in MacKinnon (1991, 1996) for the Engle–Granger (EG) procedure, they are very accurate numerically and are easy to use in practice, and they encompass and supersede comparable results in Banerjee *et al.* (1993) and Banerjee *et al.* (1998). We also provide a freely available Excel spreadsheet and a Fortran program (the latter being similar to the one in MacKinnon (1996) for the EG procedure) that compute both critical values and *p*-values for the error correction statistic. As the articles in Banerjee and Hendry (1996), Ericsson (1998), and Lütkepohl and Wolters (1998) *inter alia* highlight, conditional error correction models are ubiquitous empirically, so these tools for calculating critical values and *p*-values should be of immediate and widespread use to the empirical modeler. Finally, general distributional properties are of considerable interest. Accurate numerical approximations to the entire distribution of the error correction statistic are calculated herein and offer insights into the nature of that statistic, particularly relative to the Dickey–Fuller and EG statistics. Graphs highlight the error correction statistic's properties and relationships, and show for the first time what many of its various distributions look like. Throughout, the focus is on testing for cointegration, rather than on the complementary task of estimating the cointegrating vectors, assuming a given cointegration rank.

This paper is organized as follows. Section 2 sets the backdrop by considering the three common procedures and their relationships to each other. Section 3 outlines the structure of the Monte Carlo analysis for calculating the distributional properties of the cointegration test statistic based on the single-equation error correction model. Section 4 presents the Monte Carlo results, which include densities and finite sample critical values. Section 5 applies the finite sample critical values derived in Section 4 and the computer program for calculating *p*-values to empirical error correction models of UK narrow money demand from Hendry and Ericsson (1991) and of US federal government debt from Hamilton and Flavin (1986). Section 6 concludes.

2. AN OVERVIEW OF THREE TEST PROCEDURES

This paper focuses on finite sample inference about cointegration in a single-equation conditional error correction model (ECM).[1] To motivate the use of conditional ECMs, this section describes the analytics of and inferential methods for the three common approaches for testing cointegration: the Johansen procedure (Section 2.1), the conditional ECM (Section 2.2), and the EG procedure (Section 2.3). Differences between the three approaches turn on their various assumptions about dynamics and exogeneity (Section 2.4).

[1]Strictly speaking, the models examined herein are equilibrium correction models; see Hendry and Doornik (2001, p. 144).

2.1. The Johansen procedure

Johansen (1988, 1995) derives maximum likelihood procedures for testing for cointegration in a finite-order Gaussian vector autoregression (VAR). That system is:

$$x_t = \sum_{i=1}^{\ell} \pi_i x_{t-i} + \Phi D_t + \varepsilon_t, \qquad \varepsilon_t \sim IN(0, \Omega), \qquad t = 1, \ldots, T, \tag{1}$$

where x_t is a vector of k variables at time t; π_i is a $k \times k$ matrix of coefficients on the ith lag of x_t; ℓ is the maximal lag length; Φ is a $k \times d$ matrix of coefficients on D_t, a vector of d deterministic variables (such as a constant term and a trend); ε_t is a vector of k unobserved, sequentially independent, jointly normal errors with mean zero and (constant) covariance matrix Ω; and T is the number of observations. Throughout, x is restricted to be (at most) integrated of order one, denoted I(1), where an I(j) variable requires jth differencing to make it stationary.

The VAR in (1) may be rewritten as a vector error correction model:

$$\Delta x_t = \pi x_{t-1} + \sum_{i=1}^{\ell-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \varepsilon_t, \qquad \varepsilon_t \sim IN(0, \Omega), \tag{2}$$

where π and Γ_i are:

$$\pi = \left(\sum_{i=1}^{\ell} \pi_i \right) - I_k, \tag{3}$$

$$\Gamma_i = -(\pi_{i+1} + \cdots + \pi_\ell), \qquad i = 1, \ldots, \ell - 1, \tag{4}$$

I_k is the identity matrix of dimension k, and Δ is the difference operator.[2] For any specified number of cointegrating vectors r $(0 \le r \le k)$, the matrix π is of (potentially reduced) rank r and may be rewritten as $\alpha\beta'$, where α and β are $k \times r$ matrices of full rank. By substitution, (2) is:

$$\Delta x_t = \alpha\beta' x_{t-1} + \sum_{i=1}^{\ell-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \varepsilon_t, \qquad \varepsilon_t \sim IN(0, \Omega), \tag{5}$$

where β is the matrix of cointegrating vectors, and α is the matrix of adjustment coefficients (equivalently, the loading matrix).

Johansen (1988, 1995) derives two maximum likelihood statistics for testing the rank of π in (2) and hence for testing the number of cointegrating vectors in (2). Critical values appear in Johansen (1988, Table 1) for a VAR with no deterministic components, in Johansen and Juselius (1990, Tables A1–A3) for VARs with a constant term, and in Osterwald-Lenum (1992) and Johansen (1995, Ch. 15) for VARs with no deterministic components, with a constant term only, and with a constant term and a linear trend. Doornik (1998) derives a convenient approximation to the maximum likelihood statistics' distributions using the Gamma distribution, and MacKinnon *et al.* (1999) provide computer programs to calculate critical values and p-values for the Johansen procedure.

[2]The difference operator Δ is defined as $(1 - L)$, where the lag operator L shifts a variable one period into the past. Hence, for x_t, $Lx_t = x_{t-1}$ and so $\Delta x_t = x_t - x_{t-1}$. More generally, $\Delta_j^i x_t = (1 - L^j)^i x_t$ for positive integers i and j. If i or j is not explicit, it is taken to be unity.

2.2. Single-equation conditional error correction models

Without loss of generality, the VAR in (1) can be factorized into a pair of conditional and marginal models. If the marginal variables are weakly exogenous for the cointegrating vectors β, then inference about cointegration using the conditional model alone can be made without loss of information relative to inference using the full system (the VAR); see Johansen (1992a,b). This subsection derives a *single-equation* conditional model from the VAR and delineates two related approaches for conducting such inferences about cointegration from that conditional model. The second of those approaches is the focus of the Monte Carlo analysis in Sections 3 and 4 and of the empirical analysis in Section 5.

For expositional clarity, assume that (1) is a first-order VAR with no deterministic components. Its explicit representation as the vector error correction model (2) is:

$$\Delta y_t = \pi_{(11)} y_{t-1} + \pi_{(12)} z_{t-1} + \varepsilon_{1t} \tag{6}$$

$$\Delta z_t = \pi_{(21)} y_{t-1} + \pi_{(22)} z_{t-1} + \varepsilon_{2t}, \tag{7}$$

where $x_t' = (y_t, z_t')$, y_t is a scalar endogenous variable, z_t is a $(k-1) \times 1$ vector of potentially weakly exogenous variables, π is partitioned conformably to x_t as $\{\pi_{(ij)}\}$, and $\varepsilon_t' = (\varepsilon_{1t}, \varepsilon_{2t}')$. From (5), equations (6) and (7) may be written as:

$$\Delta y_t = \alpha_1 \beta' x_{t-1} + \varepsilon_{1t} \tag{8}$$

$$\Delta z_t = \alpha_2 \beta' x_{t-1} + \varepsilon_{2t}, \tag{9}$$

where $\alpha' = (\alpha_1, \alpha_2')$. Equations (8) and (9) may always be factorized into the conditional distribution of y_t given z_t and lags on both variables, and the marginal distribution of z_t (also given lags on both variables):

$$\Delta y_t = \gamma_0' \Delta z_t + \gamma_1 \beta' x_{t-1} + \nu_{1t} \tag{10}$$

$$\Delta z_t = \alpha_2 \beta' x_{t-1} + \varepsilon_{2t}, \tag{11}$$

where $\gamma_0' = \Omega_{12}\Omega_{22}^{-1}$, $\gamma_1 = \alpha_1 - \Omega_{12}\Omega_{22}^{-1}\alpha_2$, $\nu_{1t} = \varepsilon_{1t} - \Omega_{12}\Omega_{22}^{-1}\varepsilon_{2t}$, the expectation $\mathcal{E}(\nu_{1t}\varepsilon_{2t})$ is zero (by construction), and the error covariance matrix Ω in (1) is $\{\Omega_{ij}\}$. Equivalently, the error ε_{1t} in (8) may be partitioned into two uncorrelated components as $\varepsilon_{1t} = \nu_{1t} + \gamma_0'\varepsilon_{2t}$, and then ε_{2t} is substituted out to obtain (10).

The variable z_t is weakly exogenous for β if and only if $\alpha_2 = 0$ in (11), in which case (10) and (11) become:

$$\Delta y_t = \gamma_0' \Delta z_t + \gamma_1 \beta' x_{t-1} + \nu_{1t} \tag{12}$$

$$\Delta z_t = \varepsilon_{2t}, \tag{13}$$

where $\gamma_1 = \alpha_1$. The test of z_t being weakly exogenous for β is thus a test of $\alpha_2 = 0$; see Johansen (1992a).

If $\alpha_2 = 0$, the conditional ECM (12) by itself is sufficient for inference about β that is without loss of information relative to inference from (10) and (11) together. Two distinct approaches have evolved for testing cointegration in the conditional ECM (12): one is due to Harbo *et al.* (1998), and the other originates from the literature on ECMs. The current paper analyzes the second approach, and clarifying the distinction between the two approaches is central to understanding their respective properties.

Harbo *et al.* (1998) derive the likelihood ratio statistic for testing cointegrating rank in a conditional subsystem obtained from a Gaussian VAR when the marginal variables are weakly exogenous for β. For a single-equation conditional model such as (12), the null hypothesis being tested is $\gamma_1\beta' = 0$, i.e. that the cointegrating rank for x is zero. The alternative hypothesis is that $\gamma_1\beta' \neq 0$, implying that x has a cointegrating vector β with at least one non-zero element.

The second approach stems from the literature on error correction models and is based on transformations of (12), with an auxiliary assumption about the nature of x's cointegration. Specifically, the conditional ECM (12) can be motivated as a reparameterization of the conditional autoregressive distributed lag (ADL) model; see Davidson *et al.* (1978) and Hendry *et al.* (1984) *inter alia*. Data transformations imply reparameterizations, and two transformations are of particular interest:

$$\begin{array}{llll} \text{differencing}: & \mu_1 x_t + \mu_2 x_{t-1} & \rightarrow & \mu_1 \Delta x_t + (\mu_1 + \mu_2) x_{t-1} \\ \text{differentials}: & \mu_1 y_t + \mu_2 z_t & \rightarrow & \mu_1 (y_t - z_t) + (\mu_1 + \mu_2) z_t, \end{array}$$

for arbitrary coefficients μ_1 and μ_2. Repeatedly applying these two transformations re-arranges a conditional ADL into the conditional ECM (12):

$$y_t = \lambda_0' z_t + \lambda_1' z_{t-1} + \lambda_2 y_{t-1} + \nu_{1t} \tag{14}$$

$$y_t = \gamma_0' \Delta z_t + \lambda_3' z_{t-1} + \lambda_2 y_{t-1} + \nu_{1t} \tag{15}$$

$$\Delta y_t = \gamma_0' \Delta z_t + \lambda_3' z_{t-1} + \gamma_1 y_{t-1} + \nu_{1t} \tag{16}$$

$$\Delta y_t = \gamma_0' \Delta z_t + \gamma_1 (y_{t-1} - \delta' z_{t-1}) + \nu_{1t} \tag{17}$$

$$\Delta y_t = \gamma_0' \Delta z_t + \gamma_1 \beta' x_{t-1} + \nu_{1t}, \tag{18}$$

where λ_0, λ_1, λ_2, λ_3, and δ are various coefficients; and the cointegrating vector β has been normalized on its first coefficient (i.e. for y) such that $\beta' = (1, -\delta')$. In practice, significance testing of the error correction term typically has been based on the t-ratio for γ_1 in (16), not (17) or (18). This is the 'PcGive unit root test' in Hendry (1989, p. 149) and Hendry and Doornik (2001, p. 256), which here is denoted the ECM statistic.

When interpreted as a test for cointegration of x, this approach requires an additional assumption: namely, that the variables in z are not cointegrated among themselves. Thus, $\gamma_1 = 0$ in (16) implies (and is implied by) a lack of cointegration between y and z, whereas $\gamma_1 < 0$ implies cointegration. The t-ratio based upon the least squares estimator of γ_1 in (16) is the ECM statistic analyzed in Sections 3–5. That t-ratio is denoted $\kappa_d(k)$, where d indicates the deterministic components included in the ECM, or the number of such deterministic components, depending upon the context; and k is the total number of variables in x (not to be confused with the number of regressors in the ECM). This t-ratio is used to test the null hypothesis that $\gamma_1 = 0$, i.e. that y and z are *not* cointegrated. If weak exogeneity does not hold, critical values generally are affected; see Hendry (1995).

Campos *et al.* (1996) and Banerjee *et al.* (1998) derive the asymptotic distribution of $\kappa_d(k)$ under the null hypothesis of no cointegration:

$$\kappa_d(k) \Rightarrow \left(\int \bar{B}_v^2\right)^{-1/2} \int \bar{B}_v \, dB_v, \tag{19}$$

where B_v and B_ε are the standardized Wiener processes corresponding to v_{1t} and ε_{2t}, $\bar{B}_v$ is $B_v - \left(\int B_\varepsilon B_v\right)' \left(\int B_\varepsilon B_\varepsilon'\right)^{-1} B_\varepsilon$, '$\Rightarrow$' denotes weak convergence of the associated probability measures as $T \to \infty$, strong exogeneity of z with respect to α and β is assumed, and the ECM has no deterministic terms. If the ECM includes deterministic terms, the asymptotic distribution of $\kappa_d(k)$ is of the same form as in (19), but with the Wiener processes replaced by the corresponding Brownian bridges. Johansen (1995, Ch. 11.2) develops analogous algebra for the Johansen maximum likelihood statistic when the VAR has deterministic terms.

Kiviet and Phillips (1992) and Banerjee *et al.* (1998) discuss similarity for $\kappa_d(k)$. Notably, the asymptotic distribution in (19) depends on k and d, but not on the short-run coefficients in the ECM. That is, $\kappa_d(k)$ is asymptotically similar with respect to γ_0, and also with respect to coefficients on any lags of Δx in the ECM, provided that those parameters lie within the space satisfying the I(1) conditions for x. The statistic $\kappa_d(k)$ is *exactly* similar with respect to the constant term if the estimated ECM includes a constant term and a linear trend, and with respect to the constant term and the linear trend's coefficient if the estimated ECM includes a constant term, a linear trend, and a quadratic trend. Following Johansen (1995, p. 84), seasonal dummies with a constant term may affect the finite sample (but not asymptotic) distribution. Likewise, the choice of a fixed lag length ℓ affects the finite sample (but not asymptotic) distribution, provided ℓ is large enough to avoid mis-specification; see Banerjee *et al.* (1998, Section 5).

To summarize, the ECM statistic $\kappa_d(k)$ is designed to detect cointegration involving y in the conditional model (12). The procedure in Harbo *et al.* (1998) is designed to detect any cointegration in x in the conditional model (12), where that cointegration may include y or it may be restricted to z alone. While both statistics derive from conditional models, the two statistics are testing different hypotheses. They have different distributions—even asymptotically—and so require separate tabulation.

Harbo *et al.* (1998, Tables 2–4) present asymptotic critical values for their statistic for (typically) $k = 2, \ldots, 7$ with several choices of deterministic terms, allowing for conditional subsystems (i.e. with more than one endogenous variable) as well as conditional single equations. Pesaran *et al.* (2000, Tables 6(a)–6(e)) estimate the 5% and 10% critical values for up through five weakly exogenous variables and 12 endogenous variables. Using response surfaces, MacKinnon *et al.* (1999, Tables 2–6) extend and more precisely estimate the 5% critical values in Harbo *et al.* (1998) and Pesaran *et al.* (2000) for up through eight weakly exogenous variables and 12 endogenous variables. They also make available a program that calculates asymptotic critical values at any level and p-values. Doornik (1998, Section 9) approximates the distribution of Harbo *et al.*'s maximum likelihood trace statistic by a Gamma function. Boswijk and Franses (1992) and Boswijk (1994) analyze a Wald statistic for testing $\gamma_1\beta' = 0$. Boswijk (1994) also tabulates asymptotic critical values for this Wald statistic in the single-equation case, and they are numerically very similar to those in Harbo *et al.* (1998) for the comparable likelihood ratio statistic.

Critical values for the ECM statistic $\kappa_d(k)$ appear in Banerjee *et al.* (1993, Table 7.6) for $k = 2$ with a constant term, and in Banerjee *et al.* (1998, Table I) for $k = 2, \ldots, 6$ with a constant term and with a constant term and a linear trend. In both studies, the maximum number of variables is too small for many empirical purposes, the estimates of the critical values are relatively imprecise, and finite sample adjustments are impractical from the reported critical values. The results in Section 4 address these limitations. In the next subsection, the derivation in (14)–(18) clarifies the relationship between the ECM and EG procedures.

2.3. The Engle–Granger procedure

Engle and Granger (1987) propose testing for cointegration by testing whether the residuals of a static regression are stationary. The usual unit root test used is that of Dickey and Fuller (1981), which is based on a finite-order autoregression. Engle and Granger's procedure imposes a common factor restriction on the dynamics of the relationship between the variables involved. If that restriction is invalid, a loss of power relative to the ECM and Johansen procedures may well result. This subsection highlights the role of the common factor restriction by expressing the model for Engle and Granger's procedure as a restricted ECM.

Reconsider the conditional ECM derived from a first-order VAR:

$$\Delta y_t = \gamma_0' \Delta z_t + \gamma_1 (y - \delta' z)_{t-1} + \nu_{1t}, \tag{20}$$

where $y_t - \delta' z_t$ is the putative disequilibrium. Engle and Granger's cointegration test statistic can be formulated from (20), thus establishing the relationship between it and the ECM statistic. Specifically, subtract $\delta' \Delta z_t$ from both sides of (20) and re-arrange:

$$\Delta(y - \delta' z)_t = \gamma_1 (y - \delta' z)_{t-1} + \{(\gamma_0' - \delta') \Delta z_t + \nu_{1t}\}. \tag{21}$$

Defining the Engle–Granger residual $y_t - \delta' z_t$ as w_t, (21) may be rewritten as:

$$\Delta w_t = \gamma_1 w_{t-1} + e_t, \tag{22}$$

where, by construction, the disturbance e_t is $(\gamma_0' - \delta')\Delta z_t + \nu_{1t}$. The t-ratio on the least squares estimator of γ_1 in (22) is the EG cointegration test statistic. It is the Dickey–Fuller statistic for testing whether w has a unit root and hence whether y and z lack (or obtain) cointegration with cointegrating vector $(1, -\delta')$. Below, that t-ratio is denoted $\tau_d(k)$, paralleling Dickey and Fuller's notation.

From (21), $\tau_d(k)$ imposes $\gamma_0 = \delta$, equating the short-run and long-run elasticities (the common factor restriction). Empirically, estimated short- and long-run elasticities often differ markedly, so imposing their equality is arbitrary and hazardous. Weak exogeneity is assumed in the presentation above but is not required for the EG procedure. See Kremers *et al.* (1992) for a general derivation of the common factor restriction in the EG procedure.

If the cointegrating coefficient δ is known, then the t-ratio on γ_1 in (22) has a Dickey–Fuller distribution (equivalent to assuming $k = 1$), as originally tabulated by Dickey in Fuller (1976, Table 8.5.2). If δ is estimated by least squares prior to testing that $\gamma_1 = 0$, then other critical values are required. Engle and Granger (1987, Table II) give such critical values for the bivariate model ($k = 2$) with a constant term. The response surfaces in MacKinnon (1991, Table 1) allow construction of critical values with finite sample adjustments for $k = 1, \ldots, 6$ with a constant term and with a constant term and a linear trend. MacKinnon (1996) provides a computer program to calculate numerically highly accurate critical values at any desired level for $k = 1, \ldots, 12$ with deterministic terms up to and including a quadratic trend.

2.4. A comparison

The Johansen, ECM, and EG procedures all focus on whether or not the feedback parameters for the cointegrating vector(s) are non-zero: α for the Johansen procedure, α_1 for the ECM procedure, and γ_1 (which is α_1 under weak exogeneity) for the EG procedure. The procedures differ in their assumptions about the data generation process (DGP), and those assumptions imply both advantages and disadvantages for empirical implementation. For all three procedures, numerical computations are easy and fast for both estimation and testing.

Table 1. A comparison of the Johansen, ECM, and Engle–Granger procedures for testing cointegration.

Aspect	Procedure		
	Johansen	ECM (both types)	Engle–Granger
Statistic	Maximal eigenvalue and trace statistics.	$\kappa_d(k)$; Harbo *et al.* (1998) statistic.	$\tau_d(k)$
Assumptions	Well-specified full system.	Weak exogeneity of z_t for β.	Common factor restriction.
Advantages	Maximum likelihood of full system. Determines r (the number of cointegrating vectors), β, and α.	Starting point for ECM modeling; unrestrictive dynamics. Weak exogeneity is often valid empirically. Robust to particulars of the marginal process.	Intuitive. Super-consistent estimator of β.
Disadvantages	Full system should be well-specified.	Weak exogeneity is assumed. $r \leq 1$ is imposed (usually).	Comfac is often invalid. Inferences on β are messy. Biases in estimating β. $r \leq 1$ imposed (usually). Normalization affects estimation. Dynamics may be of interest.
Sources for critical values and p-values	Johansen (1988, 1995), Johansen and Juselius (1990), Osterwald-Lenum (1992), Doornik (1998), MacKinnon *et al.* (1999).	Banerjee *et al.* (1993), Banerjee *et al.* (1998), this paper; Harbo *et al.* (1998), MacKinnon *et al.* (1999), Pesaran *et al.* (2000).	Engle and Granger (1987), MacKinnon (1991, 1994, 1996).

Table 1 compares the assumptions of these procedures and their implied advantages and disadvantages. For the procedure using the conditional ECM, the advantages are severalfold. The conditional ECM (or, equivalently, the unrestricted ADL) is a common starting point for modeling general to specific in a single-equation context. Also, weak exogeneity is often valid empirically. And, the ECM procedure is robust to many particulars of the marginal process, e.g. specific lag lengths and dynamics involved. While the ECM procedure assumes weak exogeneity and often assumes at most a single cointegrating vector, the procedure's appeal has made it common in the literature—hence the need for a clear understanding of the procedure's distributional properties.[3] The next two sections describe the structure of the Monte Carlo analysis used for calculating such properties (Section 3) and the results obtained (Section 4).

3. THE STRUCTURE OF THE MONTE CARLO ANALYSIS

This paper's objective is to provide information on finite sample inference about cointegration in conditional error correction models. Section 2 motivated the interest in the ECM statistic by

[3]Testing for weak exogeneity in a VAR and then for cointegration in a conditional ECM need not suffer from classical pre-test problems, as the corresponding hypotheses are nested. See Hoover and Perez (1999).

clarifying its relationships to the Johansen and EG procedures. The remaining sections examine the distributional properties of the ECM statistic.

Because no analytical solution is known for even the asymptotic distribution of the ECM test statistic, distributional properties are estimated by Monte Carlo simulation. This section outlines the structure of that Monte Carlo simulation. Section 3.1 describes the focus of this paper's simulation, the DGP, and the model estimated. Sections 3.2 and 3.3 sketch the design and simulation of the Monte Carlo experiments, and Section 3.4 discusses post-simulation analysis.

3.1. The focus, the data generation process, and the model

The general object of interest is the distribution of the ECM test statistic $\kappa_d(k)$ under the null of no cointegration. Asymptotic properties are derived in Kiviet and Phillips (1992), Campos *et al.* (1996), and Banerjee *et al.* (1998), with certain invariance results appearing in Kiviet and Phillips (1992). Finite sample properties appear in Banerjee *et al.* (1993), Campos *et al.* (1996), and Banerjee *et al.* (1998), but all are very limited in their experimental design.[4] In the current paper, two aspects are of primary concern: the distribution of $\kappa_d(k)$, and critical values at common levels of significance.

To examine the properties of the ECM statistic under the null hypothesis of no cointegration, the DGP is a standardized multivariate random walk for x:

$$\Delta x_t \sim IN(0, I_k), \tag{23}$$

a common DGP for simulating the null distribution of cointegration test statistics.

The estimated model is the conditional ECM resulting from a possibly cointegrated, ℓth-order, k-variable VAR, assuming weak exogeneity of z_t for β and with y_t scalar. That is, the estimated model is:

$$\Delta y_t = \gamma_0' \Delta z_t + b' x_{t-1} + \sum_{i=1}^{\ell-1} \Gamma_{1i} \Delta x_{t-i} + \phi_1' D_t + \nu_{1t} \quad \nu_{1t} \sim IN(0, \sigma_\nu^2), \tag{24}$$

where b, Γ_{1i}, and ϕ_1 are coefficients in the conditional ECM; and σ_ν^2 is the conditional ECM's error variance. Because $b' \equiv (b_1, b_2, \ldots, b_k) = \gamma_1 \beta'$ in the notation of the ECM (18), then b_1 is γ_1, which is the coefficient of interest in the ECM statistic $\kappa_d(k)$. The deterministic component D_t may include a constant term, a constant term and a linear trend, or a constant term, a linear trend, and a quadratic trend. The corresponding ECM statistics are denoted $\kappa_c(k)$, $\kappa_{ct}(k)$, and $\kappa_{ctt}(k)$, respectively. If no variables are included in D_t, then the ECM statistic is denoted $\kappa_{nc}(k)$ (*nc* for no constant term).

[4]The current paper, like much of the literature, focuses on cointegration tests when the cointegrating vectors are unknown *a priori*. This is a reasonable approach in many situations. Economic theory may not be fully informative about the cointegrating vector, or the researcher may wish to test the implied economic restrictions. Moreover, different economic theories may imply different cointegrating vectors, as with the quantity theory and the Baumol–Tobin framework. Notably, economic theory does *not* fully specify the cointegrating vectors for the empirical applications in Section 5.

Kremers *et al.* (1992), Hansen (1995), Campos *et al.* (1996), and Zivot (2000) consider distributional properties for the ECM statistic when the cointegrating coefficients *are* known. In that case, the statistic's distribution contains nuisance parameters, even asymptotically, although those parameters can be estimated consistently. Hansen (1995) provides asymptotic critical values for such a procedure; response surfaces for finite sample properties could be developed along the lines of our paper. As Zivot (2000) shows, considerable power gains can be achieved by correctly prespecifying the cointegrating vector. Conversely, the test can be inconsistent if the cointegrating vector is incorrectly prespecified, as that prespecification induces an I(1) component in the error term. Horvath and Watson (1995) and Elliott (1995) analyze properties of cointegration tests from a VAR when the cointegrating vectors are prespecified.

3.2. Specifics of the experimental design

The analysis focuses on the finite sample properties of the ECM statistic. Three 'design parameters' are central to the statistic's distributional properties: the estimation sample size (T), the total number of variables in x_t (k), and the number of deterministic components in D_t (d). To provide results for a wide range of situations common in empirical investigations, the simulations span a full factorial design of the following T, k, and D_t:

$$\begin{aligned} T &= (20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 125, 150, 200, \\ &\qquad 400, 500, 600, 700, 1000) \\ k &= (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12) \\ D_t &= (\text{none; constant term; constant term, } t\text{; constant term, } t,\ t^2). \end{aligned} \tag{25}$$

The range of the sample size aims to provide information on both the test statistic's asymptotic properties and its finite sample deviations therefrom. The design includes all positive integer values of k up through 12, sufficient for virtually all empirical applications. The choice of D_t implies four test statistics: $\kappa_{nc}(k)$, $\kappa_c(k)$, $\kappa_{ct}(k)$, and $\kappa_{ctt}(k)$. Deterministic terms may be included in the model because they are required for adequate model specification, i.e. because the deterministic terms enter the DGP. Also, a deterministic term of one order higher than 'required' may be included in the model in order to obtain similarity to the coefficients of the lower-order deterministic terms; see Kiviet and Phillips (1992), Johansen (1994), and Nielsen and Rahbek (2000). Throughout the simulations, the model's lag length is set to unity ($\ell = 1$). However, the lag notation in (24) is useful, as $\ell > 1$ for the empirical models in Section 5.

One minor modification exists for the experimental design in (25). Because $2k-1+d$ degrees of freedom are used in the estimation of (24), some smaller values of T are not considered for larger values of k that imply $2k-1+d$ close to or exceeding T. Specifically, $T = 20$ is dropped for $k = 8$; $T = (20, 25)$ are dropped for $k = (9, 10)$; and $T = (20, 25, 30)$ are dropped for $k = (11, 12)$.

3.3. Monte Carlo simulation

This paper aims to provide numerically accurate estimates of the ECM statistic's distribution, particularly in its tails, where inference is commonly of concern. Thus, a large number of replications are simulated for each experiment in (25): specifically, 10 million replications for each pair of T and k. Such large numbers of replications do not pose difficulties for calculations of sample moments, but they are problematic for calculating quantiles—and hence densities—because the full set of replications must be stored and sorted. As a reasonably efficient second-best alternative, the adopted design divides each experiment into 50 sets of 200 000 replications apiece, determines the quantiles for each set, and then averages the estimated quantile values across the sets. Partitioning each experiment into several sets also provides an easy way to measure experimental randomness. To estimate accurately the complete densities of the ECM statistic, a large number of quantiles are calculated: 221 in total, corresponding to $p = 0.0001$, 0.0002, 0.0005, 0.001, 0.002, 0.003, ..., 0.008, 0.009, 0.010, 0.015, 0.020, 0.025, ..., 0.495, 0.500, 0.505, ..., 0.975, 0.980, 0.985, 0.990, 0.991, 0.992, ..., 0.997, 0.998, 0.999, 0.9995, 0.9998, 0.9999, where p denotes the quantile's percent level.

Because so many random numbers were generated, it was vital to use a pseudo-random number generator with a very long period. The generator used was that in MacKinnon (1994, 1996), which combines two different pseudo-random number generators recommended by L'Ecuyer (1988). The two generators were started with different seeds and allowed to run independently, so that two independent uniform pseudo-random numbers were generated at once. Each pair was then transformed into two $N(0, 1)$ variates using the modified polar method of Marsaglia and Bray (1964, p. 260). See MacKinnon (1994, p. 170) for details.

3.4. Post-simulation analysis

These Monte Carlo simulations generate a vast quantity of information: 221 estimated quantiles on 50 sets of replications for (typically) 21 sample sizes with 12 different values of k and four choices of D_t: over 10 million numbers. Graphs and regressions provide two succinct ways of conveying and summarizing such information. This paper uses both means: graphs of asymptotic and finite sample densities, and response surfaces for finite sample critical values. An explanation is helpful for interpreting both the response surfaces and the graphs.

Typically, authors have tabulated estimated critical values for several sample sizes or for one large ('close to asymptotic') sample size. Such tabulations recognize the dependence of the critical values on the estimation sample size. That dependence can be approximated by regression, regressing the Monte Carlo estimates of the critical value on functions of the sample size. Such regressions are response surfaces: see Hammersley and Handscomb (1964) and Hendry (1984) for general discussions.

Here, for each triplet defined by the quantile's percent level p, the number of variables k, and the choice of deterministic components D_t, a response surface was estimated:

$$q(T_i) = \theta_\infty + \theta_1(T_i^a)^{-1} + \theta_2(T_i^a)^{-2} + \theta_3(T_i^a)^{-3} + u_i. \tag{26}$$

The dependent variable $q(T_i)$ is the estimated finite sample pth quantile from the Monte Carlo simulation with the ith sample size T_i, which takes the values for T in the experimental design (25). The regressors are an intercept and three inverse powers of the *adjusted* sample size T_i^a (which equals $T_i - (2k - 1) - d$); θ_∞, θ_1, θ_2, and θ_3 are the corresponding coefficients; and u_i is an error that reflects both simulation uncertainty and the approximation of the quantile's true functional form by the cubic in (26).

The benefits of these response surfaces are several. First, they reduce consumption costs to the user by summarizing numerous Monte Carlo experiments in a simple regression. Second, the coefficient θ_∞ is interpretable as the *asymptotic* ($T = \infty$) pth quantile for the choice of k and D_t concerned. Estimation of that asymptotic quantile does not necessarily require very large sample sizes in the experimental design. Third, response surfaces reduce the Monte Carlo uncertainty by averaging (through regression) across different experiments. Fourth, response surfaces reduce the specificity of the simulations by allowing easy calculation of quantiles for sample sizes not included in the experimental design (25). Fifth, p-values and critical values at any level can be calculated from the response surfaces, as by the computer program accompanying MacKinnon (1996) for the EG statistic $\tau_d(k)$ and by the one accompanying this paper for the ECM statistic $\kappa_d(k)$. Finally, response surfaces for commonly used quantiles (e.g. $p = 5\%$) are easily programmed into econometrics computer packages so as to provide empirical modelers with estimated finite sample critical values directly. For instance, PcGive and EViews have incorporated the response surfaces in MacKinnon (1991) for the Dickey–Fuller critical values, and (more recently) PcGive has added the response surfaces in Tables 2–5 below for $\kappa_d(k)$; see Hendry and Doornik (2001, pp. 231, 256).

Having estimated response surfaces of the form (26) for all experiments, it is relatively easy to plot estimated asymptotic distribution functions of the ECM statistic from the estimated values of θ_∞; see Section 4.1. *Finite sample* densities may be constructed from the Monte Carlo simulations directly, or from evaluation of (26) at finite sample sizes. Details of the numerical procedures for constructing the graphs appear in MacKinnon (1994, 1996).

While the response surfaces of the form (26) are convenient for constructing graphs of the asymptotic distributions, there are too many response surfaces to report them all: 10 608 response surfaces in total, i.e. $221 \times 4 \times 12$. For testing cointegration, however, response surfaces at common levels of significance are of particular interest, so Section 4.2 reports response surfaces for 1%, 5%, and 10% levels. These response surfaces parallel those in MacKinnon (1991) for the EG test statistic $\tau_d(k)$.

Characterizing each distribution function by 221 estimated quantiles is not the only way to summarize simulation results such as ours. An alternative approach—used by MacKinnon (1994) and Doornik (1998)—is to estimate parametric approximations to the distribution functions and report the parameter estimates. Because this approach requires storing far less information to calculate quantiles and critical values, it may be more convenient for implementation in econometric software packages. However, this approach also introduces approximation errors, to the extent that the estimated functional form inadequately captures the underlying distribution. Because little is known about the distribution of the ECM statistic, and in light of the complexity of dealing with both asymptotic and finite sample distributions, we adopted the current approach and report response surface estimates that give estimated quantiles as functions of the sample size. Finding convenient, accurate distributional approximations is a topic for further research.

4. MONTE CARLO RESULTS

This section graphs estimated densities for the ECM statistic (Section 4.1) and reports response surfaces for that statistic (Section 4.2). Section 4.3 then examines critical values for the ECM statistic that were previously estimated in the literature and shows that the response surfaces in Section 4.2 encompass and supersede much of that work.

4.1. Densities for the ECM statistic

Figures 1–8 plot asymptotic and finite sample densities for the ECM statistic; see Ericsson and MacKinnon (1999) for some corresponding cumulative distribution functions. Figures 1–4 begin with the asymptotic densities for $\kappa_{nc}(k)$, $\kappa_c(k)$, $\kappa_{ct}(k)$, and $\kappa_{ctt}(k)$, respectively. Each figure graphs the densities for $k = 1, \ldots, 12$, along with the density for $N(0, 1)$. Because the ECM statistic for $k = 1$ is the Dickey–Fuller statistic, that special case is labeled explicitly on the graphs as $\tau_d(1)$.

Several features are notable in Figures 1–4. First, the density shifts systematically in the negative direction as the number of variables k increases. The shift is numerically relatively constant, about -0.2 for an incremental increase in k, although the shift appears to be gradually declining in magnitude as k increases. Second, comparing densities across figures, the magnitude of the shift appears to decline as the number of deterministic components increases. Third, the discrepancy between the density of $N(0, 1)$ and that of the ECM statistic is considerable and increases as the number of deterministic components and stochastic variables increases. Thus, inferences about cointegration when using the ECM statistic would be hazardous if (e.g.) a standardized normal distribution were assumed. Fourth, the figures highlight the unique shape of the distribution of the Dickey–Fuller statistic. Figure 1 in particular brings out the asymmetry in the density of the Dickey–Fuller statistic $\tau_{nc}(1)$, a feature apparent in MacKinnon (1994, Figure 3) and also noted by Abadir (1995), both analytically and in his Figure 1.

As discussed in Section 2, comparisons of the ECM and EG procedures are of considerable interest. MacKinnon (1994, 1996) numerically estimated the distributions for the Dickey–Fuller statistic applied to the EG cointegration residuals. Figures 5–6 plot the asymptotic densities of the ECM and EG statistics for $k = 2$ and $k = 12$, where the densities of the EG statistic $\tau_d(k)$ are derived from MacKinnon's (1996) simulations. For all choices of deterministic components, the density of $\tau_d(k)$ is shifted to the left of that for $\kappa_d(k)$, substantially so for larger values of k. The density of $\tau_d(2)$ is shifted by only a few tenths relative to $\kappa_d(2)$, whereas that for $\tau_d(12)$ is often shifted by one to two units relative to $\kappa_d(12)$.

Figures 1–6 all concern asymptotic properties. While asymptotic properties are essential for understanding the nature of the ECM statistic, empirical sample sizes are often small, so it is valuable to assess the discrepancies between asymptotic and finite sample distributions. Figures 7 and 8 plot asymptotic and finite sample densities for $\kappa_d(2)$ and $\kappa_d(12)$, where 'finite sample' is $T = 20$ for $\kappa_d(2)$ and $T = 50$ for $\kappa_d(12)$. For all choices of deterministic components, the asymptotic densities tend to be more peaked than the finite sample ones, perhaps reflecting the contribution of estimation uncertainty to the latter. While the finite sample densities tend to shift to the left as the sample size increases, this does not hold uniformly for all parts of the density. Shifts to the right are notable in the left tails in Figure 7. Figures 7 and 8 each include the densities for all four possibilities of deterministic components for a given ECM statistic. Each additional deterministic component systematically shifts the statistic's density to the left, and the incremental shift is almost invariant to the total number of deterministic components.

4.2. Response surfaces for critical values

As just discussed, the distribution of the ECM statistic $\kappa_d(k)$ depends systematically on the number of variables k, the number of deterministic components d, and the sample size T. The current subsection quantifies these dependencies through response surfaces for three quantiles: those at 1%, 5%, and 10%.

To motivate these dependencies, consider Figure 9, which plots the data to be analyzed in the response surfaces. Specifically, each 3D graph in Figure 9 plots the within-experiment average for the estimated quantile against k and $(T^a/100)^{-1}$ (a rescaled inverse of the adjusted sample size), given the choice of d and the quantile's percent level p. The previously noted dependencies on d, k, T, and p are all apparent in Figure 9. Additionally, T appears to have relatively little effect on the 5% and 10% quantiles.

Tables 2–5 list the least squares estimates of the response surface coefficients θ_∞, θ_1, θ_2, and θ_3 for the 1%, 5%, and 10% quantiles with $k = 1, \ldots, 12$.[5] The conditional ECM is estimated with no deterministic terms (Table 2), with a constant term only (Table 3), with a constant term and a linear trend only (Table 4), and with a constant term, a linear trend, and a quadratic

[5]The response surfaces reported in Tables 2–5 differ slightly from those underlying Figures 1–8. The former are estimated by OLS, whereas the latter are estimated by GMM, using methods discussed in MacKinnon (1996). The two approaches yield numerically very similar results.

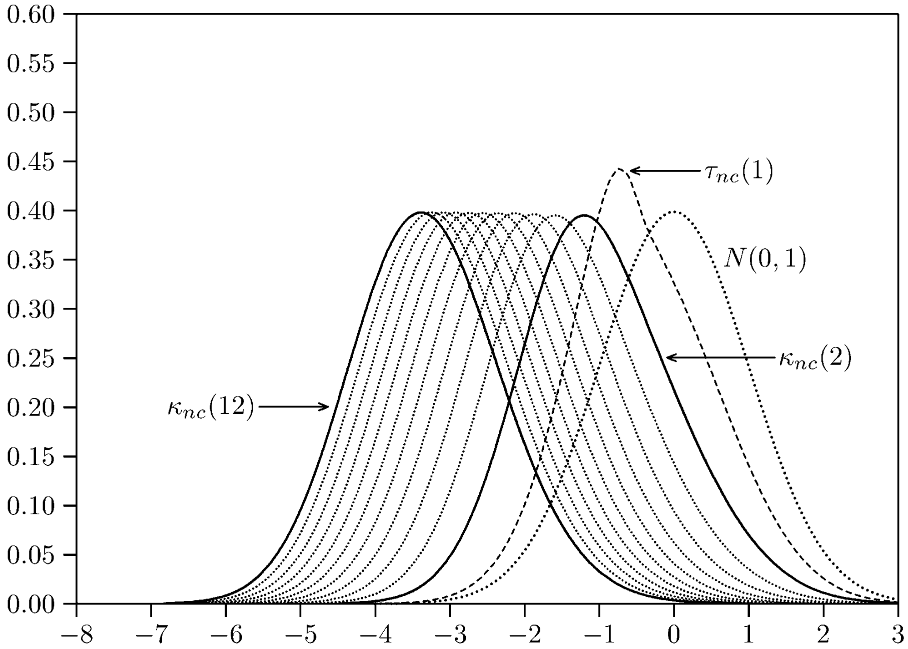

Figure 1. Asymptotic densities of the ECM statistic: no constant.

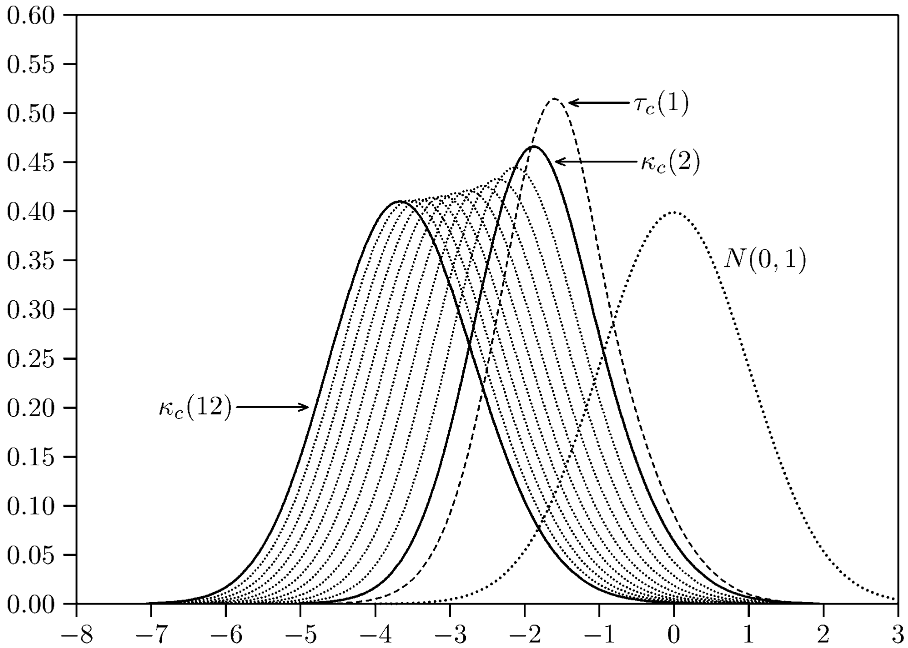

Figure 2. Asymptotic densities of the ECM statistic: constant only.

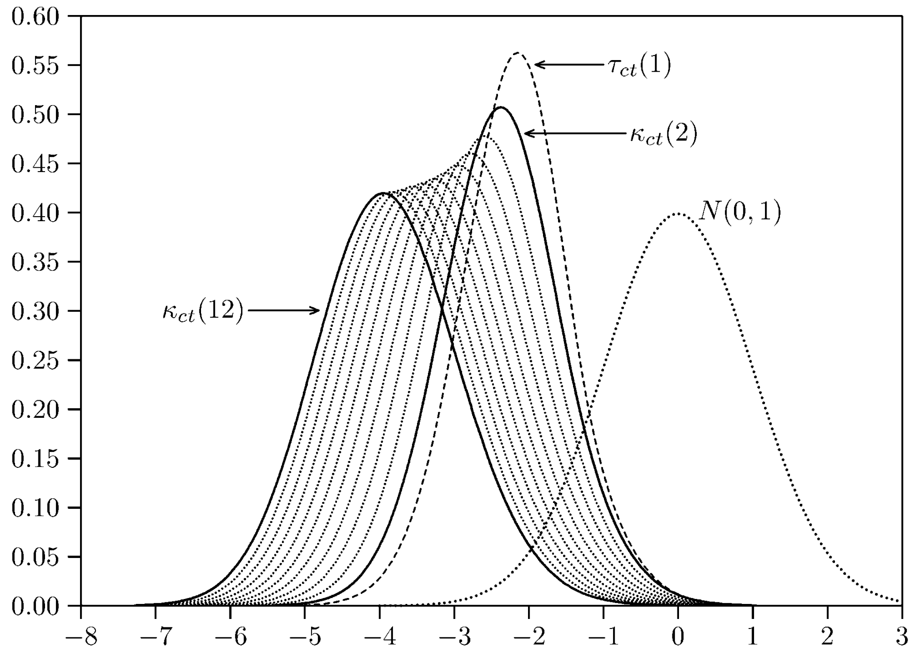

Figure 3. Asymptotic densities of the ECM statistic: constant and trend.

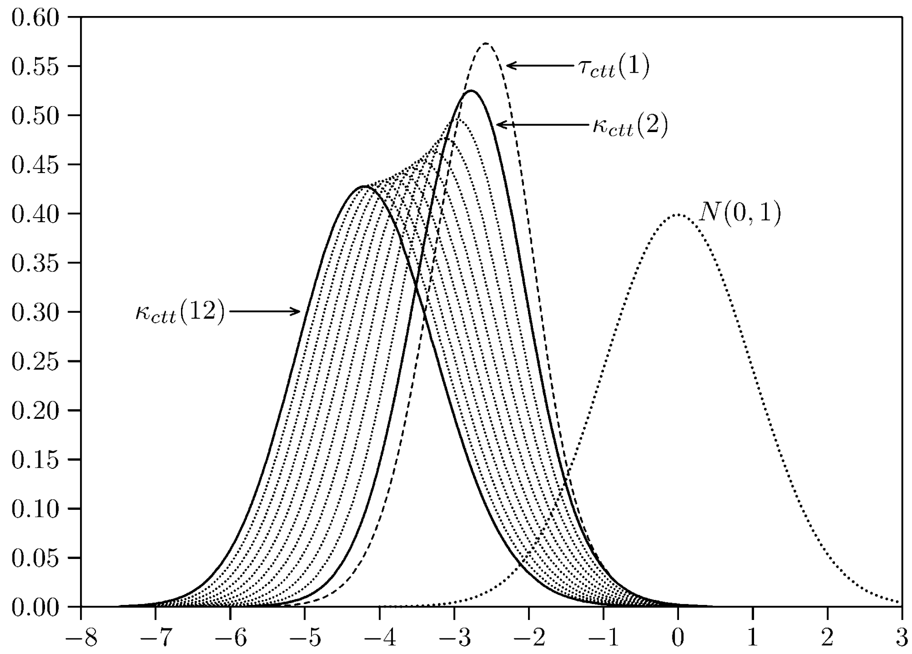

Figure 4. Asymptotic densities of the ECM statistic: constant, trend, and trend squared.

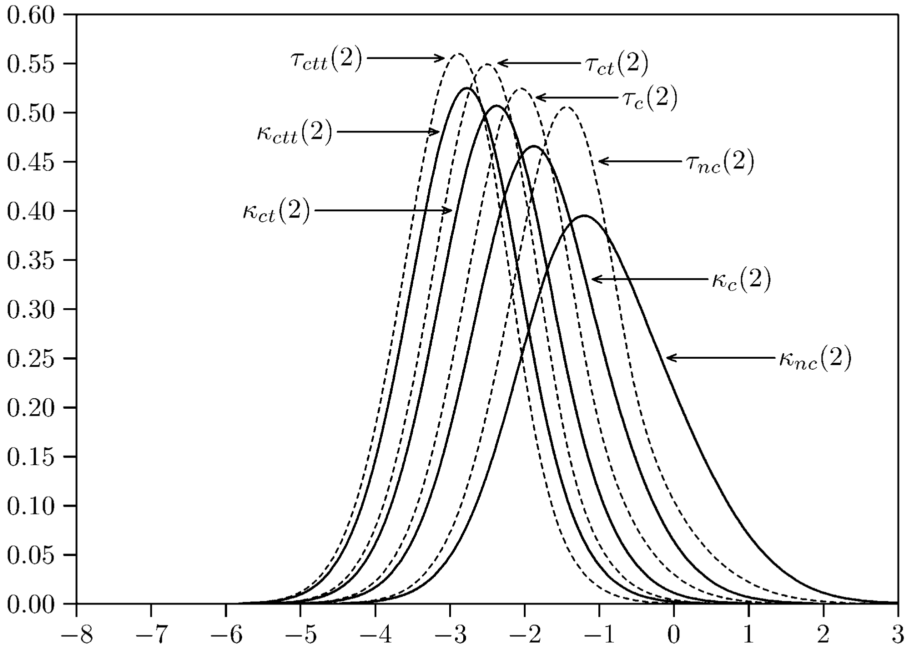

Figure 5. Asymptotic densities of the ECM and EG statistics: $k = 2$.

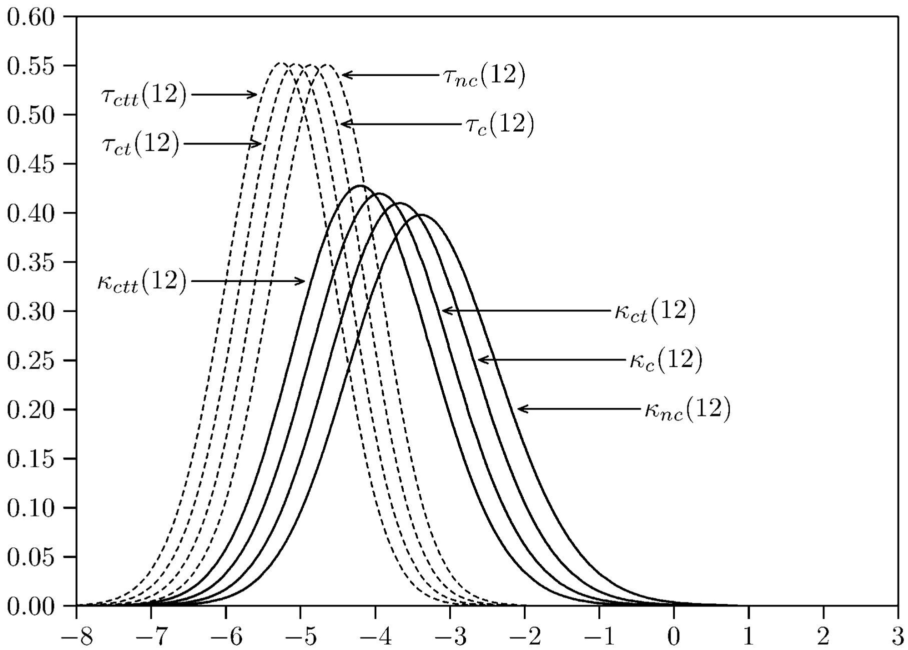

Figure 6. Asymptotic densities of the ECM and EG statistics: $k = 12$.

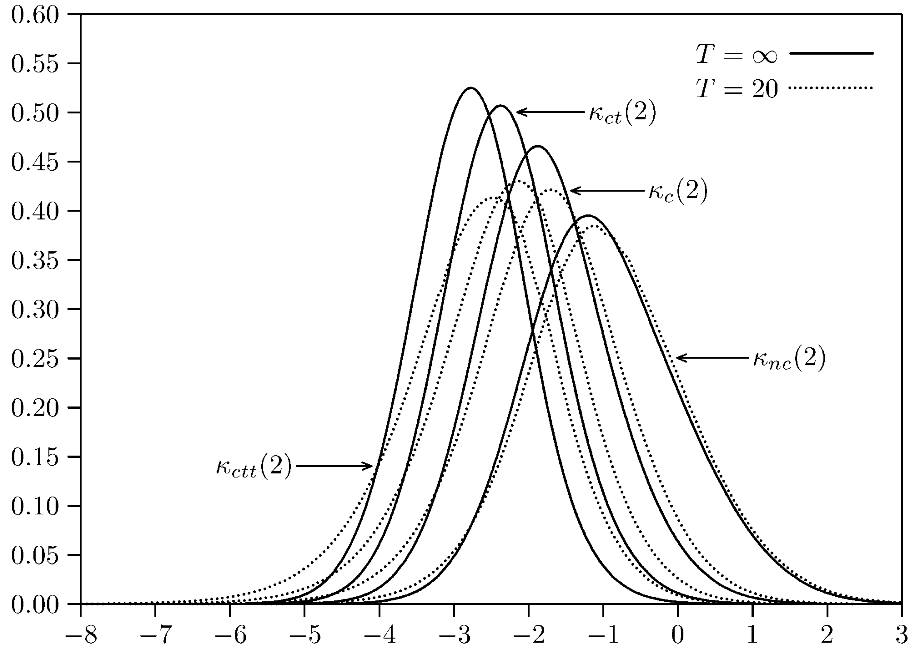

Figure 7. Asymptotic and finite sample densities of the ECM statistic: $k = 2$.

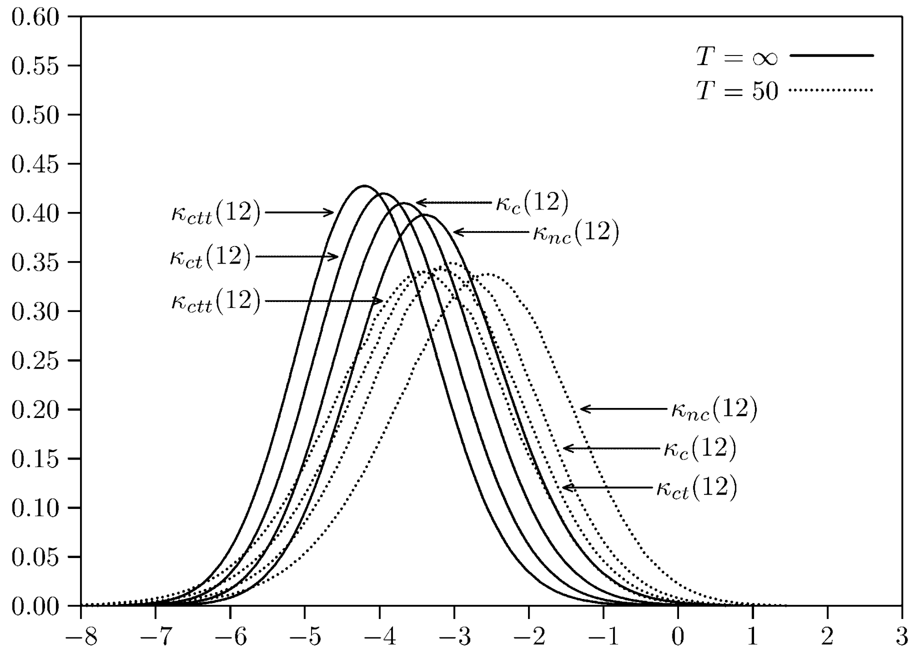

Figure 8. Asymptotic and finite sample densities of the ECM statistic: $k = 12$.

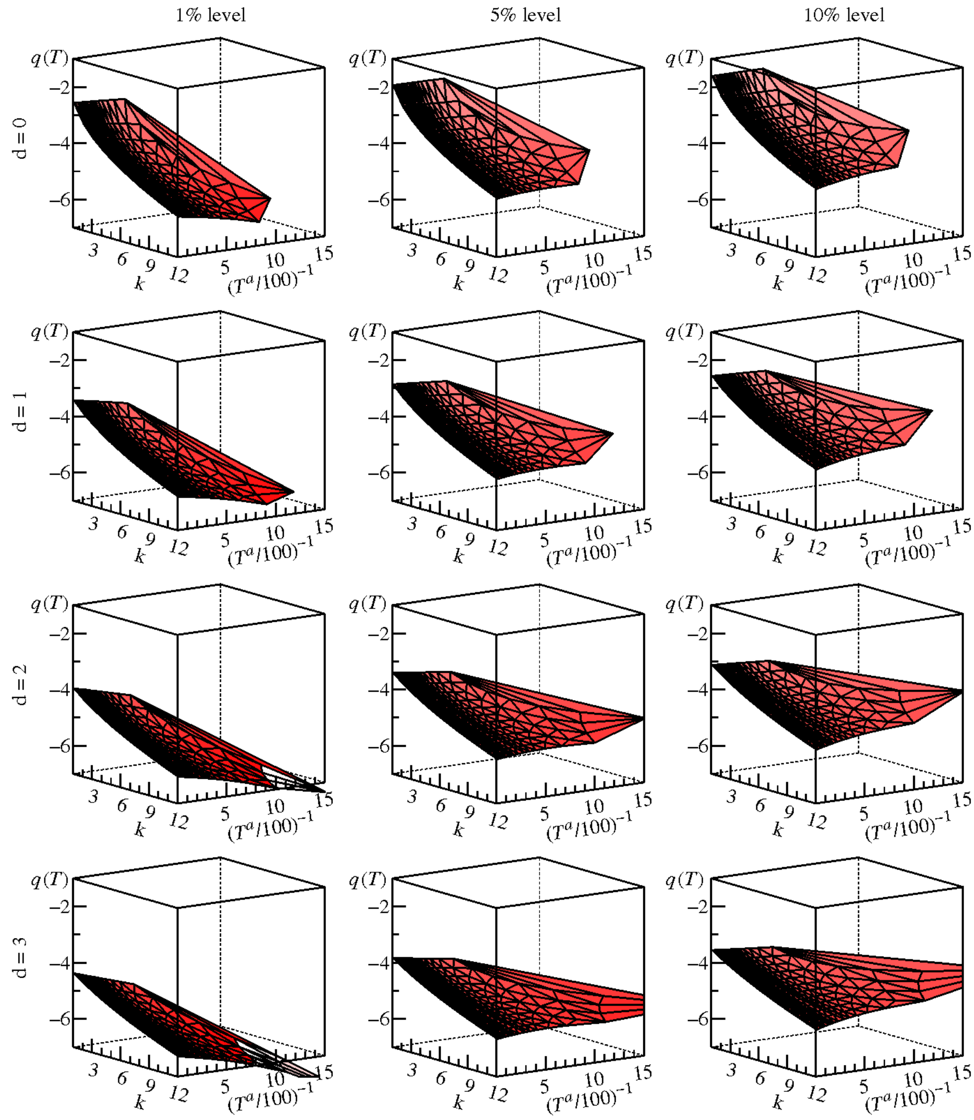

Figure 9. Estimated finite sample 1%, 5%, and 10% quantiles $q(T)$ for the ECM statistic as a function of d, k, and T^a.

trend (Table 5). The tables also include the estimated standard error ('s.e.') for θ_∞ to provide a measure of uncertainty for the estimated asymptotic quantile. This standard error is always smaller than 0.001, assuring high precision in the estimates. The estimated standard errors are jackknife heteroskedasticity consistent standard errors from MacKinnon and White (1985), as the experimental design may induce some heteroskedasticity in the estimated quantiles across different sample sizes.

Table 2. Response surface estimates for critical values of the ECM test of cointegration $\kappa_{nc}(k)$: no deterministic terms.

k	Size (%)	θ_∞	(s.e.)	θ_1	θ_2	θ_3	$\hat{\sigma}$
1	1	−2.5659	(0.0006)	−2.19	−3.6	26	0.00843
	5	−1.9408	(0.0003)	−0.35	0.6	−17	0.00430
	10	−1.6167	(0.0003)	0.23	−1.0	−6	0.00339
2	1	−3.2106	(0.0006)	−4.69	−10.5	48	0.00845
	5	−2.5937	(0.0003)	−1.53	−0.8	−24	0.00439
	10	−2.2643	(0.0003)	−0.41	−1.5	−9	0.00350
3	1	−3.6215	(0.0006)	−6.14	−5.3	−67	0.00892
	5	−3.0048	(0.0003)	−2.11	2.1	−61	0.00468
	10	−2.6744	(0.0003)	−0.57	1.2	−44	0.00372
4	1	−3.9433	(0.0006)	−7.15	−3.1	−69	0.00929
	5	−3.3268	(0.0003)	−2.04	−6.4	19	0.00455
	10	−2.9942	(0.0003)	−0.21	−5.1	13	0.00377
5	1	−4.2168	(0.0005)	−7.66	−2.1	−87	0.00920
	5	−3.5978	(0.0003)	−1.92	−3.6	−17	0.00502
	10	−3.2637	(0.0003)	0.25	−4.2	−15	0.00405
6	1	−4.4585	(0.0006)	−7.72	−7.2	−57	0.01034
	5	−3.8373	(0.0003)	−1.38	−7.7	−6	0.00519
	10	−3.5022	(0.0002)	1.15	−11.1	12	0.00397
7	1	−4.6763	(0.0005)	−7.78	−5.1	−73	0.01122
	5	−4.0535	(0.0003)	−0.76	−10.0	−7	0.00567
	10	−3.7165	(0.0002)	2.04	−14.7	15	0.00421
8	1	−4.8772	(0.0006)	−7.64	−2.4	−116	0.01035
	5	−4.2513	(0.0003)	−0.03	−12.0	−19	0.00543
	10	−3.9135	(0.0002)	3.10	−20.3	25	0.00420
9	1	−5.0634	(0.0006)	−7.13	−6.9	−113	0.01009
	5	−4.4363	(0.0003)	1.00	−18.4	−8	0.00534
	10	−4.0974	(0.0003)	4.46	−32.1	74	0.00422
10	1	−5.2381	(0.0006)	−6.68	−4.7	−149	0.01035
	5	−4.6093	(0.0003)	2.11	−25.4	10	0.00552
	10	−4.2693	(0.0003)	5.76	−38.2	72	0.00419
11	1	−5.4039	(0.0006)	−6.05	−7.1	−163	0.01038
	5	−4.7734	(0.0004)	3.37	−35.4	48	0.00556
	10	−4.4324	(0.0003)	7.33	−53.3	145	0.00426
12	1	−5.5598	(0.0006)	−5.10	−19.4	−75	0.01040
	5	−4.9279	(0.0004)	4.77	−48.8	109	0.00579
	10	−4.5864	(0.0003)	8.96	−68.0	204	0.00439

Table 3. Response surface estimates for critical values of the ECM test of cointegration $\kappa_c(k)$: with a constant term.

k	Size (%)	θ_∞	(s.e.)	θ_1	θ_2	θ_3	$\hat{\sigma}$
1	1	−3.4307	(0.0006)	−6.52	−4.7	−10	0.00790
	5	−2.8617	(0.0003)	−2.81	−3.2	37	0.00431
	10	−2.5668	(0.0003)	−1.56	2.1	−29	0.00332
2	1	−3.7948	(0.0006)	−7.87	−3.6	−28	0.00847
	5	−3.2145	(0.0003)	−3.21	−2.0	17	0.00438
	10	−2.9083	(0.0002)	−1.55	1.9	−25	0.00338
3	1	−4.0947	(0.0005)	−8.59	−2.0	−65	0.00857
	5	−3.5057	(0.0003)	−3.27	1.1	−34	0.00462
	10	−3.1924	(0.0002)	−1.23	2.1	−39	0.00364
4	1	−4.3555	(0.0006)	−8.90	−6.7	−31	0.00959
	5	−3.7592	(0.0003)	−2.92	−3.7	5	0.00484
	10	−3.4412	(0.0002)	−0.53	−4.5	4	0.00388
5	1	−4.5859	(0.0005)	−9.14	−2.5	−78	0.00970
	5	−3.9856	(0.0003)	−2.50	−1.7	−35	0.00493
	10	−3.6635	(0.0002)	0.21	−6.0	−8	0.00407
6	1	−4.7970	(0.0005)	−9.04	−5.6	−66	0.01100
	5	−4.1922	(0.0003)	−1.73	−7.8	−9	0.00514
	10	−3.8670	(0.0002)	1.26	−12.7	14	0.00402
7	1	−4.9912	(0.0005)	−8.85	−5.1	−72	0.01222
	5	−4.3831	(0.0003)	−0.90	−12.2	1	0.00606
	10	−4.0556	(0.0002)	2.39	−18.8	27	0.00437
8	1	−5.1723	(0.0006)	−8.58	−2.0	−113	0.01149
	5	−4.5608	(0.0003)	0.02	−15.4	−2	0.00571
	10	−4.2310	(0.0002)	3.59	−25.6	44	0.00427
9	1	−5.3437	(0.0006)	−7.86	−7.8	−101	0.01045
	5	−4.7287	(0.0003)	1.25	−26.0	42	0.00531
	10	−4.3975	(0.0002)	5.11	−39.2	104	0.00399
10	1	−5.5048	(0.0006)	−7.19	−9.8	−102	0.01059
	5	−4.8876	(0.0003)	2.46	−31.7	43	0.00545
	10	−4.5543	(0.0002)	6.53	−47.2	116	0.00438
11	1	−5.6588	(0.0006)	−6.39	−13.7	−105	0.01038
	5	−5.0394	(0.0004)	3.88	−45.7	117	0.00579
	10	−4.7055	(0.0003)	8.31	−66.5	222	0.00443
12	1	−5.8068	(0.0006)	−5.13	−29.2	−15	0.01060
	5	−5.1836	(0.0003)	5.33	−55.9	134	0.00555
	10	−4.8480	(0.0003)	9.94	−78.0	240	0.00431

Table 4. Response surface estimates for critical values of the ECM test of cointegration $\kappa_{ct}(k)$: with a constant term and a linear trend.

k	Size (%)	θ_∞	(s.e.)	θ_1	θ_2	θ_3	$\hat{\sigma}$
1	1	−3.9593	(0.0005)	−8.99	−4.9	39	0.00805
	5	−3.4108	(0.0003)	−4.38	4.5	−21	0.00412
	10	−3.1272	(0.0002)	−2.57	3.5	−7	0.00324
2	1	−4.2488	(0.0005)	−10.04	−4.1	−1	0.00845
	5	−3.6873	(0.0003)	−4.56	2.2	1	0.00442
	10	−3.3927	(0.0002)	−2.41	3.4	−14	0.00339
3	1	−4.4981	(0.0006)	−10.69	0.6	−58	0.00931
	5	−3.9263	(0.0003)	−4.47	5.2	−38	0.00474
	10	−3.6249	(0.0002)	−1.86	1.1	−10	0.00356
4	1	−4.7214	(0.0006)	−10.94	1.6	−77	0.00949
	5	−4.1421	(0.0003)	−3.99	2.8	−35	0.00496
	10	−3.8342	(0.0002)	−1.16	0.4	−23	0.00368
5	1	−4.9255	(0.0005)	−10.86	1.2	−94	0.01018
	5	−4.3392	(0.0003)	−3.37	1.6	−47	0.00510
	10	−4.0271	(0.0002)	−0.17	−4.4	−14	0.00406
6	1	−5.1137	(0.0005)	−10.72	1.4	−96	0.01145
	5	−4.5227	(0.0003)	−2.52	−2.8	−32	0.00536
	10	−4.2067	(0.0002)	0.94	−9.9	0	0.00415
7	1	−5.2923	(0.0005)	−10.11	−4.0	−75	0.01397
	5	−4.6952	(0.0003)	−1.43	−10.6	−5	0.00625
	10	−4.3751	(0.0002)	2.18	−16.9	18	0.00468
8	1	−5.4565	(0.0006)	−9.77	−1.5	−106	0.01202
	5	−4.8569	(0.0003)	−0.43	−14.4	−3	0.00593
	10	−4.5344	(0.0002)	3.52	−24.9	40	0.00453
9	1	−5.6149	(0.0006)	−9.11	−2.0	−126	0.01050
	5	−5.0108	(0.0003)	0.78	−21.2	12	0.00554
	10	−4.6864	(0.0003)	5.08	−37.2	88	0.00430
10	1	−5.7657	(0.0006)	−8.28	−5.3	−121	0.01180
	5	−5.1582	(0.0003)	2.12	−28.6	26	0.00558
	10	−4.8311	(0.0002)	6.62	−46.2	103	0.00438
11	1	−5.9099	(0.0006)	−7.41	−6.2	−160	0.01088
	5	−5.2992	(0.0003)	3.57	−40.0	69	0.00552
	10	−4.9707	(0.0003)	8.41	−64.7	199	0.00461
12	1	−6.0478	(0.0006)	−6.17	−20.6	−74	0.01111
	5	−5.4346	(0.0003)	5.22	−54.5	121	0.00605
	10	−5.1046	(0.0003)	10.20	−78.3	231	0.00451

Table 5. Response surface estimates for critical values of the ECM test of cointegration $\kappa_{ctt}(k)$: with a constant term, a linear trend, and a quadratic trend.

k	Size (%)	θ_∞	(s.e.)	θ_1	θ_2	θ_3	$\hat{\sigma}$
1	1	−4.3714	(0.0006)	−11.57	7.4	−66	0.00849
	5	−3.8324	(0.0003)	−5.90	9.3	−29	0.00430
	10	−3.5534	(0.0002)	−3.63	6.6	−7	0.00341
2	1	−4.6190	(0.0005)	−12.44	11.6	−130	0.00855
	5	−4.0683	(0.0003)	−5.90	9.3	−39	0.00445
	10	−3.7800	(0.0002)	−3.28	7.8	−36	0.00344
3	1	−4.8399	(0.0005)	−12.71	10.7	−136	0.00934
	5	−4.2790	(0.0003)	−5.56	9.3	−55	0.00481
	10	−3.9833	(0.0002)	−2.61	6.6	−42	0.00367
4	1	−5.0396	(0.0005)	−12.86	13.0	−149	0.01000
	5	−4.4716	(0.0003)	−4.95	6.9	−50	0.00496
	10	−4.1701	(0.0002)	−1.72	3.8	−37	0.00389
5	1	−5.2256	(0.0005)	−12.61	8.3	−121	0.01061
	5	−4.6498	(0.0003)	−4.23	5.7	−58	0.00536
	10	−4.3438	(0.0002)	−0.64	−1.0	−27	0.00401
6	1	−5.3998	(0.0005)	−12.12	4.3	−105	0.01270
	5	−4.8177	(0.0003)	−3.22	−0.4	−36	0.00590
	10	−4.5073	(0.0002)	0.60	−7.7	−7	0.00430
7	1	−5.5652	(0.0005)	−11.31	−4.0	−71	0.01776
	5	−4.9774	(0.0003)	−1.96	−9.3	−8	0.00671
	10	−4.6629	(0.0002)	2.02	−16.1	15	0.00494
8	1	−5.7181	(0.0006)	−10.97	0.8	−108	0.01143
	5	−5.1265	(0.0003)	−0.96	−10.9	−17	0.00584
	10	−4.8098	(0.0002)	3.41	−23.3	31	0.00457
9	1	−5.8656	(0.0006)	−10.32	4.3	−151	0.01103
	5	−5.2703	(0.0003)	0.33	−16.2	−17	0.00565
	10	−4.9510	(0.0003)	5.04	−35.4	74	0.00435
10	1	−6.0083	(0.0006)	−9.26	−4.0	−117	0.01218
	5	−5.4083	(0.0003)	1.80	−26.5	17	0.00589
	10	−5.0863	(0.0002)	6.63	−43.9	85	0.00440
11	1	−6.1449	(0.0006)	−8.26	−4.7	−158	0.01155
	5	−5.5415	(0.0003)	3.38	−39.0	60	0.00587
	10	−5.2176	(0.0003)	8.54	−63.7	179	0.00467
12	1	−6.2746	(0.0006)	−7.13	−13.5	−111	0.01281
	5	−5.6697	(0.0003)	5.12	−54.1	117	0.00615
	10	−5.3436	(0.0002)	10.36	−76.3	206	0.00462

The tables report an additional measure of uncertainty: $\hat{\sigma}$, the estimated equation standard error from the response surface (26). The estimate $\hat{\sigma}$ reflects both the simulation uncertainty from estimating the quantile $q(T_i)$ rather than knowing it, and the approximation error from using the cubic form in (26) rather than the true functional form. The experimental design also permits estimating the simulation uncertainty alone and so evaluating the statistical adequacy of the response surface. Specifically, the design implies 50 independent estimates of the quantile $q(T_i)$; see Section 3.3. A given response surface regression includes 50 values for $q(T_i)$ across all values of the sample size, entailing (e.g.) 1050 'observations' (50 values $\times$ 21 sample sizes) in the response surface for $\kappa_{ct}(2)$ at the 1% level. Thus, an average of the pure simulation uncertainty in a given response surface may be estimated by the equation standard error from a regression of $q(T_i)$ on a set of (e.g.) 21 dummies, one for each sample size. The response surface (26) is nested within this more general regression, and comparison of the two equations generates an F-statistic for testing the statistical significance of the error in approximating the response function's functional form; see Ericsson (1991). Even with the large number of replications, this test rejects at the 5% level for only seven of the 144 response surfaces (4.86%). No rejection increases $\hat{\sigma}$ by more than 1.1%. Thus, a cubic in the inverse of the adjusted sample size appears statistically and numerically adequate to approximate the simulated finite sample 1%, 5%, and 10% quantiles for this experimental design.

As they stand, Tables 2–5 provide a valuable tool for judging whether or not cointegration is present in empirical conditional ECMs. Econometric software packages could generate finite sample critical values for users from Tables 2–5, and PcGive already does so. Even without direct incorporation into an econometric package, calculation of critical values is trivially easy from the tables. For instance, for a conditional four-variable ($k = 4$) ECM with a constant term estimated on 47 observations, the finite sample critical value at the 5% level is -3.84, i.e. $-3.7592 - (2.92/39) - (3.7/39^2) + (5./39^3)$ from Table 3, noting that $T^a = 47 - 7 - 1 = 39$.

As with Figures 1–9, Tables 2–5 show the systematic and regular dependence of the properties of the ECM statistic on the number of variables k and the number of deterministic components d. This dependence leads to a simple rule of thumb that is captured in the following OLS 'meta' response surface for the asymptotic 1%, 5%, and 10% quantiles θ_∞ reported in Tables 2–5:

$$\theta_\infty = \underset{(0.03)}{-2.98} - \underset{(0.003)}{0.187}\,k - \underset{(0.01)}{0.33}\,(d-1) + \left\{\begin{array}{ll} \underset{(0.02)}{-0.60} & \text{at the 1\% level} \\ 0 & \text{at the 5\% level} \\ \underset{(0.02)}{+0.32} & \text{at the 10\% level} \end{array}\right\} \qquad (27)$$

$$R^2 = 0.987 \quad \hat{\sigma} = 0.109 \quad \text{number of 'observations'} = 132,$$

where R^2 is the squared multiple correlation coefficient, and $\hat{\sigma}$ is the standard deviation of the residuals. Values of θ_∞ for $k = 1$ (the Dickey–Fuller statistic) are excluded in (27), as only $k > 1$ is of interest for testing cointegration.

From (27), a crude approximation θ_{crude} to the lower 5% critical value for $\kappa_d(k)$ is:

$$\theta_{\text{crude}} = -3.0 - 0.2k - 0.3(d-1). \qquad (28)$$

The negative coefficients in (28) can be easily remembered as a '3/2/3' rule of thumb: an intercept of -3.0, a coefficient of -0.2 on the number of variables in x, and a coefficient of -0.3

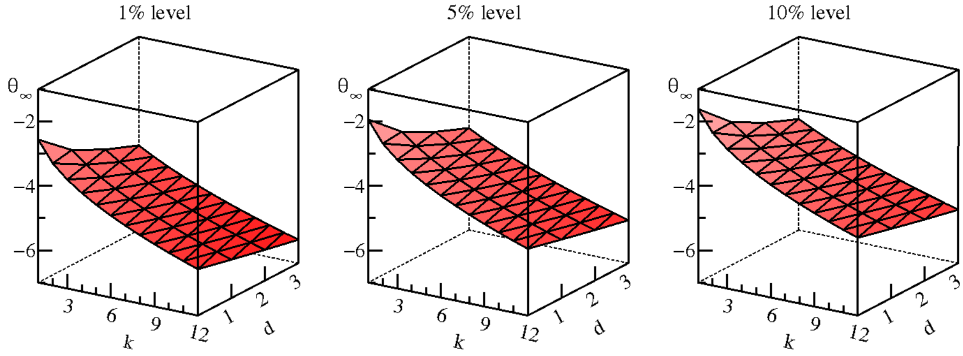

Figure 10. Estimated asymptotic 1%, 5%, and 10% quantiles θ_∞ for the ECM statistic as a function of k and d.

on the number of deterministic terms over and above a constant term. For the ECM evaluated earlier ($k = 4$, $d = 1$, $T = 47$), θ_{crude} is -3.8, deviating by only 0.04 from the value of -3.84 calculated with Table 3. While deviations between θ_{crude} and $q(T_i)$ may be larger or smaller than this for other k, d, and T, it is well worth keeping in mind that, with typical macroeconomic data, the ECM statistic itself can easily fluctuate by a few tenths, simply by adding or dropping a few observations from the sample.

Figure 10 highlights this near-linear dependence of the asymptotic quantile θ_∞ on k and d. Each 3D graph in Figure 10 plots θ_∞ against k and d, given the quantile's percent level p. The surfaces are virtually planar except for the Dickey–Fuller statistic ($k = 1$), which is excluded from (27) and (28).

The asymptotic moments of the ECM statistic also show marked regularity in the distribution's behavior. Figure 11 plots its asymptotic mean, standard deviation, skewness, and excess kurtosis as a function of k and d.[6] The asymptotic mean declines by approximately 0.2 and 0.4 respectively for unit increases in k and d, close to the estimated shifts for the critical values in (27). While the asymptotic standard deviation, skewness, and excess kurtosis also depend on k and d, those dependencies are numerically much smaller than that of the asymptotic mean. For all values of k and d examined, the asymptotic standard deviation is close to unity, and the asymptotic skewness and excess kurtosis are close to zero. These results reconfirm the visual characterization from Figures 1–8: the distribution of the ECM statistic $\kappa_d(k)$ is relatively close to normality with unit variance. In light of these observations, parametric distributional approximations to the distribution of the ECM statistic may be promising—perhaps using the normal distribution, Student's t-distribution, or an expansion thereon.

Equations (27) and (28) quantify the straightforward dependencies of the ECM statistic's quantiles on k and d, they provide a mechanism for extrapolating critical values for values of k and d outside the experimental design (25), and they offer a rough-and-ready way of assessing empirical results when Tables 2–5 are not available. Preferably, though, Tables 2–5 or the related computer program should be used.

[6]The asymptotic moments were calculated by response surfaces from a separate set of Monte Carlo experiments, following an approach like that used for the quantiles. Monte Carlo estimation of the statistic's finite sample moments does assume the existence of those moments. However, even if those moments are infinite, their Monte Carlo estimates may be close to the (finite) moments of a Nagar approximation to the statistic; see Sargan (1982).

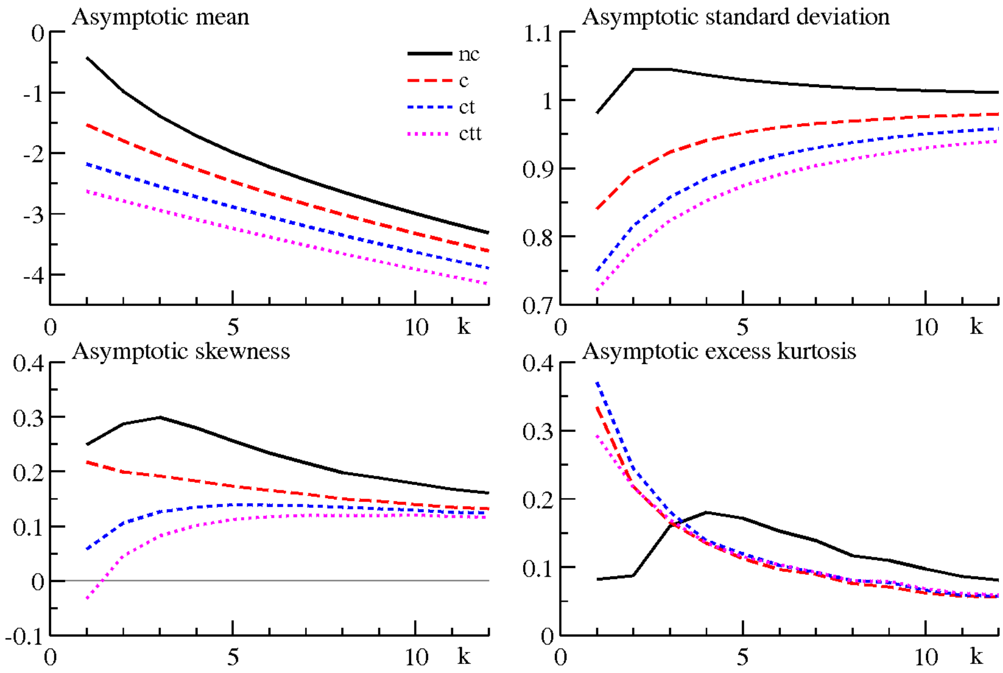

Figure 11. The asymptotic mean, standard deviation, skewness, and excess kurtosis for the ECM statistic as a function of k and d.

4.3. Encompassing previous Monte Carlo results

Two previous studies—Banerjee *et al.* (1993) and Banerjee *et al.* (1998)—report estimated critical values for the ECM statistic. This subsection shows that these previous results for the 1%, 5%, and 10% levels are superseded by the response surfaces reported in Tables 2–5. Simulation uncertainty in these two studies appears to be the dominant factor explaining discrepancies relative to the response surfaces in Tables 2–5. In this encompassing approach, many pages of existing independent Monte Carlo simulations are subsumed by the current paper's results. That is both progressive research-wise and efficient space-wise.

Pre-existing Monte Carlo studies are encompassed by evaluating the response surfaces in Tables 2–5 over the experimental designs of the past studies and comparing the critical values derived from Tables 2–5 with those reported in the studies' simulations. Deviations between the two types of critical values typically are small relative to the estimated simulation uncertainty of the pre-existing Monte Carlo studies or are simply small numerically. Hence, Tables 2–5 encompass those studies. For this purpose, the simulation uncertainty associated with the response surfaces in Tables 2–5 is treated as negligible. That assumption seems reasonable. The largest value of $\hat{\sigma}$ in Tables 2–5 is under 0.02, and each (T, d, k, p) quadruplet includes 50 estimates of the quantile, implying an associated standard error of the response surface quantile of under 0.003. Frequently, that standard error is under 0.001. The remainder of this subsection briefly describes the Monte Carlo simulations in each study and the outcomes of the encompassing exercise.

Banerjee *et al.* (1993, Table 7.6, p. 233) report estimated critical values at the 1%, 5%, and 10% levels for $\kappa_c(2)$ at $T = (25, 50, 100)$, using 5000 replications per experiment. Deviations relative to the response surfaces from Table 3 are all under 0.1 in absolute value. Using the values of $\hat{\sigma}$ in Table 3 as a benchmark and rescaling by the square root of the ratio of simulations calculated, the estimated standard errors for the three quantiles in Banerjee *et al.* (1993) are approximately 0.063, 0.032, and 0.025. The observed discrepancies between the estimated quantiles in Banerjee *et al.* (1993) and those calculated from Table 3 appear as expected, given the simulation uncertainty of the former.

Banerjee *et al.* (1998, Table I) report estimated critical values at the 1%, 5%, 10%, and 25% levels for $\kappa_c(k)$ and $\kappa_{ct}(k)$ $(k = 2, \ldots, 6)$ at $T = (25, 50, 100, 500, \infty)$, using 25 000 replications per experiment. Deviations relative to the response surfaces from Tables 3 and 4 are all under 0.2 in absolute value, and are typically 0.04 or smaller in magnitude. The estimated standard errors for the 1%, 5%, and 10% quantiles in Banerjee *et al.* (1998) are approximately 0.028, 0.014, and 0.011.

5. TWO EMPIRICAL APPLICATIONS

This section applies the finite sample critical values derived earlier and the computer program for calculating p-values to two empirical ECMs. Section 5.1 considers a model of UK narrow money demand from Hendry and Ericsson (1991), and Section 5.2 a model of US federal government debt from Hamilton and Flavin (1986). (Ericsson and MacKinnon (1999) also assess the model of UK consumers' expenditure from Davidson *et al.* (1978).) The model in Hendry and Ericsson (1991) has played a significant role in the literature on ECMs and cointegration, and Hamilton and Flavin (1986) was one of the early papers to employ unit root statistics for testing economic hypotheses. Each subsection briefly reviews the estimated equation and considers corresponding conditional ECM tests. Tables summarize the results, reporting the empirical t-values for testing cointegration, along with critical values and p-values. Use of the critical values from Tables 2–5 for the ECM statistic affects the economic inferences drawn.

Several issues arise in testing for cointegration in these models. First, the ECM for money demand was derived from an unrestricted ADL. Both the ADL and the ECM allow testing of cointegration, although the ECM requires slight modification to apply the critical values from Tables 2–5. Second, dynamic specification affects the degrees of freedom used in estimation. Hence, when computing critical values, the adjusted sample size T^a is calculated as $T - h$ (rather than as $T - (2k + d - 1)$), where h is the total number of regressors, including deterministic variables. The calculation of p-values utilizes h similarly. Third, the choice of deterministic variables affects the t-values and the corresponding critical values and p-values, so potentially affecting inference. Finally, nonlinearity of the deterministic trend and lack of weak exogeneity are important in the model of government debt. Throughout this section, capital letters denote both the generic name and the level of a variable, logarithms are in lowercase, and OLS standard errors are in parentheses.

5.1. UK narrow money demand

Hendry and Ericsson (1991, equation (6)) model UK narrow money demand as a conditional

ECM, whose final parsimonious form is as follows:

$$\Delta(m-p)_t = -\underset{(0.125)}{0.687}\,\Delta p_t - \underset{(0.058)}{0.175}\,\Delta(m-p-i)_{t-1} - \underset{(0.060)}{0.630}\,R_t^{\text{net}}$$

$$-\underset{(0.0085)}{0.0929}\,(m-p-i)_{t-1} + \underset{(0.0040)}{0.0234} \qquad (29)$$

$$T = 100 \text{ (1964Q3–1989Q2)} \quad R^2 = 0.76 \quad \hat{\sigma} = 1.313\%.$$

The data are nominal narrow money M_1 (M, in £ millions), real total final expenditure (TFE) at 1985 prices (I, in £ millions), the TFE deflator (P, 1985 = 1.00), and the net interest rate (R^{net}, in percent per annum expressed as a fraction). The last series is the difference between the three-month local authority interest rate and a learning-adjusted retail sight-deposit interest rate.

While the t-value on the error correction term $(m-p-i)_{t-1}$ in (29) is very large and negative (-10.87), significance levels are not known, given the presence of nuisance parameters; see Kremers *et al.* (1992) and Kiviet and Phillips (1992). This difficulty arises because one of the coefficients in the cointegrating vector—the long-run income elasticity—is constrained. One solution is to estimate that coefficient unrestrictedly, as occurs when estimating (29) with i_{t-1} added:

$$\Delta(m-p)_t = -\underset{(0.128)}{0.702}\,\Delta p_t - \underset{(0.058)}{0.178}\,\Delta(m-p-i)_{t-1} - \underset{(0.067)}{0.611}\,R_t^{\text{net}}$$

$$-\underset{(0.0113)}{0.0882}\,(m-p-i)_{t-1} + \underset{(0.0104)}{0.0065}\,i_{t-1} - \underset{(0.117)}{0.049} \qquad (30)$$

$$T = 100 \text{ (1964Q3–1989Q2)} \quad R^2 = 0.76 \quad \hat{\sigma} = 1.317\%.$$

The t-value on $(m-p-i)_{t-1}$ in (30) is -7.78, which is significant at the 1% level for $\kappa_c(4)$, with critical value of -4.45. In fact, the finite sample p-value for -7.78 is 0.0000.

Equations (29) and (30) can be derived from an unrestricted fifth-order ADL model in $m-p$, Δp, i, and R^{net}. The ECM statistic for that ADL is -5.17, also significant at the 1% level for $\kappa_c(4)$, with critical value of -4.47. Its finite sample p-value is 0.0014, suggesting a minor loss in power from estimating additional coefficients on dynamics relative to (30).

Both this fifth-order ADL and the ECM in (30) include one deterministic component: a constant term. Table 6 reports the statistic $\kappa_d(4)$ for the four choices of deterministic components considered in the sections earlier; the value of h; the finite sample, asymptotic, and crude critical values at the 1%, 5%, and 10% levels; finite sample and asymptotic p-values; the estimated equation standard error $\hat{\sigma}$; and an F-statistic for testing the significance of omitted deterministic components. The symbols $^+$, *, and ** denote rejection at the 10%, 5%, and 1% levels, respectively. With a constant term, linear trend, and quadratic trend included, the statistic $\kappa_{ctt}(4)$ is insignificant at the 10% level for both the ADL and the ECM: their p-values are 0.3859 and 0.4544. With fewer deterministic components, cointegration is detected at the 0.5% level or smaller in the ADL and the ECM, as the statistics $\kappa_{ct}(4)$, $\kappa_c(4)$, and $\kappa_{nc}(4)$ show.

The final column in Table 6 lists the F-statistics for testing the significance of the omitted deterministic components in the corresponding regressions, relative to the regressions for obtaining $\kappa_{ctt}(4)$: degrees of freedom for the F-statistics appear in parentheses as $F(\cdot,\cdot)$, and the statistics' p-values are in brackets $[\cdot]$. These F-statistics indicate that the constant term, linear trend, and quadratic trend are statistically insignificant, so all the reported ECM statistics in Table 6 make statistically justifiable assumptions about these deterministic components. The

Table 6. Empirical t-values, critical values, and p-values for the ECM statistic: models of UK narrow money demand.

Statistic Model or calculation	Empirical t-value	h	Critical value 1%	Critical value 5%	Critical value 10%	p-value Finite sample	p-value Asymptotic	$\hat{\sigma}$ (%)	F-statistic vs. the model for $\kappa_{ctt}(4)$
$\kappa_{ctt}(4)$									
ADL	−3.29	26	−5.21	−4.54	−4.19	0.3859	0.4140	1.313	—
ECM	−3.14	8	−5.18	−4.52	−4.19	0.4544	0.4819	1.326	—
Asymptotic	—	—	−5.04	−4.47	−4.17	—	—	—	—
Crude	—	—	−5.0	−4.4	−4.1	—	—	—	—
$\kappa_{ct}(4)$									
ADL	-5.14^{**}	25	−4.87	−4.19	−3.85	0.0047	0.0024	1.306	$F(1, 74) = 0.11$ [0.74]
ECM	-6.53^{**}	7	−4.84	−4.18	−3.85	0.0000	0.0000	1.320	$F(1, 92) = 0.16$ [0.69]
Asymptotic	—	—	−4.72	−4.14	−3.83	—	—	—	—
Crude	—	—	−4.7	−4.1	−3.8	—	—	—	—
$\kappa_{c}(4)$									
ADL	-5.17^{**}	24	−4.47	−3.80	−3.45	0.0014	0.0006	1.301	$F(2, 74) = 0.28$ [0.76]
ECM	-7.78^{**}	6	−4.45	−3.79	−3.45	0.0000	0.0000	1.317	$F(2, 92) = 0.37$ [0.69]
Asymptotic	—	—	−4.36	−3.76	−3.44	—	—	—	—
Crude	—	—	−4.4	−3.8	−3.5	—	—	—	—
$\kappa_{nc}(4)$									
ADL	-6.10^{**}	23	−4.04	−3.35	−3.00	0.0000	0.0000	1.297	$F(3, 74) = 0.36$ [0.78]
ECM	-10.57^{**}	5	−4.02	−3.35	−3.00	0.0000	0.0000	1.311	$F(3, 92) = 0.31$ [0.82]
Asymptotic	—	—	−3.94	−3.33	−2.99	—	—	—	—
Crude	—	—	−4.1	−3.5	−3.2	—	—	—	—

statistics $\kappa_{nc}(4)$, $\kappa_c(4)$, and $\kappa_{ct}(4)$ reject at standard levels, but $\kappa_{ctt}(4)$ does not, pointing to the value of parsimony in deterministic components for obtaining increased power of the cointegration test, when parsimony is merited. The insignificance of a linear trend is particularly interesting. In a system analysis of this dataset, Hendry and Mizon (1993) find a second cointegrating vector, which includes a linear trend; but in their system model, that cointegrating vector does not enter the equation for money.

Table 6 lists the asymptotic and crude critical values at the 1%, 5%, and 10% levels, and these differ by at most 0.21 from the calculated finite sample critical values. Likewise, the finite sample and asymptotic p-values in the table differ by only modest amounts. These numerically small discrepancies are not surprising because the sample size is relatively large ($T = 100$).

5.2. US federal government debt

The second model is an ADL from Hamilton and Flavin (1986, p. 816), relating real US federal government debt to a deterministic nonlinear trend or 'bubble' $(1 + r)^t$ and the budget surplus:

$$
\begin{aligned}
B_t = \underset{(26.40)}{48.41} - \underset{(21.29)}{22.68}\,(1+r)^t + \underset{(0.21)}{0.69}\,B_{t-1} + \underset{(0.24)}{0.20}\,B_{t-2} \\
- \underset{(0.13)}{1.30}\,S_t - \underset{(0.31)}{0.63}\,S_{t-1} \qquad (31) \\
T = 23\ (1962\text{–}1984) \quad R^2 = 0.98 \quad \hat{\sigma} = 7.405.
\end{aligned}
$$

The data are the adjusted debt (B) for the end of the fiscal year and the adjusted surplus (S) for the fiscal year (both in \$ millions, 1967 prices). The variable r is set to 0.0112, the average *ex post* real interest rate on US government bonds over 1960–84. The coefficient on $(1 + r)^t$ is statistically insignificant, consistent with the absence of a speculative bubble. From this and related evidence, Hamilton and Flavin (1986, pp. 816–817) conclude that '. . . the data appear quite consistent with the assertion that the government has historically operated subject to the constraint that expenditures not exceed receipts in expected present-value terms'.

This interpretation of the evidence assumes a long-run solution to (31) relating debt and surplus. That is equivalent to assuming both cointegration between B and S, and the presence of the corresponding cointegrating vector in (31). Empirically, however, (31) does not support cointegration of B and S. Rewriting (31) as an unrestricted ECM yields the following equation:

$$
\begin{aligned}
\Delta B_t = \underset{(26.40)}{48.41} - \underset{(21.29)}{22.68}\,(1+r)^t - \underset{(0.076)}{0.104}\,B_{t-1} - \underset{(0.24)}{0.20}\,\Delta B_{t-1} \\
- \underset{(0.13)}{1.30}\,\Delta S_t - \underset{(0.36)}{1.92}\,S_{t-1} \qquad (32) \\
T = 23\ (1962\text{–}1984) \quad R^2 = 0.94 \quad \hat{\sigma} = 7.405.
\end{aligned}
$$

The t-value on B_{t-1} is -1.36, which is insignificant at the 10% level for $\kappa_{ct}(2)$, with critical value of -3.53. Using the critical value for $\kappa_{ct}(2)$ assumes that $(1 + r)^t$ is well approximated by a linear trend, which, visually, it is. Alternatively, the 10% critical value for $\kappa_{ctt}(2)$ is -3.95, again with no rejection. The finite sample p-values under these two alternative assumptions are 0.8386 and 0.9247. Notably, estimating (32) (or (31)) with t and t^2 rather than with $(1 + r)^t$ obtains a statistically significantly better fitting model, pointing to mis-specification in (32).

Table 7 reports the t-values and critical values for (32) with various choices of deterministic components. The bubble $(1 + r)^t$ is statistically insignificant in (32), whereas a linear trend and quadratic trend in its stead are statistically significant. Even so, the resulting t-value for $\kappa_{ctt}(2)$ is -2.96, which is insignificant at the 10% level, having a p-value of 0.3689. Cointegration does not appear to hold in this conditional model, undercutting the economic inferences drawn by Hamilton and Flavin (1986).

The sample size in (32) is small: $T = 23$. Correspondingly, the finite sample adjustments for critical values are typically larger numerically in Table 7 than in Table 6, with the largest adjustment being -0.72 at the 1% level, i.e. about two thirds of a standard error in the t-value. The p-values have small finite sample adjustments, which mainly reflect each reported t-value being far from the lower tail of the associated density; cf. Figure 7.

The single-equation results in Table 7 all assume that S is weakly exogenous, whereas S does not appear to be so empirically. Starting with a second-order VAR in B and S, a single cointegrating vector is apparent from the Johansen procedure when $(1 + r)^t$ or a linear trend is restricted to lie in the cointegration space. Weak exogeneity of S is rejected, as is that of B, invalidating cointegration analysis in a conditional single equation such as (31). Without weak exogeneity,

Table 7. Empirical t-values, critical values, and p-values for the ECM statistic: models of US federal government debt.

Statistic Model or calculation	Empirical t-value	h	Critical value 1%	5%	10%	p-value Finite sample	Asymp-totic	$\hat{\sigma}$	F-statistic vs. the model for $\kappa_{ctt}(2)$
$\kappa_{ctt}(2)$									
ADL + bubble	−1.36	6	−5.34	−4.39	−3.95	0.9247	0.9651	7.40	—
ADL	−2.96	7	−5.38	−4.41	−3.96	0.3689	0.4121	6.37	—
Asymptotic	—	—	−4.62	−4.07	−3.78	—	—	—	—
Crude	—	—	−4.6	−4.0	−3.7	—	—	—	—
$\kappa_{ct}(2)$									
ADL + bubble	−1.36	6	−4.85	−3.95	−3.53	0.8386	0.8947	7.40	—
ADL	−1.38	6	−4.85	−3.95	−3.53	0.8308	0.8886	7.38	$F(1, 16) = 6.82$ [0.02]
Asymptotic	—	—	−4.25	−3.69	−3.39	—	—	—	—
Crude	—	—	−4.3	−3.7	−3.4	—	—	—	—
$\kappa_{c}(2)$									
ADL + bubble	—	—	—	—	—	—	—	—	—
ADL	−1.50	5	−4.25	−3.40	−2.99	0.5944	0.6458	7.43	$F(2, 16) = 4.26$ [0.03]
Asymptotic	—	—	−3.79	−3.21	−2.91	—	—	—	—
Crude	—	—	−4.0	−3.4	−3.1	—	—	—	—
$\kappa_{nc}(2)$									
ADL + bubble	—	—	—	—	—	—	—	—	—
ADL	+2.58	4	−3.48	−2.68	−2.29	0.9984	0.9992	7.94	$F(3, 16) = 4.52$ [0.02]
Asymptotic	—	—	−3.21	−2.59	−2.26	—	—	—	—
Crude	—	—	−3.7	−3.1	−2.8	—	—	—	—

single equation inference about cointegration is hazardous at best; and testing the implied exogeneity assumptions is clearly important. For example, in the Johansen procedure, the coefficient on the bubble $(1 + r)^t$ or on the linear trend is statistically significant and negatively related to B, whichever type of trend is included. That contrasts with the statistical insignificance of the coefficient on $(1 + r)^t$ in (31). Furthermore, the negative coefficient on the trend is economically surprising and puzzling, although it may be indicative of certain non-ergodic features of the data: see Kremers (1988) *inter alia*.

In summary, the first empirical analysis illustrates the importance of parsimony, both in the choice of deterministic terms and in the reduction from an ADL to a simpler ECM. The second analysis shows that mis-specification can render inference hazardous, even when the mis-specification is indirect, as with a violation of weak exogeneity. Imposition of valid restrictions on the cointegrating vector may increase power, although asymptotically correct critical values for such ECM statistics have been derived only for the case when all cointegrating coefficients are known; see Hansen (1995, Table 1).

6. CONCLUSIONS

This paper has assessed the distributional properties of the ECM statistic for testing cointegration. Graphs and response surfaces provide complementary summaries of the vast array of results from the Monte Carlo study undertaken. Both the graphs and the response surfaces highlight some simple dependencies of the quantiles on the number of variables in the ECM, the choice of deterministic components, and the estimation sample size. The reported response surfaces provide a computationally convenient way for calculating finite sample critical values at the 1%, 5%, and 10% levels. The response surfaces also encompass and supersede much of the literature's previous estimates of critical values for the ECM statistic. A computer program, freely available over the Internet, can be used to calculate p-values and critical values at any level. Empirical conditional ECMs are ubiquitous in the cointegration literature, so these tools should be of immediate use to the empirical modeler. Two previous empirical studies illustrate how critical values and p-values for the ECM statistic can be employed in practice, and how their use may affect economic inferences.

Several limitations of the current study come to mind, thereby suggesting some possible extensions. First, the model's lag order is assumed to be (and is) unity throughout the Monte Carlo analysis. For longer lags, the adjusted sample size may be corrected for additional degrees of freedom lost in estimation and thence used to calculate critical values from a response surface, as in Section 5. This refinement may not be sufficient in itself, so an extended analysis, such as in Cheung and Lai (1995) for the Dickey–Fuller statistic, may be required. Second, all of the ECM statistics with deterministic components have those components fully unconstrained in estimation. In analyzing similar statistics, Harbo *et al.* (1998) and Doornik *et al.* (1998) argue strongly for constraining the highest-order deterministic component to lie in the cointegration space, so distributional properties for so constrained versions of $\kappa_c(k)$, $\kappa_{ct}(k)$, and $\kappa_{ctt}(k)$ are of interest. That said, virtually all empirically calculated ECM statistics to date have been with unconstrained deterministic components. Finally, the current paper has considered the properties of the ECM statistic only under the null of no cointegration. While Banerjee *et al.* (1993), Campos *et al.* (1996), and Banerjee *et al.* (1998) present some calculations on the power of the ECM statistic, further analysis could be illuminating, particularly comparisons with the Johansen procedure and the EG procedure under various assumptions about weak exogeneity and common factor restrictions.

ACKNOWLEDGEMENTS

The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System. JGM's research was supported in part by grants from the Social Sciences and Humanities Research Council of Canada. An early version of this paper appeared under the title 'Finite Sample Properties of Error Correction Tests for Cointegration', and an intermediate version was circulated as Ericsson and MacKinnon (1999). We are grateful to Shaghil Ahmed, David Bowman, Jon Faust, David Hendry, Søren Johansen, Fred Joutz, Andy Levin, Jaime Marquez, Bent Nielsen, John Rogers, Neil Shephard, an editor, and an anonymous referee for helpful comments and discussion; to Hayden Smith and Sebastian Thomas for research assistance; and to Jurgen Doornik and David Hendry for providing us with a beta-test version of GiveWin Version 2. Monte Carlo simulations and the graphs of the densities were obtained from modified versions of programs for MacKinnon (1994, 1996). Response

surfaces were obtained using PcGive Professional Version 9.2, and 3D graphics were generated from GiveWin Version 2.02: see Doornik and Hendry (2001). The paper's tables of response surface coefficients and the Excel spreadsheet and Fortran program for calculating critical values and p-values are available from JGM's home page.

REFERENCES

Abadir, K. M. (1995). The limiting distribution of the t ratio under a unit root. *Econometric Theory 11*, 775–93.

Banerjee, A., J. J. Dolado, J. W. Galbraith and D. F. Hendry (1993). *Co-integration, Error Correction, and the Econometric Analysis of Non-stationary Data*. Oxford: Oxford University Press.

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*, 267–83.

Banerjee, A. and D. F. Hendry (eds) (1996). The econometrics of economic policy, Special issue. *Oxford Bulletin of Economics and Statistics 58*, 567–819.

Boswijk, H. P. (1994). Testing for an unstable root in conditional and structural error correction models. *Journal of Econometrics 63*, 37–60.

Boswijk, H. P. and P. H. Franses (1992). Dynamic specification and cointegration. *Oxford Bulletin of Economics and Statistics 54*, 369–81.

Campos, J., N. R. Ericsson and D. F. Hendry (1996). Cointegration tests in the presence of structural breaks. *Journal of Econometrics 70*, 187–220.

Cheung, Y.-W. and K. S. Lai (1995). Lag order and critical values of the augmented Dickey-Fuller test. *Journal of Business and Economic Statistics 13*, 277–80.

Davidson, J. E. H., D. F. Hendry, F. Srba and S. Yeo (1978). Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *Economic Journal 88*, 661–92.

Dickey, D. A. and W. A. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica 49*, 1057–72.

Doornik, J. A. (1998). Approximations to the asymptotic distributions of cointegration tests. *Journal of Economic Surveys 12*, 573–93. Erratum in (1999) *13*, i.

Doornik, J. A. and D. F. Hendry (2001). *PcGive Version 10 for Windows*. London: Timberlake Consultants Press.

Doornik, J. A., D. F. Hendry and B. Nielsen (1998). Inference in cointegrating models: UK M1 revisited. *Journal of Economic Surveys 12*, 533–72.

Elliott, G. (1995). Tests for the correct specification of cointegrating vectors and the error correction model. Mimeo. La Jolla, CA: Department of Economics, University of California at San Diego.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica 55*, 251–76.

Ericsson, N. R. (1991). Monte Carlo methodology and the finite sample properties of instrumental variables statistics for testing nested and non-nested hypotheses. *Econometrica 59*, 1249–77.

Ericsson, N. R. (ed.) (1998). Exogeneity, cointegration, and economic policy analysis. Special section, *Journal of Business and Economic Statistics 16*, 369–449.

Ericsson, N. R. and J. G. MacKinnon (1999). Distributions of error correction tests for cointegration. International finance discussion paper no. 655, Board of Governors of the Federal Reserve System, Washington, D.C. December, available on the WorldWide Web at `www.federalreserve.gov/pubs/ifdp/1999/655/default.htm`.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*. New York: John Wiley and Sons.

Hamilton, J. D. and M. A. Flavin (1986). On the limitations of government borrowing: a framework for empirical testing. *American Economic Review 76*, 808–19.

Hammersley, J. M. and D. C. Handscomb (1964). *Monte Carlo Methods*. London: Chapman and Hall.

Hansen, B. E. (1995). Rethinking the univariate approach to unit root testing: using covariates to increase power. *Econometric Theory 11*, 1148–71.

Harbo, I., S. Johansen, B. Nielsen and A. Rahbek (1998). Asymptotic inference on cointegrating rank in partial systems. *Journal of Business and Economic Statistics 16*, 388–99.

Hendry, D. F. (1984). Monte Carlo experimentation in econometrics. In Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics*, Ch. 16, vol. 2, pp. 937–76. Amsterdam: North-Holland.

Hendry, D. F. (1989). *PC-GIVE: An Interactive Econometric Modelling System*, Version 6.0/6.01. Oxford: Institute of Economics and Statistics and Nuffield College, University of Oxford.

Hendry, D. F. (1995). On the interactions of unit roots and exogeneity. *Econometric Reviews 14*, 383–419.

Hendry, D. F. and J. A. Doornik (2001). *Empirical Econometric Modelling Using PcGive 10*, vol. 1. London: Timberlake Consultants Press.

Hendry, D. F. and N. R. Ericsson (1991). Modeling the demand for narrow money in the United Kingdom and the United States. *European Economic Review 35*, 833–81. With discussion.

Hendry, D. F. and G. E. Mizon (1993). Evaluating dynamic econometric models by encompassing the VAR. In P. C. B. Phillips (ed.), *Models, Methods, and Applications of Econometrics: Essays in Honor of A. R. Bergstrom*, Ch. 18, pp. 272–300. Cambridge, MA: Blackwell Publishers.

Hendry, D. F., A. Pagan and J. D. Sargan (1984). Dynamic specification. In Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics*, Ch. 18, vol. 2, pp. 1023–1100. Amsterdam: North-Holland.

Hoover, K. D. and S. J. Perez (1999). Data mining reconsidered: encompassing and the general-to-specific approach to specification search. *Econometrics Journal 2*, 167–91. With discussion.

Horvath, M. T. K. and M. W. Watson (1995). Testing for cointegration when some of the cointegrating vectors are prespecified. *Econometric Theory 11*, 984–1014.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control 12*, 231–54.

Johansen, S. (1992a). Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics 52*, 389–402.

Johansen, S. (1992b). Testing weak exogeneity and the order of cointegration in UK money demand data. *Journal of Policy Modeling 14*, 313–34.

Johansen, S. (1994). The role of the constant and linear terms in cointegration analysis of nonstationary variables. *Econometric Reviews 13*, 205–29.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Johansen, S. and K. Juselius (1990). Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and Statistics 52*, 169–210.

Kiviet, J. F. and G. D. A. Phillips (1992). Exact similar tests for unit roots and cointegration. *Oxford Bulletin of Economics and Statistics 54*, 349–67.

Kremers, J. J. M. (1988). Long-run limits on the US federal debt. *Economics Letters 28*, 259–62.

Kremers, J. J. M., N. R. Ericsson and J. J. Dolado (1992). The power of cointegration tests. *Oxford Bulletin of Economics and Statistics 54*, 325–48.

L'Ecuyer, P. (1988). Efficient and portable combined random number generators. *Communications of the ACM 31*, 742–9, 774.

Lütkepohl, H. and J. Wolters (eds) (1998). Money demand in Europe. *Empirical Economics 23*, 263–524 (special issue).

MacKinnon, J. G. (1991). Critical values for cointegration tests. In R. F. Engle and C. W. J. Granger (eds), *Long-run Economic Relationships: Readings in Cointegration*, Ch. 13, pp. 267–76. Oxford: Oxford University Press.

MacKinnon, J. G. (1994). Approximate asymptotic distribution functions for unit-root and cointegration tests. *Journal of Business and Economic Statistics 12*, 167–76.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics 11*, 601–18.

MacKinnon, J. G., A. A. Haug and L. Michelis (1999). Numerical distribution functions of likelihood ratio tests for cointegration. *Journal of Applied Econometrics 14*, 563–77.

MacKinnon, J. G. and H. White (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics 29*, 305–25.

Marsaglia, G. and T. A. Bray (1964). A convenient method for generating normal variables. *SIAM Review 6*, 260–4.

Nielsen, B. and A. Rahbek (2000). Similarity issues in cointegration analysis. *Oxford Bulletin of Economics and Statistics 61*, 5–22.

Osterwald-Lenum, M. (1992). A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. *Oxford Bulletin of Economics and Statistics 54*, 461–72.

Pesaran, M. H., Y. Shin and R. J. Smith (2000). Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics 97*, 293–343.

Phillips, A. W. (1954). Stabilisation policy in a closed economy. *Economic Journal 64*, 290–323.

Sargan, J. D. (1964). Wages and prices in the United Kingdom: a study in econometric methodology. In P. E. Hart, G. Mills and J. K. Whitaker (eds), *Econometric Analysis for National Economic Planning*, vol. 16 of *Colston Papers*, pp. 25–54. London: Butterworths (with discussion).

Sargan, J. D. (1982). On Monte Carlo estimates of moments that are infinite. In R. L. Basmann and G. F. Rhodes, Jr (eds), *Advances in Econometrics: A Research Annual*, vol. 1, pp. 267–99. Greenwich, CT: JAI Press.

Zivot, E. (2000). The power of single equation tests for cointegration when the cointegrating vector is prespecified. *Econometric Theory 16*, 407–39.