9-1971

Comparison of k-Class Estimators When the Disturbances Are Small

Joseph B. Kadane
Carnegie Mellon University, kadane@stat.cmu.edu

Comparison of k-Class Estimators When the Disturbances Are Small
Author(s): Joseph B. Kadane
Source: *Econometrica*, Vol. 39, No. 5 (Sep., 1971), pp. 723-737
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/1909575

Econometrica, Vol. 39, No. 5 (September, 1971)

COMPARISON OF k-CLASS ESTIMATORS WHEN THE DISTURBANCES ARE SMALL

BY JOSEPH B. KADANE[1]

A new approach to the choice of econometric estimators, called small-sigma asymptotics, is introduced and applied to the choice of k-class estimators of the parameters of a single equation in a system of linear simultaneous stochastic equations. I find that when the degree of overidentification is no more than six, the two stage least squares estimator uniformly dominates the limited information maximum likelihood estimator in a certain sense. The small sigma method can be used on many problems in statistics and econometrics.

1. INTRODUCTION

THE STUDY OF simultaneous equation econometric models has led to many alternative estimators to ordinary least squares: single equation limited information maximum likelihood, and two stage least squares, for example. The behavior of these estimators has been difficult to describe, however, and it has been difficult to choose among these estimators. The work described in this paper explores this problem for the case in which lagged dependent variables are not permitted.

To be most useful for normative purposes, a description must be detailed enough to give a good approximation and expose differences between estimators, and yet be simple enough to strengthen intuition and yield easily described comparisons. Since detail and simplicity are in conflict, approaches may differ in this respect.

This paper introduces a new approach, based on asymptotic series in a scalar multiple, σ, of the variance of the disturbance in the model. As $\sigma \to 0$ the regression function is an increasingly good description of the random variables generated. Intuitively this is suggested by Gauss' "Theory of Errors"—the errors were never intended to be so large as to swamp the regression function.

One important approach used in the past is large sample asymptotic theory. This reveals a persistent bias in ordinary least squares, and a large sample asymptotic equivalence between two stage least squares and single equation limited information maximum likelihood. Additionally, Nagar [**13**] found the $1/T$ term in the large sample asymptotic bias and the $1/T$ and $1/T^2$ terms of the moment matrix of two stage least squares. Economists have been uneasy, however, about application of large sample theory to samples which may not be "large" in the relevant sense. Additionally large sample asymptotic results often depend on an assumption about the asymptotic behavior of the moment matrix of exogenous variables which is difficult to justify.

[1] This paper is based on the author's Ph.D. dissertation written under the direction of Herman Chernoff at Stanford University. Other helpful comments were contributed by Kenneth Arrow, David Grether, G. S. Maddala, Marc Nerlove, and John G. Ramage. This research was supported at Stanford by a grant from the National Science Foundation, and at the Cowles Foundation by grants from the National Science Foundation and the Ford Foundation. An earlier version was given at the Econometric Society meetings of December, 1966.

Monte Carlo experiments [**7**, **8**, **14**, **15**, **22**], a third approach to the problem, have only rarely been used to explore the domain of validity of large sample approximations to the mean and variance of an estimator. They have provided some insights into the behavior of estimators under a variety of circumstances, but hypotheses generated by Monte Carlo experiments are difficult to place in a general theory unless they have analytic confirmation of some kind. The parameter space is so large that Monte Carlo results often fail to provide a reasonably comprehensive picture. Corollary 1 below gives an example of a result that is true for small models, like those used in most Monte Carlo work, but which appears to be reversed in large models, which are often encountered in practice.

A fourth approach, introduced by Basmann [**2**], finds fixed sample exact densities and moments. More recent work on this line has been done by Basmann [**3**, **4**], Bergstrom [**5**], Kabe [**9**, **10**], Richardson [**16**, **17**], Takeuchi [**23**], Sawa [**19**, **20**, **21**] and Mariano [**12**]. All of these papers have been limited to the case of two endogenous variables in the equation being estimated. An important result of this work has been the finding (in a special case by Basmann [**2**], in a more general form by Takeuchi [**23**] and Sawa [**19**]) that for two stage least squares, moments of order less than K exist, and those larger do not, where K is the number of exogenous variables in the system. The method is difficult, however, as it involves integrating a noncentral Wishart distribution, and the results for the exact moments and densities have been so complicated as not to be very illuminating.

The major results of this paper are the computation of the bias (to order σ^2) and matrices of second moments about the true values (to order σ^4) for all k-class estimators (for fixed k) and for single equation limited information maximum likelihood. In doing these computations, all predetermined variables are assumed to be exogenous, and the disturbances are assumed normal and uncorrelated over time. The results are given in Section 2, and proved in the Appendices. They provide the basis for a number of interesting corollaries.

One corollary is that for equations in which the degree of overidentification is less than or equal to six, two stage least squares *uniformly* dominates the limited information maximum likelihood estimators (in the sense that the difference between the moment matrices of these estimators is asymptotically (as $\sigma \to 0$) positive semidefinite regardless of the values taken by parameters or exogenous variables). This is unexpected on the basis of the considerations introduced by Chow [**6**]. Interpreted in his context, this result means that allowing the data to choose the direction of minimization (limited information maximum likelihood) introduces too much variability into the estimator, compared to the benefit gained by fixing an arbitrary direction (two stage least squares). There are some indications that this preference for two stage least squares is reversed as the degree of overidentification gets large.

A second corollary shows that for sufficiently small sample sizes and degree of overidentification, ordinary least squares dominate two stage least squares in the same sense. This has been suspected by econometricians for some time, I believe. (In assessing the sampling properties of ordinary least squares, the simultaneous equation model leads to different means and variances than does the standard single equation model for regression.) Finally, and a surprise, a third corollary shows that the k-class estimator with smallest asymptotic variance occurs when k is negative (of course, this takes no account of the bias, which can be considerable). These are examples of the usefulness of reasonably simple approximations.

The results obtained bear an interesting relationship to each of the three other approaches. Since the sample size, T, is a parameter in small σ asymptotics, a natural way to compare large sample and small σ results is to allow $T \to \infty$ in the latter. Remarkably, in each case in which large sample results are available, the limit of the small σ expression (as $T \to \infty$) is the large sample asymptotic expression. Thus the results of Nagar [**13**] are obtained for the special case $k = 1 + \alpha/T$ (for constant α) in the computation of the bias and moment matrix (Theorems 2 and 3). Also the results of Anderson and Rubin [**1**] are obtained on the distribution of the root, λ, of the determinental equation appearing in the theory of limited information maximum likelihood (see also [**11**]). Thus small σ asymptotics can be thought of as a reasonably good approximation to the behavior of k-class estimators whenever some combination of large sample (i.e. large T) and low phenomenon variability (small σ) ought to lead to reasonably good estimation. Small σ asymptotics have the important advantage over large sample theory of being able to "correct" for sample size. Therefore, whenever an econometrician is prepared to trust large sample theory, he should be willing to trust small σ theory more.

In conclusion, small σ asymptotics have the following advantages: (i) they are as simple as, and a generalization of, large sample theory; (ii) they can provide definite answers to normative choice of estimator questions.

Whether ultimately small σ asymptotics proves to be the best compromise between simplicity and detail remains to be seen. It is interesting, however, to note that this approach can be applied to many other econometric and statistical questions [**11**]. A similar idea was recently used by Samuelson [**18**] in a discussion of the foundations of portfolio theory.

2. STATEMENT OF RESULTS

Let the complete system

$$(1) \qquad YB + Z\Gamma + \sigma U = 0$$

have a first equation

$$(2) \qquad y = Y_1\beta + Z_1\gamma + \sigma u$$

where Y is a $T \times G$ matrix of endogenous variables, partitioned $Y = (y, Y_1, Y_2)$; y is $T \times 1$, Y_1 is $T \times G_1$, and Y_2 is $T \times G_2$ $(G = G_1 + G_2 + 1)$; Z is a $T \times K$ matrix of exogenous variables, partitioned $Z = (Z_1, Z_2)$; Z_1 is $T \times K_1$, Z_2 is $T \times K_2$ $(K = K_1 + K_2)$, and Z is assumed to have rank K; B is a nonsingular $G \times G$ matrix of parameters with first column $(-1, \beta', 0')'$ where -1 is a scalar, β is $G_1 \times 1$, and 0 is a $G_2 \times 1$ vector of zeros; Γ is a $K \times G$ matrix of parameters with first column $(\gamma', 0')'$, where γ is $K_1 \times 1$ and 0 is a $K_2 \times 1$ vector of zeros;

U is a $T \times G$ matrix of jointly normal residuals with zero means and covariances $Eu_{ti}u_{t'j} = \sigma_{ij}\delta_{tt'}$ and with first column u; $\sigma_{11} = 1$ and σ is a (small) positive number. The general k-class estimator of

$$\begin{pmatrix} \beta \\ \gamma \end{pmatrix}$$

is

$$\begin{pmatrix} \beta \\ \gamma \end{pmatrix}_k = \begin{bmatrix} Y_1'Y_1 - kV^{*\prime}V^* & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{bmatrix} (Y_1 - kV^*)' \\ Z_1' \end{bmatrix} y \tag{3}$$

where $V^* = \bar{P}_Z Y_1$ and where, for any matrix X, $\bar{P}_X = I - X(X'X)^{-1}X'$ is the projection onto the space orthogonal to the columns of X. As is well known, the two stage least squares estimate corresponds to $k = 1$, ordinary least squares corresponds to $k = 0$, and limited information (single equation) maximum likelihood corresponds to $k = \lambda$, where

$$\lambda = \min_{\beta_*} \frac{\beta_*' Y_*' \bar{P}_{Z_1} Y_* \beta_*}{\beta_*' Y_*' \bar{P}_Z Y_* \beta_*} = \frac{\hat{\beta}_*' Y_*' \bar{P}_{Z_1} Y_* \hat{\beta}_*}{\hat{\beta}_*' Y_*' \bar{P}_Z Y_*' \hat{\beta}_*} \tag{4}$$

and $Y_* = (y, Y_1)$. The term $\hat{\beta}_*'$ in (4), when normalized, can be written as $(-1, \hat{\beta}_\lambda')$ where $\hat{\beta}_\lambda$ is the limited information maximum likelihood estimator of β.

To write the reduced form of the system, partition B^{-1} conformably with $Y = (y, Y_1, Y_2)$ as $B^{-1} = (b, B_1, B_2)$ where b is $G \times 1$, B_1 is $G \times G_1$, and B_2 is $G \times G_2$. Then

$$\begin{aligned} y &= -Z\Gamma b - \sigma U b, \\ Y_1 &= -Z\Gamma B_1 - \sigma U B_1, \\ Y_2 &= -Z\Gamma B_2 - \sigma U B_2. \end{aligned} \tag{5}$$

From (5), write

$$[Y_1, Z_1] = [-ZB_1, Z_1] + \sigma[-UB_1, 0] = X + \sigma V.$$

If the first equation is identified, X is of (full) rank $G_1 + K_1$. Let $Q = (X'X)^{-1}$. Also let the first column of Σ be σ_1. Then, following Nagar [13], define

$$q = \operatorname{cov}(V, u) = \frac{E(V'u)}{T} = \begin{bmatrix} -B_1'\sigma_1 \\ \cdots\cdots \\ 0 \end{bmatrix}, \qquad \text{a } G \times 1 \text{ vector,}$$

$$C_1 = qq',$$

$$C_2 = \begin{pmatrix} B_1'(\Sigma - \sigma_1\sigma_1')B_1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \text{a } G \times G \text{ matrix.}$$

The matrix TC_2 is the variance-covariance matrix of $V' - qu'$, so C_2 is positive semidefinite.

Finally, let $r_k = (1 - k)T + kK - G_1 - K_1 - 1$, so that, for ordinary least squares,

$$r_0 = T - G_1 - K_1 - 1;$$

for two stage least squares,

$$r_1 = K - G_1 - K_1 - 1 = K_2 - G_1 - 1 = L - 1$$

where $L = K_2 - G_1$ is the degree of overidentification.

THEOREM 1 (ASYMPTOTIC BIAS):

$$\begin{aligned} &E(e_k) = \sigma^2 r_k Qq + 0(\sigma^3) \quad \textit{for fixed } k, \textit{ and} \\ (6) \qquad &E(e_\lambda) = -\sigma^2 Qq + 0(\sigma^3) \quad \textit{for limited information maximum likelihood,} \end{aligned}$$

where

$$e_k = \begin{pmatrix} \beta \\ \gamma \end{pmatrix}_k - \begin{pmatrix} \beta \\ \gamma \end{pmatrix}.$$

Substituting $k = 1 + \alpha/T$ in (6) yields Nagar's [**13**] bias result as $T \to \infty$. Also notice that

$$k = \frac{T - G_1 - K_1 - 1}{T - K} = 1 + \frac{L - 1}{T - K}$$

yields a small σ asymptotically unbiased estimator. Lack of bias, however, is not a very attractive property for econometric problems, so this is not recommended as an estimator. (The small σ asymptotic bias of Nagar's "large sample unbiased" estimator can be evaluated using (6).)

THEOREM 2: *Let* $s_k = k(k - 1)(T - K)$. *Then*

$$\begin{aligned} (7) \qquad E(e_k e_k') = \sigma^2 Q + \sigma^4[&\{(1 - 2r_k)\operatorname{tr}(C_1 Q) + \operatorname{tr}(C_2 Q)\}Q \\ &+ ((r_k - 1)^2 + 2s_k + 2)QC_1Q + (s_k - r_k + 1)QC_2Q] + 0(\sigma^6) \end{aligned}$$

for fixed k, and, provided $T > K + 2$,

$$\begin{aligned} E(e_\lambda e_\lambda') = \sigma^2 Q + \sigma^4 \Bigg[&\{3\operatorname{tr}(C_1 Q) + \operatorname{tr}(C_2 Q)\}Q \\ &+ 6QC_1Q + \left\{\frac{(L + 2)(T - K + L - 2)}{T - K - 2}\right\}QC_2Q \Bigg] + 0(\sigma^5). \end{aligned}$$

Again notice that substituting $k = 1 + \alpha/T$ in (7) yields the same expression as Nagar [**13**] found as $T \to \infty$.

In order to compare moment matrices, a strong criterion is adopted: $A \geqslant B$ if and only if $A - B$ is nonnegative definite (since A and B are loss functions, this implies that B is preferred). The following lemma is useful.

LEMMA: *If $A \geqslant 0$, then $(tr\ A)I \geqslant A$.*

PROOF: Let H be an orthogonal matrix which diagonalizes A. Then

$$e_i H\{(\operatorname{tr} A)I - A\}H'e_i' = \sum_{j=1}^{n} \lambda_j - \lambda_i = \sum_{j \neq i} \lambda_j \geqslant 0$$

where $\{\lambda_i\}$ are the eigenvalues of A and e_i is a unit vector.

COROLLARY 1: $E(e_\lambda e_\lambda') \geqslant E(e_1 e_1')$ if $L \leqslant 6$.

PROOF:

$$\begin{aligned} E(e_\lambda e_\lambda') - E(e_1 e_1') &= \sigma^4[2L(\operatorname{tr}(C_1Q)Q + (4 - L)LQC_1Q \\ &\quad + L\left(2 + \frac{L+2}{T-K-2}\right)QC_2Q\} + 0(\sigma^5) \\ &\geqslant \sigma^4\left\{(-L^2 + 6L)QC_1Q\right. \\ &\quad \left. + L\left(2 + \frac{L+2}{T-K-2}\right)QC_2Q\right\} + 0(\sigma^5) \\ &\geqslant 0 + 0(\sigma^5) \quad \text{if} \quad L \leqslant 6 \end{aligned}$$

since

$$\begin{aligned} QC_1Q &= Q^{\frac{1}{2}}(Q^{\frac{1}{2}}C_1Q^{\frac{1}{2}})Q^{\frac{1}{2}} \leqslant Q^{\frac{1}{2}}(\operatorname{tr}(Q^{\frac{1}{2}}C_1Q^{\frac{1}{2}}))Q^{\frac{1}{2}} \\ &= (\operatorname{tr}(C_1Q))Q \quad \text{using the Lemma.} \end{aligned}$$

This means that the expected mean squared error of any linear combination of parameters is smaller (for $L \leqslant 6$) estimated by two stage least squares than by limited information maximum likelihood. Also it suggests that for really large systems the reverse might be true; Monte Carlo results from small systems, where two stage least squares is better, may mislead when applied to large systems.

COROLLARY 2: $E(e_0 e_0') \leqslant E(e_1 e_1')$ if $0 \leqslant T - K \leqslant 2(3 - L)$.

PROOF:

$$\begin{aligned} E(e_1 e_1') - E(e_0 e_0') &= \sigma^4(r_0 - r_1)\{2\operatorname{tr}(C_1Q)Q + (2 - r_0 - r_1)QC_1Q \\ &\quad + QC_2Q\} + 0(\sigma^5) \\ &\geqslant \sigma^4(r_0 - r_1)\{(4 - r_0 - r_1)QC_1Q + QC_2Q\} + 0(\sigma^5) \\ &\geqslant 0 + 0(\sigma^5) \end{aligned}$$

if the above condition is satisfied. *Q.E.D.*

COROLLARY 3:

$$\frac{d}{dk}\{E(e_k e_k') - (Ee_k)(Ee_k')\} = 2(T - K)\sigma^4\{(\operatorname{tr} C_1 Q)Q + k(2QC_1Q + QC_2Q)\} + 0(\sigma^5).$$

Therefore the minimum variance k-class estimator occurs when k is negative.

Carnegie-Mellon University

Manuscript received October, 1969.

APPENDIX A: PROOF OF THEOREMS

Throughout these proofs it is important to distinguish the cases when k is fixed from those when it may be random. When k can be either, it is denoted k^*; when fixed, it is denoted k.

Some additional notation is useful here: for any matrix F, $P_F = F(F'F)^{-1}F'$ is the projection onto the space spanned by F. $\bar{P}_F = I - P_F$, which was introduced in Section 2 following Equation (3), is the projection onto the space orthogonal to the space spanned by F. Also notice that

$$\bar{P}_{Z_1}\bar{P}_X = \bar{P}_X,$$
$$\bar{P}_{Z_1}\bar{P}_Z = \bar{P}_Z,$$

and

$$\bar{P}_X\bar{P}_Z = \bar{P}_Z.$$

Finally, let

$$S = V'X + X'V.$$

LEMMA A1: *If* $k^* = 0_p(1)$, *then*

$$e_{k^*} = \sigma QX'u + \sigma^2 Q(V'N^*u - SQ'u) + \sigma^3 Q(SQSQX'u - V'N^*VQX'u - SQV'N^*u) + 0_p(\sigma^4)$$

where $N^* = I - k^*\bar{P}_Z$.

PROOF:

$$\begin{aligned} e_{k^*} &= \begin{bmatrix} Y_1'Y_1 - k^*V^{*\prime}V^* & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{bmatrix} (Y_1 - k^*V^*)' \\ Z_1' \end{bmatrix}\sigma u \\ &= \sigma\{(X'X) + \sigma S + \sigma^2 V'N^*V)^{-1}(X' + \sigma V'N^*)u \\ &= \sigma\{I + Q(\sigma S + \sigma^2 V'N^*V)\}^{-1}Q(X' + \sigma V'N^*)u. \end{aligned}$$

Use of the standard geometric expansion for the inverse of a matrix,

$$(I + h\Delta)^{-1} = I - h\Delta + h^2\Delta^2 - h^3\Delta^3 + \dots,$$

yields

$$\begin{aligned} e_{k^*} &= \sigma Q(X' + \sigma V'N^*)u - \sigma Q[\sigma S + \sigma^2 V'N^*V]Q(X' + \sigma V'N^*)u \\ &\quad + \sigma Q(\sigma S + \sigma^2 V'N^*V)Q(\sigma S + \sigma^2 V'N^*V)Q(X' + \sigma V'N^*)u + \dots \\ &= \sigma QX'u + \sigma^2 Q(V'N^*u - SQX'u) + \sigma^3 Q(SQSQX'u - V'N^*VQX'u \\ &\quad - SQV'N^*u) + 0_p(\sigma^4). \end{aligned}$$

Q.E.D.

LEMMA A2: $\lambda = 0_p(1)$.

PROOF: See Lemma 1 of Kadane [11].

LEMMA A3:

$$\lambda = \frac{u'\bar{P}_X u}{u'\bar{P}_Z u} + 2\sigma\left[\frac{(u'XQV'\bar{P}_Z u)(u'\bar{P}_X u) - (u'XQV'\bar{P}_X u)(u'\bar{P}_Z u)}{(u'\bar{P}_Z u)^2}\right] + 0_p(\sigma^2).$$

(The first-order term of Lemma A3 is discussed in relation to the previous literature in [11].)

PROOF: Let X^* be the first G_1 rows of Q and X^{**} the last K_1 rows, so that

$$Q = \begin{bmatrix} X^* \\ X^{**} \end{bmatrix}.$$

Then, recalling $\bar{P}_{Z_1} Z_1 = 0$,

$$\begin{aligned}
\bar{P}_{Z_1} Y_* \hat{\beta}_* &= \bar{P}_{Z_1}(y, Y_1)\left[\begin{pmatrix} -1 \\ \beta \end{pmatrix} + \sigma\begin{pmatrix} 0 \\ X^* \end{pmatrix} X'u + \sigma^2\begin{pmatrix} 0 \\ X^* \end{pmatrix}(V'N^*u - SQX'u) + 0_p(\sigma^3)\right] \\
&= \bar{P}_{Z_1}[-\sigma u + \sigma(X + \sigma V)QX'u + \sigma^2(X + \sigma V)Q(V'N^*u - SQX'u) + 0_p(\sigma^3)] \\
&= \bar{P}_{Z_1}[-\sigma \bar{P}_X u + \sigma^2 \bar{P}_X VQX'u + \sigma^2 XQV'(N^* - P_X)u + 0_p(\sigma^3)] \\
&= -\sigma \bar{P}_X u + \sigma^2 \bar{P}_X VQX'u + \sigma^2 \bar{P}_{Z_1} XQV'(N^* - P_X)u + 0_p(\sigma^3).
\end{aligned}$$

Writing

$$\begin{aligned}
\lambda &= \frac{N}{D} = \frac{\sigma^2 N_2 + \sigma^3 N_3 + \dots}{\sigma^2 D_2 + \sigma^3 D_3 + \dots} \\
&= \frac{N_2}{D_2} + \sigma\left[\frac{N_3 D_2 - D_3 N_2}{D_2^2}\right] + 0_p(\sigma^2); \\
N &= \hat{\beta}'_* Y'_* \bar{P}_{Z_1} Y_* \hat{\beta}_* = (\hat{\beta}'_* Y'_* \bar{P}_{Z_1})(\bar{P}_{Z_1} Y_* \hat{\beta}_*) \\
&= \{-\sigma u'\bar{P}_X + \sigma^2 u'XQV'\bar{P}_X + \sigma^2 u'(N^* - P_X)VQX' + 0_p(\sigma^3)\} \\
&\quad \times \{-\sigma \bar{P}_X u + \sigma^2 \bar{P}_X VQX'u + \sigma^2 \bar{P}_{Z_1} XQV'(N^* - P_X)u + 0_p(\sigma^3)\} \\
&= \sigma^2 u'\bar{P}_X u - 2\sigma^3 u'\bar{P}_X VQX'u + 0_p(\sigma^4).
\end{aligned}$$

Therefore $N_2 = u'\bar{P}_X u$ and $N_3 = -2u'\bar{P}_X VQX'u$.
Similarly

$$\begin{aligned}
D &= \hat{\beta}'_* Y'_* \bar{P}_Z Y_* \hat{\beta}_* = (\hat{\beta}'_* Y'_* \bar{P}_{Z_1})\bar{P}_Z(\bar{P}_{Z_1} Y_* \hat{\beta}_*) \\
&= \{-\sigma u'\bar{P}_X + \sigma^2 u'XQV'\bar{P}_X + \sigma^2 u'(N^* - P_X)VQX'\bar{P}_{Z_1} + 0_p(\sigma^3)\}\bar{P}_Z \\
&\quad \times \{-\sigma \bar{P}_X u + \sigma^2 \bar{P}_X VQX'u + \sigma^2 \bar{P}_{Z_1} XQV'(N^* - P_X)u + 0_p(\sigma^3)\} \\
&= \sigma^2 u'\bar{P}_Z u - 2\sigma^3 u'\bar{P}_Z VQX'u + 0_p(\sigma^4).
\end{aligned}$$

Hence $D_2 = u'\bar{P}_Z u$ and $D_3 = -2\sigma u'\bar{P}_Z VQX'u$, and

$$\lambda = \frac{u'\bar{P}_X u}{u'\bar{P}_Z u} + 2\sigma\left[\frac{(u'XQV'\bar{P}_Z u)(u'\bar{P}_X u) - (u'XQV'\bar{P}_X u)(u'\bar{P}_Z u)}{(u'\bar{P}_Z u)^2}\right] + 0_p(\sigma^2),$$

Q.E.D.

The following numbers are useful in the proofs of Theorems 1 and 2:

$$\begin{aligned}
a &= \operatorname{tr} \bar{P}_Z = T - K, \\
b &= \operatorname{tr}(\bar{P}_X - \bar{P}_Z) = \operatorname{tr}(P_Z - P_X) = K_2 - G_1 = L, \\
c &= \operatorname{tr} P_X = K_1 + G_1 = T - a - b.
\end{aligned}$$

Also the notation $[\dots]^*$ is used to denote a single expression, valid for both $k^* = k$ and $k^* = \lambda$, the upper part referring to k, the lower part to λ. For example

$$N^* = I - k^*\bar{P}_Z = I - \begin{bmatrix} k \\ \lambda \end{bmatrix}^* \bar{P}_Z = \begin{bmatrix} (1-k)I + kP_Z \\ I - \lambda\bar{P}_Z \end{bmatrix}^*.$$

The lemmas on expectation in Appendix B are used without special comment in the proofs to follow.

PROOF OF THEOREM 1: Recalling Lemma A1,

$$E(e_{k^*}) = \sigma E(QX'u) + \sigma^2 E(QV'N^*u - Q(V'X + X'V)QX'u) + 0(\sigma^3), \tag{A1}$$

$$E(QX'u) = QX'E(u) = 0. \tag{A2}$$

Let $W' = V' - qu'$. The expression W' is obviously independent of u', by construction. Then

$$\begin{aligned} E(QV'N^*u) &= E(QV'(I - k^*\bar{P}_Z)u) \\ &= QE(W'(I - k^*\bar{P}_Z)u) + QqE(u'(I - k^*\bar{P}_Z)u) \\ &= Qq(T - E(k^*u'\bar{P}_Z u)) + 0(\sigma) \qquad \text{(A3)} \\ &= Qq\left(T - \begin{bmatrix} ka \\ a + b + 0(\sigma) \end{bmatrix}^*\right); \end{aligned}$$

$$E(QV'XQX'u) = QE(W'XQX'u + qu'XQX'u) = Qq \operatorname{tr} P_X = cQq; \tag{A4}$$

$$E(QX'VQX'u) = QX'E(WQX'u + uq'QX'u) = QX'Euu'XQq = Qq. \tag{A5}$$

Substitution of (A2), (A3), (A4), and (A5) in (A1) yields Theorem 1.

PROOF OF THEOREM 2: Again return to Lemma A1:

$$Q^{-1}e_{k^*}e_{k^*}Q^{-1} = B_1 + B_2 + B_2' + B_3 + B_4 + B_4' + 0_p(\sigma^5) \tag{A6}$$

where

$$\begin{aligned} B_1 &= \sigma^2(X'uu'X), \\ B_2 &= \sigma^3(X'u(u'N^*V - u'XQS)), \\ B_3 &= \sigma^4(V'N^*u - SQX'u)(u'N^*V - u'XQS), \\ B_4 &= \sigma^4(SQSQX'u - V'N^*VQX'u - SQV'N^*u)(u'X). \end{aligned} \tag{A7}$$

The expectation of each of the above terms is evaluated below.

$$E(B_1) = \sigma^2 X'Euu'X = \sigma^2 X'X = \sigma^2 Q^{-1}. \tag{A8}$$

For B_2, since expectations of products with an odd number of factors of any normal variables with zero mean are zero, the only possible non-zero contribution may come from N^* when $k^* = \lambda$.

$$EX'uu'[I - \lambda\bar{P}_Z]V = -E\lambda X'uu'\bar{P}_Z V. \tag{A9}$$

Because the computation of $e_{k^*}e_{k^*}'$ is sought in this proof to order σ^4, both σ^0 and σ^1 order terms of λ are relevant here. Hence

$$\begin{aligned} -E\lambda X'uu'\bar{P}_Z V = &-E\frac{u'\bar{P}_X u}{u'\bar{P}_Z u}X'uu'\bar{P}_Z V \\ &+ 2\sigma E\left\{\left[\frac{(u'XQV'\bar{P}_X u)(u'\bar{P}_Z u) - (u'XQV'\bar{P}_Z u)(u'\bar{P}_X u)}{(u'\bar{P}_Z u)^2}\right]X'uu'\bar{P}_Z V\right\} + 0(\sigma^2). \end{aligned} \tag{A10}$$

Again using the independence of $W = V - uq'$ and u,

$$E\left\{\frac{u'\bar{P}_X u}{u'\bar{P}_Z u} X'uu'\bar{P}_Z V\right\} = E\left\{\frac{u'\bar{P}_X u}{u'\bar{P}_Z u} X'uu'\bar{P}_Z uq'\right\} = Eu'\bar{P}_X uX'uq' = 0,$$

$$2\sigma X'Eu\left[\frac{(u'XQV'\bar{P}_X u)(u'\bar{P}_Z u) - (u'XQV'\bar{P}_Z u)(u'\bar{P}_X u)}{(u'\bar{P}_Z u)^2}\right]u'\bar{P}_Z V$$

$$= 2\sigma X'Eu\left[\frac{(u'XQV'\bar{P}_X u)(u'\bar{P}_Z u) - (u'XQV'\bar{P}_Z u)(u'\bar{P}_X u)}{(u'\bar{P}_Z u)^2}\right]u'\bar{P}_Z uq'$$

$$+ 2\sigma X'Eu\left[\frac{(u'XQV'\bar{P}_X u)(u'\bar{P}_Z u) - (u'XQV'\bar{P}_Z u)(u'\bar{P}_X u)}{(u'\bar{P}_Z u)^2}\right]u'\bar{P}_Z W$$

$$= 2\sigma X'Eu\left[\frac{u'XQqu'\bar{P}_X uu'\bar{P}_Z u - u'XQqu'\bar{P}_Z uu'\bar{P}_X u}{(u'\bar{P}_Z u)}\right]q'$$

(A11) $$+ 2\sigma X'Eu\left[\frac{(u'XQW'\bar{P}_X u)(u'\bar{P}_Z u) - (u'XQW'\bar{P}_Z u)(u'\bar{P}_X u)}{(u'\bar{P}_Z u)^2}\right]u'\bar{P}_Z W$$

$$= 2\sigma X'Eu\left[\frac{(u'\bar{P}_Z u)u'XQW'\bar{P}_X u - (u'\bar{P}_X u)u'XQW'\bar{P}_Z u}{(u'\bar{P}_Z u)^2}\right]u'\bar{P}_Z W$$

$$= 2\sigma X'Eu\left[\frac{(u'\bar{P}_Z u)u'XQ\operatorname{tr}(\bar{P}_X uu'\bar{P}_Z) - (u'\bar{P}_X u)(u'XQ)\operatorname{tr}(\bar{P}_Z uu'\bar{P}_Z)}{(u'\bar{P}_Z u)^2}\right]C_2$$

$$= 2\sigma X'Eu\left[u'XQ - \frac{(u'\bar{P}_X u)u'XQ}{u'\bar{P}_Z u}\right]C_2$$

$$= -2\sigma X'Eu\left[\frac{u'(\bar{P}_X - \bar{P}_Z)u}{u'\bar{P}_Z u} u'XQC_2\right]$$

$$= -2\sigma X'\left[\frac{b+2}{a-2}(\bar{P}_X - \bar{P}_Z) + \frac{b}{a}\bar{P}_Z + \frac{b}{a-2}P_X\right]XQC_2 = -2\sigma\frac{b}{a-2}C_2;$$

(A12) $$EB_2 = \sigma^4\begin{bmatrix} 0 \\ 2b \\ -\overline{a-2} \end{bmatrix}^* C_2 + 0(\sigma^5),$$

using (A9), (A10), and (A11).

Moving now to B_3,

(A13)
$$E(V'N^*u - SQX'u)(u'N^*V - u'XQS$$
$$= E\{V'N^*uu'N^*V\} - E\{V'N^*uu'XQS\} - E\{SQX'uu'N^*V\} + E\{SQX'uu'XQS\}.$$

(A14)
$$E\{V'N^*uu'N^*V\} = E\{W'N^*uu'N^*W\} + E\{qu'N^*uu'N^*uq'\}$$
$$= E(u'(N^*)^2u)C_2 + E(u'N^*u)^2C_1.$$

$$u'(N^*)^2u = u'[I - k^*\bar{P}_Z][I - k^*\bar{P}_Z]u$$
$$= u'[(P_Z) + (k^* - 1)^2\bar{P}_Z]u.$$

$$E\{u'N^2u\} = T - a + (k-1)^2a.$$

$$E\{u'[P_Z + (\lambda - 1)^2\bar{P}_Z]u\} = T - a + E\frac{(u'(\bar{P}_X - \bar{P}_Z)u)^2}{u'\bar{P}_Z u} + 0(\sigma)$$
$$= T - a + \frac{b(b+2)}{a-2} + 0(\sigma).$$

$$u'N^*u = u'[P_Z + (1 - k^*)\bar{P}_Z]u = u'P_Z u + (1 - k^*)u'\bar{P}_Z u.$$

$$E(u'Nu)^2 = (T-a)(T-a+2) + 2(1-k)(T-a)a + (1-k)^2a(a+2).$$

$$E(u'(I - \lambda\bar{P}_Z)u)^2 = E(u'P_X u)^2 + 0(\sigma) = c(c+2) + 0(\sigma).$$

Summarizing, from (A14),

$$
\text{(A15)}\quad
\begin{aligned}
E\{V'N^*uu'N^*V\} &= \left[\frac{(T-a)(T-a+2)+2(1-k)(T-a)a+(1-k)^2a(a+2)}{c(c+2)+0(\sigma)}\right]^* C_1\\
&\quad + \left[\frac{T-a+(k-1)^2a}{T-a+\dfrac{b(b+2)}{a-2}+0(\sigma)}\right]^* C_2 .
\end{aligned}
$$

Next, from (A13),

$$
\text{(A16)}\quad
\begin{aligned}
E\{SQX'uu'N^*V\} &= E\{(V'X+X'V)QX'uu'N^*V\}\\
&= E\{(W'X+X'W)QX'uu'N^*W\} + E\{(qu'X+X'uq')QX'uu'N^*uq'\}\\
&= E\{u'N^*XQX'u\}C_2 + E\{X'N^*uu'XQ\}C_2 + Eu'[P_X]uu'N^*uC_1\\
&\quad + E[X'uu'N^*uu'XQC_1]
\end{aligned}
$$

(since $N^*X = X$)

$$
\begin{aligned}
&= (c+1)C_2 + \left[\frac{(T+2-ka)c}{c(c+2)+0(\sigma)}\right]^* C_1 + \left[\frac{T+2-ka}{c+2+0(\sigma)}\right]^* C_1\\
&= (c+1)C_2 + (c+1)\left[\frac{T+2-ka}{c+2+0(\sigma)}\right]^* C_1 .
\end{aligned}
$$

Finally, the last term from (A13),

$$
\text{(A17)}\quad
\begin{aligned}
ESQX'uu'XQS &= E(V'X+X'V)QX'uu'XQ(V'X+X'V)\\
&= E(W'X+X'W)QX'uu'XQ(W'X+X'W)\\
&\quad + E\{(qu'X+X'uq')QX'uu'XQ(qu'X+X'uq')\}\\
&= C_2[QX'IXQX']X + X'(XQX'IXQ)C_2 + \operatorname{tr}(XQX'IXQX')C_2\\
&\quad + X'(\operatorname{tr}(QX'IXQC_2)X + E\{qu'XQX'uu'XQX'uq'\}\\
&\quad + E\{X'uq'QX'uu'XQqu'X\} + EX'uq'QX'uu'XQX'uq'\\
&\quad + Equ'XQX'uu'XQqu'X\\
&= 2C_2 + \operatorname{tr}(P_X)C_2 + \operatorname{tr}(QC_2)Q^{-1} + E(u'P_Xu)^2C_1\\
&\quad + X'E(uu'XQC_1QX'uu')X + X'Euu'P_Xuu'XQC_1 + C_1QX'uu'P_Xuu'X\\
&= (c+2)C_2 + \operatorname{tr}(QC_2)Q^{-1} + c(c+2)C_1 + \operatorname{tr}(QC_1)Q^{-1}\\
&\quad + 2(c+2)C_1 + 2C_1\\
&= (c+2)C_2 + \operatorname{tr}(QC_2)Q^{-1} + ((c+2)^2+2)C_1 + \operatorname{tr}(QC_1)Q^{-1}.
\end{aligned}
$$

Returning to (A13) and using (A15), (A16), and (A17),

$$
\text{(A18)}\quad
\begin{aligned}
E(B_3) = \sigma^4\Bigg\{&\left[\frac{[b+(1-k)a]^2+2a(k-1)k+2}{2+0(\sigma)}\right]^* C_1 + \operatorname{tr}(QC_1)Q^{-1}\\
&+\left(b+\left[\frac{(k-1)^2a}{\dfrac{b(b+2)}{a-2}+0(\sigma)}\right]^*\right)C_2 + (\operatorname{tr}QC_2)Q^{-1}\Bigg\}.
\end{aligned}
$$

Finally, B_4 must be computed:

$$\text{(A19)} \quad \begin{aligned} &E(SQSQX'u - V'N^*VQX'u - SQV'N^*u)(u'X) \\ &\quad = E(SQSQX'uu'X) - E(V'N^*VQX'uu'X) - E(SQV'N^*uu'X). \end{aligned}$$

$$\begin{aligned} E(SQSQX'uu'X) &= E(V'X + X'V)Q(V'X + X'V)QX'uu'X \\ &= E(W'X + X'W)Q(W'X + X'W)QX'uu'X \\ &\quad + E(qu'X + X'uq')Q(qu'X + X'uq')QX'uu'X \\ &= E(W'XQW'X) + E(W'XQX'W) + E(X'WQW'X) + E(X'WQX'W) \\ &\quad + E(qu'XQqu'XQX'uu'X) + E(qu'XQX'uq'QX'uu'X) \\ &\quad + E(X'uq'Qqu'XQX'uu'X) + E(X'uq'QX'uq'QX'uu'X) \\ \text{(A20)} \quad &= C_2QX'X + \operatorname{tr}(P_X)C_2 + \operatorname{tr}(C_2Q)Q^{-1} + X'XQC_2 \\ &\quad + E(qq'QX'uu'P_Xuu'X) + Eqq'QX'uu'P_Xuu'X \\ &\quad + E(\operatorname{tr}(C_1Q)X'uu'P_Xuu'X) + E(X'uu'XQqq'QX'uu'X) \\ &= (c + 2)C_2 + \operatorname{tr}(C_2Q)Q^{-1} + 2(c + 2)C_1 + (c + 2)\operatorname{tr}(C_1Q)Q^{-1} \\ &\quad + \operatorname{tr}(C_1Q)Q^{-1} + 2C_1 \\ &= (c + 2)C_2 + \operatorname{tr}(C_2Q)Q^{-1} + 2(c + 3)C_1 + (c + 3)\operatorname{tr}(C_1Q)Q^{-1}. \end{aligned}$$

$$\begin{aligned} E\{V'N^*VQX'uu'X\} &= E\{W'N^*WQX'uu'X\} + E\{qu'N^*uq'QX'uu'X\} \\ &= E(\operatorname{tr} N^*)C_2QX'uu'X + Eqq'QX'uu'N^*uu'X \\ &= E\left[\frac{T - ka}{T - \dfrac{u'\bar{P}_Xu}{u'\bar{P}_Zu}a + 0(\sigma)}\right]^* C_2QX'uu'X + \left[\frac{T + 2 - ka}{c + 2 + 0(\sigma)}\right]^* C_1. \end{aligned}$$

Now

$$\begin{aligned} EC_2QX'u\left(T - \frac{u'\bar{P}_Xu}{u'\bar{P}_Zu}a\right)u'X &= TC_2 - aEC_2QX'uu'X - EaC_2QX'u\left(\frac{u'(\bar{P}_X - \bar{P}_Z)u}{u'\bar{P}_Zu}\right)u'X \\ &= (T - a)C_2 - \frac{abC_2}{a - 2} = \left(T - a - \frac{ab}{a - 2}\right)C_2. \end{aligned}$$

Then

$$\text{(A21)} \quad E\{V'N^*VQX'uu'X\} = \left[\frac{T + 2 - ka}{c + 2 + 0(\sigma)}\right]^* C_1 + \left[\frac{T - ka}{T - a - \dfrac{ab}{a - 2} + 0(\sigma)}\right]^* C_2.$$

Finally,

$$\begin{aligned} E(SQV'N^*uu'X) &= E(V'X + X'V)QV'N^*uu'X \\ &= E(W'X + X'W)QW'N^*uu'X + Equ'XQqu'N^*uu'X \\ &\quad + EX'uq'Qqu'N^*uu'X \\ \text{(A22)} \quad &= EC_2QX'N^*uu'X + E(\operatorname{tr} C_2Q)X'N^*uu'X \\ &\quad + C_1QX'Euu'N^*uu'X + \operatorname{tr}(C_1Q)X'Euu'N^*uu'X \\ &= C_2 + \operatorname{tr}(C_2Q)Q^{-1} + \left[\frac{T + 2 - ka}{c + 2 + 0(\sigma)}\right]^* C_1 \\ &\quad + \left[\frac{T + 2 - ka}{c + 2 + 0(\sigma)}\right]^* \operatorname{tr}(C_1Q)Q^{-1}; \end{aligned}$$

$$EB_4 = \sigma^4\left\{\begin{bmatrix} 2\{a(k-1)-b+1\} \\ 2+0(\sigma) \end{bmatrix}^* C_1 + \begin{bmatrix} (k-1)a-b+1 \\ -b+\dfrac{ab}{a-2}+1+0(\sigma) \end{bmatrix}^* C_2 \right. \tag{A23}$$

$$\left. + \begin{bmatrix} a(k-1)-b+1 \\ 1+0(\sigma) \end{bmatrix}^* (\operatorname{tr} C_1 Q) Q^{-1} \right\}.$$

Finally, Theorem 2 follows using (A6), (A8), (A12), (A18), and (A23).

APPENDIX B: CERTAIN EXPECTATIONS

LEMMA B1: *Let A be a $T \times T$ constant matrix. Then*

$$E(W'AW) = (\operatorname{tr} A)C_2.$$

PROOF:

$$(EW'AW)_{ij} = E\sum_{t,t'} w_{ti}a_{tt'}w_{t'j} = E\sum w_{it}w_{tj}a_{tt}$$
$$= (C_2)_{ij}\sum_t a_{tt} = (\operatorname{tr} A)(C_2)_{ij}.$$

LEMMA B2: *Let A be a $c \times c$ constant matrix. Then*

$$E(WAW') = \operatorname{tr}(C_2 A)I.$$

PROOF:

$$[EWAW']_{tt'} = E\sum_{j,j'} w_{tj}a_{jj'}w_{t'j'} = E\sum (C_2)_{jj'}\delta_{tt'}a_{jj'}$$
$$= \operatorname{tr}(C_2 A)\delta_{tt'}.$$

LEMMA B3: *Let A be a $c \times T$ constant matrix. Then*

$$E(WAW) = A'C_2.$$

PROOF:

$$E(WAW)_{tj} = Ew_{tk}a_{kt'}w_{t'j} = Ea_{kt}\sum w_{tk}w_{tj} = (A'C_2)_{tj}.$$

LEMMA B4: *Let P_1 and P_2 be projections such that $P_1P_2 = 0$. Then*

$$Eu\frac{u'P_1u}{u'P_2u}u' = \left(\frac{\operatorname{tr} P_1 + 2}{\operatorname{tr} P_2 - 2}\right)P_1 + \left(\frac{\operatorname{tr} P_1}{\operatorname{tr} P_2}\right)P_2 + \left(\frac{\operatorname{tr} P_1}{\operatorname{tr} P_2 - 2}\right)[I - P_1 - P_2].$$

PROOF: Since P_1 and P_2 commute, let Γ (orthogonal) simultaneously diagonalize them:

$$\Gamma P_1 \Gamma' = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = I_1; \qquad \Gamma P_2 \Gamma' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{bmatrix} = I_2.$$

Also let $v' = u'\Gamma$ and $v = (v_1, v_2, v_3)$ be partitioned conformably with I_1 and I_2. Then

$$Eu\frac{u'P_1u}{u'P_2u}u' = \Gamma E\left[v\frac{v_1'v_1}{v_2'v_2}v'\right]\Gamma'$$
$$= \Gamma\left[\frac{\operatorname{tr} P_1 + 2}{\operatorname{tr} P_2 - 2}I_1 + \frac{\operatorname{tr} P_1}{\operatorname{tr} P_2}I_2 + \frac{\operatorname{tr} P_1}{\operatorname{tr} P_2 - 2}(I - (I_1 + I_2))\right]\Gamma'$$
$$= \frac{\operatorname{tr} P_1 + 2}{\operatorname{tr} P_2 - 2}P_1 + \frac{\operatorname{tr} P_1}{\operatorname{tr} P_2}P_2 + \frac{\operatorname{tr} P_1}{\operatorname{tr} P_2 - 2}(I - P_1 - P_2). \qquad Q.E.D.$$

In particular,

$$Eu\frac{u'(\bar{P}_X - \bar{P}_Z)u}{u'\bar{P}_Z u}u' = \frac{b+2}{a-2}(\bar{P}_X - \bar{P}_Z) + \frac{b}{a}\bar{P}_Z + \frac{b}{a-2}P_X.$$

LEMMA B5: *Let R be a symmetric, constant matrix*

$$Euu'Ruu' = (\operatorname{tr} R)I + 2R.$$

PROOF: Diagonalize R:

$$\Gamma R\Gamma' = D_\lambda, \qquad v = \Gamma u, \qquad v \sim N(0, I).$$

$$\begin{aligned} Euu'Ruu' &= \Gamma' E(vv'D_\lambda vv')\Gamma \\ &= \Gamma'(Ev \textstyle\sum \lambda_i v_i^2 v')\Gamma \end{aligned}$$

has the jth diagonal element

$$\begin{aligned} E(\textstyle\sum \lambda_i v_i^2)v_j^2 &= E\left\{\sum_{i \neq j} \lambda_i v_i^2 v_j^2 + \lambda_j v_j^4\right\} \\ &= \sum_{i \neq j} \lambda_i + 3\lambda_j = \sum_i \lambda_i + 2\lambda_j, \qquad \text{i.e., } \operatorname{tr}(R)I + 2D_\lambda. \end{aligned}$$

So the whole expectation is $(\operatorname{tr} R)I + 2R$. *Q.E.D.*

COROLLARY: *Let R_1 be a symmetric, constant matrix and R_2 a constant matrix; then*

$$Eu'R_1uu'R_2u = (\operatorname{tr} R_1)(\operatorname{tr} R_2) + 2\operatorname{tr}(R_1R_2).$$

PROOF:

$$\begin{aligned} Eu'R_1uu'R_2u &= E\{\operatorname{tr} u'R_1uu'R_2u\} \\ &= E \operatorname{tr} uu'R_1uu'R_2 = \operatorname{tr}(\{Euu'R_1uu'\}R_2) \\ &= \operatorname{tr}(((\operatorname{tr} R_1)I + 2R_1)R_2) = \operatorname{tr} R_1 \operatorname{tr} R_2 + 2\operatorname{tr} R_1R_2. \end{aligned}$$

Q.E.D.

$$Euu'N^*uu' = \begin{bmatrix} (T + 2 - ka)I - 2k\bar{P}_Z \\ (c+2)I - 2\bar{P}_X + 0(\sigma) \end{bmatrix}^*;$$

$$Eu'Ruu'N^*u = \begin{bmatrix} (T + 2 - ka)\operatorname{tr} R - 2k\operatorname{tr}(R\bar{P}_Z) \\ (c+2)\operatorname{tr} R - 2\operatorname{tr}(R\bar{P}_X) + 0(\sigma) \end{bmatrix}^*.$$

REFERENCES

[1] ANDERSON, T. W., AND H. RUBIN: "The Asymptotic Properties of Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, 21 (1950), 570–582.

[2] BASMANN, R. L.: "A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in Two Leading Over-identified Cases," *Journal of the American Statistical Association*, 56 (1961), 619–636.

[3] ———: "A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in a Leading Three Equation Case," *Journal of the American Statistical Association*, 58 (1963), 161–171.

[4] ———: "Remarks Concerning the Application of Exact Finite Sample Distribution Functions of GCL Estimators in Econometric Statistical Inference," *Journal of the American Statistical Association*, 58 (1963), 943–976.

[5] BERGSTROM, A. R.: "The Exact Sampling Distributions of Least Squares and Maximum Likelihood Estimators of the Marginal Propensity to Consume," *Econometrica*, 30 (1962), 480–490.

[6] CHOW, G. C.: "A Comparison of Alternative Estimators for Simultaneous Equations," *Econometrica*, 32 (1964), 532–553.

[7] CRAGG, J. G.: "Small-Sample Properties of Various Simultaneous Equation Estimators: The Results of Some Monte-Carlo Experiments," Princeton University, Econometric Research Program, Research Memorandum No. 68, 1964.

[8] ———: "On the Relative Small-Sample Properties of Several Structural-Equation Estimators," *Econometrica*, 35 (1967), 89–110.

[9] KABE, D. G.: "A Note on the Exact Distributions of the GCL Estimators in Two Leading Overidentified Cases," *Journal of the American Statistical Association*, 58 (1963), 535–537.

[10] ———: "A Note on the Exact Distributions of the GCL Estimators in a Leading Three-Equation Case," *Journal of the American Statistical Association*, 59 (1964), 881–894.

[11] KADANE, J. B.: "Testing Overidentifying Restrictions When the Disturbances are Small," *Journal of the American Statistical Association*, 65 (1970), 182–185.

[12] MARIANO, R. S.: "On Distributions and Moments of Single-Equation Estimators in a Set of Simultaneous Linear Stochastic Equations," Technical Report No. 27, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1969.

[13] NAGAR, A. L.: "The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations," *Econometrica*, 27 (1959), 575–595.

[14] QUANDT, R.: "Some Small Sample Properties of Certain Structural Equation Estimators," Princeton University, Econometric Research Program, Research Memorandum No. 48, 1962.

[15] ———: "On Certain Small Sample Properties of k-Class Estimators," *International Economic Review*, 6 (1965), 92–104.

[16] RICHARDSON, D. H.: "An Investigation of the Sampling Distribution of Estimators and Test Statistics Associated with a Simultaneous Equations Model," unpublished dissertation, Purdue University, 1966.

[17] ———: "The Exact Distribution of a Structural Coefficient Estimator," *Journal of the American Statistical Association* 63 (1968), 1214–1226.

[18] SAMUELSON, P.: "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments," *Review of Economic Studies*, 37 (1970), 537–542.

[19] SAWA, T.: "The Exact Sampling Distribution of Ordinary Least Squares and Two Stage Least Squares Estimators," *Journal of the American Statistical Association*, 64 (1969), 923–937.

[20] ———: "The Exact Finite-Sampling Distribution of the Limited-Information Maximum Likelihood Estimator," Discussion Paper, Series No. 20, Kyoto Institute of Economic Research, Kyoto University, 1969.

[21] ———: "Exact Finite-Sample Moments of the k-Class Estimators of Parameters in Simultaneous Equation Systems," mimeographed, 1970.

[22] SUMMERS, R. L.: "A Capital Intensive Approach to the Small Sample Properties of Various Simultaneous Equation Estimators," *Econometrica*, 33 (1965), 1–41.

[23] TAKEUCHI, K.: "Exact Sampling Moments of the Ordinary Least Squares, Instrumental Variable, and Two-Stage Least Squares Estimators," *International Economic Review*, 11 (1970), 1–12.