7-1-1976

The Small-Disturbance-Asymptotic Moments of the Instrumental Variables and Ordinary Least Squares Estimators for a Dynamic Equation with Correlated Errors

Jon K. Peck

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 433

THE SMALL-DISTURBANCE-ASYMPTOTIC MOMENTS OF THE INSTRUMENTAL VARIABLES AND ORDINARY LEAST SQUARES ESTIMATORS FOR A DYNAMIC EQUATION WITH CORRELATED ERRORS

Jon K. Peck

July 19, 1976

THE SMALL-DISTURBANCE-ASYMPTOTIC MOMENTS OF THE INSTRUMENTAL VARIABLES AND ORDINARY LEAST SQUARES ESTIMATORS FOR A DYNAMIC EQUATION WITH CORRELATED ERRORS*

by

Jon K. Peck

This paper presents an analysis of some finite-sample properties of the instrumental variables (IV) and ordinary least squares (OLS) estimators for a single equation that includes the lagged dependent variable as a regressor. The disturbances in the equation are assumed to be normally distributed with an arbitrary nonsingular covariance matrix. Approximations are found for the mean and mean squared error of the estimators. The approximations improve in accuracy as the disturbance variance becomes small. Comparisons are presented with the conventional large-sample-asymptotic approximations for this model; some comparisons of IV and OLS are presented.

The model to be estimated is

$$(1) \qquad \underset{T\times 1}{y} = \underset{T\times 1}{y_{-1}}\alpha + \underset{T\times K}{X}\ \underset{K\times 1}{\beta} + \sigma \underset{T\times 1}{u}$$

where $y = (y_1, \ldots, y_T)'$, $y_{-1} = (y_0, y_1, \ldots, y_{T-1})'$,

*The research described in this paper was undertaken by grants from the National Science Foundation and Ford Foundation. The author would like to thank without implicating David F. Hendry and J. Kadane for their assistance.

$X = (x_1, x_2, \ldots, x_K)$, $x_i = (x_{i1}, \ldots, x_{iT})'$, and $u = (u_1, \ldots, u_T)$. The disturbances, u , are assumed to be normally distributed independently of X with mean zero and a normalized nonsingular covariance matrix Ω . The regressors are of full rank, and X is assumed to be nonstochastic.

Since finite-sample results are to be obtained, some assumption must be made about the initial observation on y , y_0 . Perhaps the most natural assumption would be that it is stochastic and drawn from the same distribution as the other observations, y_t, but this is a conditional distribution depending on the regressor values in the pre-sample period. These values are unknown by assumption. If a distribution were assumed for the exogenous variables, a marginal distribution could be computed for y_0 if the equation were stable, but this is not consistent with the assumptions made about the exogenous variables. It is, therefore assumed that y_0 is fixed and the results obtained below are thus conditional on the value of y_0 .

The approach taken in this paper is to determine the approximate bias and mean squared error of the estimators where the approximation is accurate up to terms of order σ^{2k} , the disturbance variance, rather than the customary approximation up to terms of order T^k , where T is the sample size. T is a parameter in the sigma expansions. This approach has been applied in Brown [2], Kadane [4], and Peck [6].

The approximations to be presented are referred to as the (small-sigma)-asymptotic moments of the estimator, but it is not guaranteed that these approximations converge to the exact finite-sample moments any more than the moments of a large-sample limiting distribution are necessarily the limits of the finite-sample moments. Indeed, the exact finite sample moments of IV do not even exist. (See e.g. Hatanaka [3])

Even in such a case it can be argued that the (finite) limiting distribution moments are a useful approximate characterization of the behavior of the distribution in finite samples and are indications of where probability is concentrated. They are, perhaps, more useful than the infinite "exact" moments. The validity of this type of expansion is analyzed further in Ramage [8] and Sargan [11].

The calculation of the small-sigma moments is performed by expressing the error and squared error of the estimator, $e = \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix}$, as infinite series in powers of σ . Expectations of these series are taken term by term until sufficient accuracy of the approximation has been achieved. These approximations, then, are more accurate as the disturbance variance is smaller.[1]

In deriving the small disturbance moments of the estimators it is useful to distinguish between the original equation (1) and the "final form" of that equation in which y_{-1} does not appear. The final form is given by

$$
\begin{aligned}
(2) \qquad y &= (W + \sigma V)\alpha + X\beta + \sigma u \\
&= (Z + \sigma V^{*})\begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \sigma u
\end{aligned}
$$

where W and V are $T\mathrm{x}1$ column vectors:

$$W = (w_0, w_1, \ldots, w_{T-1})' \quad \text{and}$$

$$V = (v_0, v_1, \ldots, v_{T-1})'$$

[1] In the dynamic model, the error variance relative to the dispersion of X is relevant, i.e., the noise-signal ratio. This method cannot be used to study the case where the equation contains no exogenous variables (an infinite ratio).

with elements

$$
(3) \qquad w_t = \begin{cases} \sum_{j=1}^{t} \alpha^{t-j} X_j \beta + \alpha^t y_0 , & t = 1, 2, \ldots, T-1 \\ y_0 , & t = 0 \end{cases}
$$

and

$$
(4) \qquad v_t = \begin{cases} \sum_{j=1}^{t} \alpha^{t-j} u_j , & t = 1, 2, \ldots, T-1 \\ 0 , & t = 0 . \end{cases}
$$

Z and V^* are $T\times(k+1)$ matrices. $Z = (\underset{T\times 1}{W} \quad \underset{T\times k}{X})$ and $V^* = (\underset{T\times 1}{V} \quad \underset{T\times k}{0})$.[2] Z is the nonstochastic part of the regressors including the systematic component of y_{-1} and V^* is the remaining stochastic component. It should be noted that W and V correspond to the fixed and random parts of y_{-1} ; not to y itself.

The following definitions of expectations are used below. $Euu' \equiv \Omega$, $EuV' \equiv C$, and $EVV' = G$. Ω is given by assumption. C and G depend in turn on Ω. The theorems to be presented are valid for any nonsingular Ω provided that C and G are appropriately computed. From expression (4) above it is clear that $V = Au$ where A is the lower triangular matrix

$$
A = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ \alpha^0 & 0 & \ldots & 0 \\ \alpha^1 & \alpha^0 & 0 \ldots & 0 \\ \vdots & & & \\ \alpha^{T-2} & \alpha^{T-3} & \ldots \alpha^0 & 0 \end{bmatrix}
$$

[2] A derivation of this form is given in Peck [6].

Therefore, for any Ω, $C = EuV' = Euu'A' = \Omega A'$ and $G = EVV' = EAuu'A = A\Omega A' = AC$.

The difficulty in the estimation of equation (1) is that the lagged dependent variable is correlated with u unless $\Omega = I$. Thus ordinary least squares is an inconsistent estimator. IV is consistent in this case for appropriately chosen instruments although biased but is asymptotically inefficient compared with estimators which take the covariance structure into account (See Sargan [10].) IV is the simplest consistent estimator to compute and can be used even when $\Omega = I$, therefore, a preliminary test for autocorrelation or other nonrandomness in the disturbances is not required although it may be beneficial (see Peck [7]). We compute first the small disturbance moments of IV and then compute the small disturbance moments of (inconsistent) OLS for comparison.

It is assumed that a $T\times r$ matrix, N, of r nonstochastic instruments for y_{-1} is available. Let $D = (N \vdots X)$. The instruments N are assumed to satisfy $E(u|N) = 0$ and $D'Z$ is of full rank $k+1$. N is thus correlated with W, the nonstochastic part of y_{-1}, and uncorrelated with V, the stochastic part. Lagged values of the exogenous variables X will generally satisfy these assumptions (except for the constant term and a time trend). The exogenous variables are their own instruments.[3]

Define P_R as the orthogonal projector into the space spanned by the columns of R, $P_R = R(R'R)^{-1}R'$ and $\bar{P}_R = I - P_R$. Then the instrumental variables estimator is

[3] This formulation rules out IV procedures not using all variables in X as instruments.

$$(5) \qquad \widehat{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}} = [(y_{-1}X)'P_D(y_{-1}X)]^{-1}(y_{-1}X)'P_Dy \ .$$

Since $(y_{-1}X) = Z + \sigma V^*$, the error of the estimator $e = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$ can be written as

$$(6) \qquad e = \sigma[(Z+\sigma V^*)'P_D(Z+\sigma V^*)]^{-1}(Z+\sigma V^*)'P_Du \ .$$

It is assumed that the matrix inverted in expressions (5) and (6) is non-singular. The presence of the random quantity y_{-1} or V^* in these expressions means that this is not always true. However, $Z'P_DZ$ is non-singular by assumption. The bracketed matrix can be written as $Z'P_DZ + \sigma M$, where M contains the random components. This is a continuous function of σ, and, therefore, there exists a neighborhood for σ around zero in which the bracketed matrix is nonsingular. Then for σ sufficiently small the necessary inverse matrix will exist.

Writing the inverse matrix in (6) as

$$Q[I+\sigma(S+\sigma V^{*\prime}P_DV^*)Q]^{-1}$$

where $Q = (Z'P_DZ)^{-1}$ and $S = Z'P_DV^* + V^{*\prime}P_DZ$, and expanding the inverse as a power series in σ gives an approximate expression for the error of IV as

<u>Lemma 1</u>.

$$\begin{aligned} e = \sigma Q\{ & \hat{Z}'u+\sigma(-SQ\hat{Z}'u+V^{*\prime}P_Du) \\ & -\sigma^2(V^{*\prime}P_DV^*Q\hat{Z}'u - SQSQ\hat{Z}'u + SQV^{*\prime}P_Du) + \sigma^3(SQV^{*\prime}P_DV^*Q\hat{Z}'u + V^{*\prime}P_DV^*QSQ\hat{Z}'u) \\ & - SQSQSQ\hat{Z}'u - V^{*\prime}P_DV^*QV^{*\prime}P_Du + SQSQV^{*\prime}P_Du)\} + o_p(\sigma^5) \end{aligned}$$

where $\hat{Z} = P_D Z = (\hat{W}\ X)$ since $X \subset D$.

Lemma 1 will be used to find the small disturbance bias and mean squared error of IV. The bias is found in the Appendix as

Theorem 1.

$$Ee_{IV} = \frac{\sigma^2}{W'P_1W}\left\{\begin{bmatrix} 1 \\ -(X'X)^{-1}X'W \end{bmatrix} \operatorname{tr} P_2 C - Q\hat{Z}'C'P_1W\right\}$$

$$+ \frac{\sigma^4}{(W'P_1W)^2} Q\ \{[\operatorname{tr} P_2 C \operatorname{tr}(P_1 - P_2)G + 2 \operatorname{tr} P_2 C(P_1 - P_2)G + 2 \operatorname{tr} P_1 C \operatorname{tr} P_2 G$$

$$+ 4 \operatorname{tr} P_1 C P_2 G] \begin{bmatrix} 1 \\ -(X'X)^{-1}X'W \end{bmatrix}$$

$$+ 2\hat{Z}'[G \operatorname{tr} P_2 C + G(P_2 - P_1)C' - \tfrac{1}{2}C' \operatorname{tr}(P_1 - P_2)G + CP_2C + C'P_2G]P_1W\} + O(\sigma^6)$$

where $P_1 = P_N \bar{P}_X P_W$ and $P_2 = P_N \bar{P}_X \bar{P}_W$. Roughly, the P_1 space contains the useful contribution of the instruments and the P_2 space the instrumental variation uncorrelated with the systematic part of y_{-1} .

Considering only terms of order σ^2 , the bias can be written as

$$(7) \qquad \frac{\sigma^2}{(W'P_1W)^2}\begin{bmatrix} \operatorname{tr}[(W'P_1W)P_2 - P_1WW'P_1]C \\ (X'X)^{-1}X'W \operatorname{tr}(P_1WW'P_1 - W'P_1WP_2)C - W'P_1W(X'\bar{P}_{\hat{W}}X)^{-1}X'C'P_1W \end{bmatrix} + O(\sigma^4)$$

where the first line gives the bias for $\hat{\alpha}$ and the second line is the bias for $\hat{\beta}$. The size of the bias depends not only on the number of instrumental variables, but on their statistical characteristics and the parameters of the model as well.

An immediate corollary of the theorem is

<u>Corollary 1</u>. The IV estimator will be unbiased to order σ^4 if $P_N C = 0$, i.e., the instruments are orthogonal to the columns of $C = \Omega A'$.

This condition can be met in principle if only one instrument is used since the rank of C is at most $T-1$ and the null space of P_N is then of dimension $T-1$. For a larger number of instruments this condition would be satisfied only by chance.

The sign of the bias generally depends on several factors. Considering only the bias in α and making the simplifying assumptions that N and X are orthogonal,[4] expression (7) becomes

$$E(\hat{\alpha}-\alpha) = \sigma^2 \frac{1}{(W'P_N W)^2} \operatorname{tr}[(W'P_N W)\bar{P}_W - P_N WW']P_N C + O(\sigma^4) . \tag{8}$$

Since C is indefinite, the bias can have either sign. If W were used as an instrument, expression (8) becomes

$$E(\hat{\alpha}-\alpha) = -\sigma^2 \frac{W'CW}{(W'W)^2} + O(\sigma^4) . \tag{9}$$

Ignoring the initial value of y, y_0, expression (9) is $-\sigma^2\beta'X'A'\Omega A'AX\beta$. Thus the bias can be either positive or negative even if $\Omega = I$, since A is indefinite.

From formula (8), the effect of adding an instrument uncorrelated with W, X and N can be found. The change in bias from adding such an instrument, n, to N is

$$\sigma^2 \frac{\operatorname{tr} P_n C}{(W'P_N W)^2} + O(\sigma^4) . \tag{10}$$

---

[4] This is true if X is serially uncorrelated and X_{-1} is used as instruments.

In a case where the sign of the bias is known, it may be possible to add a random instrument which will reduce the finite sample bias of the IV estimator. Of course, this may not improve the overall performance of the estimator.

The next corollary gives the bias of $\hat{\alpha}$ when the most obvious set of instruments, X_{-1}, is used.

Corollary 2. If X is not autocorrelated and $N = X_{-1}$, the bias of $\hat{\alpha}$ is

$$(11) \qquad E(\hat{\alpha}-\alpha) = \frac{\sigma^2}{(\beta' X'_{-1} X_{-1}\beta)} tr(P_{X_{-1}} \bar{P}_W - P_{W_{-1}})C + 0(\sigma^4)$$

assuming X_0 is known and y_0 is zero.

Proof: This corollary follows from the observation that $P_{X_{-1}} W = X_{-1}\beta$, if X is serially independent.

Using the same procedure the bias of $\hat{\alpha}$ when a subset of the lagged exogenous variables is used is found. Assume for simplicity that X is serially independent and $X_{-1} = (X_{11} \vdots X_{12})$ where X_{11} is the set of instruments for y_{-1}. Correspondingly $W = (W_1 \vdots W_2) = A(X_1 \vdots X_2)\begin{pmatrix}\beta_1 \\ \beta_2\end{pmatrix}$. Suppose X_1 and X_2 are orthogonal and $W_{-1} = (W_{11} \vdots W_{12})$. Then for these instruments

$$(12) \qquad E(\hat{\alpha}-\alpha) = \frac{1}{W'_{11}W_{11}} tr(P_{X_{11}} \bar{P}_W - P_{W_{11}})C + 0(\sigma^4) .$$

The difference between the expectations in (11) and (12) is the change in bias due to reducing the set of instruments. This difference is

$$(13) \qquad \frac{1}{(W_{-1}'W_{-1})}[\mathrm{tr}(P_{X_{12}}\bar{P}_W - P_{W_{12}})C - \frac{W_{12}'W_{12}}{W_{11}'W_{11}}\mathrm{tr}(P_{X_{11}}\bar{P}_W - P_{W_{11}})C] \; .$$

We consider now the mean squared error of the IV estimator. The mean squared error is found in the appendix as

Theorem 2

$$Ee' = \sigma^2 Q\hat{Z}'\Omega\hat{Z}Q + \sigma^4 q_{11}Q\{B_{11} - b - b' + \hat{Z}'(B_2 + B_2')\hat{Z}\}Q + O(\sigma^6)$$

where $B_{11} = \begin{pmatrix} b_{11} & 0 \\ 0 & 0 \end{pmatrix}$, $b_{11} = q_{11}^{-1}(\mathrm{tr}\, P_2C\; \mathrm{tr}\, P_2C + \mathrm{tr}\, P_2CP_2C + \mathrm{tr}\, P_2GP_2\Omega)$,

$$b = \hat{Z}'(b_1 + b_1')[P_1W \,\vdots\, 0] \; ,$$

$$b_1 = CP_2C + CP_2C' + C\,\mathrm{tr}\, P_2C + GP_2\Omega + \tfrac{1}{2}\Omega\,\mathrm{tr}\, P_2G \; ,$$

$$B_2 = C(P_1 - P_2)C + C'P_1C - C\,\mathrm{tr}\, P_2C + \Omega(P_1 - P_2)G + \tfrac{1}{2}G\,\mathrm{tr}\, P_1\Omega \; ,$$

and $q_{11} = (W'P_1W)^{-1}$.

Choosing different instruments affects the MSE through $\hat{Z}$, P_1 and P_2 . The term of order σ^2 in this formula agrees with the large sample result for fixed regressors except for the use of only Z , the nonstochastic part of the regressors (y_{-1}, X) a difference which is T-asymptotically zero. The first three elements of the σ^4 term contribute only to the variance and covariances of $\hat{\alpha}$ while the remaining terms affect the entire covariance matrix. Using the same method employed for theorems 1 and 2, the bias and MSE for OLS are found.

Theorem 3.

$$Ee_{OLS} = \frac{\sigma^2}{W'\overline{P}_X W}\left[\begin{pmatrix} 1 \\ -(X'X)^{-1}X'W \end{pmatrix} \mathrm{tr}\, \overline{P}_X \overline{P}_W C + \frac{1}{W'\overline{P}_X W}\begin{pmatrix} -W'\overline{P}_X C \overline{P}_X W \\ ((X'X)^{-1}X'WW' - W'\overline{P}_X W(X'\overline{P}_W X)^{-1}X')C'\overline{P}_X W \end{pmatrix}\right]$$

$$+ 0(\sigma^4)$$

where e_{OLS} is $\begin{pmatrix} \hat{\alpha} \\ \beta \end{pmatrix}_{OLS} - \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$.

Proof: $$e_{OLS} = \sigma[(Z+\sigma V^*)'(Z+\sigma V^*)]^{-1}(Z+\sigma V^*)'u$$

$$= \sigma Q_0[I - \sigma(V^{*\prime}Z + Z'V^*)Q_0](Z+\sigma V^*)'u + 0_p(\sigma^3)$$

$$= \sigma Q_0 Z'u + \sigma^2 Q_0[V^{*\prime}u - (V^{*\prime}Z + Z'V^*)Q_0 Z'u] + 0_p(\sigma^3) \;, \quad \text{where} \quad Q_0 = (Z'Z)^{-1}$$

$$EQ_0 Z'u = 0 \quad \text{and} \quad EV^{*\prime}u = E\begin{pmatrix} V'u \\ 0 \end{pmatrix} = \begin{pmatrix} \mathrm{tr}\ C \\ 0 \end{pmatrix}.$$

$$EV^{*\prime}ZQ_0Z'u = \begin{pmatrix} \mathrm{tr}\ P_Z C \\ 0 \end{pmatrix} \text{and} \quad EV^*Q_0Z'u = EVq_0^r Z'u = C'Zq_0^c \;,$$

where q_0^r is the first row of Q_0 and q_0^c the first column. Collecting terms gives the result:

Theorem 4: $Ee_{OLS}e'_{OLS} = \sigma^2 Q_0 Z'\Omega Z Q_0 + 0(\sigma^4)$.

Proof: $Eee' = E\sigma^2 Q_0 Z'uu'ZQ_0 = \sigma^2 Q_0 Z'\Omega Z Q_0 + 0(\sigma^4)$.

The terms of order σ^4 in these two expressions are exceedingly lengthy and are, therefore, omitted.

From Theorems 2 and 4 we find:

Corollary 3: If $\Omega = I$ the difference of the covariance matrices of OLS and IV is

$$Ee_{OLS}e'_{OLS} - Ee_{IV}e'_{IV} = -\sigma^2 kff' + O(\sigma^4)$$

where

$$f = \begin{bmatrix} 1 \\ -(X'X)^{-1}X'W \end{bmatrix}$$

and

$$k = W'\bar{P}_D W(W'P_N\bar{P}_X W)^{-1}(W'\bar{P}_X W)^{-1} = \left(\frac{W^{*\prime}W^*}{W^{*\prime}P_N W^*} - 1\right)(W^{*\prime}W^*)^{-1} ,$$

where $W^* = \bar{P}_X W$.

Proof:

$$\begin{aligned} & Ee_{OLS}e'_{OLS} - e_{IV}e'_{IV} \\ &= \sigma^2(Z'Z)^{-1}[I - (Z'Z)(Z'P_D Z)^{-1}] + O(\sigma^4) \\ &= -\sigma^2(Z'Z)^{-1}Z'\bar{P}_D Z(Z'P_D Z)^{-1} + O(\sigma^4) \\ &= W'\bar{P}_D W(W'P_N\bar{P}_X W)^{-1}(W'\bar{P}_X W)^{-1} \left[\begin{array}{c|c} 1 & -W'X(X'X)^{-1} \\ \hline -(X'X)^{-1}X'W & \_\_ \end{array}\right] \left[\begin{array}{c|c} 1 & -\hat{W}'X(X'X)^{-1} \\ \hline 0 & 0 \end{array}\right] \end{aligned}$$

using the formula for a partitioned inverse. Since $\hat{W}'X = W'P_D X = W'X$, the result follows.

Since the difference of MSE matrices is negative definite unless D contains W , e.g. W is its own instrument, OLS is always superior in the dynamic model as long as the errors satisfy classical assumptions. Thus this result holds not only asymptotically in T (see e.g. Theil [12, p. 412]) but in finite samples for σ small. Since k is the only element in this difference which is affected by the choice of instruments, more valid instruments correlated with W always lead to an improvement in the IV estimator, or at least no deterioration, for σ sufficiently small.

Comparing the bias of OLS and IV for the estimate of α we find

$$(14) \qquad \frac{Ee_{OLS}}{Ee_{IV}} = \left(\frac{\mathrm{tr}\ \bar{P}_X P_N WW'}{\mathrm{tr}\ \bar{P}_X WW'}\right)^2 \frac{\mathrm{tr}(W'\bar{P}_X W\bar{P}_W - \bar{P}_X WW')\bar{P}_X C}{\mathrm{tr}(W'P_N\bar{P}_X W\bar{P}_W - \bar{P}_X P_N WW')\bar{P}_X P_N C} + O(\sigma^4)$$

assuming that the denominator is nonzero. Thus the relative bias of the two estimators depends on the relationship of the instruments to W and to $C = EuV'$.

Finally, we illustrate some bias calculations in detail for a special case. Consider the equation

$$(15) \qquad y = \alpha y_{-1} + x + \gamma + u$$

containing one exogenous variable x with unit coefficient and a constant term. Assume $x_t = \lambda x_{t-1} + v$, and $u_t = \rho u_{t-1} + \eta_t$ with $Ev = E\eta = 0$, $E(v|x, u, v_{-1}) = E(\eta|x, v, \eta_{-1}) = 0$, $Ev^2 = \sigma_v^2$, $E\eta^2 = \sigma^2$. Using x_{-1} as the instrument for y_{-1} , the bias expressions for the OLS and IV estimates of α can be written

$$(16) \qquad E(\hat{\alpha}-\alpha)_{OLS} = \frac{\sigma^2}{W^{*\prime}W^*}\left[\mathrm{tr}\, C - \mathrm{tr}\,(X'X)^{-1}X'CX - 2\,\frac{W^{*\prime}CW^*}{W^{*\prime}W^*}\right] + O(\sigma^4)$$

and

$$(17) \qquad E(\hat{\alpha}-\alpha)_{IV} = \frac{\sigma^2}{\tilde{W}'\tilde{W}}\left[\mathrm{tr}(R'R)^{-1}R'CR - 2\,\frac{\tilde{W}'C\tilde{W}}{\tilde{W}'\tilde{W}}\right] + O(\sigma^4)$$

where $W^* = \bar{P}_X W$, the residuals from the regression of W on x ,

$\tilde{W} = P_N\bar{P}_X W$, the fitted values of the regression of W^* on the instruments,

and $R = P_X N$, the fitted values of the instrument (here x_{-1}) on x .

With autoregressive errors it is straightforward to show that

$$\operatorname{tr} C = \frac{\rho}{1-\alpha\rho}\left[T - 1 - \frac{\alpha\rho - (\alpha\rho)^T}{1-\alpha\rho}\right] ,$$

which is $O(T)$.

All of the terms in square brackets in (16) and (17) are $O(1)$ except $\operatorname{tr} C$, and $W^{*\prime}W^{*}$ and $\tilde{W}'\tilde{W}$ are $O(T)$, assuming $(X'X)^{-1}$ tends to a finite limit, $|\alpha| < 1$, and $|\rho| < 1$. Therefore, the large sample value of the small-sigma IV bias is zero and (16) tends to

$$\text{(18)} \qquad \frac{\sigma^2 \operatorname{tr} C}{\lim W^{*\prime}W^{*}} ,$$

which is nonzero unless $\rho = 0$. If further, $\lambda = 0$, (17) is T-asymptotically

$$\text{(19)} \qquad \frac{\rho(1-\alpha^2)\sigma^2}{(1-\alpha\rho)\sigma_X^2(1-\rho^2)} .$$

This differ from the T-asymptotic formula (Malinvaud [5], p. 560) which is, in our notation,

$$\text{(20)} \qquad \operatorname{plim}(\hat{\alpha}-\alpha)_{OLS} = \rho\left[\frac{1+\alpha\rho}{1-\alpha^2} + \frac{(1-\alpha\rho)(1-\rho^2)}{(1-\alpha^2)}\,\frac{\sigma_X^2}{\sigma^2}\right]^{-1} ,$$

which includes the additional term $(1+\alpha\rho)/(1-\alpha^2)$. As σ^2 becomes small in (20) the second term predominates and (20) tends to (19). The additional term in (20) is of higher order in σ^2 and appears in the next term in the small-sigma bias approximation, which was not presented. No such asymptotic disagreements occur with the IV expressions.

BIBLIOGRAPHY

[1] Anderson, T. W. Introduction to Multivariate Statistical Analysis, New York: John Wiley and Sons, 1958.

[2] Brown, George F., Joseph B. Kadane and John G. Ramage. "The Asymptotic Bias and Mean-Squared Error of Double K-Class Estimators When the Disturbances Are Small," International Economic Review, Vol. 15 (1974), pp. 667-679.

[3] Hatanaka, M. "An Efficient Two-Step Estimator for the Dynamic Adjustment Model with Autoregressive Errors," Journal of Econometrics, Vol. 2 (1974), pp. 199-220.

[4] Kadane, Joseph B. "Comparison of k-Class Estimators When the Disturbances Are Small," Econometrica, Vol. 34 (1971), pp. 723-737.

[5] Malinvaud, E. Statistical Methods of Econometrics, 2nd revised ed. New York: American Elsevier Publishing Co., Inc., 1970.

[6] Peck, Jon K. "A Comparison of Alternative Estimators for a Dynamic Relationship Estimated from a Time Series of Cross-Sections When the Disturbances are Small," Cowles Foundation Discussion Paper 325, 1972.

[7] __________. "The Estimation of a Dynamic Equation Following a Preliminary Test for Autocorrelation," Cowles Foundation Discussion Paper 404R, 1976.

[8] Ramage, John G. "A Perturbation Study of the k-Class Estimators in the Presence of Specification Error," Unpublished Ph.D. Dissertation, Yale University, 1971.

[9] Rao, C. R. Linear Statistical Inference and Its Applications, 2nd ed. New York: John Wiley and Sons, 1973.

[10] Sargan, J. D. "The Estimation of Economic Relationships Using Instrumental Variables," Econometrica, Vol. 26 (1958), pp. 393-415.

[11] __________. "The Validity of Nagar's Expansion for the Moments of Econometric Estimators," Econometrica, Vol. 42 (1974), pp. 169-76.

[12] Theil, Henri. Principles of Econometrics. New York: John Wiley, 1971.

APPENDIX

The derivation of the bias and mean squared error of the estimators requires the following lemmas.

We record some relationships among projections as

Lemma A1.

a) $M = q_{11}(P_{\hat{Z}} - P_X) = q_{11}P_1$, where $M = \hat{Z}Q^c Q^{c'}\hat{Z}'$ and Q^c is the first column of Q ,

b) $P_{\hat{Z}} = P_X + P_1$,

c) $P_D = P_X + \overline{P}_X P_N$,

d) $P_D - P_{\hat{Z}} = P_2$,

e) $\overline{P}_X P_N = P_1 + P_2$.

Proof:

a) The matrix M is equal to $q_{11}(P_{\hat{Z}} - P_X)$. This follows from expressing $Q^c Q^{c'}$ as

$$(\text{A-14}) \qquad q_{11}\left\{\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & q_{22} - \dfrac{q_{21}q_{12}}{q_{11}} \end{bmatrix}\right\}$$

Then, using the formula for the partitioned inverse, $q_{22} - \dfrac{q_{21}q_{12}}{q_{11}} = (X'X)^{-1}$.

Hence $M = q_{11}[\hat{Z}Q\hat{Z}' - X(X'X)^{-1}X'] = q_{11}(P_{\hat{Z}} - P_X)$. Since $\hat{Z} = (\hat{W}\ X)$, $P_{\hat{Z}} - P_X$ is the projection operator $P_{\widetilde{W}}$ where $\widetilde{W}$ is obtained by regressing the residuals of the regression of W on X in turn on N , i.e. $\widetilde{W} = P_N\overline{P}_X W$. Therefore $M = q_{11}P_N\overline{P}_X P_W = q_{11}P_1$.

b) $$P_{\hat{Z}} = \hat{Z}(\hat{Z}'\hat{Z})^{-1}\hat{Z}' = P_D Z[Z'Z - Z'\overline{P}_D Z]^{-1} Z' P_D \ .$$

Using an inversion formula from Rao ([9], p. 33), this expression is

$$P_D Z[(Z'Z)^{-1} - (Z'Z)^{-1} Z'\overline{P}_D(\overline{P}_D P_Z \overline{P}_D - I)^{-1} P_D Z(Z'Z)^{-1}] Z' P_D$$

$$= P_D P_Z P_D - P_D P_Z \overline{P}_D(\overline{P}_D P_Z - I)^{-1} P_D P_Z P_D = P_D P_Z \ ,$$

interpreting the inverse matrix as a generalized inverse. Using (c),

$$P_D P_Z = (P_X + \overline{P}_X P_N)(P_X + \overline{P}_X P_W)$$

$$= P_X + \overline{P}_X P_N P_W = P_X + P_1 \ .$$

c) $$P_D = P_X + P_N - P_X P_N = P_X + \overline{P}_X P_N \ .$$

d) $$P_D - P_{\hat{Z}} = P_X + \overline{P}_X P_N - (P_X + P_1) = \overline{P}_X P_N - \overline{P}_X P_N P_W = P_2 \ .$$

e) $$\overline{P}_X P_N = \overline{P}_X P_N (P_W + \overline{P}_W) = P_1 + P_2 \ .$$

Q.E.D.

The group of lemmas below are derived from the following result (see Anderson [1], p. 39). Let X_i , $i = 1, \ldots, 4$ be random variables with a joint normal distribution $N(0,\Sigma)$. Then $EX_i X_j X_k X_\ell = \sigma_{ij}\sigma_{k\ell} + \sigma_{ik}\sigma_{j\ell} + \sigma_{i\ell}\sigma_{jk}$ where $\Sigma = (\sigma_{ij})$. Recall that $V^* = (\underset{T\times 1}{V} \quad \underset{T\times K}{0})$, and let D , F , and L be arbitrary conformable constant matrices, the elements of D are d_{ij} . Let d^c be the first column and d^r the first row of D with similar definitions for F and L .

Lemma A2: $$EV^{*\prime}DV^*V^{*\prime}Fu = \begin{bmatrix} \mathrm{tr}\ GD\ \mathrm{tr}\ CF + \mathrm{tr}\ GFCD' + \mathrm{tr}\ GFCD \\ 0 \end{bmatrix}$$

Proof: $$V^{*\prime}DV^*V^{*\prime}Fu = \begin{bmatrix} V'DVV'Fu \\ 0 \end{bmatrix} .$$

The expectation of $V'DVV'Fu$ is

$$E \sum_{ijk\ell} v_i d_{ij} v_j v_k f_{k\ell} u_\ell$$

$$= \sum_{ijk\ell} E(v_i v_j) E(u_\ell v_k) d_{ij} f_{k\ell} + \sum_{ijk\ell} E(v_i v_k) E(u_\ell v_j) d_{ij} f_{k\ell}$$

$$+ \sum_{ijk\ell} E(u_\ell v_i) E(v_j v_k) d_{ij} f_{k\ell}$$

$$= \sum_{ijk\ell} g_{ij} c_{\ell k} d_{ij} f_{k\ell} + \sum_{ijk\ell} g_{ik} f_{k\ell} c_{\ell k} d_{ij} + \sum_{ijk\ell} c_{\ell i} d_{ij} g_{jk} f_{k\ell}$$

which is the element-by-element expression for the stated expectation.

The proofs of the remaining lemmas are similar to the proof of Lemma A2 and are omitted but can be found in Peck [6].

Lemma A3: $EV^*V^{*\prime}DV^*Fu = (GD + GD' + I\ \mathrm{tr}\ GD)C'f^{r'}$.

Lemma A4: $EV^*V^{*\prime}DV^{*\prime}Fu = (G\ \mathrm{tr}\ CF + CFC + C'F'G)d^c$.

Lemma A5: $EV^{*\prime}Duu'FV^* = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ where $a = \mathrm{tr}\ CD\ \mathrm{tr}\ C'F + \mathrm{tr}\ CDCF' + \mathrm{tr}\ GD\Omega F$.

Lemma A6: $EV^*Duu'FV^* = (C'\ \mathrm{tr}\ CF' + C'FC' + GF'\Omega)[d^{r'}\ 0]$.

Lemma A7: $EV^*Duu'FV^* = C'f^c d^r C + C'd^{r'} f^{c'} C + G\ \mathrm{tr}\ f^c d^r \Omega$.

<u>Lemma A8</u>: $Euu'DV^{*}{}'FV^{*}{}' = Cf^{c}d^{c'}C + C\,\mathrm{tr}\,f^{c}d^{c'}C + \Omega d^{c}f^{c'}G$.

<u>Lemma A9</u>: $Euu'DV^{*}{}'FV^{*} = (\Omega\,\mathrm{tr}\,GF' + CFC' + CF'C')[d^{c}0]$.

<u>Lemma A10</u>: $Euu'DV^{*}V^{*}{}' = \Omega DG + CD'C + C\,\mathrm{tr}\,CD'$.

<u>Lemma A11</u>: $Euu'DV^{*}FV^{*} = (\Omega DG + CD'C + C\,\mathrm{tr}\,CD')[f^{r'}0]$.

With the aid of these lemmas Theorem 1 is now proved.

<u>Proof of Theorem 1</u>

All odd power terms in the sigma power series expansion of e are products of odd numbers of zero-mean normally distributed random variables and, therefore, have expectation zero. The terms of order σ^2 are

(A-1) $$-QSQ\hat{Z}'u + QV^{*}{}'P_{D}u \ .$$

The second of these is $\begin{pmatrix} V'P_{D}u \\ 0 \end{pmatrix}$ which has expectation $\begin{pmatrix} \mathrm{tr}\ P_{D}C \\ 0 \end{pmatrix}$.

The first term can be written

(A-2) $$-Q\left\{\hat{Z}'Vu'\hat{Z}Q^{c} + \begin{pmatrix} \mathrm{tr}\ V'P_{\hat{Z}}u \\ 0 \end{pmatrix}\right\}$$

which gives for the expectation of (A-1)

(A-3) $$Q[I\,\mathrm{tr}\,(P_{D} - P_{\hat{Z}})\,C - \hat{Z}'C'\hat{Z}]Q^{c} \ .$$

Omitting the common leading factor of $\sigma^4 Q$, the terms of order σ^4 are

$$\text{(A-4)} \qquad -V^{*\prime}P_D V^* Q V^{*\prime} P_D u + SQSQV^{*\prime}P_D u - SQSQSQ + SQV^{*\prime}P_D V^* Q\hat{Z}'u$$
$$+ V^{*\prime}P_D V^* QSQ\hat{Z}'u$$
$$= -B_1 + B_2 - B_3 + B_4 + B_5 \ .$$

Applying Lemma A2,

$$\text{(A-5)} \qquad EB_1 = q_{11}\begin{pmatrix} \operatorname{tr} P_D G \operatorname{tr} P_D C + 2\operatorname{tr} P_D G P_D C \\ 0 \end{pmatrix} .$$

Expanding B_2 gives

$$B_2 = q_{11}\hat{Z}'V^* Q\hat{Z}'V^* V^{*\prime}P_D u + q_{11}\hat{Z}'V^* V^{*\prime}\hat{Z}QV^{*\prime}P_D u$$
$$\text{(A-6)} \qquad + q_{11}V^{*\prime}P_{\hat{Z}}V^* V^{*\prime}P_D u + V^{*\prime}\hat{Z}QV^{*\prime}\hat{Z}QV^{*\prime}P_D u$$
$$\text{(A-7)} \qquad = 2q_{11}\hat{Z}'V^* V^{*\prime}\hat{Z}QV^{*\prime}P_D u + V^{*\prime}(M + q_{11}P_{\hat{Z}})V^* V^{*\prime}P_D u \ .$$

Applying Lemma A2 and A4 gives

$$\text{(A-8)} \qquad EB_2 = 2q_{11}\hat{Z}'[G \operatorname{tr} P_D C + CP_D C + C'P_D G]\hat{Z}Q^c$$
$$+ \begin{bmatrix} \operatorname{tr}(M + q_{11}P_{\hat{Z}})G \operatorname{tr} P_D C + 2\operatorname{tr}(M + q_{11}P_{\hat{Z}})GP_D C \\ 0 \end{bmatrix}$$

Multiplying out the terms in B_3 gives

$$\text{(A-9)} \qquad \hat{Z}'V^*Q\hat{Z}'V^*Q\hat{Z}'V^*Q\hat{Z}u + q_{11}\hat{Z}'V^*Q\hat{Z}'V^*V^{*\prime}P_{\hat{Z}}u$$
$$+ q_{11}\hat{Z}'V^*V^{*\prime}P_{\hat{Z}}V^*Q\hat{Z}'u + q_{11}\hat{Z}'V^*V^{*\prime}\hat{Z}QV^{*\prime}P_{\hat{Z}}u$$
$$+ V^{*\prime}P_{\hat{Z}}V^*Q\hat{Z}'V^*Q\hat{Z}'u + q_{11}V^{*\prime}P_{\hat{Z}}V^*V^{*\prime}P_{\hat{Z}}u$$
$$+ V^{*\prime}\hat{Z}QV^{*\prime}P_{\hat{Z}}V^*Q\hat{Z}'u + V^{*\prime}\hat{Z}QV^{*\prime}\hat{Z}QV^{*\prime}P_{\hat{Z}}u$$
$$= \hat{Z}'V^*V^{*\prime}(M + q_{11}P_{\hat{Z}})V^*Q\hat{Z}'u + 2q_{11}\hat{Z}'V^*V^{*\prime}\hat{Z}QV^{*\prime}P_{\hat{Z}}u$$
$$+ 2V^{*\prime}P_{\hat{Z}}V^*V^{*\prime}Mu + V^{*\prime}(M + q_{11}P_{\hat{Z}})V^*V^{*\prime}P_{\hat{Z}}u \ .$$

The expectation of B_3 is found by applying Lemmas A2, A3, and A4 to be

$$\text{(A-10)} \qquad EB_3 = \hat{Z}'[2G(M + q_{11}P_{\hat{Z}}) + I \ \mathrm{tr}(M + q_{11}P_{\hat{Z}})G]C'\hat{Z}Q^c$$
$$+ 2q_{11}\hat{Z}'[G \ \mathrm{tr}\ P_{\hat{Z}}C + CP_{\hat{Z}}C + C'P_{\hat{Z}}G]\hat{Z}Q^c$$
$$+ \begin{bmatrix} 2\,\mathrm{tr}\ P_{\hat{Z}}G\ \mathrm{tr}\ MC + 4\,\mathrm{tr}\ P_{\hat{Z}}GMC + \mathrm{tr}(M + q_{11}P_{\hat{Z}})G\ \mathrm{tr}\ P_{\hat{Z}}C \\ \quad + 2\,\mathrm{tr}\ (M + q_{11}P_{\hat{Z}})GP_{\hat{Z}}C \\ 0 \end{bmatrix}$$

Expanding B_4 gives

$$B_4 = q_{11}\hat{Z}'V^*V^{*\prime}P_D V^*Q\hat{Z}'u + V^{*\prime}\hat{Z}QV^{*\prime}P_D V^*Q\hat{Z}'u \ ,$$

and EB_4 is found from Lemmas A2 and A3 as

$$\text{(A-11)} \qquad EB_4 = q_{11}\hat{Z}'[2GP_D C' + C'\ \mathrm{tr}\ P_D G]\hat{Z}Q^c + \begin{bmatrix} \mathrm{tr}\ P_D G\ \mathrm{tr}\ MC + 2\,\mathrm{tr}\ P_D GMC \\ 0 \end{bmatrix} .$$

Finally, expanding B_5 gives

$$B_5 = V^{*\prime}P_D V^*V^{*\prime}(M + q_{11}P_{\hat{Z}})u$$

which has expectation

$$(\text{A-12})\qquad EB_5 = \begin{bmatrix} \operatorname{tr} P_D G \operatorname{tr}(M + q_{11}P_{\hat{Z}})C + 2 \operatorname{tr} P_D G(M + q_{11}P_{\hat{Z}})C \\ 0 \end{bmatrix}$$

from an application of Lemma A2. Then collecting the σ^4 scalar terms in (A-3), (A-5), (A-8), (A-10), (A-11), and (A-12) and simplifying using Lemma A1 gives

$$\begin{aligned}
(\text{A-13})\qquad & q_{11}[\operatorname{tr}(P_X + \bar{P}_X P_N)C \operatorname{tr}(P_1 - P_2)G + \operatorname{tr} P_1 C \operatorname{tr} P_2 G - \operatorname{tr} P_1 C \operatorname{tr}(P_X + P_1)G \\
& - \operatorname{tr} P_1 G \operatorname{tr}(P_X + P_1)C - \operatorname{tr}(P_X + P_1)G \operatorname{tr}(P_X + P_1)C + \operatorname{tr} P_1 C \operatorname{tr}(P_X + \bar{P}_X P_N)G \\
& + \operatorname{tr}(P_X + \bar{P}_X P_N)G \operatorname{tr}(P_X + P_1)C \\
& + 2\operatorname{tr}(P_X + \bar{P}_X P_N)C(P_1 - P_2)G + 2 \operatorname{tr} P_1 C P_2 G - 2 \operatorname{tr} P_1 C(P_X + P_1)G \\
& - 2\operatorname{tr}(P_X + P_1)C P_1 G - 2\operatorname{tr}(P_X + P_1)C(P_X + P_1)G + 2\operatorname{tr}(P_X + \bar{P}_X P_N)G P_1 C \\
& + 2\operatorname{tr}(P_X + P_1)C(P_X + \bar{P}_X P_N)G]
\end{aligned}$$

$$\begin{aligned}
(\text{A-14})\qquad & = q_{11}[\operatorname{tr}(P_1 + P_2)C \operatorname{tr}(P_1 - P_2)G - \operatorname{tr} P_1 C \operatorname{tr}(P_1 - P_2)G - \operatorname{tr} P_1 C \operatorname{tr} P_1 G \\
& + 2 \operatorname{tr} P_1 C \operatorname{tr}(P_1 + P_2)G \\
& + 2[\operatorname{tr} P_1 C(P_1 - P_2)G + \operatorname{tr} P_2 C(P_1 - P_2)G + \operatorname{tr} P_1 C P_2 G - \operatorname{tr} P_1 C P_1 G \\
& + \operatorname{tr} P_1 C P_2 G + \operatorname{tr} P_1 C P_2 G]
\end{aligned}$$

$$(\text{A-15})\qquad = q_{11}[\operatorname{tr} P_2 C \operatorname{tr}(P_1 - P_2)G + 2 \operatorname{tr} P_1 C \operatorname{tr} P_2 G + 2[2 \operatorname{tr} P_1 C P_2 G + \operatorname{tr} P_2 C(P_1 - P_2)G]] ,$$

using Lemma A1 repeatedly.

A similar but simpler process gives the nonscalar terms. Explicit evaluation of $Q = (\hat{Z}'\hat{Z})^{-1}$ shows that

$$\text{(A-16)} \qquad Q = (\hat{W}'\bar{P}_X\hat{W})^{-1}\begin{bmatrix} I & -\hat{W}'X(X'X)^{-1} \\ -(X'X)^{-1}X'\hat{W} & \hat{W}'\bar{P}_X\hat{W}(X'\bar{P}_{\hat{W}}X)^{-1} \end{bmatrix} .$$

Applying (A-16) to (A-15) gives Theorem 1.

The proof of Theorem 2 procedes by computing ee' from Lemma 1; then taking expectations term by term in that expression. The term of orders which are odd powers of σ all consist of products of odd powers of normally distributed zero-mean random variables which contribute nothing to the expectation of ee' and they are, therefore, omitted throughout. Except for the σ^3 terms,

$$\text{(A-17)} \qquad ee' = \sigma^2 Q\hat{Z}'uu'\hat{Z}Q + \sigma^4 Q\{H_1 + H_2 + H_2'\}Q + O_p(\sigma^5)$$

where $H_1 = (-SQ\hat{Z}'u + V^{*\prime}P_D u)(-u'\hat{Z}QS + u'P_D V^*)$, and
$H_2 = -\hat{Z}'u(u'\hat{Z}QV^{*\prime}P_D V^* - u'\hat{Z}QSQS + u'P_D V^* QS)$.

$$\text{(A-18)} \qquad E\sigma^2 Q\hat{Z}'uu'\hat{Z}Q = \sigma^2 Q\hat{Z}'\Omega\hat{Z}Q .$$

Multiplying out H_1 gives

$$\text{(A-19)} \qquad SQ\hat{Z}'uu'\hat{Z}QS - SQ\hat{Z}'uu'P_D V^* - V^{*\prime}P_D uu'\hat{Z}QS + V^{*\prime}P_D uu'P_D V^* .$$

Let (A-19) be $H_{11} - H_{12} - H_{12}' + H_{13}$. Then multiplying out H_{11} gives

$$\text{(A-20)} \qquad \hat{Z}'V^*Q\hat{Z}'uu'P_{\hat{Z}}V^* + \hat{Z}'V^*Q\hat{Z}'uu'\hat{Z}QV^{*\prime}\hat{Z}$$
$$+ V^{*\prime}P_{\hat{Z}}uu'P_{\hat{Z}}V^* + V^{*\prime}P_{\hat{Z}}uu'\hat{Z}QV^{*\prime}\hat{Z} .$$

The last term is the transpose of the first. Applying Lemmas A5, A6, and A7, the expectation is computed as

$$(A\text{-}21) \qquad EH_{11} = \hat{Z}'[C' \,\mathrm{tr}\, P_{\hat{Z}}C + C'P_{\hat{Z}}C' + GP_{\hat{Z}}\Omega][\hat{Z}Q^{c} \;\; 0]$$

$$+ \begin{pmatrix} Q^{c'}\hat{Z}' \\ 0 \end{pmatrix} [C \,\mathrm{tr}\, P_{\hat{Z}}C + CP_{\hat{Z}}C + \Omega P_{\hat{Z}}G]\hat{Z} + \hat{Z}'[2C'MC + G \,\mathrm{tr}\, M\Omega]\hat{Z} + \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$$

where $a = (\mathrm{tr}\; P_{\hat{Z}}C)^2 + \mathrm{tr}\; P_{\hat{Z}}CP_{\hat{Z}}C + \mathrm{tr}\; P_{\hat{Z}}GP_{\hat{Z}}\Omega$. H_{12} is $\hat{Z}'V^*Q\hat{Z}'uu'P_DV^* + V^{*\prime}P_{\hat{Z}}uu'P_DV^*$ which has expectation

$$(A\text{-}22) \qquad EH_{12} = \hat{Z}'[C' \,\mathrm{tr}\, P_DC + C'P_DC' + GP_D\Omega][\hat{Z}Q^{c} \;\; 0]$$

$$+ \begin{bmatrix} \mathrm{tr}\; P_{\hat{Z}}C \,\mathrm{tr}\, P_DC + \mathrm{tr}\, P_DCP_{\hat{Z}}C + \mathrm{tr}\, P_DGP_{\hat{Z}}\Omega & 0 \\ 0 & 0 \end{bmatrix}$$

from Lemmas A5 and A6. Then

$$(A\text{-}23) \qquad EH_{13} = \begin{bmatrix} \mathrm{tr}\; P_DC \,\mathrm{tr}\, P_DC + \mathrm{tr}\, P_DCP_DC + \mathrm{tr}\, P_DGP_D\Omega & 0 \\ 0 & 0 \end{bmatrix}$$

using Lemma A5. Turning next to the second term H_2 ,

$$(A\text{-}24) \qquad H_2 = -\hat{Z}'uu'\hat{Z}QV^{*\prime}P_DV^* + \hat{Z}'uu'\hat{Z}QSQS - \hat{Z}'uu'P_DV^*QS$$

$$= -H_{21} + H_{22} - H_{23} \;.$$

Using Lemma A9 we find

$$(A\text{-}25) \qquad EH_{21} = \hat{Z}'[\Omega \,\mathrm{tr}\, P_DG + 2CP_DC'][\hat{Z}Q^{c} \;\; 0]$$

$$(A\text{-}26) \qquad H_{22} = \hat{Z}'uu'P_{\hat{Z}}V^*Q\hat{Z}'V^* + q_{11}\hat{Z}'uu'P_{\hat{Z}}V^*V^{*\prime}\hat{Z}$$

$$+ \hat{Z}'uu'\hat{Z}QV^{*\prime}P_{\hat{Z}}V^* + \hat{Z}'uu'\hat{Z}QV^{*\prime}\hat{Z}QV^{*\prime}\hat{Z} \;.$$

The expectation is found by applying Lemmas A8, A9, A10, and A11 to be

$$\text{(A-27)} \qquad EH_{22} = \hat{Z}'[\Omega P_{\hat{Z}}G + CP_{\hat{Z}}C + C\,\mathrm{tr}\,P_{\hat{Z}}C][\hat{Z}Q^c \;\; 0]$$

$$+ q_{11}\hat{Z}'[\Omega P_{\hat{Z}}G + CP_{\hat{Z}}C + C\,\mathrm{tr}\,P_{\hat{Z}}C]\hat{Z}$$

$$+ \hat{Z}'[\Omega\,\mathrm{tr}\,P_{\hat{Z}}G + 2CP_{\hat{Z}}C'][\hat{Z}Q^c \;\; 0]$$

$$+ \hat{Z}'[CMC + C\,\mathrm{tr}\,MC + \Omega MG]\hat{Z} .$$

Finally, $H_{23} = \hat{Z}'uu'P_D V^* Q\hat{Z}'V^* + q_{11}\hat{Z}'uu'P_D V^* V^{*\prime}\hat{Z}$ which has expectation

$$\text{(A-28)} \qquad EH_{23} = \hat{Z}'[\Omega P_D G + CP_D C + C\,\mathrm{tr}\,P_D C][\hat{Z}Q^c \;\; 0]$$

$$+ q_{11}\hat{Z}'[\Omega P_D G + CP_D C + C\,\mathrm{tr}\,P_D C]\hat{Z} .$$

from Lemmas A10 and A11.

Collecting the scalar terms of the expectations above simplifies to

$$\text{(A-29)} \qquad -\mathrm{tr}\,P_{\hat{Z}}C\,\mathrm{tr}\,P_2C + \mathrm{tr}\,P_DC\,\mathrm{tr}\,P_2C - \mathrm{tr}\,P_{\hat{Z}}CP_2C + \mathrm{tr}\,P_DCP_2C$$

$$- \mathrm{tr}\,P_{\hat{Z}}\Omega P_2G + \mathrm{tr}\,P_DGP_2\Omega$$

$$\text{(A-30)} \qquad = \mathrm{tr}\,P_2C\,\mathrm{tr}\,P_2C + \mathrm{tr}\,P_2CP_2C + \mathrm{tr}\,P_2GP_2\Omega .$$

The column vector terms (and the transpose of the row vector terms) are

$$\text{(A-31)} \qquad \hat{Z}'\{C'\,\mathrm{tr}\,P_{\hat{Z}}C + C'P_{\hat{Z}}C' + GP_{\hat{Z}}\Omega - C'\,\mathrm{tr}\,P_DC - C'P_DC' - GP_D\Omega$$

$$- \Omega\,\mathrm{tr}\,P_DG - 2CP_DC' + \Omega P_{\hat{Z}}G + CP_{\hat{Z}}C + C\,\mathrm{tr}\,P_{\hat{Z}}C$$

$$+ \Omega\,\mathrm{tr}\,P_{\hat{Z}}G + 2CP_{\hat{Z}}C' - \Omega P_DG - CP_DC - C\,\mathrm{tr}\,P_DC\}\hat{Z}Q^c$$

$$\text{(A-32)} \qquad = -\hat{Z}'\{[C\,\mathrm{tr}\,P_2C + CP_2C + GP_2\Omega + CP_2C' + \tfrac{1}{2}\Omega\,\mathrm{tr}\,P_2G]$$

$$+ [C'\,\mathrm{tr}\,P_2C + C'P_2C' + \Omega P_2G + CP_2C' + \tfrac{1}{2}\Omega\,\mathrm{tr}\,P_2G]\}\hat{Z}Q^c .$$

Finally the full matrix terms are

(A-33) $$\hat{Z}'\{2C'MC + G\,\mathrm{tr}\,M\Omega + q_{11}(\Omega P_{\hat{Z}}G + GP_{\hat{Z}}\Omega + CP_{\hat{Z}}C + C'P_{\hat{Z}}C' + (C + C')\mathrm{tr}\,P_{\hat{Z}}C)$$

$$+ CMC + C'MC' + (C + C')\mathrm{tr}\,MC + \Omega MG + GM\Omega$$

$$- q_{11}(\Omega P_D G + GP_D\Omega + CP_D C + C'P_D C' + (C + C')\mathrm{tr}\,P_D C\}\hat{Z}$$

(A-34) $$= q_{11}\hat{Z}'\{G\,\mathrm{tr}\,P_1\Omega - \Omega P_2 G - GP_2\Omega - CP_2C - C'P_2C' - (C + C')\mathrm{tr}\,P_2C$$

$$+ \Omega P_1 G + GP_1\Omega + CP_1C + C'P_1C' + 2C'P_1C\}\hat{Z} \;.$$

Theorem 2 now follows by collecting these expectations.