A Modified Stein-like Estimator for the Reduced Form Coefficients of Simultaneous Equations

Esfandiar Maasoumi

Econometrica, Volume 46, Issue 3 (May, 1978), 695-703.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197805%2946%3A3%3C695%3AAMSEFT%3E2.0.CO%3B2-Q

Econometrica, Vol. 46, No. 3 (May, 1978)

A MODIFIED STEIN-LIKE ESTIMATOR FOR THE REDUCED FORM COEFFICIENTS OF SIMULTANEOUS EQUATIONS

BY ESFANDIAR MAASOUMI[1]

In this paper a reduced form estimator is developed which combines the corresponding restricted 3SLS and the unrestricted LS estimators. This estimator is similar to the 'positive part' Stein-like estimators proposed by Baranchik [2] and S. Sclove [16] in the classical multivariate regression context. It is shown that, whereas the restricted (derived) 3SLS and 2SLS reduced form estimates possess no finite moments (hence have unbounded risk), the modified Stein-like reduced form (MSRF) estimator has finite moments of up to order $(T-n-m)$, where T is the sample size, n and m are the number of the endogenous and the non-stochastic exogenous variables in the system. Furthermore it is argued that, asymptotically, the difference between the MSRF and the 3SLS estimators is negligible.

1. INTRODUCTION

THE OBSERVED PHENOMENON of 'outliers' in the distribution of almost all reduced form estimators has been a cause for concern. In this connection it has been shown that, whereas the computationally tedious FIML estimators possess finite moments of order $(T-n-m)$ the corresponding 3SLS and 2SLS reduced form estimators have no integral moments (see Sargan [**13, 14**] and McCarthy [**11**]). However, it is possible to develop estimators which have almost all the asymptotic properties of the 3SLS but possess superior finite sample properties. In particular, combining least squares (LS) and 3SLS has obvious computational advantages over FIML estimators. Combined estimators have been proposed by various authors in a variety of situations. Use of extraneous information and stochastic parameter restrictions lead to "mixed regression" and "minimum expected loss" (MELO) estimators which can be expressed as combined restricted–unrestricted estimators (see Goldberger [**5**], Zellner [**19**], Maasoumi [**9**], and Sawa [**15**]).

Moreover, since the publication of Stein [**17**] and James–Stein [**18**], various Stein-like improved estimators have appeared in the literature which are based on the known estimators.[2] In their survey, Zellner and Vandaele [**20**] discuss various Bayesian and non-Bayesian interpretations associated with the Stein-like estimators which suggest their optimality under a variety of quadratic criteria. Also favorable reports have been made concerning their performance in some applied studies.[3] It would seem reasonable to expect that such estimators will be used more widely than has been the case in the past.

In the next section a "positive rule" Stein-like estimator and its modified variant are presented. In Section 3, the existence of finite moments is formally proved. The final section is concerned with the asymptotic distribution of the MSRF estimator in relation to that of the 3SLS estimator.

[1] I am indebted to Professor J. D. Sargan for much help and encouragement. I would also wish to thank Dr. D. F. Hendry, Dr. C. Wymer, and other members of the L.S.E. Econometric Workshop. Comments from three anonymous referees are gratefully acknowledged.

[2] E.g., see Baranchik [**2**] and S. Sclove [**16**].

[3] See D. J. Aigner and J. Judge [**1**] and Effron and Morris [**4**].

2. THE MSRF ESTIMATOR

Consider the structural model:

$$BY' + CZ' = AX' = U' \tag{1}$$

where $A = [B : C]$ is the $n \times (n+m)$ matrix of unknown coefficients, $X = [Y : Z]$ is the matrix of T observations on the n endogenous and the m non-stochastic exogenous variables such that $\lim_{T\to\infty} (Z'Z/T) = M$ is finite and non-negative definite. U represents the T values of n serially independent disturbance terms such that $U_{.t} \sim IIN(0, \Omega_u)$ and Ω_u is non-singular. The corresponding reduced form model is:

$$Y_t = PZ_t + V_t \qquad (t = 1, \ldots, T) \tag{2}$$

where $P = -B^{-1}C$ and $E(V_tV_t') = B^{-1}\Omega_u B'^{-1} = \Omega_v$. Denote the 3SLS estimate of A by $A^\dagger$ (with parameter constraints imposed). The following Wald type asymptotic test is employed to test the validity of the parameter constraints (and specification) on A:[4]

$$\operatorname{tr}\,(\Omega_2^{-1}A^\dagger(X'Z)(Z'Z)^{-1}(Z'X)A'^\dagger) \underset{a}{\sim} \chi^2_N. \tag{3}$$

Where Ω_2 is a consistent estimate of Ω_u (usually the 2SLS) and N is the total number of over-identifying degrees in (1). However, the following, asymptotically equivalent, test is developed in terms of the 3SLS reduced form estimates $P^\dagger = -B^{\dagger-1}C^\dagger$:[5]

$$\phi^\dagger = \operatorname{tr}\,[W^{-1}(\hat{P} - P^\dagger)(Z'Z)(\hat{P} - P^\dagger)'] \underset{a}{\sim} \chi^2_N \tag{4}$$

where $W = Y'[I - Z(Z'Z)^{-1}Z]Y/T$ is a consistent estimate of Ω_v and $\hat{P} = (Y'Z)(Z'Z)^{-1}$ is the LS estimate of P.[6]

For large sample sizes, the specifying restrictions are rejected if $\phi^\dagger$ or (3) exceed an appropriate critical value of the test. However, for small samples, asymptotic tests such as (3) and (4) lead to unduly high rates of rejection even for reasonably specified models.[7] Moreover, in this uncertain situation the unrestricted LS ($\hat{P}$) estimator may perform quite well. Given the above small sample problem we may wish to combine the unrestricted with the restricted estimator and allow the test result to determine the weights attached. Consequently the following estimator is proposed:

$$\begin{aligned} P^* &= \lambda P^\dagger + (1-\lambda)\hat{P} \\ &= P^\dagger + (1-\lambda)(\hat{P} - P^\dagger). \end{aligned} \tag{5}$$

[4] E.g., see E. Malinvaud [**10**, Ch. 9, pp. 358–360].

[5] Ibid., f.n. (3); also refer to Section 4 of this paper.

[6] Also note that: $A^\dagger(X'Z)(Z'Z)^{-1} = B^\dagger(Y'Z - P^\dagger Z'Z)(Z'Z)^{-1} = B^\dagger(\hat{P} - P^\dagger)$.

[7] R. L. Basmann [**3**] reports on this phenomena using a similar 2SLS test. Also the author has observed close to 35 per cent rate of rejection when small samples were applied to (4) in Monte Carlo experiments.

Note that:

$$\phi^* = \mathrm{tr}\,[W^{-1}(\hat{P}-P^*)(Z'Z)(\hat{P}-P^*)'] = \lambda^2\phi^\dagger. \tag{6}$$

Then if C_p is the chosen critical value of the test, we choose λ such that

$$\lambda = \begin{cases} 1 \quad \text{if} \quad \phi^\dagger \leq C_p \quad \text{(hypothesis accepted)}, \\ \left(\dfrac{\phi_2}{\phi^\dagger}\right)^{\frac{1}{2}} \quad \text{or} \quad \left(\dfrac{\phi_2}{\phi^\dagger}\right) \quad \text{if} \quad \phi^\dagger > C_p, \end{cases}$$

where $\phi_2 \leq C_p$ and may be chosen so as to minimize a desired quadratic loss measure. The similarity of P^* to the Stein-like estimators is seen from the following:

$$P^* = P^\dagger + I_{(C_p,\infty)}\left(1-\frac{\phi_2}{\phi^\dagger}\right)(\hat{P}-P^\dagger) \tag{7}$$

in which $I_{(\ )} = 1$ if $\phi^\dagger > C_p$ and zero otherwise.

3. THE MOMENTS

The following theorem demonstrates the existence of the first $(T-n-m)$ moments of the MSRF estimator. The methodology is similar to Sargan **[14]** and could be applied to obtain comparable results for other Stein-like estimators.[8]

THEOREM 1. *The integral moments, up to order $r \leq T-n-m$, of the MSRF reduced form estimator are uniformly bounded as $T \to \infty$.*

PROOF: From the definition of P^* it follows that:

$$\phi^* = \mathrm{tr}\left[TW^{-1}(\hat{P}-P^*)\left(\frac{Z'Z}{T}\right)(\hat{P}-P^*)'\right] \leq C_p. \tag{8}$$

The left-hand side of (8) is expanded using the identity $P^*-\hat{P} = (P^*-P)-(\hat{P}-P)$. We also note that by Cauchy's inequality:

$$\begin{aligned} &\mathrm{tr}\,[TW^{-1}(P^*-P)M(\hat{P}-P)'] \\ &\quad \leq \{\mathrm{tr}\,[TW^{-1}(P^*-P)M(P^*-P)']\}^{\frac{1}{2}} \cdot \{\mathrm{tr}\,[TW^{-1}(\hat{P}-P)M(\hat{P}-P)']\}^{\frac{1}{2}}. \end{aligned}$$

Then:

$$\{\mathrm{tr}\,[TW^{-1}(P^*-P)M(P^*-P)']\}^{\frac{1}{2}} \leq \{\mathrm{tr}\,[TW^{-1}(\hat{P}-P)M(\hat{P}-P)']\}^{\frac{1}{2}} + C_p^{\frac{1}{2}}. \tag{9}$$

Let $\gamma = \sqrt{T}f'\,\mathrm{vec}\,(P^*-P)$ be an arbitrary linear function of the elements of (P^*-P). Then we note that $[(\Omega_v^{-1}\otimes M)-ff'/\lambda_M]$ is non-negative definite where $\lambda_M = f'(\Omega_v \otimes M^{-1})f$ is the only non-zero (largest) root of $[ff'-\lambda(\Omega_v^{-1}\otimes M)]Z = 0$. Let $\phi = \mathrm{tr}\,[T\Omega_v^{-1}(P^*-P)M(P^*-P)']$; it follows that:

$$\gamma^2 \leq \lambda_M\phi \quad \text{and} \quad E(|\gamma|^r) \leq \lambda_M^{\frac{1}{2}r}E(\phi^{\frac{1}{2}r}). \tag{10}$$

[8] This would certainly be true for the test-based variety such as the "positive rule" and "pre-test" estimators.

In Appendix A, inequality (9) is utilized to prove that:

$$\lim_{T\to\infty} E(\phi^{\frac{1}{2}r}) \leq s^r e^{n\mu r}(1-2\mu r)^{-\frac{1}{2}nm}$$

which is obtained for $r \leq T-n-m$ and $\mu r < \frac{1}{2}$. *Q.E.D.*

4. ASYMPTOTICS

In what follows, we examine the limiting moments and distribution of the MSRF and compare them with the asymptotic moments and distribution of the 3SLS estimator ($P^\dagger$). It is also shown that, asymptotically, $\phi^\dagger$ of expression (4) is in fact a χ^2_N.

It is known that 3SLS deviations could be asymptotically expressed in terms of LS deviations, $\Delta\hat{P} = \hat{P} - P$. We note that:[9]

$$\Delta P^\dagger = P^\dagger - P = -B^{\dagger-1}C^\dagger + B^{-1}C \simeq B^{-1}\Delta B^\dagger B^{-1}C - B^{-1}\Delta C^\dagger$$

(first order approximation). Hence

$$\Delta P^\dagger \simeq -B^{-1}\Delta A^\dagger Q \quad \text{where} \quad Q = \begin{pmatrix} P \\ I_m \end{pmatrix}$$

and

$$\text{vec}\,\Delta P^\dagger \simeq -(B^{-1} \otimes Q')\,\text{vec}\,\Delta A^\dagger. \tag{11}$$

$S\alpha - s = \text{vec}\,A$ (or $S\Delta\alpha = \text{vec}\,\Delta A^\dagger$) is a parameterization of the linear restrictions and α is the vector of unrestricted parameters.[10]

From the 3SLS formulation:

$$\begin{aligned}\text{vec}\,\Delta A^\dagger &= -\frac{1}{T}F(\Omega_u^{-1} \otimes Q)\,\text{vec}\,(U'Z) \\ &= -\frac{1}{T}F(\Omega_u^{-1}B \otimes Q(Z'Z))\,(\text{vec}\,\Delta\hat{P})\end{aligned}$$

where $F = S[S'(\Omega_2^{-1} \otimes \hat{R})S]^{-1}S'$ and $\hat{R} = (X'Z)(Z'Z)^{-1}(Z'X)/T$. Therefore, asymptotically, (11) may be rewritten as follows:

$$\sqrt{T}(\text{vec}\,\Delta P^\dagger) = (B^{-1} \otimes Q')F^*(\Omega_u^{-1}B \otimes QM)(\sqrt{T}\,\text{vec}\,\Delta\hat{P}) \tag{12}$$

where F^* is F evaluated with Ω_u and $R = QMQ'$. Also

$$\begin{aligned}\sqrt{T}\,\text{vec}\,(\hat{P} - P^\dagger) &= \sqrt{T}(\text{vec}\,\Delta P - \text{vec}\,\Delta P^\dagger) \\ &= \sqrt{T}[I - (B^{-1} \otimes Q')F^*(\Omega_u^{-1}B \otimes QM)](\text{vec}\,\Delta\hat{P}).\end{aligned} \tag{13}$$

The right-hand side matrix is idempotent but not symmetric. Hence we consider

[9] E.g., see E. Malinvaud [**10**] and note that $\Delta B^\dagger = B^\dagger - B$ and $\Delta C^\dagger = C^\dagger - C$.

[10] S is an appropriate block diagonal matrix which selects the right-hand side variables in each equation and s is a vector selecting the corresponding dependent variables; $\Delta\alpha = \alpha^\dagger - \alpha$ and $\Delta A^\dagger = A^\dagger - A$.

suitable functions of $\Delta\hat{P}$ and $\Delta P^{\dagger}$ as follows:

$$(14) \qquad \xi = T^{\frac{1}{2}}(\Omega_v^{-\frac{1}{2}} \otimes M^{\frac{1}{2}}) \operatorname{vec} \Delta\hat{P} \underset{a}{\sim} N(0, I_{nm})$$

and

$$\eta = T^{\frac{1}{2}}(\Omega_v^{-\frac{1}{2}} \otimes M^{\frac{1}{2}}) \operatorname{vec} \Delta P^{\dagger}.$$

From (12), it follows that:

$$(15) \qquad \eta = (\Omega_v^{-\frac{1}{2}} B^{-1} \otimes M^{\frac{1}{2}} Q') F^* (B'^{-1} \Omega_v^{-\frac{1}{2}} \otimes Q M^{\frac{1}{2}}) \xi = K\xi$$

where K is idempotent and symmetric with rank $nm - N$. Also:

$$(16) \qquad T^{\frac{1}{2}}(\Omega_v^{-\frac{1}{2}} \otimes M^{\frac{1}{2}}) \operatorname{vec} 4(\hat{P} - P^{\dagger}) \simeq (I - K)\xi$$

where $I - K$ is of rank N (degrees of over-identification). Since W is a consistent estimate of Ω_v, we conclude that asymptotically:

$$\phi^{\dagger} \underset{a}{\simeq} \xi'(I - K)\xi \underset{a}{\sim} \chi^2$$

with N degrees of freedom. To examine the limiting distribution of P^* we note that:

$$\begin{aligned} (17) \qquad \zeta &= T^{\frac{1}{2}}(\Omega_v^{-\frac{1}{2}} \otimes M^{\frac{1}{2}}) \operatorname{vec} (P^* - P) \\ &= (1 - \lambda)\xi + \lambda\eta \\ &= (1 - \lambda)(I - K)\xi + \eta. \end{aligned}$$

From (15), η represents the asymptotic distribution of the 3SLS estimator. The first term in (17) is negligible since $\operatorname{pr}(\lambda = 1) = 100(1 - p)$ where p may be selected to be very small if the sample size is very small. In our simulation experiments, $\lambda \neq 1$ was observed in only 3–5 per cent of the replications with reasonably large samples. Even in such cases λ was extremely close to unity. As the limiting moments of the MSRF estimator exist (see Theorem 1), it is reasonable to expect that asymptotically, $E(\zeta) \doteqdot E(\eta)$ and $\operatorname{var}(\zeta) \doteqdot \operatorname{var}(\eta)$.

This can be seen by considering the moments of simpler functions of the MSRF as follows:

$$\begin{aligned} \phi = \zeta'\zeta &= T \operatorname{tr} [\Omega_v^{-1}(P^* - P)M(P^* - P)'] \\ &= (1 - \lambda)^2[\xi'(I - K)\xi] + [\xi' K \xi] = (1 - \lambda)^2 \phi^{\dagger} + \xi' K \xi. \end{aligned}$$

$\phi^{\dagger}$ and $\xi' K \xi$ are two independent χ^2 with N and $(nm - N)$ degrees of freedom, respectively. The second χ^2 corresponds to 3SLS alone. Hence

$$E(\phi) = E(\phi^{\dagger}(1 - \lambda)^2) + nm - N.$$

Denote by E^* the expectation taken over the range $\phi^{\dagger} \geqslant C_p$; hence

$$\lambda = \left(\frac{\phi_2}{\phi^{\dagger}}\right)^{\frac{1}{2}}:$$

$$E(\phi) = E^*(\phi^{\dagger}) - 2\sqrt{\phi_2} E^*(\phi^{\dagger\frac{1}{2}}) + \phi_2 E^*(1) + (nm - N).$$

Using

$$\int_{C_p}^{\infty} X^{k-1}\, e^{-X}\, dX = \Gamma(k|C_p)$$

$$\simeq e^{-C_p} C_p^{k-1}\,[1+(k-1)/C_p+(k-1)(k-2)/C_p^2],$$

$$E(\phi)=\frac{2\Gamma[(N/2+1)|C_p/2]-2\sqrt{2\phi_2}\,\Gamma[(N/2+\tfrac{1}{2})|C_p/2]+\phi_2\Gamma(N/2|C_p/2)}{\Gamma(N/2)}+(nm-N),$$

we obtain:

$$\grave{E}(\phi)=(nm-N)+(\sqrt{C_p}-\sqrt{\phi_2})^2\frac{\Gamma(N/2|C_p/2)}{\Gamma(N/2)}. \tag{18}$$

And if $\phi_2 = C_p$ is selected we approximate (18) by:

$$E(\phi)=(nm-N)+\frac{\Gamma(N/2|C_p/2)}{C_p\Gamma(N/2)}=(nm-N)+g \tag{19}$$

which can be seen to differ negligibly from the 3SLS component $(nm-N)$. As an example, take $N=10$ and $\phi_2 = C_p$:

Significance level	.05	.01	.001
C_p	18.3	23.2	29.6
g	.0027	.00043	.000034

Hence the difference, g, is negligible indeed.

It would seem that the limiting moments of the MSRF approximate the moments of the asymptotic distribution of the 3SLS very closely.

5. CONCLUSION

We have obtained a combined OLS-3SLS estimator for the coefficients of the reduced form of simultaneous equations which can be seen to be the same as 3SLS when the model is correctly specified but moves closer to the unrestricted OLS estimates the less reliable our specification becomes. A Wald type χ^2 test of specification is used for this purpose but has been transformed from the structural into an asymptotically equivalent χ^2 test in terms of the reduced form parameters.

It would seem that the MSRF estimator is preferable to the corresponding 3SLS since it possesses finite moments (and finite risk) and has the edge on the LS as it is nearly asymptotically equivalent to the 3SLS estimators.

Monte Carlo experiments reported in [**8**] have shown that the MSRF estimates are often closer to the true values and confirm the theoretical observations at the end of the last section. In particular, no outliers occur in its sampling distribution. Also, computationally, one needs no more work than is needed for the calculation of the 3SLS estimates and far less than FIML would require. This estimator is recommended when sample size is small and the specification problem is as described in Section 2.

London School of Economics

Manuscript received September, 1976; *final revision received August,* 1977.

APPENDIX A

THEOREM 2: *There is a finite uniform bound on $E(\phi^{\frac{1}{2}r})$ as $T \to \infty$.*

PROOF: We note that:

$$\phi = \operatorname{tr}[T\Omega_v^{-\frac{1}{2}}(P^*-P)M(P^*-P)'\Omega_v^{-\frac{1}{2}}] \leq \lambda_M \operatorname{tr}[TW^{-1}(P^*-P)M(P^*-P)'] \tag{A1}$$

or

$$\phi = \operatorname{tr}(A) \leq \lambda_M \operatorname{tr}(W^{*-1}A)$$

where λ_M is the largest latent root of $W^* = \Omega_v^{-\frac{1}{2}}W\Omega_v^{-\frac{1}{2}}$; A and W^* are non-negative definite.[11] From inequality (9), Section 3, it follows that:

$$\phi \leq \lambda_M\{\operatorname{tr}[TW^{-1}(\hat{P}-P)M(\hat{P}-P)']^{\frac{1}{2}} + C_p^{\frac{1}{2}}\}^2 = \lambda_M\{(\operatorname{tr} TW^{*-1}D)^{\frac{1}{2}} + C_p^{\frac{1}{2}}\}^2 \tag{A2}$$

where $D = \tilde{P}\tilde{P}'$ and $\tilde{P} = \Omega_v^{-\frac{1}{2}}(\hat{P}-P)M^{\frac{1}{2}} \sim N(0, I/T)$. From (A2) it can be seen that:

$$\phi^{\frac{1}{2}}(\det W^*)^{\frac{1}{2}}e^{-\mu \operatorname{tr}(W^*+TD)} \leq \lambda_M^{\frac{1}{2}}[\operatorname{tr}(TW^{*-1}D)^{\frac{1}{2}} + C_p^{\frac{1}{2}}](\det W^*)^{\frac{1}{2}}e^{-\mu \operatorname{tr}(W^*+TD)}.$$

The right-hand side is a continuous function of W^* and D which tends to zero as $T \to \infty$. Hence it must be bounded above for all W^* and D, for a given T. Supremum of this function (S) is found in Appendix B. From (A2) it follows that:

$$E(\phi^{\frac{1}{2}r}) \leq S^r E\{(\det W^*)^{-\frac{1}{2}r}\exp[\mu r \operatorname{tr}(W^*+TD)]\}. \tag{A3}$$

Noting the definition of D and that W^* has an unadjusted Wishart distribution, Sargan [**14**] derives the limit, as $T \to \infty$, of the expectation on the r.h.s. of (A3). Hence:

$$E(\phi^{\frac{1}{2}r}) \leq S^r e^{\mu nr}(1-2\mu r)^{-\frac{1}{2}nm}$$

where

$$\mu r < \tfrac{1}{2} \quad \text{and} \quad r \leq T-n-m.$$

Q.E.D.

APPENDIX B

To find the supremum of $\lambda_M^{\frac{1}{2}}[\operatorname{tr}(TW^{*-1}D)^{\frac{1}{2}} + C_p^{\frac{1}{2}}](\det W^*)^{\frac{1}{2}}\exp[-\mu \operatorname{tr}(W^*+TD)]$, taking latent root transformation, $W^* = H\Lambda H'$, the above expression can be rewritten, with $C = TH'(W^*)^{-\frac{1}{2}}D(W^*)^{-\frac{1}{2}}H$, as follows:

$$\lambda_m \prod_{i\neq m} \lambda_i^{\frac{1}{2}}[(\operatorname{tr} C)^{\frac{1}{2}} + C_p^{\frac{1}{2}}]\exp[-\mu \operatorname{tr}(I+C)\Lambda], \quad \text{independent of } T.$$

Maximization with respect to λ_i gives $\lambda_i = 1/2\mu d_i$ and $\lambda_m = 1/\mu dm$ where $d_i = 1 + C_{ii}$, with the requirement that $d_m \leq 2d_i$, $\forall i \neq m$ (since $\lambda_m \geq \lambda_i$, where m refers to the largest root).

Maximization with respect to off-diagonal elements of C gives

$$2^{-\frac{1}{2}(n-1)}e^{-\frac{1}{2}(n+1)}\mu^{-\frac{1}{2}(n+1)}dm^{-1}\left(\prod^n d_i\right)^{-\frac{1}{2}}[C_p^{\frac{1}{2}} + (\sum d_i - n)^{\frac{1}{2}}]. \tag{B1}$$

[11] E.g., see C. R. Rao [**12**] and note that $(W^{*-1} - \lambda_M^{-1}I)$ is non-negative definite.

Let $\prod_{i\neq m} d_i = D$ and keep d_m and D constant; let d_i vary to maximize. There are two cases.

Case I: $1 \leqslant dm \leqslant 2$. Then the only restrictions are: $D \geqslant 1$ and $d_i \geqslant 1$.

Clearly one maximizes (B1) if d_i is maximized. This is done if $d_i = 1$, all $i \neq m$ except for one i (which implies $D = 1, \ldots, 1 d_j = d_j$ for some j), and $\Sigma\, d_i = (dm + n - 2 + D)$. But

$$2^{-\frac{1}{2}(n-1)}(e\mu)^{-\frac{1}{2}(n+1)}\frac{C_p^{\frac{1}{2}} + (dm + D - 2)^{\frac{1}{2}}}{D^{\frac{1}{2}}dm}$$

is increasing in D and dm near unity. Also $D \geqslant 1$ can be shown to be binding. Then maximize with respect to d_m:

(B2) $\qquad \max(B1) = 2^{-\frac{1}{2}(n-1)}(e\mu)^{-\frac{1}{2}(n+1)}[C_p^{\frac{1}{2}} + (C_p + 1)^{\frac{1}{2}}]/2.$

Case II: $dm > 2$. In this case we must have $d_i \geqslant dm/2$, all i, and so $D \geqslant (dm/2)^{n-1}$. Once more we maximize $\Sigma\, d_i$ by taking all $d_i = (dm/2)$ except one which is $D/(dm/2)^{n-2}$. (B1) can now be written as follows:

$$2^{-\frac{1}{2}(n-1)}(e\mu)^{-\frac{1}{2}(n+1)}\frac{C_p^{\frac{1}{2}} + [dm + D/(dm/2)^{n-2} + (n-2)dm/2 - n]^{\frac{1}{2}}}{D^{\frac{1}{2}}dm}$$

or

$$2^{-\frac{1}{2}(n-1)}(e\mu)^{-\frac{1}{2}(n+1)}\frac{C_p^{\frac{1}{2}} + [ndm/2 + D/(dm/2)^{n-2} - n]^{\frac{1}{2}}}{D^{\frac{1}{2}}dm}.$$

For $C_p^{\frac{1}{2}}/D^{\frac{1}{2}}dm$ maximum is attained when $D \geqslant (dm/2)^{n-1}$ has its minimum which is at $dm = 2 \Rightarrow D = 1$. Hence $C_p^{\frac{1}{2}}/D^{\frac{1}{2}}dm \leqslant C_p^{\frac{1}{2}}/2$. Also

$$\frac{[(n(dm/2-1)/D) + (dm/2)^{-(n-2)}]^{\frac{1}{2}}}{dm} \leqslant \frac{1}{2}\left[\frac{n(dm/2-1) + dm/2}{(dm/2)^{n+1}}\right]^{\frac{1}{2}}.$$

Since $dm/2 > 1$ and $D \geqslant (dm/2)^{n-1}$, the maximum of this as a function of dm is attained at $dm/2 = 1$, at which it takes the value $\frac{1}{2}$. Then:

(B3) $$\frac{C_p^{\frac{1}{2}} + [ndm/2 + D/(dm/2)^{n-2} - n]^{\frac{1}{2}}}{D^{\frac{1}{2}}dm} \leqslant \frac{1}{2} + C_p^{\frac{1}{2}}/2$$

and the supremum over both Cases I and II is as follows:

(B4) $\qquad S = 2^{-\frac{1}{2}(n-1)}(e\mu)^{-\frac{1}{2}(n+1)}[C_p^{\frac{1}{2}} + (C_p + 1)^{\frac{1}{2}}]/2.$

Note that S is the constrained maxima (at $D = 1$) of Case I obtained at:

$$1 < d_m = \frac{2(C_p + 1)^{\frac{1}{2}}}{(C_p + 1)^{\frac{1}{2}} + C_p^{\frac{1}{2}}} < 2.$$

REFERENCES

[1] AIGNER, D. J., AND G. JUDGE: "Some Applications of Pre-Test and Stein-Rule Estimators to Economic Data," SSRI Workshop Series, University of Wisconsin, Madison, 1975.

[2] BARANCHIK, A. J.: "A Family of Minimax Estimators of the Mean of a Multivariate Normal Distribution," *Annals of Mathematical Statistics*, 41 (1970), 642–645.

[3] BASMANN, R. L.: "Exact Finite Sample Distributions for Some Econometric Estimators and Test Statistics: A Survey and Appraisal," *Frontiers of Quantitative Economics*, Vol. II, ed. by M. D. Intriligator and D. A. Kendrick. Amsterdam: North Holland Publishing Co., 1974.

[4] EFFRON, B., AND C. MORRIS: "Stein's Estimation Rule and its Competitors—An Empirical Bayes Approach," *Journal of American Statistical Association*, 68 (1973), 117–130.

[5] GOLDBERGER, A.: "Efficient Estimation in Overidentified Models: An Interpretive Analysis," *Structural Equation Models in the Social Sciences*, ed. by A. S. Goldberger and O. D. Duncan. Seminar Press, 1973.

[12] This result is not sensitive to the choice of μ.

[6] JAMES, W., AND C. STEIN: "Estimation with Quadratic Loss," *Fourth Berkeley Symposium on Mathematics, Statistics and Probability*, Vol. I. Berkeley: University of California Press, 1961.

[7] JUDGE, G., M. E. BOCK, AND T. A. YANCEY: "On Post Data Model Evaluation," Faculty Working Papers, University of Illinois at Urbana, Champaign, October 1972; *Review of Economics and Statistics*, 56 (1974), 245–253.

[8] MAASOUMI, E.: "A Study of Improved Methods of Estimating Reduced Form Coefficients Based Upon 3SLS," unpublished Ph.D. Thesis, London School of Economics, 1977.

[9] ———: "Stochastic Restrictions and a Quasi-Maximum Likelihood Reduced Form Estimator," Econometric Workshop, University of Birmingham, 1976.

[10] MALINVAUD, E.: *Statistical Methods of Econometrics*, Second Edition. Amsterdam: North-Holland Publishing Co., 1970.

[11] MCCARTHY, M. D.: "A Note on the Forecasting Properties of 2SLS Restricted Reduced Forms," *International Economic Review*, 13 (1972), 757–761.

[12] RAO, C. R.: *Linear Statistical Inference and its Applications*, Second Edition. New York: John Wiley and Sons, 1973.

[13] SARGAN, J. D.: "On the Existence of the Moments of Estimates of the Reduced Form Coefficients of Simultaneous Equation Models," London School of Economics, 1974.

[14] ———: "The Tails of the FIML Estimates of the Reduced Form Coefficients," London School of Economics, 1973.

[15] SAWA, T.: "Almost Unbiased Estimator in Simultaneous Equations Systems," *International Economic Review*, 14 (1972), 97–106.

[16] SCLOVE, S. L.: "Improved Estimators for Coefficients in Linear Regressions," *Journal of American Statistical Association*, 63 (1968), 596–606.

[17] SCLOVE, S., C. MORRIS, AND R. RADHAKRISHNAN: "Non-Optimality of Preliminary-Test Estimators for the Mean of a Multivariate Normal Distribution," *The Annals of Mathematical Statistics*, 43 (1972), 1481–1490.

[18] STEIN, C.: "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," *Third Berkeley Symposium on Mathematics, Statistics and Probability*, Vol. I. Berkeley: University of California Press, 1961.

[19] ZELLNER, A.: "A Note on the Relationship of MELO and Other Structural Coefficient Estimates," mimeo., University of Chicago, 1976.

[20] ZELLNER, A., AND W. VANDAELE: "Bayes-Stein Estimators for k-Means, Regression and Simultaneous Equation Models," in *Studies in Bayesian Econometrics and Statistics*, ed. by S. E. Fienberg and A. Zellner. Amsterdam: North-Holland Publishing Co., 1975.