7-1-1988

Estimation and Inference in Models of Cointegration: A Simulation Study

Bruce E. Hansen

Peter C.B. Phillips

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station

New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 881

ESTIMATION AND INFERENCE IN MODELS OF COINTEGRATION: A SIMULATION STUDY

by

Bruce E. Hansen

&

Peter C. B. Phillips

July 1988

ESTIMATION AND INFERENCE IN MODELS OF COINTEGRATION: A SIMULATION STUDY

by

Bruce E. Hansen and Peter C. B. Phillips[1]

Cowles Foundation for Research in Economics
Yale University

0. ABSTRACT

This paper studies the finite sample distributions of estimators of the cointegrating vector of linear regression models with I(1) variables. Attention is concentrated on the least squares (OLS) and instrumental variables (IV) methods analyzed in other recent work (Phillips and Hansen (1988)). The general preference of OLS to IV techniques suggested by asymptotic theory is reinforced by our simulations. An exception arises for cases of low signal to noise, where spurious IV techniques (so named for their use of instruments that are structurally unrelated to the model) outperform uncorrected least squares. We verify the presence of a small sample estimation bias and show that the Park-Phillips bias correction does reduce the magnitude of this problem. We also find that there is substantial distributional divergence of t-statistics from the normal, unless the Phillips-Hansen endogeneity correction is used. Finally, we apply these methods to aggregate consumption and income data. Our empirical results indicate that the endogeneity and serial dependence corrections are important and lead to intuitively plausible changes in the estimated coefficients.

July 1988

---

[1]We are grateful to Glena Ames for her skill and effort in keyboarding the manuscript of this paper and to the NSF for research support.

1. INTRODUCTION

In our recent paper (Phillips and Hansen (1988)) we studied the asymptotic distributions of a large class of estimators of the "cointegrating vector" of linear regression models with I(1) variables. These estimators included ordinary least squares (OLS), standard instrumental variables (IV) and "spurious" instrumental variables. All were found to be "super-consistent" under quite general assumptions, including endogeneity in the regressors and serial correlation in the innovations. It was shown that neither bias corrections--see Phillips (1987a) and Park and Phillips (1987a, 1987b)--nor IV techniques could overcome substantial problems of nuisance parameter dependencies and non-normalities in the asymptotic distributions of the standardized statistics, except in special leading cases. Instead, a semi-parametric endogeneity correction which is asymptotically equivalent to maximum likelihood--see Phillips (1988b)--was derived which solves these problems. These "fully modified" estimators have asymptotic mixed normal distributions. This permits quite general hypothesis tests using conventional techniques.

This paper attempts a systematic investigation of the small sample properties of these methods through Monte Carlo simulations. Of course, due to the wealth of possible data generating processes it might be unwise to make strong general claims from these results. Nevertheless, we feel that several conclusions can be drawn from this study. First, if uncorrected estimates are compared, IV estimation may beat OLS under strong endogeneity or high noise. Second, if the true residuals are known (for instance, if the true regression coefficients are known from the null hypothesis) then the fully modified OLS estimates perform extremely well, and are the preferred estimation method. Third, feasible modified estimates (based on estimated residuals) unfortunately work much less well, although better than the uncorrected estimates. Fourth, conventional t-statistics are quite misleading in the sense that their distributions are far from the standard normal while fully modified t-statistics are well approximated by a normal distribution, although with a variance slightly greater than unity.

We also report the results of a simple application of these methods to the aggregate "consumption function." We find that fully modified statistics cannot reject a unit coefficient on income.

Our notation follows that of Phillips and Hansen (1988). We use the symbol " $\Rightarrow$ " to signify weak convergence, and the symbol " $\equiv$ " to signify equality in distribution. Stochastic processes such as the Brownian motion $B(r)$ on $[0,1]$ are frequently written as B to achieve notational economy. Similarly, we write integrals with respect to Lebesgue measure such as $\int_0^1 B(s)ds$ more simply as $\int_0^1 B$. Vector Brownian motion with covariance matrix Ω is written " $BM(\Omega)$." We use I(1) and I(0) to signify time series of order one and zero, respectively.

2. PRELIMINARY THEORY

The results presented here are substantially simplified from the theory in Phillips and Hansen (1988) in order to ease presentation and focus on the particular model used in the simulation. Consider the process

$$y_t = a_1 + a_2 t + a_3' x_t + u_{1t}$$

$$x_t = x_{t-1} + u_{2t} , \quad m \times 1$$

$$z_t = z_{t-1} + u_{3t} , \quad m \times 1 .$$

The innovation vector $u_t = (u_{1t}, u_{2t}', u_{3t}')'$ is assumed to be strictly stationary and ergodic with zero mean, finite covariance matrix $\Sigma > 0$ and continuous spectral density matrix $f_{uu}(\lambda)$ with $\Omega = 2\pi f_{uu}(0)$. We also assume that the partial sum process constructed from u_t satisfies the multivariate invariance principle

(1) $$T^{-1/2}\Sigma_1^{[Tr]} u_j \Rightarrow B(r) \equiv BM(\Omega) , \quad 0 < r \leq 1 .$$

We decompose the "long-run" covariance matrix as follows:

$$\Omega = \Sigma + \Lambda + \Lambda'$$

where

$$\Sigma = E(u_0 u_0') , \quad \Lambda = \sum_{k=1}^{\infty} E(u_0 u_k')$$

and we define

$$\Delta = \Sigma + \Lambda .$$

See Phillips (1987b) for a review of the conditions under which (1) holds. We partition B , Ω , Σ , Λ and Δ conformably with u_t . For example, in the case of Ω we write

$$(2) \qquad \Omega = \begin{bmatrix} \omega_{11} & \omega_{21}' & \omega_{31}' \\ \omega_{21} & \Omega_{22} & \Omega_{32}' \\ \omega_{31} & \Omega_{32} & \Omega_{33} \end{bmatrix} .$$

We also make the strong assumption that z_t and x_t are cointegrated. Then the Brownian motions B_2 and B_3 are related linearly

$$B_2 = \Omega_{23}\Omega_{33}^{-1}B_3$$

and the time series can be written in a reduced form as

$$(3) \qquad x_t = \Omega_{23}\Omega_{33}^{-1}z_t + v_t \ , \quad v_t \equiv I(0) \ .$$

This condition may be equivalently expressed as a restriction upon (2)

$$\Omega_{22} = \Omega_{23}\Omega_{33}^{-1}\Omega_{32} \ .$$

Relationship (3) suggests that we can use z_t as an instrument for x_t , analogous to classic 2SLS. See Phillips and Hansen (1988) for a further discussion.

We are interested in estimates of $a = (a_1, a_2, a_3')'$. A linear combination of y_t and x_t are trend stationary, yet each are I(1), thus y_t and x_t are cointegrated in the terminology of Engle and Granger (1987). Following Phillips and Hansen (1988) there are four natural estimators to consider:

[1] OLS

$$\hat{a} = (X'X)^{-1}(X'Y)$$

$$X = [1,\ t,\ x_t]_{t=1,\ldots,T}$$

$$Y = [y_t]_{t=1,\ldots,T}$$

[2] "Standard" IV

$$\tilde{a} = (Z'X)^{-1}(Z'Y)$$

$$Z = [1,\ t,\ z_t]_{t=1,\ldots,T}$$

[3] "Spurious" IV with stochastic instruments

$$\tilde{\tilde{a}} = (X'P_S X)^{-1}(X'P_S Y)$$

$$P_S = S(S'S)^{-1}S$$

$$S = [1,\ t,\ s_t]_{t=1,\ldots,T}$$

$$s_t = s_{t-1} + \zeta_t \ , \quad \{\zeta_t\}_1^T \text{ iid } N(0,\ I_{n_1})$$

[4] "Spurious" IV with deterministic instruments

$$\bar{a} = (X'P_K X)^{-1}(X'P_K Y)$$

$$P_K = K(K'K)^{-1}K'$$

$$K = [1,\ t,\ k_{Tt}]_{t=1,\ldots,T}$$

where k_{Tt} is a deterministic n_2-dimensional function of time (t) and possibly sample size (T) . We consider both polynomials in time:

$$t^{p_1},\ t^{p_2},\ \ldots$$

and sinusoids

$$\sin(2\pi\lambda_1 t/T),\ \cos(2\pi\lambda_1 t/T),\ \text{etc.}$$

For further discussion, see our earlier paper. The relevant condition which these variables must satisfy is absence of asymptotic collinearity.

Estimators [3] and [4] defined above are "spurious" since there is no structural relationship between the regressors and the instruments. A surprising result of Phillips and Hansen (1988) is that these estimators are consistent. A brief digression on this point seems warranted. Consider the standard just-identified IV estimator under conditions of stationarity, orthogonality, and identification:

$$\hat{a} - a = \frac{\Sigma_1^T z_t u_t}{\Sigma_1^T z_t x_t} = \frac{0_p(T^{1/2})}{0_p(T)} \equiv 0_p(T^{-1/2}) \ .$$

Under stationarity, both orthogonality between z_t and u_t and identification (relevance of z_t to x_t) are required for the denominator and numerator to have the stochastic orders indicated above. If either orthogonality or identification fails, the numerator and denominator are of the same stochastic order and the estimate converges to the "wrong" value in the first case, or a Cauchy-type distribution in the second (see Phillips (1987c)). If both conditions fail simultaneously, then the estimates diverge at rate $T^{1/2}$.

In our model, however, the story is quite different. As long as x_t and s_t are I(1), and $u_t \equiv I(0)$, then

$$\Sigma_1^T s_t u_t = 0_p(T)$$

$$\Sigma_1^T s_t x_t = 0_p(T^2)$$

regardless of any other assumption, yielding consistent estimation. The reason why spurious deterministic instruments work is quite analogous. In fact, one may regard these results as beneficial artifacts of the problem of "spurious regressions"--see Phillips (1986). The generalizations to multivariate regression with deterministic components is straightforward and is presented in our earlier paper.

The consistency of spurious IV estimation may appear to conflict with standard approaches to identification in simultaneous equation systems. Under the 2SLS interpretation of IV estimation, we have the reduced form (in matrix notation):

$$(y,X) = S(\pi_1, \pi_2) + (v_1 , v_2) .$$

We are accustomed to think of $\pi = (\pi_1, \pi_2)$ as parameterizing the conditional mean of (y,X) given S. Then the coefficient vector a is identified by the relation

$$\pi_1 - \pi_2 a = 0$$

when (and only when) π_2 has full rank.

This does not have a meaningful interpretation in our model since S is independent of the data. We can instead interpret π in terms of linear projections:

$$\begin{aligned}(y,X) &= P_S(y,X) + (I - P_S)(y,X) \\ &= S\hat{\pi} + \hat{v}\end{aligned}$$

where

$$P_S = S(S'S)^{-1}S'$$

$$\hat{\pi} = (\hat{\pi}_1, \hat{\pi}_2)$$

$$\hat{\pi}_1 = (S'S)^{-1}S'y$$

$$\hat{\pi}_2 = (S'S)^{-1}S'X \ .$$

Since

$$\hat{\pi}_1 \Rightarrow \pi_1^0 \equiv \left[\int_0^1 B_s B_s'\right]^{-1} \int_0^1 B_s dB_2' a$$

$$\hat{\pi}_2 \Rightarrow \pi_2^0 \equiv \left[\int_0^1 B_s B_s'\right]^{-1} \int_0^1 B_s dB_2'$$

we see that cointegration of y and X guarantees that the identifiability relationship holds in the limit. That is

$$(\hat{\pi}_1 - \hat{\pi}_2 a) \not\to_p 0$$

and a is therefore asymptotically identified using the instrument set S if π_2^0 is of full rank (almost surely), which is shown by Phillips and Hansen (1988, Lemma A3). In this framework, the limit representation π^0 is not fixed but is a random matrix--see Phillips (1986). Identification is therefore a beneficial artifact of the spurious regression phenomenon as indicated above.

Although it is known that OLS and IV are consistent under substantial endogeneity, it is also known that correlation between u_t and lagged values of Δx_t introduces a second-order bias effect. Specifically, if $x_t \equiv I(1)$ and $u_t \equiv I(0)$, then

(4) $$T^{-1}\Sigma_1^T x_t u_t \Rightarrow \int_0^1 B_2 dB_1 + \Delta_{21} .$$

To derive (4) rigorously, see Phillips (1988a). More intuitively, the presence of Δ_{21} can be explained by noting the correlation

$$\Delta_{21} = \sum_{k=0}^{\infty} E(\Delta x_{-k} u_0) = E(x_0 u_0) .$$

For example, consider the simple process

$$x_t = x_t^* + u_t$$

$$x_t^* = x_{t-1}^* + \epsilon_t$$

where u_t and ϵ_t are mutually independent white noise. Then

$$T^{-1}\Sigma_1^T x_t u_t = T^{-1}\Sigma_1^T x_t^* u_t + T^{-1}\Sigma_1^T u_t^2$$

(5) $$\Rightarrow \int_0^1 B_2 dB_1 + \sigma_u^2 .$$

Conditional on $\mathcal{F}_2 = \sigma(B_2(r),\ 0 < r \leq 1)$, (5) is distributed as

$$N(\sigma_u^2,\ \sigma_u^2 \int_0^1 B_2^2) \equiv \sigma_u N(\sigma_u,\ \int_0^1 B_2^2) .$$

Although the non-centrality is eliminated asymptotically since both OLS and IV are consistent, one would expect some evidence of bias to appear in finite samples. This has, in fact, been shown to be the case in the simulations of Banerjee *et al.* (1986).

To correct this bias, and permit inference via asymptotic distribution theory, Phillips (1987a) and Park and Phillips (1987a, 1987b) have proposed semi-parametric estimation of Δ_{21} . This eliminates the bias effect asymptotically. We observe that these corrections are not relevant for spurious IV methods, as the instruments are strictly exogenous for the regression errors.

Specifically, using the residuals $\hat{u}_t$ from a (consistent) first-stage regression, we can estimate Δ_{21} and Δ_{31} by

$$\hat{\Delta}_{21} = T^{-1}\sum_{k=0}^{\ell}\sum_{t=k+1}^{T}\Delta x_{t-k}\hat{u}_t$$

$$\hat{\Delta}_{31} = T^{-1}\sum_{k=0}^{\ell}\sum_{t=k+1}^{T}\Delta z_{t-k}\hat{u}_t$$

where $\ell \to \infty$ as $T \to \infty$ such that $\ell = o(T^{1/4})$. We then define the "bias-corrected" estimators

$$\hat{a}^* = (X'X)^{-1}[X'Y - e_m T\hat{\Delta}_{21}]$$

$$\tilde{a}^* = (Z'X)^{-1}[Z'Y - e_m T\hat{\Delta}_{31}]$$

where

$$e_m = \overset{m}{\begin{bmatrix} 0 \\ I \end{bmatrix}}\begin{matrix} 2 \\ m \end{matrix} .$$

One can show that $\hat{a}^*$ and $\tilde{a}^*$ are consistent and asymptotically unbiased up to $O_p(T^{-1})$ --see Phillips and Hansen (1988). The asymptotic distributions, however, are generally non-normal and dependent upon nuisance parameters. Both arise from the long-run endogeneity of the regressors. We may write

$$(6) \qquad \int_0^1 B_2 dB_1 \equiv \int_0^1 B_2 dB_2 \Omega_{22}^{-1}\Omega_{21} + \int_0^1 B_2 dW\omega_{11\cdot 2}^{1/2}$$

where

$W \equiv BM(1)$ and independent of B_2 and

$$\omega_{11\cdot 2} = \omega_{11} - \omega_{12}\Omega_{22}^{-1}\omega_{21}$$

is the conditional long-run variance of u, given Δx. In (6), the first stochastic integral on the right-hand-side is of the "unit root" form, while the second stochastic integral has a mixture normal distribution. Consider a "bias-corrected" u_t :

$$u_t^+ = u_t - \omega_{12}\Omega_{22}^{-1}\Delta x_t$$

which has zero long-run correlation with Δx_t. Then

$$T^{-1}\Sigma_1^T x_t u_t^+ - \hat{\Delta}_{21}^+ \Rightarrow \omega_{11\cdot 2}^{1/2}\int_0^1 B_2 dW \equiv \int_{G>0} N(0,\ \omega_{11\cdot 2}G)dP(G)\ ,\quad G \equiv \int_0^1 B_2 B_2'$$

where

$$\hat{\Delta}_{21}^+ = \hat{\Delta}_{21} - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\hat{\Delta}_{22}$$

is a consistent estimate of

$$\Delta_{21}^+ = E(x_0 u_0^+)\ .$$

We now define the "fully modified" estimator $\hat{a}^+$ using

$$y_t^+ = y_t - \hat{\omega}_{12}\hat{\Omega}_{22}^{-1}\Delta x_t$$

as

$$\hat{a}^+ = (X'X)^{-1}[X'Y^+ - e_m T\hat{\Delta}_{21}^+]\ .$$

Also define

$$\delta_T = \begin{bmatrix} T^{1/2} & 0 & 0 \\ 0 & T^{3/2} & 0 \\ 0 & 0 & I_m T \end{bmatrix}$$

and

$$J(r) = \begin{bmatrix} 1 \\ r \\ B_2(r) \end{bmatrix} , \quad 0 < r \le 1 .$$

Then

$$\delta_T(\hat{a}^+ - a) \Rightarrow \left[\int_0^1 JJ'\right]^{-1} (\int_0^1 JdV)\omega_{11\cdot 2}^{1/2} .$$

Conditional on $\mathcal{F}_2 = \sigma(B_2(r), 0 < r \le 1)$, this has the distribution

$$N\left[0, \ \omega_{11\cdot 2}\left[\int_0^1 JJ'\right]^{-1}\right] \qquad (7)$$

which is analogous to conventional asymptotic theory. Unconditionally, of course, the limit distribution is the mixture of normals

$$\int_{G>0} N(0, \ \omega_{11\cdot 2}G^{-1})dP(G) \ , \quad G \equiv (\int_0^1 JJ') \ .$$

Turning to inference, we consider the linear hypothesis

$$H_0 : R'a = r \ , \quad \text{rank}(R) = q$$

and the test statistic

$$G_R(\hat{a}^+, \ \hat{\omega}_{11\cdot 2}) = (R'\hat{a}^+ - r)'\left[\hat{\omega}_{11\cdot 2}R'(X'X)^{-1}R\right]^{-1}(R'\hat{a}^+ - r)$$

where $\hat{\omega}_{11\cdot 2}$ is a consistent estimate of $\omega_{11\cdot 2}$. We see from (7) that under H_0 ,

(8) $$G_R(\hat{a}^+, \hat{\omega}_{11\cdot 2}) \Rightarrow \chi^2_q \quad \text{as} \quad T \to \infty ,$$

which is not generally true if $\hat{a}$ or $\hat{a}^*$ were used in the construction of $G_R(\cdot)$.

One common application is the single coefficient test, i.e.

$$H_0 : a_i = a_i^0 .$$

Then we can construct a modified t-statistic

(9) $$t(\hat{a}_i^+ - a_i^0, \hat{\omega}_{11\cdot 2}) = \frac{(\hat{a}_i - a_i^0)}{\left[(X'X)^{-1}_{ii}\hat{\omega}_{11\cdot 2}\right]^{1/2}}$$

and from (7) we have under H_0

$$t(\hat{a}_i^+ - a_i^0, \hat{\omega}_{11\cdot 2}) \Rightarrow N(0,1) \quad \text{as} \quad T \to \infty .$$

This permits us to define "fully modified standard errors" by the quantity

$$\left[(X'X)^{-1}_{ii}\hat{\omega}_{11\cdot 2}\right]^{1/2}$$

to replace the conventional "standard errors" calculated by statistical packages.

These fully modified estimators have several advantages over the unadjusted and "bias-corrected" estimators. First, the standard distributional results for fully modified statistics permits inference to proceed conventionally. The bias-corrected estimators, on the other hand, produce test statistics with highly complicated limiting distributions, making inference problematic. Second, since $\omega_{11\cdot2} \leq \omega_{11}$, it would seem that the fully modified estimator has smaller dispersion asymptotically than the bias-corrected estimator. In fact, under certain conditions the fully modified estimator is asymptotically efficient and equivalent to full maximum likelihood (see Phillips (1988b) on optimal inference).

3. SIMULATION RESULTS

The Data Generating Process (DGP) used for the simulation is based on that used by Banerjee, *et al.* (1986) and Engle and Granger (1987)

(10a) $\quad y_t - 2x_t = u_t \;, \quad (1-\rho L)u_t = \epsilon_{1t}$

(10b) $\quad -y_t + 3x_t = z_t \;, \quad (1-L)z_t = \epsilon_{2t}$

$$\left\{ \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \right\}_1^T \equiv \text{iid } N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma\theta \\ \sigma\theta & \sigma^2 \end{bmatrix} \right]$$

$$|\rho| < 1 \;, \quad |\theta| < 1 \;.$$

(10) states that one linear combination of y_t and x_t is stationary, while another is I(1). We may rewrite these two equations into a conventional simultaneous equation system:

(11a) $\quad y_t = 2x_t + u_t$

(11b) $\quad x_t = z_t + u_t \;.$

The fact that the error terms in (11a) and (11b) are perfectly correlated seems unusual, but could have been generalized without substantially affecting the results if additional parameters were introduced.

Equation (11a) resembles many equations in macroeconometrics: y_t and x_t are cointegrated with a serially correlated error term. Equation (11b) shows that the long-run behavior of x_t is governed by the random walk z_t. Thus x_t will be long-run independent of u_t when (and only when) Δz_t is long-run uncorrelated with u_t. This occurs when $\theta = 0$, which was the parameterization implicitly chosen by Banerjee *et al.*

In most macroeconomic applications, analysts include a time trend. This is the correct procedure if any variables display "drift." The distributions tabulated by Sam Ouliaris in the appendix of Park and Phillips (1987a) show that this yields a significantly "fatter" distribution for OLS under long-run endogeneity than the simple case without a trend. To ensure the relevance of our simulations, all our results are based upon estimates of the regression

$$(12) \qquad y_t = \hat{a}_t + \hat{a}_2 t + \hat{a}_3 x_t + \hat{u}_t .$$

However, we report only the distributions of estimates of a_3, as this is the parameter of typical economic interest. We tried four estimation techniques, which we will refer to as OLS, IVZ, IVS and IVK for brevity:

OLS – standard least squares

IVZ – instrument x_t with z_t

IVS – instrument x_t with s_t, an independent Gaussian random walk

IVK – instrument x_t with k_{Tt}, some deterministic function.

In some sense, z_t is an "ideal" instrument for x_t *a priori* since they are cointegrated, yet z_t is not "contaminated" with u_t, as is

x_t. This is idealized as well because z_t may not be observable in an actual application.

In all our simulations, we generated 2000 series of length 200, starting with $u_0 = 0$ and $z_0 = 0$, and then discarding the initial 100 observations, generating a sample of size 100. The GAUSS matrix programming language and its RNDN function were used to generate the psuedo-normal innovations. The latter function uses the fast acceptance--rejection algorithm proposed by Kinderman and Ramage (1976). Start-up seeds were randomized by the clock. In our opinion the exact properties of the pseudo-normal numbers are unlikely to be very important in studies of this nature, since the theory is asymptotic and does not require normality or serial independence.

We need to select a suitable choice of instruments for IVS and IVK. Unfortunately, asymptotic theory provides little guidance. Our first simulation compares a variety of choices, under the parameterization $\rho = .8$, $\theta = .5$, and allowing σ to vary among $\{.5, 1, 2\}$. The results are reported in Table I (all tables are in the appendix). For IVK (trended instruments) a variety of time polynomials and sinusoids are compared. Average bias $(\tilde{a}-2)$ and square root of the average mean squared error $(\tilde{a}-2)^2$ are reported, the latter in parentheses. All the choices fared reasonably well. Based on MSE we selected the fourth option, consisting of two sinusoids and two cosinusoids.

For IVS (spurious stochastic instruments) we tried from one to eight independent random walks with psuedo-normal increments. The MSE seems to decline with the number of instruments, although the differences are fairly small after five instruments. Nevertheless we chose to use eight in the subsequent simulations.

Table II reports the performance of the five (uncorrected) estimation techniques. Twenty parameter settings are reported. As expected, OLS displays considerable bias for low signal/noise ratios. In fact, OLS is generally beaten by the other techniques for low σ . For large σ OLS continues to generally display the highest bias, yet beats the IV techniques in MSE for $\theta = 0$, and performs similarly in MSE for $\theta = .5$.

Notice the high MSE for IVZ under $\sigma = .5$, $\theta = 0$, $\rho = .85$. This occurred in unreported simulations under different parameter settings for low σ , and seems to have been caused by low frequency outliers. This "fat-tailed" property can be explained as follows. As σ approaches zero, the behavior of z_t approaches that of a constant. Since a constant and time trend are the other two instruments, the subspace spanned by the instruments approaches two dimensions as $\sigma \to 0$ and identification fails. As discussed by Phillips (1987c), partially unidentified IV estimators have non-degenerate Cauchy-type distributions, which seems a likely explanation of this aberration in the simulation.

In addition, it is interesting to note that IVS and IVK are upwardly biased, although this is not predicted by asymptotic theory.

The estimation of bias corrections, as noted earlier, is complicated by the dependence of these corrections upon preliminary coefficient estimates which may possess considerable bias in small samples, as shown in Table II. In order to focus on the potential value of bias corrections, we first side-step this issue by using the "true" residuals to calculate the bias correction terms. Throughout, we used the estimation equations:

$$\hat{\Delta} = T^{-1} \sum_{k=0}^{\ell} \sum_{t=k+1}^{T} \eta_{t-k}\eta_t'$$

$$\hat{\Omega} = T^{-1}\Sigma_1^T \eta_t\eta_t' + T^{-1} \sum_{k=1}^{\ell} \omega_{k\ell} \sum_{t=k+1}^{T} (\eta_{t-k}\eta_t' + \eta_t\eta_{t-k}')$$

$$\eta_t = (\hat{u}_t, \Delta x_t, \Delta z_t)'$$

$$\omega_{k\ell} = 1 - k/(\ell+1) .$$

The triangular weights used to estimate Ω constrain $\hat{\Omega}$ to be positive definite--see Newey and West (1987). The bias corrections $\hat{\Delta}_{21}$ and $\hat{\Delta}_{22}$ do not need this constraint, and are therefore estimated without the weights. The lag truncation number, ℓ , was set arbitrarily at seven.

Table III reports "bias-corrected" OLS and IVZ. As indicated above, the corrections were calculated using the true residuals. IVS and IVK, of course, do not need corrections according to the asymptotic theory. The small average bias and MSE in Table III are encouraging for bias-corrected OLS. Bias-corrected IVZ, however, in general did not improve over its uncorrected performance.

Moving to Table IV, which displays the fully modified estimators, we see continued improvement in OLS. The other estimators, however, perform quite poorly when compared to their uncorrected counterparts. The message from this simulation is clear: for small samples (in this case, $T = 100$) bias corrections and fully modified statistics only work well on OLS, additionally, if the true coefficients are known then the fully modified least squares estimator will be highly accurate.

The knowledge of the true coefficients is not completely impossible, if, say, the entire coefficient vector is specified in the null hypothesis to be tested, as in, for example, the test of Purchasing Power Parity by Corbae and Ouliaris (1988). The use of the null specification to calculation the correction term is analogous to the Lagrange Multiplier (LM) statistic in tests of linear restrictions in the standard linear model, where the constrained parameter values are used to calculate the variance of the error term.

In general, of course, most, if not all, of the coefficients will be unknown. The bias corrections will be calculated from coefficients estimated by a preliminary OLS or IV regression. Table V presents simulation results for "feasible" fully modified OLS. Three methods are used for the first stage regression: OLS, IVS and IVK. Compare the bias and MSE of the estimators in Table V with unadjusted OLS (Table II). In general, feasible fully modified OLS performs better than unadjusted OLS, but not by much. When compared to the fully modified OLS estimates using the true coefficients (Table IV), we see that the use of first stage regression coefficients significantly reduces the effectiveness of the modifications. Regardless, the results of Table V indicate that these are the best feasible estimation methods. When comparing among choices of first stage estimation, it appears that IVS may be the best choice for low signal to noise ratios, while OLS may be more appropriate for high values. This is surprising at first glance since we know from Table II that IVK generally performed better than OLS and IVS both in bias and MSE. A close reading of those figures shows that IVK has a high variance. This presumably increases the variance of the estimated bias corrections. Looking back at Table V, we see that feasible OLS using IVK first stage estimates has lower bias, but higher variance, than the estimators using OLS or IVS first stage estimates.

Thus bias-corrections using IVK are more accurate, but more variable, than those based on OLS and IVS.

We now turn to the problem of inference by examining the distribution of the standard and modified t-statistics for a_3. For brevity, only OLS techniques are examined. We consider three t-statistics:

$$\hat{t} = t(\hat{a}_3 - 2,\ \hat{\sigma}_u^2)\ ,\quad \hat{\sigma}_u^2 = T^{-1}\Sigma_1^T \hat{u}_t^2$$

$$t^* = t(\hat{a}_3^* - 2,\ \hat{\omega}_{11})$$

$$t^+ = t(\hat{a}_3^+ - 2,\ \hat{\omega}_{11\cdot 2})\ .$$

(See equation (9) for the definition of $t(\cdot,\cdot)$.)

$\hat{t}$ is a conventional t-ratio as printed by standard statistical packages, t^* is the "bias-corrected" t-ratio (a signed square root of the Wald statistic proposed by Park and Phillips (1987a)), and t^+ is the fully modified t-ratio of Phillips and Hansen (1988). t-statistics are commonly compared against the standard normal distribution for inferential purposes. As discussed earlier, asymptotic theory demonstrates that this will not generally yield valid inferences, except for the fully modified statistics. Of course, the tables in Park and Phillips (1987a) allow t^* to be used, although this is a cumbersome procedure, requiring estimation of a nuisance parameter and only permitting block tests.

Although many other parameterizations were run, we only report the results for $\rho = .7$, $\sigma = \{2,10\}$ and $\theta = \{0, .5\}$ because these summarize the main effects. 8000 replications of samples of size 100 were generated. Table VI reports the first four cumulants of the data. Figures 1 through 4 display non-parametric estimates of the probability density function (pdf's). Each figure displays all three statistics. In each case, the estimated pdf of the fully modified statistic is the closest to the standard normal density. Its variance ranges from 2 to 3, suggesting that there will be size distortion in moderate samples. The variances of standard least squares (14 to 21) and bias corrected least squares (8 to 11) are substantially higher.

While the asymptotic approximation to the distribution of t^+ is $N(0,1)$, the asymptotic distribution of t^* is a mixture of $N(0,1)$ and a unit root distribution, depending only upon the parameter θ (the degree of long-run correlation). It is interesting to note that this dependence is not indicated by the simulated small sample densities. We note that when $\theta = 0$, the distribution of t^* is far from the asymptotic $N(0,1)$, and is decidedly inferior to the asymptotically equivalent t^+. Moreover, while the estimated density of t^* is similar for $\theta = 0$ and $\theta = .5$, it does change shape as the signal to noise ratio, σ, is varied. This is in contrast to the prediction of asymptotic theory and is therefore an important finite sample effect. Similar but less dramatic effects occur for the statistic t^+ (see Table VI).

These results suggest that the bias corrected statistics of Park and Phillips are not well approximated by a standard $N(0,1)$ even when this is their asymptotic distribution (the case $\theta = 0$). On the other hand the fully modified statistics are much better approximated by the standard $N(0,1)$. Moreover, this is true even when the long run endogeneity correction is not required. It would therefore seem that little is lost in finite samples by employing the endogeneity correction.

4. THE AGGREGATE CONSUMPTION FUNCTION

A perennially examined macroeconomic relationship is the postulated linear dependence of aggregate consumption upon aggregate disposable income:

$$c_t = \alpha_1 + \alpha_2 y_t + u_t \; . \tag{13}$$

Even though the microfoundations of (13) have never been well established, it appears in one form or another in many theoretical and applied macro models. Since c_t and y_t are both believed to be I(1) processes--see Hall (1978) and Perron and Phillips (1987)--(13) makes sense as a long-run relationship if and only if $u_t \equiv I(0)$. This is emphasized in a recent test of the Permanent Income Hypothesis by Campbell (1987). Given cointegration, we can estimate α using the methods discussed in this paper. We use quarterly real ($1982) per capita personal consumption expenditure and personal disposable income, 1941/1 to 1987/4, using the consumption deflator for both series. Data is in thousands of dollars. Since both series display a mild trend, we estimate

$$c_t = \hat{\alpha}_1 + \hat{\alpha}_2 t + \hat{\alpha}_3 y_t + \hat{u}_t \; .$$

We only consider OLS techniques.

The "standard" OLS coefficients and standard errors are

$$c_t = \underset{(14.7)}{9.99} - \underset{(13.9)}{0.24}\, t + \underset{(.034)}{.897}\, y_t \, .$$

(The coefficients for the constant and time trend have been scaled by 10^8 and 10^{10}, respectively, throughout this presentation.)

Despite the apparent precision, no modern-trained econometrician would trust these estimates, as the consumption function is a classic example of a simultaneous equation system. The theory of cointegration tells us that we can consistently estimate α by least squares, but that we should place no faith in the standard error estimates. We will try and improve upon these estimates by first calculating bias-corrected estimates, and second employing the fully modified techniques.

To construct bias-corrected estimates we first examine the cross-correlagram between Δy_t and the fitted residuals, $\hat{u}_t$:

$$\gamma(k) = \frac{\sum_{t=k+1}^{T} \hat{u}_t \Delta y_{t-k}}{\left[\left[\sum_1^T \hat{u}_t^2\right]\left[\sum_1^T \Delta y_t^2\right]\right]^{1/2}}$$

k	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
$\gamma(k)$	-0.94	-1.32	-1.16	-1.17	-1.40	-0.60	1.37	1.42	1.60	3.44

k	0	1	2	3	4	5	6	7	8	9	10
$\gamma(k)$	-2.17	-0.89	0.01	-0.69	0.07	0.25	0.16	0.02	-0.27	0.03	-0.89

We notice that the cross-correlagram is small for positive values of k , and rather large for negative values. This suggests that $\hat{\Delta}_{21}$ may not need many lags to estimate Δ_{21} , but ω_{21} may need more. By calulating the parameter $\hat{\Delta}_{21}$ for several lag truncations, we can see explicitly how the bias-corrected estimates are sensitive to this choice:

ℓ	$\hat{\alpha}_1^*$	$\hat{\alpha}_2^*$	$\hat{\alpha}_2^*$
1	19.5	8.77	.875
2	23.4	12.46	.866
5	25.6	14.57	.861
10	24.8	13.76	.863

The coefficient of interest, $\hat{\alpha}_3^*$, is stable for $\ell \geq 2$, and is slightly less than unadjusted OLS.

To construct fully modified estimates, we use an estimate of $\hat{\Omega}$ using a lag truncation of ten due to the large values of the cross correlagram for $k < 0$ and the large values of the correlogram for $\hat{u}_t$ defined by

$$\hat{\gamma}_{uu}(k) = \frac{\sum_t \hat{u}_t \hat{u}_{t-k}}{\sum_1^T \hat{u}_t^2} .$$

k	1	2	3	4	5	6	7	8	9	10
$\hat{\gamma}_{uu}(k)$	1	.80	.70	.60	.50	.46	.41	.39	.36	.34

We vary the lag truncation number, ℓ , used to estimate $\hat{\Delta}_{21}^+$, separately, finding the estimates

ℓ	$\hat{\alpha}_1^+$	$\hat{\alpha}_2^+$	$\hat{\alpha}_3^+$
1	-4.81	-14.94	.933
2	-8.61	-18.54	.942
5	-11.02	-20.82	.947
10	-9.11	-19.01	.942

Again, the estimates of a_3 are not very sensitive for the choice of $\ell \geq 2$. It is evident, however, that these values are quite different from the bias-corrected estimates. The final statistics needed are the fully modified standard errors. We calculate $\hat{\omega}_{11\cdot2} = 1.398 \times 10^{-13}$ from $\hat{\Omega}$ and can write in the conventional format

$$c_t = \underset{(37.93)}{-8.61} - \underset{(36.95)}{18.54}\, t + \underset{(.088)}{.942}\, y$$

(We used $\ell = 2$ for $\hat{\Delta}^+_{21}$.) If a_3 is the coefficient of economic interest, we could write down the 95% confidence interval [.770, 1.114] which is unfortunately quite large. We can easily accept, for example, a unit coefficient (which is implied by some versions of the Permanent Income Hypothesis).

5. CONCLUSION

This study set out to explore the small sample properties of OLS and IV estimators in cointegrating regressions--unadjusted, bias corrected, and fully modified--in order to evaluate the usefulness of the asymptotic theory developed recently by Phillips and Hansen (1988). We discovered that IV estimators work quite well, including the "spurious" procedures which use instruments structurally unrelated to the DGP. The fully modified estimators, shown in our earlier work to possess limiting mixed normal distributions, also worked well, especially when the true coefficient vector was known *a priori*. Feasible corrections, based on preliminary regressions fared reasonably well, but did not eliminate the problem of small sample bias in estimation. The approximation of the distribution of fully modified t-statistics by the normal density is excellent, and despite a higher variance than unity, performs substantially better than conventional inference procedures.

The DGP used in this exercise was designed to decompose possible problems into three categories: signal/noise, serial correlation, and endogeneity. The critical factor, it appears, is the signal/noise ratio (σ in the simulation), not the degree of long-run endogeneity. If the variance of the increments of the random walk which drives the long-run behavior of the variables is high relative to the variance of the short-term dynamics, the bias problem is negligible, and OLS works well, and fully modified estimates will permit inference and testing to proceed in a conventional fashion. If, however, the relative signal is low, spurious IV techniques may be necessary to obtain preliminary estimates for the modified least squares estimator.

TABLE I: TRENDED/SPURIOUS
BIAS (ROOT MSE)

Trended Instruments P_i	λ_j	σ = 0.5	1.0	2.0
2,3,4	ϕ	.385 (.525)	.178 (.375)	.063 (.173)
1/2,2,3,4	ϕ	.411 (.498)	.190 (.298)	.071 (.152)
2,3	ϕ	.340 (.777)	.163 (.537)	.067 (.392)
ϕ	1,2	.422 (.495)	.193 (.279)	.075 (.148)
ϕ	1,2,4	.466 (.510)	.225 (.281)	.091 (.138)
ϕ	2,4,8	.581 (.615)	.319 (.371)	.143 (.195)
2,3	2,4	.466 (.510)	.225 (.279)	.091 (.138)

$$k_{Tt} = \left\{t^{P_1}, t^{P_2}, \ldots, t^{P_N}, \sin(2\pi\lambda_1 t/T), \cos(2\pi\lambda_1 t/T), \ldots, \sin(2\pi\lambda_I t/T), \cos(2\pi\lambda_I t/T)\right\}$$

Spurious Instruments	σ = 0.5	1.0	2.0
1	25.737 (1170)	.522 (12.53)	-1.240 (83.2)
2	.397 (.810)	.189 (.603)	.084 (.401)
3	.432 (.609)	.208 (.378)	.081 (.212)
4	.450 (.550)	.209 (.322)	.087 (.167)
5	.460 (.530)	.218 (.300)	.088 (.155)
6	.470 (.526)	.227 (.297)	.093 (.148)
7	.479 (.528)	.234 (.295)	.096 (.145)
8	.486 (.529)	.240 (.293)	.099 (.145)

TABLE II: UNCORRECTED ESTIMATES
BIAS (ROOT MSE)

θ ρ	σ = 0.5		1.0		2.0		5.0		10.0	
0 .7										
OLS	.545	(.566)	.241	(.268)	.075	(.100)	.012	(.023)	.003	(.013)
IVZ	.083	(.383)	-.017	(.138)	-.004	(.063)	-.001	(.025)	-.001	(.013)
IVS	.357	(.435)	.127	(.212)	.037	(.100)	.005	(.038)	.001	(.018)
IVK	.262	(.410)	.081	(.207)	.020	(.105)	.002	(.045)	.000	(.023)
0 .85										
OLS	.654	(.673)	.326	(.359)	.108	(.145)	.020	(.046)	.005	(.022)
IVZ	1.036	(47.2)	-.059	(.283)	-.015	(.110)	-.001	(.043)	-.000	(.021)
IVS	.563	(.628)	.246	(.345)	.075	(.345)	.014	(.064)	.004	(.032)
IVK	.498	(.611)	.202	(.349)	.059	(.349)	.011	(.074)	.003	(.037)
.5 .7										
OLS	.533	(.546)	.280	(.297)	.121	(.134)	.037	(.044)	.016	(.020)
IVZ	.187	(.255)	.112	(.148)	.062	(.084)	.025	(.034)	.013	(.018)
IVS	.373	(.422)	.164	(.217)	.064	(.100)	.017	(.037)	.007	(.018)
IVK	.284	(.381)	.110	(.197)	.041	(.100)	.011	(.042)	.005	(.022)
.5 .85										
OLS	.611	(.621)	.358	(.374)	.170	(.187)	.058	(.069)	.026	(.032)
IVZ	.246	(.359)	.160	(.210)	.093	(.122)	.042	(.056)	.022	(.029)
IVS	.541	(.577)	.287	(.339)	.126	(.173)	.040	(.067)	.018	(.033)
IVK	.488	(.553)	.238	(.326)	.098	(.167)	.031	(.071)	.014	(.036)

TABLE III: BIAS-CORRECTED USING TRUE COEFFICIENTS
BIAS (ROOT MSE)

θ	ρ		σ = 0.5		1.0		2.0		5.0		10.0
0	.7										
OLS		-.029	(.155)	-.015	(.114)	-.005	(.063)	-.002	(.029)	-.001	(.015)
IVZ		-.036	(.365)	-.006	(.148)	-.001	(.071)	-.002	(.030)	-.001	(.015)
0	.85										
OLS		.042	(.217)	.018	(.182)	.006	(.118)	.002	(.048)	.001	(.024)
IVZ		-.298	(5.59)	-.024	(.318)	-.004	(.138)	.001	(.049)	.001	(.025)
.5	.7										
OLS		-.114	(.164)	-.094	(.130)	-.060	(.084)	-.025	(.036)	-.013	(.019)
IVZ		-.235	(.333)	-.126	(.176)	-.066	(.089)	-.026	(.038)	-.013	(.019)
.5	.85										
OLS		-.017	(.179)	-.040	(.155)	-.035	(.110)	-.034	(.057)	-.018	(.030)
IVZ		-.215	(2.19)	-.111	(.279)	-.052	(.134)	-.037	(.061)	-.019	(.031)

TABLE IV: FULLY MODIFIED USING TRUE COEFFICIENTS
BIAS (ROOT MSE)

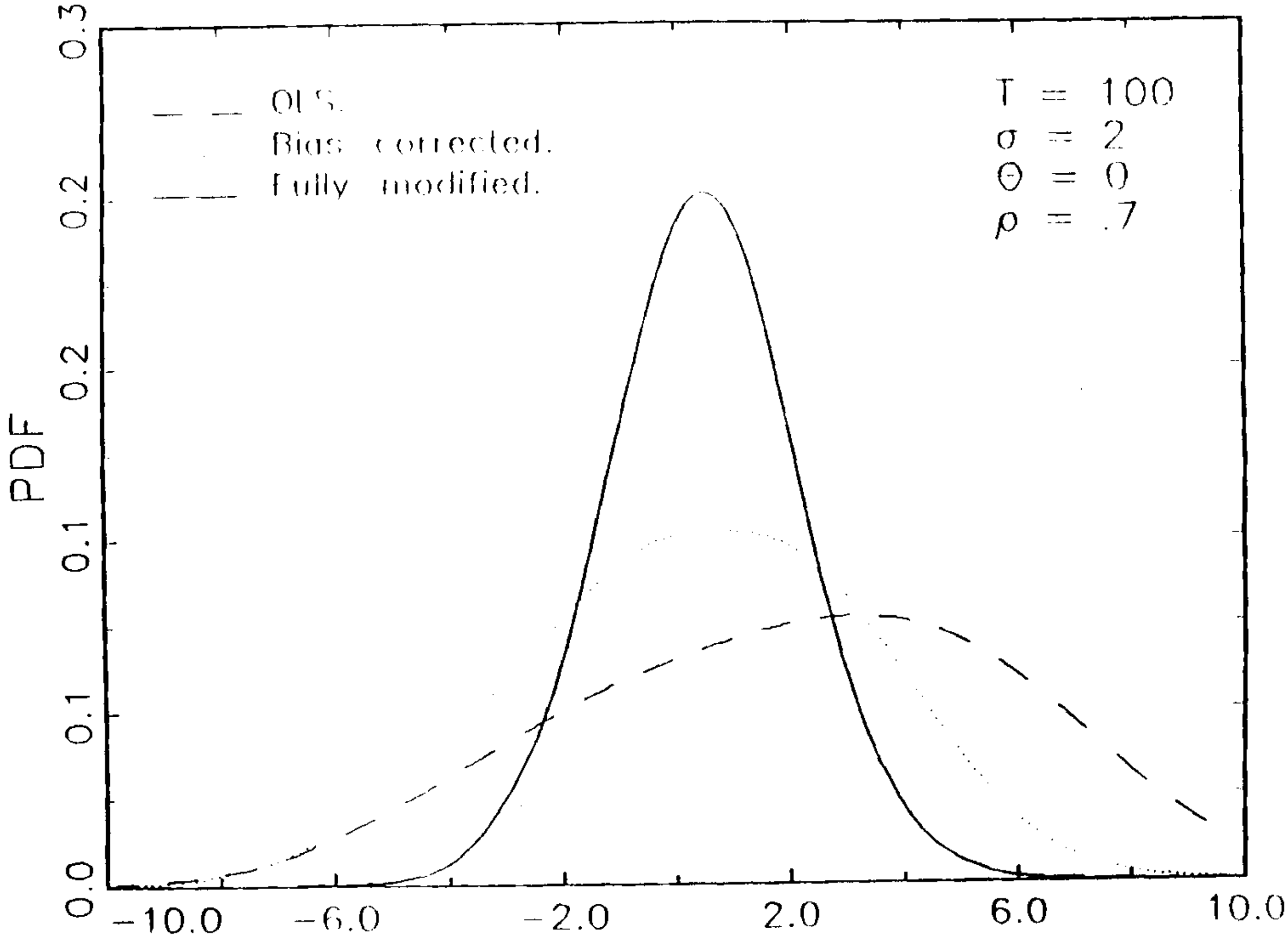
θ	ρ		σ = 0.5	1.0	2.0	5.0	10.0
0	.7						
		OLS	-.017 (.138)	-.006 (.095)	-.001 (.055)	-.001 (.023)	-.001 (.012)
		IVZ	-.076 (.582)	-.014 (.152)	-.003 (.071)	.011 (.030)	.003 (.014)
		IVS	.333 (.435)	.122 (.272)	.042 (.207)	.006 (.181)	.002 (.177)
		IVK	.247 (.449)	.080 (.310)	.022 (.272)	.001 (.243)	-.001 (.237)
0	.85						
		OLS	.059 (.200)	.029 (.152)	.010 (.089)	.003 (.037)	.001 (.019)
		IVZ	-.536 (13.2)	-.044 (.303)	-.008 (.126)	.020 (.051)	.005 (.024)
		IVS	.537 (.615)	.234 (.374)	.071 (.235)	.008 (.186)	-.002 (.180)
		IVK	.467 (.611)	.201 (.414)	.065 (.293)	.007 (.232)	-.002 (.229)
.5	.7						
		OLS	-.078 (.134)	-.052 (.095)	-.026 (.055)	-.009 (.022)	-.004 (.011)
		IVZ	-.018 (.192)	.017 (.105)	.011 (.055)	.014 (.028)	.005 (.013)
		IVS	.333 (.404)	.140 (.245)	.050 (.195)	.015 (.181)	.007 (.178)
		IVK	.256 (.405)	.094 (.286)	.030 (.253)	.007 (.230)	.002 (.230)
.5	.85						
		OLS	-.007 (.152)	-.027 (.118)	-.021 (.077)	-.012 (.036)	-.007 (.019)
		IVZ	-.142 (9.95)	.064 (.187)	.041 (.105)	.030 (.049)	.012 (.024)
		IVS	.499 (.545)	.260 (.335)	.110 (.214)	.033 (.168)	.013 (.164)
		IVK	.414 (.561)	.214 (.371)	.086 (.272)	.027 (.229)	.009 (.223)

TABLE V: FEASIBLE FULLY MODIFIED OLS
BIAS (ROOT MSE)

θ	ρ		σ = 0.5	1.0	2.0	5.0	10.0
0	.7						
		OLS	.455 (.501)	.178 (.232)	.052 (.095)	.008 (.031)	.002 (.015)
		IVS	.298 (.411)	.092 (.213)	.025 (.108)	.005 (.035)	.001 (.018)
		IVK	.200 (.424)	.049 (.236)	.140 (.124)	.004 (.039)	.001 (.019)
0	.85						
		OLS	.585 (.629)	.266 (.339)	.084 (.155)	.016 (.051)	.004 (.025)
		IVS	.518 (.608)	.217 (.340)	.068 (.174)	.013 (.058)	.004 (.029)
		IVK	.463 (.638)	.183 (.369)	.055 (.201)	.013 (.065)	.004 (.032)
.5	.7						
		OLS	.419 (.452)	.191 (.230)	.075 (.105)	.021 (.035)	.009 (.016)
		IVS	.278 (.361)	.096 (.185)	.030 (.091)	.013 (.033)	.006 (.016)
		IVK	.188 (.367)	.047 (.203)	.008 (.104)	.010 (.038)	.005 (.019)
.5	.85						
		OLS	.533 (.559)	.281 (.319)	.119 (.158)	.045 (.063)	.020 (.030)
		IVS	.470 (.524)	.221 (.298)	.086 (.153)	.038 (.064)	.017 (.031)
		IVK	.433 (.534)	.185 (.311)	.067 (.173)	.035 (.069)	.016 (.033)

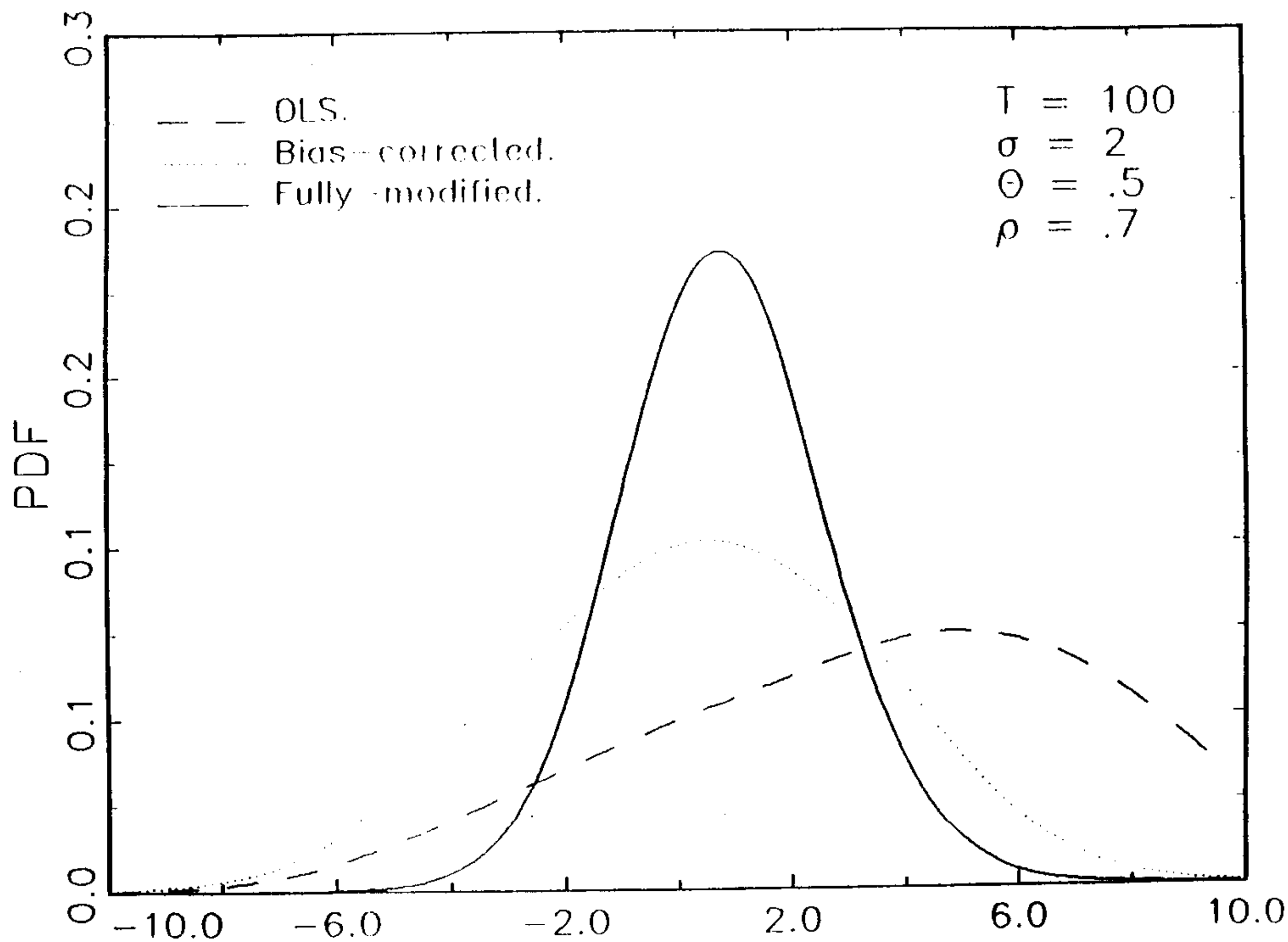
TABLE VI: DISTRIBUTION OF t-STATISTICS

	$\hat{t}$	t^*	t^+
$\sigma = 2,\ \theta = 0$			
Mean	2.046	.304	.469
Variance	14.193	8.301	2.552
Skewness	-.163	-.175	.044
Kurtosis	-.595	-.376	.221
$\sigma = 2,\ \theta = .5$			
Mean	3.642	.326	.818
Variance	16.389	8.833	2.999
Skewness	-.212	-.064	.218
Kurtosis	-.512	-.323	.473
$\sigma = 10,\ \theta = 0$			
Mean	.270	-.273	-.172
Variance	21.114	10.747	2.077
Skewness	-.040	-.098	-.183
Kurtosis	-1.074	-.727	.076
$\sigma = 10,\ \theta = .5$			
Mean	.694	-.457	-.122
Variance	21.259	10.736	2.149
Skewness	-.013	-.041	-.133
Kurtosis	-1.053	-.730	.150



OLS.
Bias corrected.
Fully modified.
T = 100
σ = 2
Θ = 0
ρ = .7
PDF
0.0
0.1
0.1
0.2
0.2
0.3
-10.0
-6.0
-2.0
2.0
6.0
10.0

Figure 1


OLS.
Bias-corrected.
Fully-modified.
T = 100
σ = 2
Θ = .5
ρ = .7
PDF
0.0
0.1
0.1
0.2
0.2
0.3
−10.0
−6.0
−2.0
2.0
6.0
10.0
Figure 2

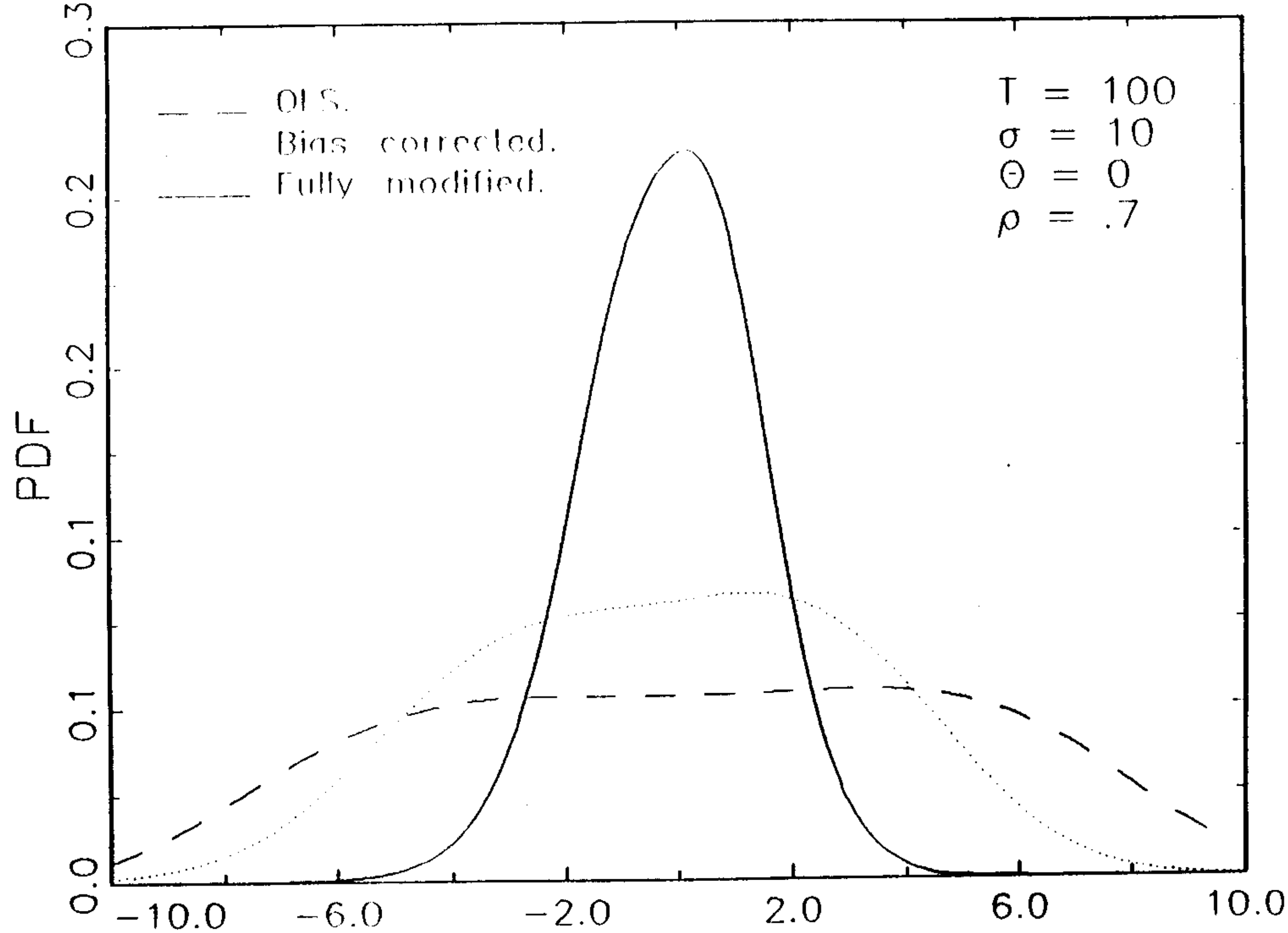
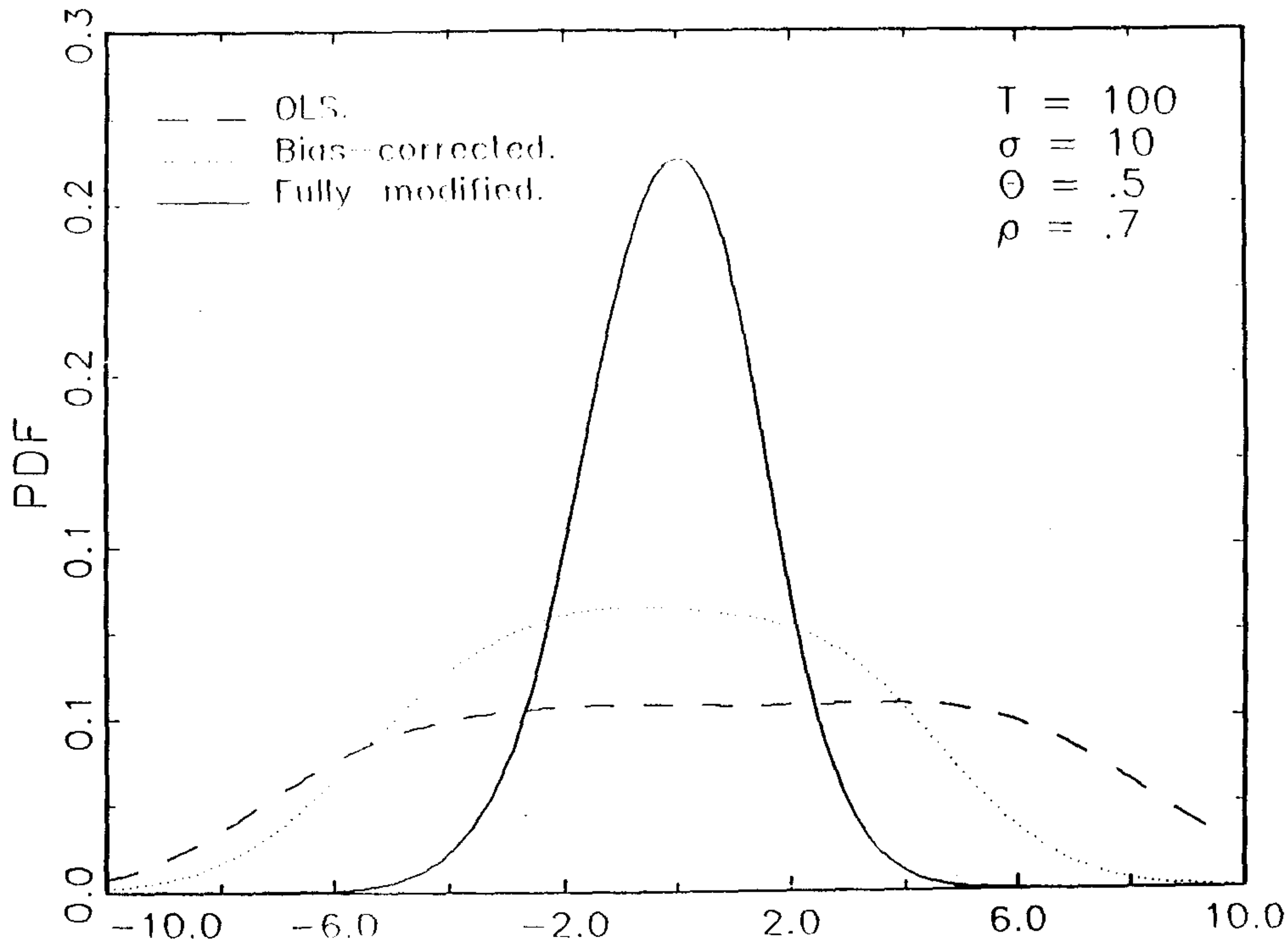

OLS.
Bias corrected.
Fully modified.
T = 100
σ = 10
Θ = 0
ρ = .7
PDF
0.0
0.1
0.1
0.2
0.2
0.3
−10.0
−6.0
−2.0
2.0
6.0
10.0

Figure 3



OLS.
Bias-corrected.
Fully modified.
T = 100
σ = 10
Θ = .5
ρ = .7
PDF
0.0
0.1
0.1
0.2
0.2
0.3
-10.0
-6.0
-2.0
2.0
6.0
10.0

Figure 4

REFERENCES

Banerjee, A., J. J. Dolado, D. F. Hendry and G. W. Smith (1986). "Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence," *Oxford Bulletin of Economics and Statistics*, 48, 253-277.

Campbell, J. Y. (1987). "Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis," *Econometrica*, 55, 1249-1273.

Corbae, P. D. and S. Ouliaris (1988). "Cointegration and tests of purchasing power parity," *Review of Economics and Statistics*, forthcoming.

Engle, R. F. and C. W. J. Granger (1987). "Cointegration and error correction: Representation, estimation, and testing," *Econometrica*, 55, 251-276.

Hall, R. E. (1978). "Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence," *Journal of Political Economy*, 89, 974-1009.

Kinderman, A. J. and J. G. Ramage (1976). "Computer Generation of Normal Random Numbers," *Journal of the American Statistical Association*, 71, 893-896.

Newey, W. and K. D. West (1987). "A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703-708.

Park, J. Y. and P. C. B. Phillips (1987a). "Statistical inference in regressions with integrated processes: Part 1," Cowles Foundation Discussion Paper No. 811R, Yale University.

________ (1987b). "Statistical inference in regressions with integrated processes: Part 2," Cowles Foundation Discussion Paper No. 819R, Yale University.

Perron, P. and P. C. B. Phillips (1987). "Does GNP have a unit root? A re-evaluation," *Economic Letters*, 23, 139-145.

Phillips, P. C. B. (1986). "Understanding spurious regressions in econometrics," *Journal of Econometrics*, 33, 311-340.

________ (1987a). "Time series regression with a unit root," *Econometrica*, 55, 277-301.

________ (1987b). "Multiple regression with integrated processes," *AMS/IMS/SIAM Conference on Stochastic Processes Proceedings* (forthcoming).

_________ (1987c). "Partially identified econometric models," Cowles Foundation Discussion Paper No. 845.

_________ (1988a). "Weak convergence to the matrix stochastic integral $\int_0^1 BdB'$," *Journal of Multivariate Analysis*, 24, 252-264.

_________ (1988b). "Optimal inference in cointegrated systems," Cowles Foundation Discussion Paper No. 866.

_________ and S. N. Durlauf (1986). "Multiple time series with integrated variables," *Review of Economic Studies*, 53, 473-496.

Phillips, P. C. B. and B. E. Hansen (1988). "Statistical inference in instrumental variables regression with I(1) processes," Cowles Foundation Discussion Paper No. 869.