1980

The Finite Sample Properties of OLS and IV Estimators in Special Rational Distributed Lag Models

R. A. L. Carter
Aman Ullah

Follow this and additional works at: https://ir.lib.uwo.ca/economicsresrpt

Part of the Economics Commons

Citation of this paper:
RESEARCH REPORT 8013

THE FINITE SAMPLE PROPERTIES OF OLS AND IV ESTIMATORS IN SPECIAL RATIONAL DISTRIBUTED LAG MODELS

by
R. A. L. Carter
and
Aman Ullah

June 1980
THE FINITE SAMPLE PROPERTIES OF OLS AND IV ESTIMATORS IN SPECIAL RATIONAL DISTRIBUTED LAG MODELS

R. A. L. Carter and Aman Ullah

University of Western Ontario

1. Introduction

A popular model in economics is the rational distributed lag model proposed by Jorgenson (1966). In this paper we deal with the version of this model in which the polynomial in the denominator is of degree one. Popular special cases are the geometric distributed lag model of Koyck (1954) and Nerlove's (1958) adaptive expectations model. The coefficients of such models are sometimes estimated by ordinary least squares (OLS), which ignores the correlation between the right-hand lagged dependent variable and the autocorrelated disturbances and is, hence, inconsistent. Alternatively, Liviatan's (1961) instrumental variables estimator (IV) may be used to obtain consistent estimates. Some small sample properties of IV have been derived by Nagar and Gupta (1968) and by Scadding (1973). The aim of this paper is to derive exact and small $\sigma$ asymptotic properties for OLS and small $\sigma$ asymptotic properties of IV.

The main results of the paper can be summarized as follows. The exact, and approximate bias and mean squared error of the OLS estimator are derived and it is shown that, with the sample size fixed, OLS converges to the true value of the parameter if the noncentrality parameter of its distribution increases indefinitely; that is, as $\sigma$ grows small. With regard to the IV estimator we note that its exact moments, to any order, do not exist. An approximation to the exact distribution has been obtained which is centered on the
true parameter. When this approximation is valid we are able to give the distribution of the IV estimates of the coefficients of the exogenous variables.

In section 2 we present the model and its assumptions. Then in section 3 we analyze the exact and approximate moments of the OLS estimator. Finally, in section 4 we consider the distribution of the IV estimator.

2. The Model

We begin with the assumption that the values of $y_t$ are independent drawings from normal populations with constant variances but varying means.

\[(2.1) \quad y_t \sim N(\mu_t, \sigma^2) \quad \text{for} \quad t = 1, \ldots, T\]

Next, we assume that $\mu_t$ is determined by a linear function analogous to a regression equation

\[(2.2) \quad \mu_t = \gamma \mu_{t-1} + X_t' \beta\]

where $X_t'$ is a non-random $1 \times K$ vector, $\beta$ is a $K \times 1$ vector of unknown coefficients, $\gamma$ is an unknown scalar coefficient, and $\mu_t$ and $\mu_{t-1}$ are unknown means of $y_t$ and $y_{t-1}$. To ensure that the process described by (2.2) is stable we assume

\[(2.3) \quad |\gamma| < 1.\]

Now $y_t$ can be written as

\[(2.4) \quad y_t = \mu_t + \eta_t = \gamma \mu_{t-1} + X_t' \beta + \eta_t = \gamma (\mu_{t-1} + \eta_{t-1}) + X_t' \beta + \eta_t - \gamma \eta_{t-1} = \gamma y_{t-1} + X_t' \beta + \eta_t, \quad t = 2, \ldots, T,\]
where \( \varepsilon_t = \eta_t - \gamma \eta_{t-1} \) and \( \eta_t \) is an independent drawing from \( N(0, \sigma^2) \). Therefore, \( \varepsilon_t \) follows a first order moving average scheme with

\[
(2.5) \quad E \varepsilon_t = 0 \quad \text{and}
\]

\[
(2.6) \quad \text{var}(\varepsilon_t) = E(\eta_t - \gamma \eta_{t-1})^2 = (1 + \gamma^2)\sigma^2
\]

The covariance of \( \varepsilon_t \) and \( \varepsilon_s \) (\( s < t \)) is

\[
(2.7) \quad E(\varepsilon_t \varepsilon_s) = E(\eta_t - \gamma \eta_{t-1})(\eta_s - \gamma \eta_{s-1})
\]

\[
= \begin{cases} 
\gamma \sigma^2 & \text{if } s = t - 1 \\
0 & \text{if } s < t - 1
\end{cases}
\]

Therefore, the coefficient of autocorrelation is

\[
(2.8) \quad r(\varepsilon_t \varepsilon_{t-1}) = \frac{E(\varepsilon_t \varepsilon_{t-1})}{\sqrt{E(\varepsilon_t^2)E(\varepsilon_{t-1})}} = \frac{-\gamma}{1 + \gamma^2}
\]

This moving average process may seem to be unduly arbitrary.

However, it is solely the result of assumptions (2.1) and (2.2), and (2.2) can be obtained in several appealing ways. First, consider a Koyck (1954) type distributed lag model which has been specified in terms of \( \mu_t \) (instead of in terms of the random \( \eta_t \))

\[
(2.9) \quad \mu_t = \alpha_0 + \alpha_1 z_t + \alpha_1 \lambda z_{t-1} + \alpha_1 \lambda^2 z_{t-2} + \ldots
\]

This model says that the mean response, \( \mu_t \), is a function of the present value and all past values of an exogenous variable, \( z_t \), where \( \alpha_1 \) and \( \lambda \), \( 0 < \lambda < 1 \), are unknown coefficients. Then, applying the Koyck transformation we have

\[
(2.10) \quad \mu_t = \lambda \mu_{t-1} + [1, z_t] \begin{bmatrix} (1-\lambda)\alpha_0 \\ \alpha_1 \end{bmatrix}
\]

which is of the form (2.2) with \( \gamma = \lambda \), \( X'_t = [1, z_t] \) and \( \beta = \begin{bmatrix} (1-\lambda)\alpha_0 \\ \alpha_1 \end{bmatrix} \)
A second justification for (2.2) is the adaptive expectations model of Nerlove (1958).

\[(2.11) \quad \mu_t = \alpha_0 + \alpha_1 \gamma_t^*\]

\[(2.12) \quad \gamma_t^* - \gamma_{t-1}^* = \delta (\gamma_{t-1} - \gamma_{t-1}^*) \quad 0 < \delta < 1\]

which says that the mean response, \(\mu_t\), depends on expectations about the future, \(\gamma_t^*\), and that these expectations are adjusted by some fraction of the extent to which past expectations were in error. By substitution and transformation, this model can be reduced to

\[(2.13) \quad \mu_t = (1-\delta)\mu_{t-1} + [1 \quad \gamma_{t-1}^*] \begin{bmatrix} \alpha_0 \delta \\ \alpha_1 \delta \end{bmatrix}\]

which is of the form of (2.2) with \(\gamma = (1-\delta), \quad X_t^* = [1 \quad \gamma_{t-1}^*]\) and \(\beta = \begin{bmatrix} \alpha_0 \delta \\ \alpha_1 \delta \end{bmatrix}\).

More generally, consider the rational distributed lag model, Jorgenson (1966), in the case that the polynomial in the denominator is of degree one.

\[(2.14) \quad \mu_t = \frac{\sum_{j=1}^{J} \alpha_j(L)z_{tj}}{1-\gamma L}, \quad 0 < \gamma < 1,\]

where \(\alpha_j(L)\) is a polynomial in the lag operator \(L\) and \(z_{tj}\) is observation \(t\) on exogenous variable number \(j\). This is of the form (2.2) with \(X_t^\gamma \beta = \sum_{j=1}^{J} \alpha_j(L)z_{tj}\).

3. Least Squares Estimation

It is convenient, at this point, to rewrite equation (2.4) in matrix notation

\[(3.1) \quad y = \gamma y_{-1} + X \beta + \varepsilon\]

where \(y\) is an \(n \times 1\) vector of independent random variables \(y_t \sim N(\mu_t, \sigma^2), \quad n = T - 1\), \(y_{-1}\) is an \(n \times 1\) vector of independent random variables \(y_{t-1} \sim N(\mu_{t-1}, \sigma^2), \quad \mu_t\) is an \(n \times K\) matrix with non-random rows \(X_t^\prime\), and \(\varepsilon\) is an \(n \times 1\) vector of moving average disturbances. The OLS estimates of \(\gamma\) and \(\beta\) are obtained by solving the equations
\[(3.2) \quad c y'_{-1} y_{-1} + y'_{-1} X b = y'_{-1} y\]

\[(3.3) \quad c X' y_{-1} + X' X b = X' y\]

OLS will not be consistent in this application because

\[(3.4) \quad y'_{-1} \varepsilon = \gamma \eta'_{-1} \eta_{-1} - \gamma \mu'_{-1} \eta_{-1} + \mu'_{-1} \eta + \eta'_{-1} \eta,
\]

(where \(\mu_{-1}, \eta\) and \(\eta_{-1}\) are \(n \times 1\) vectors of elements \(\mu_t, \eta_t\) and \(\eta_{t-1}\)) and we cannot reasonably expect to have \(\text{plim} n^{-1} \eta_{-1} \eta_{-1} = 0\). Nevertheless, we will consider the OLS estimator of \(\gamma, c,\) and present its exact moments.

Equation (3.3) can be solved for \(b\) which can then be substituted into (3.2) to obtain:

\[(3.5) \quad c = \frac{y'_{-1} M y}{y'_{-1} M y_{-1}} = \frac{z' \frac{N z}{z' D'_{1} M D_{1} z}}{z' D'_{1} M D_{1} z} \]

where \(M = I - X(X'X)^{-1}X'\) is an idempotent matrix with rank \(h = n - K,\)

\[z_t = \frac{y_t}{\sigma_t} \sim N(\mu_t/\sigma_t, 1) \quad (t = 1, \ldots, T),\]

and \(N = \frac{1}{2} (D'_{1} M D_{1} + D'_{2} M D_{2}),\) with

\(D_{1} = [I_n 0]\) and \(D_{2} = [0 I_n].\) That is, \(D_{1}\) and \(D_{2}\) are \(n \times n\) identity matrices bordered by columns of zeros. \(N\) and \(D'_{1} M D_{1}\) are both symmetric and \(D'_{1} M D_{1}\) is idempotent of rank \(h\) so that we can find a \(T \times T\) orthogonal matrix \(P\) such that

\[(3.6) \quad P D'_{1} M D_{1} P' = \begin{bmatrix} I_h & 0 \\ 0 & 0 \end{bmatrix}, \]

where \(I_h\) is an \(h \times h\) identity matrix. Now let

\[(3.7) \quad Pz = \begin{bmatrix} s \\ t \end{bmatrix} \sim N(Pz, I_T) = N \begin{bmatrix} (\bar{s}) \\ (\bar{t}) \end{bmatrix}, I_T \]

where \(\bar{s} = Ez, \bar{s} = Es, \bar{t} = Et\) and \(s\) and \(t\) are vectors of \(h\) and \(m = T - h = K + 1,\) elements respectively. Also let
\[ P \cdot N \cdot P' = \begin{bmatrix} A & C' \\ C & B \end{bmatrix} \]

where \( A \) and \( B \) are square symmetric matrices of order \( h \) and \( m \), respectively, and \( C \) is \( mxh \). Then we can write

\[ c = \frac{z' P' P N P' P z}{z' P' P D_1 MD_1 P' P z} = \frac{s' A s + 2 t' C s + \tilde{t}' B \tilde{t}}{s' s} \]

Now \( s' s \) has a non-central \( \chi^2 \) distribution with \( h \) degrees of freedom and a non-centrality parameter

\[ \phi = \frac{\bar{z} s' \tilde{z}}{\bar{z} z} = \frac{\bar{z}' D_1 MD_1 \bar{z}}{\bar{z}' \bar{z}} = \frac{\mu' M \mu}{2 \sigma^2} \]

Also

\[ [\bar{z}' \tilde{t}] \begin{bmatrix} A & C' \\ C & B \end{bmatrix} \begin{bmatrix} s' \\ \tilde{t} \end{bmatrix} = \bar{z}' N \tilde{z} = \frac{\mu' M \mu}{\sigma^2} = \frac{\gamma M \mu}{\sigma^2} = 2 \gamma \phi \]

since \( M \mu = \gamma M \mu \).

Remembering that the elements of \( t \) are independent of those of \( s \), we expand (3.9) to obtain

\[ Ec = \sum_{i} a_{ii} E\left(\frac{s_i^2}{s_s}\right) + \sum_{i} a_{ii} E\left(\frac{s_i^2}{s_s}\right) + 2 \sum_{i} C_{ij} E\left(\frac{s_i^2}{s_s}\right) + 2 \sum_{i} C_{ij} E\left(\frac{s_i^2}{s_s}\right) \]

Using (A.5), (A.7), (A.8) and (A.9) from Appendix A gives
\[ (3.13) \quad E_c = \frac{h}{m} \sum_i a_{ii} \left( s_i^2 f_{1,2} + f_{0,1} \right) + \frac{h}{m} \sum_{i \neq j} \left( s_i f_{1,2} + s_j f_{1,2} \right) \\
+ \frac{f_{1,0}}{2} \left[ \sum_i b_{ii} \left( 1 + \tilde{t}_i^2 \right) + \sum_{i \neq j} b_{ij} \tilde{t}_i \tilde{t}_j \right] \\
= \frac{h}{m} \left( s^\prime A s \right) f_{1,2} + \frac{1}{m} \left( \text{tr} A + 2t^\prime C s \right) f_{0,1} + \frac{1}{m} \left( t^\prime B \tilde{t} + \text{tr} B \right) f_{-1,0} \\
+ \frac{1}{m} \left( \text{tr} A \right) f_{0,1} + \frac{1}{m} \left( \text{tr} B \right) f_{-1,0} \\
= \gamma \theta f_{1,2} + t^\prime C s \left( f_{0,1} - f_{1,2} \right) + \frac{1}{m} t^\prime B \tilde{t} \left( f_{-1,0} - f_{1,2} \right) \\
+ \frac{1}{m} \left( \text{tr} A \right) f_{0,1} + \frac{1}{m} \left( \text{tr} B \right) f_{-1,0} \\
\]

Using the asymptotic expansion of the confluent hypergeometric function in (A.6) we can write functions like \( f_{0,1}, f_{1,2} \) etc., up to order \( \frac{1}{\theta^2} \) (or order \( \sigma^4 \)), as

\[ (3.14) \quad f_{\delta, \lambda} = \theta^{-\delta/2} \left[ 1 + \frac{(\lambda - \delta)(1-\frac{h}{2}) \theta}{\delta} + \frac{(\lambda - \delta)(\lambda - \delta + 1)(2-\frac{h}{2}) \theta}{\delta} \right]. \]

Substituting (3.14) into (3.13) gives, up to order \( \sigma^2 \),

\[ (3.15) \quad E(c - y) = \left[ \frac{1}{2} \text{tr} N - \frac{h}{2 \gamma} \right] + \frac{z^\prime Q_1 z}{\theta^2} \]

where \( \text{tr} A + \text{tr} B = \text{tr} P N P^\prime = \text{tr} N \) and \( Q_1 = P \begin{bmatrix} 0 & 0 \\ C & B \end{bmatrix} \). Since \( z^\prime Q_1 z \) is of order \( \sigma^{-2} \), and all other terms inside the square brackets are of order 1, \( E_c \to \gamma \) as \( \theta \to \infty \) by \( \sigma \to 0 \), with \( n \) fixed. That is, for a given \( n \), as the random component of \( y \) shrinks \( E_c \) approaches \( \gamma \).
To find \( Ec^2 \) we first expanded \( c^2 \) (see Appendix B) then took expectations and simplified to obtain

\[
(3.16) \quad Ec^2 = \gamma^2 \theta^2 f_{2,4} + [ (\tilde{\xi}' \cdot C \tilde{s}) (\tilde{s}' \cdot A \tilde{s}) ] (f_{1,3} - f_{2,4}) \\
+ [ \theta (\tilde{\xi}' \cdot B \tilde{s}) (\tilde{s}' \cdot A \tilde{s}) + (\tilde{\xi}' \cdot C \tilde{s}) ] (f_{0,2} - f_{2,4}) \\
+ [ (\tilde{\xi}' \cdot B \tilde{s}) (\tilde{s}' \cdot C \tilde{s}) ] (f_{1,1} - f_{2,4}) + [ \frac{4}{\kappa} (\tilde{\xi}' \cdot B \tilde{A}^2 \tilde{s} + \tilde{\xi}' \cdot A \tilde{s} - \tilde{s}' \cdot A \tilde{s} )^2 \tilde{s} ] f_{1,3} \\
+ [\frac{2}{\kappa} (\tilde{s}' \cdot A \tilde{s}) (\text{tr} A) + \tilde{s}' \cdot A^2 \tilde{s} + \frac{2}{\kappa} (\tilde{s}' \cdot A \tilde{s}) (\text{tr} B) + \tilde{s}' \cdot C \tilde{s} + \frac{2}{\kappa} (\text{tr} A)^2 ] f_{0,2} \\
+ \frac{4}{\kappa} (\text{vec} A)' (\text{vec} A) f_{0,2} \\
+ \frac{2}{\kappa} (\tilde{s}' \cdot B \tilde{A}^2 \tilde{s} + \tilde{s}' \cdot A \tilde{s} - \tilde{s}' \cdot A \tilde{s} ) (\text{tr} B) + \tilde{s}' \cdot C \tilde{s} + (\tilde{s}' \cdot C \tilde{s}) (\text{tr} B) + \text{tr}(C^2) \\
+ \frac{2}{\kappa} (\text{tr} B) (\text{tr} A) f_{1,1} \\
+ [2 \tilde{\xi}' \cdot B \tilde{A}^2 \tilde{s} + \tilde{s}' \cdot B \tilde{A}^2 \tilde{s} + 2 \tilde{s}' \cdot B \tilde{s} + (\tilde{s}' \cdot B \tilde{s}) (\text{tr} B) + (\text{tr} B)^2 ] f_{-2,0}
\]

where \( A_\times \) is a diagonal matrix whose non zero elements are from the main diagonal of \( A_\times \), \( A_{\times \times} \) is a matrix with the vector \([a_{11} \ a_{22} \ldots a_{hh}]\) for each column and \( \text{vec} A \) (or \( \text{vec} B \)) is the \( h^2 \times 1 \) vector of all the elements of \( A \) (or \( B \)).

Using (3.14) we can write the series expansion of \( Ec^2 \), up to order \( c^2 \), as (see Appendix B for details)
\[(3.17) \quad E \mathcal{C}^2 = \gamma^2 - \left( (h + 2) \gamma \right)^2 \frac{1}{\bar{\theta}} + \left[ \bar{\theta} (\bar{z}^2 \cdot Q_2 \cdot \bar{z}) \cdot (\text{tr} \; N) \right] \]

\[+ \bar{z} \cdot N \cdot \bar{z} + (\text{tr} \; B + 2 \gamma) \bar{z} \cdot Q_1 \cdot \bar{z} + \bar{z} \cdot (2A_{\alpha} \cdot A - A_{\alpha}^2) \bar{z} + c \cdot A_{\alpha} \cdot \bar{z} \]

\[+ c' \cdot \left[ 3B^2 + \frac{1}{\bar{\theta}^2} \cdot \text{tr} (B \cdot B) \cdot \bar{\epsilon} \right] \]

where

\[Q_2 = P' \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} P. \]

The term \((h + 2) \gamma^2\) inside the first set of curly brackets on the right side of \((3.17)\) is of order one. The terms within the second are of order \(\sigma^{-2}\). Since \(\theta\) and \(\theta^2\) are of orders \(\sigma^2\) and \(\sigma^4\), respectively, \(E \mathcal{C}^2 \rightarrow \gamma^2\) as \(\sigma \rightarrow 0\), that is as \(\theta \rightarrow \infty\).

We conclude from \((3.15)\) and \((3.17)\) that

\[(3.18) \quad \lim_{\sigma \rightarrow 0} E(c - \gamma) = 0 \]

and

\[(3.19) \quad \lim_{\sigma \rightarrow 0} E(c - \gamma)^2 = 0 \]

from which it follows that \(c\) converges in probability to \(\gamma\) as \(\sigma \rightarrow 0\).

We now consider the first moment of \(b\). Using \((3.3)\) and \((3.1)\) we have

\[(3.20) \quad b = (X'X)^{-1} X' \gamma - c (X'X)^{-1} X' \gamma_{-1} = \beta + (\gamma - c)(X'X)^{-1} X' \gamma_{-1} + (X'X)^{-1} X' \epsilon \]

\[= \beta + \sigma (\gamma - c)(X'X)^{-1} X' D_1 P' P z + (X'X)^{-1} X' \epsilon \]

\[= \beta + \sigma (\gamma - c) L \begin{bmatrix} s \\ t \end{bmatrix} + (X'X)^{-1} X' \epsilon \]

where \(L = (X'X)^{-1} X' D_1 P'\). Consider any element from this vector, say the first. Its expectation is

\[(3.21) \quad Eb_1 = \beta_1 + \sigma \gamma \ell'_1 \begin{bmatrix} s \\ t \end{bmatrix} - \sigma E[c' \ell'_1 \begin{bmatrix} s \\ t \end{bmatrix}] \]

where \(\ell'_1 = [d', e']\) is the first row of \(L\) which is partitioned to conform to
\[ s \quad \text{Ec}[d', e'] \begin{bmatrix} s \\ t \end{bmatrix} \] is obtained in detail in Appendix B. The expectation which results is,

\[(3.22) \quad Eb_1 = \beta_1 + \sigma \gamma (d' \tilde{s} + e' \tilde{t}) - \sigma (d' \tilde{s} \otimes f_{2,3} + e' \tilde{t} \otimes f_{1,2}) \]

\[- \sigma (t' C \tilde{s})(f_{1,2} - f_{2,3}) + (t' C \tilde{s})(e' \tilde{t})(f_{0,1} - f_{1,2}) \]

\[+ \frac{1}{2}(t' B \tilde{t})(d' \tilde{s})(f_{0,1} - f_{2,3}) + \frac{1}{2}(t' B \tilde{t})(e' \tilde{t})(f_{1,0} - f_{1,2}) \]

\[+ [d' A \tilde{s} + \frac{1}{2}(\text{tr } A)(d' \tilde{s})] \tilde{f}_{1,2} + [t' C d + e' C \tilde{s} + \frac{1}{2}(\text{tr } A)(e' \tilde{t}) \]

\[+ \frac{1}{2}(\text{tr } B)(d' \tilde{s})] f_{0,1} + [e' B \tilde{t} + \frac{1}{2}(\text{tr } B)(e' \tilde{t})] f_{1,0}. \]

We can now use (A.6) to write the terms, up to order \( \theta^{-1} \) of the series, like \( f_{2,3}, f_{1,2} \) etc., which appear in (3.22). After collecting terms we have

\[(3.23) \quad Eb_1 = \beta_1 + \sigma [\frac{Y}{2}(d' \tilde{s} + e' \tilde{t}) + \gamma d' \tilde{s} - \frac{1}{2}(\text{tr } N)(d' \tilde{s} + e' \tilde{t}) \]

\[+ [d' e'] P N P' \begin{bmatrix} \tilde{s} \\ \tilde{t} \end{bmatrix} \frac{1}{\theta} - \sigma ((1 - h)z' Q_1 \tilde{z} (d' \tilde{s} + e' \tilde{t}) \]

\[+ 2(z' Q_1 \tilde{z})(e' \tilde{t}) + (z' B \tilde{t})(d' \tilde{s} + e' \tilde{t})) \frac{1}{\theta^2} + O(\theta^{-2}). \]

The terms inside the first curly bracket are, after multiplication by \( \sigma \), of order one. The terms inside the second curly bracket are, after multiplication by \( \sigma \), of order \( \sigma^{-2} \), while \( \theta^{-1} \) is of order \( \sigma^{-2} \). Therefore if \( \theta \rightarrow 0 \) by \( \sigma \rightarrow 0 \)

\[ Eb_1 \rightarrow \beta_1. \] That is, for a given \( n \), as the random component of the model shrinks \( Eb_1 \) approaches \( \beta \).

If we wish to consider the distribution of \( b \) as \( \sigma \rightarrow 0 \) we must standardize equation (3.20) to obtain
\[ (3.24) \quad \frac{1}{\sigma} (b - \beta) = (\gamma - \sigma) (X'X)^{-1} X' D_1 z + \frac{1}{\sigma} (X'X)^{-1} X' \varepsilon . \]

Now \((X'X)^{-1} X' D_1 z\) is multivariate normal with means vector \((X'X)^{-1} X' \mu_{-1}\) and covariance matrix \((X'X)^{-1} X' D_1 D_1' X(X'X)^{-1} = (X'X)^{-1}\). Also \(\frac{1}{\sigma} (X'X)^{-1} X' \varepsilon\) is multivariate normal with a zero means vector and a covariance matrix \((X'X)^{-1} X' \Sigma X(X'X)^{-1}\), where

\[
(3.25) \quad \Sigma = \begin{bmatrix}
1 + \gamma^2 & -\gamma & 0 & 0 \\
-\gamma & 1 + \gamma^2 & -\gamma & 0 \\
0 & -\gamma & 1 + \gamma^2 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

Neither of these two multivariate normal distributions depend on \(\sigma^2\). Furthermore, from (3.18) and (3.19), \((c - \gamma)\) converges in probability to zero as \(\sigma^2 \to 0\). Therefore, \(\frac{1}{\sigma} (b - \beta)\) converges in distribution to \(\frac{1}{\sigma} (X'X)^{-1} X' \varepsilon\) as \(\sigma^2 \to 0\). Hence, as \(\sigma^2\) grows small \(b\) is approximately normal and unbiased with approximate covariance matrix \(\sigma^2 (X'X)^{-1} X' \Sigma X(X'X)^{-1}\).

4. Consistent Estimation of \(\gamma\)

Liviatan (1963) has proposed two consistent estimators for models like (2.4). The simplest of them uses a lagged exogenous variable, \(\hat{w}\), as an instrument to produce the normal equations.

\[ (4.1) \quad \gamma \hat{w}'y_{-1} + \hat{w}'\hat{\beta} = \hat{w}'y \]

\[ (4.2) \quad \hat{\gamma} X'y_{-1} + X'\hat{\beta} = X'y \]

from which

\[ (4.3) \quad \hat{\gamma} = \frac{w'M_y}{w'My_{-1}} = \frac{u}{u_1} \]

where \(u \sim N(w'M \mu, \sigma^2 w'Mw) = N(u, \omega^2), u_1 \sim N(w'M \mu_{-1}, \sigma^2 w'Mw) = N(u_1, \omega^2)\) and
\[ E(u - \bar{u})(u_1 - \bar{u}_1)' = \sigma^2 \omega M D_2 M w = \rho \omega^2 \] where \( \rho \) is the coefficient of correlation between \( u \) and \( u_1 \). Under (2.1), this ratio has a distribution of the type described by Fieller (1932) which has no moments of any order. However, if \( \frac{\omega}{\bar{u}_1} < 1/3 \) the distribution of \( \hat{\gamma} \) can be approximated by (Scadding (1973)).

\[
(4.4) \quad f(\hat{\gamma}) = \frac{-2 \{ (\bar{u}_1 - \bar{u}) + (\bar{u}_1 - \bar{u})\hat{\gamma} \}}{\sqrt{2\pi \omega^2 (\hat{\gamma}^2 - 2\hat{\gamma} \rho + 1)^3}} \exp \left\{ \frac{- (\bar{u}_1 - \bar{u} \hat{\gamma})^2}{2\omega^2 (\hat{\gamma}^2 - 2\hat{\gamma} \rho + 1)} \right\}.
\]

When this approximation is valid \( \hat{\gamma} \) is nearly unbiased (Nagar and Gupta (1968), Carter (1976)). Higher moment can be obtained from Merrill (1928). Scadding (1973) and Nagar and Gupta (1968) analyzed the distribution of \( \hat{\gamma} \) only. We wish to consider the distribution of \( \hat{\beta} \). So we use (4.2) to obtain

\[
(4.5) \quad \hat{\beta} = (X' X)^{-1} X' (\gamma - \hat{\gamma} y_1).
\]

Its sampling error can be written as

\[
(4.6) \quad \frac{1}{\sigma} (\hat{\beta} - \beta) = (\gamma - \hat{\gamma})(X' X)^{-1} X' D_1 z_1 + \frac{1}{\sigma} (X' X)^{-1} X' \varepsilon.
\]

If \( \sigma \to 0 \) (hence \( \frac{\omega}{\bar{u}_1} \to 0 \)), \( \gamma \to \gamma \) and the distribution of \( \hat{\beta} - \beta \) approaches to that of \( \frac{1}{\sigma} (X' X)^{-1} X' \varepsilon \) which is the same as the limiting distribution of \( \frac{1}{\sigma}(\beta - \beta) \) as seen at the end of section 3. Then, like \( b \), as \( \sigma \to 0 \) the distribution of \( \hat{\beta} \) is approximately normal with means vector \( \beta \) and covariance matrix \( \sigma^2 (X' X)^{-1} X' \varepsilon (X' X)^{-1} \).
5. Conclusions

The coefficients of a rational distributed lag model with a first degree polynomial in the denominator can be estimated by least squares or by instrumental variables. This paper presents some exact and asymptotic properties of these estimators.

We find that the OLS estimator of the coefficient of the lagged dependent variable converges in probability to the true value of the coefficient as $\theta$, the non-centrality parameter of its distribution, grows large; that is as $\sigma$, the standard deviation of the errors, grows small. Also, as $\sigma$ grows small, the OLS estimator of the coefficients vector of the exogenous variables becomes unbiased. In addition, for small $\sigma$, the IV coefficient estimators possess these same properties. These findings suggest that when both the sample size and the error variance are small OLS is a useful estimator which is not inferior to IV.
APPENDIX

A. Expectations Required in Section 3

Let \( z_1, \ldots, z_T \) be independent normal variates with

\[
E z_i = \tilde{z}_i \quad \text{and} \quad \text{Var} \ z_i = 1 \quad i = 1, \ldots, T
\]

Then we know that the distribution of

\[
W = z' B z,
\]

where \( z' = [z_1, \ldots, z_T] \) and \( B \) is idempotent with rank \( h \), is 'noncentral chi-square' with \( h \) degrees of freedom and the parameter of noncentrality

\[
\theta = \frac{1}{2} z' B z
\]

The density function of \( W \) is given by

\[
f(W) = e^{-\theta} \sum_{m=0}^{\infty} \frac{\theta^m}{m!} \frac{\Gamma(h+2m)}{2^{h+2m} \Gamma(h+2m/2)} e^{-\frac{1}{2}W} \quad 0 < W < \infty.
\]

Therefore, if \( h/2 > r \) [see Ullah (1974, p. 147)]

\[
E W^{-r} = \int_0^\infty W^{-r} f(W) \ dW = 2^{-r} f_{-r,0}^{-} \quad r=1,2,\ldots
\]

where

\[
f_{\delta,\nu} = \frac{\Gamma(h/2 + \delta)}{\Gamma(h/2 + \nu)} e^{-\theta} \ \text{I}_F(h/2 + \delta; h/2 + \nu; \theta)
\]

and writing \( \delta = -r, \nu = 0 \), we get \( f_{-r,0}^{-} \) and so on. The function \( \text{I}_F(\ ) \) represents the confluent hypergeometric function.

The expectations required in Section 3 may now be stated as follows:

\[
E(z_i W^{-r}) = 2^{-r} \tilde{z}_i f_{-r+1,1}^{-}
\]
\[(A.8)\] \(E(z_i^2 W^{-r}) = 2^{-r} [z_i^2 f_{-r+2,2} + f_{-r+1,1}]\);

\[(A.9)\] \(E(z_i z_j W^{-r}) = 2^{-r} \bar{z}_i \bar{z}_j f_{-r+2,2}\)

\[(A.10)\] \(E(z_i^3 W^{-r}) = 2^{-r} z_i^3 f_{-r+3,3} + 3 \times 2^{-r} \bar{z}_i f_{-r+2,2}\)

\[(A.11)\] \(\sum_{i \neq j} E(z_i z_j W^{-r}) = 2^{-r} z_i \bar{z}_j f_{-r+3,3} + 2^{-r} \bar{z}_j f_{-r+2,2}\)

\[(A.12)\] \(E(z_i z_j z_k W^{-r}) = 2^{-r} \bar{z}_i \bar{z}_j \bar{z}_k f_{-r+3,3}\)

\[(A.13)\] \(E(z_i^4 W^{-r}) = 2^{-r} z_i^4 f_{-r+4,4} + 6 \times 2^{-r} \bar{z}_i^2 f_{-r+3,3} + 3 \times 2^{-r} f_{-r+2,2}\)

\[(A.14)\] \(\sum_{i \neq j} E(z_i z_j W^{-r}) = 2^{-r} z_i \bar{z}_j f_{-r+4,4} + 3 \times 2^{-r} \bar{z}_i \bar{z}_j f_{-r+3,3}\)

\[(A.15)\] \(\sum_{i \neq j} E(z_i^2 z_j W^{-r}) = 2^{-r} z_i \bar{z}_j f_{-r+3,3} + 2^{-r} (z_i^2 + \bar{z}_j^2) f_{-r+3,3} + 2^{-r} f_{-r+2,2}\)

\[(A.16)\] \(E(z_i z_j z_k W^{-r}) = 2^{-r} z_i \bar{z}_j \bar{z}_k f_{-r+4,4} + 2^{-r} \bar{z}_j \bar{z}_k f_{-r+3,3}\)

\[(A.17)\] \(\sum_{i \neq j \neq k} E(z_i z_j z_k W^{-r}) = 2^{-r} z_i \bar{z}_j \bar{z}_k f_{-r+4,4}\)
B. The Evaluation of Useful Expectations

The first step in finding $E c^2$ is to expand $c^2$ as:

\[ c^2 = \frac{1}{(s's')^2} [(s'A s)^2 + 4(t' C s)(s'A s) + 2(t'B t)(s'A s) \]
\[ + 4(t' C s)^2 + 4(t'B t)(t' C s) + (t'B t)^2]. \]

Keeping in mind the symmetry of $A$ and $B$, the terms inside the square brackets can be expanded as:

\[ (s'A s)^2 = \sum_i a_{ii} s_i^4 + 4 \sum_i \sum_j a_{ij} s_i s_j + 2 \sum_i \sum_j \sum_k a_{ijk} s_i s_j s_k \]
\[ + 4 \sum_i \sum_j \sum_k a_{ij} s_i s_j s_k + 2 \sum_i \sum_j \sum_k a_{ij} s_i s_j s_k \]
\[ + \sum_i \sum_j \sum_k a_{ijk} s_i s_j s_k. \]

\[ (t'C s)(s'A s) = \sum_i \sum_k c_{ik} a_{kk} t_i s_k^3 + \sum_i \sum_j c_{ij} a_{kk} t_i s_k s_j \]
\[ + 2 \sum_i \sum_j \sum_k c_{ij} a_{ij} t_i s_j s_k + \sum_i \sum_j \sum_k c_{ij} a_{kk} t_i s_j s_k \]
\[ + \sum_i \sum_j \sum_k c_{ij} a_{ij} t_i s_j s_k. \]

\[ (t'C s)^2 = \sum_i \sum_k c_{ik} t_i s_k^2 + \sum_i \sum_k c_{ik} c_{ik} t_i s_k \]
\[ + \sum_i \sum_j \sum_k c_{ik} c_{jk} t_i j s_j s_k + \sum_i \sum_j \sum_k c_{ij} c_{ij} t_i j s_j s_k. \]

\[ (t'B t)(t'C s) = \sum_i \sum_k b_{ik} c_{ik} t_i s_k^3 + \sum_i \sum_j b_{ik} c_{ij} t_i j s_j s_k \]
\[ + 2 \sum_i \sum_j \sum_k b_{ij} c_{ij} t_i j s_j s_k + \sum_i \sum_j \sum_k b_{ij} c_{ij} t_i j s_j s_k. \]
Since $t_i \sim N(\tilde{\xi}_i, 1)$, the moments of $t_i$ are

(B.6) \hspace{1cm} E t_i^2 = 1 + \tilde{\xi}_i^2

(B.7) \hspace{1cm} E t_i^3 = \tilde{\xi}_i^3 + 3 \tilde{\xi}_i

(B.8) \hspace{1cm} E t_i^4 = \tilde{\xi}_i^4 + 6 \tilde{\xi}_i^2 + 3.

Using (B.1) to (B.8) together with (A.5) to (A.17) we can write the exact second moment of $c$ for $h > 4$.

(B.9) \hspace{1cm} c^2 = \frac{1}{h} \sum_i \frac{1}{i} a_{i1}(s_i, f_{2,4} + 6 s_i f_{1,3} + 3 f_{0,2}) +

\hspace{1cm} + \frac{1}{h} \sum_{i \neq j} a_{ij}(s_i^2 s_j f_{2,4} + 3 s_i s_j f_{1,3})

\hspace{1cm} + \frac{1}{h} \sum_{i \neq j \neq k} (a_{ii} a_{jk} + 2 a_{ij} a_{ik})(s_i s_j s_k f_{2,4} + s_i s_j s_k f_{1,3})

\hspace{1cm} + \frac{1}{h} \sum_{i \neq j \neq k \neq l} (a_{ij} + 2 a_{ij})(s_i s_j s_k s_l f_{2,4} + (s_i^2 + s_j^2)f_{1,3} + f_{0,2})

\hspace{1cm} + \frac{1}{h} \sum_{i \neq j \neq k} a_{ij} a_{ik} \tilde{\xi}_i (s_j^3 f_{1,3} + 3 s_j f_{0,2}) +

\hspace{1cm} + \frac{1}{h} \sum_{i \neq j \neq k} (a_{kk} + 2 a_{jk}) \tilde{\xi}_i (s_j s_k f_{1,3} + s_k f_{0,2})

\hspace{1cm} + \frac{1}{h} \sum_{i \neq j \neq k \neq l} a_{ij} a_{ik} \tilde{\xi}_i \tilde{\xi}_1 (s_j s_k s_l f_{1,3})

\hspace{1cm} + \frac{1}{h} \left[ \sum_{i \neq j} \frac{1}{i} (1 + \tilde{\xi}_i^2) + \sum_{i \neq j} \frac{1}{i} \frac{m}{i} b_i (\tilde{\xi}_i \tilde{\xi}_j \tilde{\xi}_1) \right] \sum_{i \neq j} \frac{1}{i} (s_i^2 f_{0,2} + f_{-1,1})
\[ \sum_i \sum_j a_{ij} \tilde{s}_i \tilde{s}_j f_{0,2} \]

\[ + \sum_i \sum_j a_{ij} (1 + \tilde{\epsilon}_j^2) + \sum_i \sum_j b_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j (\tilde{s}_i^2 f_{0,2} + f_{-1,1}) \]

\[ + \sum_i \sum_k c_{ik} (1 + \tilde{\epsilon}_i^2) + \sum_i \sum_j c_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j (\tilde{s}_i^2 f_{0,2} + f_{-1,1}) \]

\[ + \sum_i \sum_k c_{ik} (1 + \tilde{\epsilon}_i^2) + \sum_i \sum_j c_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j (\tilde{s}_i^2 f_{0,2} + f_{-1,1}) \]

\[ + \sum_i \sum_j c_{ij} (1 + \tilde{\epsilon}_i^2 + 3\tilde{\epsilon}_i) + \sum_i \sum_j b_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j (1 + \tilde{\epsilon}_i^2) \tilde{\epsilon}_j \]

\[ + \sum_i \sum_j b_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j \tilde{\epsilon}_k \]

\[ + \sum_i \sum_j b_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j (1 + \tilde{\epsilon}_i^2) \tilde{\epsilon}_j \]

\[ + \sum_i \sum_j b_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j (1 + \tilde{\epsilon}_i^2 + 3\tilde{\epsilon}_i) \tilde{\epsilon}_j \]

\[ + \sum_i \sum_j b_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j \tilde{\epsilon}_k \]

\[ + \sum_i \sum_j b_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j (1 + \tilde{\epsilon}_i^2 + 3\tilde{\epsilon}_i) \tilde{\epsilon}_j \]

\[ + \sum_i \sum_j b_{ij} \tilde{\epsilon}_i \tilde{\epsilon}_j (1 + \tilde{\epsilon}_i^2) \tilde{\epsilon}_j \]

In moving from (B.9) to (3.16) we have used (B.2) to (B.5) as well as the following equalities:
\( (\tilde{s}' A \tilde{s})(tr\ A) = \sum_{i} a_{ii} \tilde{s}_{i}^{2} + \sum_{i \neq j} a_{ij} \tilde{s}_{i} \tilde{s}_{j} + 2 \sum_{i \neq j} a_{i} a_{j} \tilde{s}_{i} \tilde{s}_{j} \)

\( + \sum_{i \neq j = k} a_{ik} \tilde{s}_{i} \tilde{s}_{j} \)

\( (\tilde{s}' A^{2} \tilde{s} = \sum_{i} a_{ik} \tilde{s}_{i}^{2} + \sum_{i \neq j} a_{ik} a_{jk} \tilde{s}_{i} \tilde{s}_{j} \)

\( (\tilde{s}' A_{x} \tilde{s} = \sum_{i} a_{ii} \tilde{s}_{i}^{2} + \sum_{i \neq j} a_{ij} \tilde{s}_{i} \tilde{s}_{j} \)

\( \tilde{c}' CA \tilde{s} = \sum_{i} c_{ij} a_{jj} \tilde{s}_{i} \tilde{s}_{j} + \sum_{i \neq j} c_{ij} a_{ik} \tilde{s}_{i} \tilde{s}_{k} \)

\( \tilde{c}' CA_{x} \tilde{s} = \sum_{i} c_{ij} a_{jj} \tilde{s}_{i} \tilde{s}_{j} + \sum_{i \neq j} c_{ij} a_{ik} \tilde{s}_{i} \tilde{s}_{k} \)

\( \tilde{c}' C \tilde{s} = \sum_{j} c_{ij} \tilde{s}_{j}^{2} + \sum_{i \neq j} c_{ij} c_{ik} \tilde{s}_{j} \tilde{s}_{k} \)

\( (\text{vec } A)' (\text{vec } A) = \sum_{i} a_{ii}^{2} + \sum_{i \neq j} a_{ij} \)

\( (\tilde{c}' C \tilde{s})(tr\ B) = \sum_{i} b_{i} c_{ik} \tilde{s}_{i} \tilde{s}_{k} + \sum_{i \neq j} b_{i} c_{jk} \tilde{s}_{i} \tilde{s}_{k} \)

The terms \( (\tilde{s}' A \tilde{s})(tr\ B) \), \( (\tilde{c}' B \tilde{c})(tr\ A) \) and \( (\tilde{c}' B \tilde{c})(tr\ B) \) are all similar in form to \( (B.10) \), \( \tilde{c}' B C \tilde{s} \) is similar to \( (B.13) \), \( \tilde{c}' C C' \tilde{c} \) and \( \tilde{c}' B B' \tilde{c} \) are similar to \( (B.15) \) and \( (\text{vec } B)' (\text{vec } B) \) is similar to \( (B.16) \). \( A_{x} \) is the diagonal matrix formed by setting all off diagonal elements of \( A \) to zero. \( A_{x,x} \) is an h\times h matrix whose every column is the vector \( [a_{11} a_{22} \ldots a_{hh}]' \) and \( \text{vec } A \) is the h^2\times 1 vector of all elements of \( A \).
In moving from (3.16) to (3.17) the following equality was used:

\[(B.18)\quad \tilde{z}^* N^2 \tilde{z} = s' A^2 s + s' C'C s + 2 \tilde{z}' C A \tilde{s} + 2 \tilde{z}' B C \tilde{s} + \tilde{t}' C C' \tilde{t} + \tilde{t}' B^2 \tilde{t}\]

We consider now \(E c[d' e'] \begin{bmatrix} s \\ t \end{bmatrix}\) which is a part of \(E b_1\). The first component of this expectation is

\[(B.19)\quad E c d' s = E \frac{(s' A s + 2 \tilde{z}' C s + \tilde{t}' B t)(d' s)}{s' s}\]

The terms in the numerator inside the square bracket must be expanded to give:

\[(B.20)\quad (s' A s)(d' s) = \sum_{i} a_{ii} d_{i} s_{i}^{3} + \sum_{i} a_{ii} d_{i} s_{i}^{2} s_{k} + \sum_{i} a_{ij} d_{i} s_{i} s_{j} s_{k}\]

\[+ \sum_{i} \sum_{j} a_{ij} d_{i} s_{i} s_{j} s_{k}\]

\[(B.21)\quad (\tilde{c}' C s)(d' s) = \sum_{i} c_{ij} d_{i} t_{i} s_{j} s_{k}\]

\[(B.22)\quad (t' B t)(d' s) = (\sum_{i} b_{ii} t_{i}^{2} + \sum_{i} \sum_{j} b_{ij} t_{i} t_{j})(s_{k})\]

When we combine the expansions with the results of Appendix A we obtain

\[(B.23)\quad E c d' s = \frac{h}{6} \sum_{i} a_{ii} d_{i} (s_{i}^{3} f_{2,3} + 3 s_{i} f_{1,2}) + \frac{h}{6} \sum_{i} a_{ii} d_{i} (s_{i}^{2} \tilde{s}_{k} f_{2,3}\]

\[+ s_{k} f_{1,2})\]

\[+ \sum_{i} \sum_{j} a_{ij} d_{i} (s_{i} s_{j} f_{2,3} + s_{j} f_{1,2}) + \sum_{i} \sum_{j} a_{ij} s_{i} s_{j} s_{k} f_{2,3}\]
\[ m \ h \ + \ \Sigma \Sigma \ c_{ij} d_{j} \tilde{t}_{i}(s_{j} f_{1,2} + f_{0,1}) + \Sigma \Sigma \Sigma \ c_{ij} d_{k} \tilde{t}_{i} s_{i} s_{j} f_{1,2} \]
\[ + \ \frac{m}{h} (\Sigma b_{ii}(\tilde{t}_{i}^{2} + 1) + \Sigma b_{ij} \tilde{t}_{i} \tilde{t}_{j}) \Sigma d_{k} s_{k} f_{0,1}. \]

Noting that
\[ (B.24) \quad d' A \tilde{s} = \Sigma a_{ii} d_{i} \tilde{s}_{i} + \Sigma a_{ij} d_{i} \tilde{s}_{j}, \]
\[ (B.25) \quad (\text{tr } A)(d' \tilde{s}) = \Sigma a_{ii} d_{i} \tilde{s}_{i} + \Sigma a_{ij} d_{j} \tilde{s}_{j} \quad \text{and} \]
\[ (B.26) \quad \tilde{c'} C d = \Sigma c_{ij} \tilde{t}_{i} d_{j} \]

we can simplify (B.23) to
\[ (B.27) \quad E c d' s = \gamma \Theta d' \tilde{s} f_{2,3} + (\tilde{c'} C \tilde{s})(d' \tilde{s})(f_{1,2} - f_{2,3}) + \frac{1}{h} (\tilde{c'} B \tilde{c})(d' \tilde{s})(f_{0,1} - f_{2,3}) \]
\[ + [d'A \tilde{s} + \frac{1}{h} (\text{tr } A)(d' \tilde{s})] f_{1,2} + [(\tilde{c'} C d) + \frac{1}{h} (\text{tr } B)(d' \tilde{s})] f_{0,1}. \]

The second component of \( E c[d' e'] \)
\[ \begin{bmatrix} s \\ t \end{bmatrix} \]
is
\[ (B.28) \quad E c e' t = E[\tilde{e}' A s](e' t) + 2(t' C s)(e' t) + (t' B t)(e' t)] \]
\[ = \frac{h}{m} \Sigma a_{ii}(s_{i} f_{1,2} + f_{0,1}) + \frac{h}{m} \Sigma a_{ij} \tilde{s}_{i} \tilde{s}_{j} f_{1,2} e' \tilde{t} \]
\[ + \Sigma[\Sigma c_{ij} e_{i}(1 + \tilde{t}_{i}^{2}) + \Sigma c_{ij} e_{k} \tilde{t}_{i} \tilde{t}_{k}] s_{j} f_{0,1}. \]
\[ + \sum_{i=1}^{m} b_{ii} e_i (\bar{t}_i^3 + \bar{t}_i^3 + 3\bar{t}_i^1) + \sum_{i \neq k}^{m} e_k (1 + \bar{t}_i^2) \bar{t}_k \]
\[ + 2 \sum_{i \neq j}^{m} b_{ij} e_i (1 + \bar{t}_i^2) \bar{t}_j + \sum_{i \neq j \neq k}^{m} e_k \bar{t}_i \bar{t}_j \bar{t}_k \frac{f_{-1,0}}{2} \]
\[ = \gamma \theta (e' \bar{t}) f_{1,2} + (\bar{t}' C \bar{s})(e' \bar{t})(f_{0,1} - f_{1,2}) + \frac{1}{2} \theta (\bar{t}' B \bar{t})(e' \bar{t})(f_{-1,0} - f_{1,2}) \]
\[ + \left[ \frac{1}{2} (\text{tr} A)(e' \bar{t}) + C \bar{s} \right] f_{0,1} + \left[ e' B \bar{t} + \frac{1}{2} \theta (\text{tr} B)(e' \bar{t}) \right] f_{-1,0} \]

Using similar expansions and simplifications as were used in deriving E c d's, equation (3.22) is obtained by substituting (B.27) and (B.28) into (3.21).
Footnotes

1The authors are grateful to D. Hendry, R. D. Terrell, R. Kohn and D. F. Nicholls for extremely useful suggestions on our earlier draft of this paper which was entitled "The Finite Sample Properties of OLS and IV Estimators in Regression Models with a Lagged Dependent Variable." An earlier version of this paper was presented to the European meetings of the Econometric Society in Vienna in September 1977.

2At first glance equation (3.5) resembles the analogous equation for two-stage least squares and so one might hope to use the results of Richardson (1968) and Sawa (1969) to analyze the behavior of \( c \). However, these earlier results do not apply here directly because the elements of \( y_{-1} \) are not independent of those of \( y \) and there is no cannonical form of the model for which independence holds. In fact the covariance matrix of \( y \) and \( y_{-1} \) is singular. Another complication is the presence of \( y_T \) in the numerator of (3.5) but not in the denominator.

3If \( \theta > 0 \) and \( a, c > 0 \), then, using Sawa's (1972, p. 667) results we have

\[
 _1 \text{F}_1 (a;c;\theta) = \frac{\Gamma c}{\Gamma a} \theta^{-(c-a)} \left[ \sum_{j=0}^{p-1} \frac{(c-a)_j (1-a)_1}{j!} \theta^{-j} + O(\theta^{-p}) \right].
\]

\( \theta \) will grow if \( M \mu_1 \rightarrow \infty \) or if \( \sigma^2 \rightarrow 0 \). Kadane (1970), (1971) has analyzed the behavior of estimators as \( \sigma^2 \rightarrow 0 \).

One can also obtain the bias up to order \( 1/n \) by using the following result. For large \( a \) and \( b \), with \( \theta > 0 \),
\[ _1F_1(a; b; bx) = e^{bx} (1+bx)^{a-b} \left[ 1 - \frac{(b-a)(b-a+1)}{2b} \left( \frac{x}{1+x} \right)^2 + o \left( \frac{1}{|b|^2} \right) \right] \]

so long as \((b-a)\) and \(x\) are bounded; Slater (1960, p. 66). \(\theta\) is also a concentration parameter because

\[ p[|c - \gamma| > \varepsilon] = 0 \text{ as } \theta \to \infty \]

using (3.18) and (3.19).

Liviatan considered the case where \(X\) has only one column so the choice of \(w\) was obvious.

\[ _1F_1(a; c; x) = \frac{c}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{\Gamma(a+n)}{n!} \frac{x^n}{\Gamma(c+n)} \]

The results in (A.7), (A.8), (A.10) and (A.13) are given in Ullah (1974). (A.7) and (A.8) also follow by using (A.5) in the results of Bock (1975, p. 216). The remaining expectations can be obtained from Nagar and Ullah (1973).

We have not expanded the product \((t' B t)(s' A s)\) because \(t\) is independent of \(s\). Also the expansion of \((t' B t)^2\) is of the same form as (B.2).
References


Merrill, A. S. (1928), "Frequency Distribution of an Index When Both the Components Follow the Normal Law," *Biometrika* 20, pp. 53-63.


