MSE Dominance of the Positive-Part Shrinkage Estimator when Each Individual Regression Coefficient is Estimated

(Preliminary Version)

Akio Namba*

Graduate School of Economics
Kobe University
Rokko, Nada-ku, Kobe 657-8501
Japan

Abstract

In this paper we consider a regression model and a general family of shrinkage estimators of regression coefficients. We derive the formula for the mean squared error (MSE) of the estimators for each individual regression coefficient. It is shown analytically that the usual shrinkage estimators are dominated by their positive-part variants in terms of MSE.

Key Words: Pre-test, Shrinkege estimator, Positive-part estimator, Mean squared error, Dominance

*Akio Namba
Graduate School of Economics, Kobe University
Rokko, Nada-ku, Kobe, 657-8501, Japan
Phone & Fax: +81-78-803-6821
e-mail: namba@kobe-u.ac.jp

1 Introduction

In the context of a linear regression, the Stein-rule (SR) estimator proposed by Stein (1956) and James and Stein (1961) dominates the ordinary least squares (OLS) estimator in terms of predictive mean squared error (PMSE) if the number of the regression coefficients is larger than or equal to three. Though the SR estimator dominates the OLS estimator, Baranchik (1970) showed that the SR estimator is further dominated by the positive-part Stein-rule (PSR) estimator. As is shown in Judge and Bock (1978), the PSR estimator is considered as a pre-test estimator after a pre-test for the null hypothesis that all the regression coefficients are zeros. Nickerson (1988) and Namba (2003) considered general classes of shrinkage estimators which includes the SR and PSR estimator as special cases, and showed that shrinkage estimators are dominated by their positive-part variants.

Though the SR and PSR estimators dominate the OLS estimator when all the regression coefficients are estimated simultaneously, the dominance does not necessarily hold when each individual regression coefficient is estimated separately (e.g., Efron and Morris (1973) and Rao and Shinozaki (1978)). Ohtani and Kozumi (1996) examined the mean squared error (MSE) performance of the SR and PSR estimators when our concern is to estimate each individual regression coefficient, and showed that the SR and PSR estimators do not necessarily dominate the OLS estimator while the MSE dominance of the PSR estimator over the SR estimator still holds.

In this paper, we suppose that our concern is to estimate each individual regression coefficient separately, and consider a general class of pre-test shrinkage estimator considered by Namba (2003). We derive the explicit formulae for the moments of the estimators and show that shrinkage estimators are dominated by their positive-part variants in terms of MSE even when our concern is to estimate each individual regression coefficient.

2 Model and the estimators

Consider a linear regression model,

$$y = X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I_n), \tag{1}$$

where y is an $n \times 1$ vector of observations on a dependent variable, X is an $n \times k$ matrix of full column rank of observations on nonstochastic independent variables, β is a $k \times 1$ vector of regression coefficients, and ϵ is an $n \times 1$ vector of normal error terms.

Following Judge and Yancey (1986, p. 11.), we reparameterize the model (1) and work with the following orthonormal counterpart:

$$y = Z\gamma + \epsilon, \tag{2}$$

where $Z = XS^{-1/2}$, $\gamma = S^{1/2}\beta$, and $S^{1/2}$ is the symmetric matrix such that $S^{-1/2}SS^{-1/2} = Z'Z = I_k$, where

$S = X'X$. Then the ordinary least squares (OLS) estimator of γ is

$$\hat{\gamma} = Z'y. \tag{3}$$

In the context of reparameterized model, the Stein-rule (SR) estimator proposed by Stein (1956) is defined as

$$\hat{\gamma}_{\mathrm{SR}} = \left(1 - \frac{ae'e}{\hat{\gamma}'\hat{\gamma}}\right)\hat{\gamma}, \tag{4}$$

where $e = y - Z\hat{\gamma}$ and a is a constant such that $0 \leqslant a \leqslant 2(k-2)/(\nu+2)$, where $\nu = n - k$. As is shown by Stein (1956), the SR estimator dominates the OLS estimator in terms of predictive mean squared error (PMSE) when $k \geqslant 3$. Also, James and Stein (1961) showed that the PMSE of the SR estimator is minimized when $a = (k-2)/(\nu+2)$.

Although the SR estimator dominates the OLS estimator, Baranchik (1970) showed that the SR estimator is further dominated by the positive-part Stein-rule (PSR) estimator defined as

$$\hat{\gamma}_{\mathrm{PSR}} = \max\left[0,\ 1 - \frac{ae'e}{\hat{\gamma}'\hat{\gamma}}\right]\hat{\gamma}. \tag{5}$$

Ohtani and Kozumi (1996) showed that the PSR estimator for each individual regression coefficient dominates the SR estimator in terms of mean squared error (MSE).

After the findings of Stein (1956) and James and Stein (1961), lots of estimators which may have smaller PMSE than the OLS estimator have been proposed. These estimators are called shrinkage estimators because they are usually obtained by shrinking the OLS estimator towards the origin. In particular, Baranchik (1970) proposed a general family of shrinkage estimators:

$$\hat{\gamma}_B = \left[1 - \frac{r(F)}{F}\right]\hat{\gamma}, \tag{6}$$

where $F = (\hat{\gamma}'\hat{\gamma})/(e'e)$. If we define $F_1 = \nu F/k$, then F_1 is the test statistic for the null hypothesis $\mathrm{H}_0 : \gamma = 0$ against the alternative $\mathrm{H}_1 : \gamma \neq 0$. This estimator dominates the OLS estimator in terms of PMSE if

(i). $r(\cdot)$ is monotone, non-decreasing,

(ii). $0 \leqslant r(\cdot) \leqslant 2(k-2)/(\nu+2)$.

Also, Namba (2003) considered a general class of shrinkage estimators:

$$\hat{\gamma}_\phi = (1 - \phi(F))\hat{\gamma}, \tag{7}$$

and its positive-part variant:

$$\hat{\gamma}_\phi^+ = \max[0,\ (1 - \phi(F))]\hat{\gamma}, \tag{8}$$

where $\phi(F)$ is any real value function of F. In general, $\phi(\cdot)$ is positive and continuous. Hereafter, we call these estimators the shrinkage estimators and the positive-part shrinkage estimator respectively. The shrinkage and positive-part shrinkage estimators include the most of the estimators proposed so far as special cases. Namba (2003) showed that the positive-part shrinkage estimator dominates the shrinkage estimator in terms of PMSE. However, the MSE dominance when our concern is to estimate each individual regression coefficient has not been examined so far. Thus, in this paper, we derive the MSE of these estimators for each individual coefficient and show that the MSE of the positive-part shrinkage estimator is smaller than that of the shrinkage estimator.

3 MSE and dominance

Consider a following pre-test shrinkage estimators:

$$\tilde{\gamma}_\phi(c) = I(F \geqslant c)(1 - \phi(F))\hat{\gamma}, \tag{9}$$

where $I(A)$ is an indicator function such that $I(A) = 1$ if an event A occurs and $I(A) = 0$ otherwise, and c is the critical value of the pre-test. This estimator reduces to $\hat{\gamma}_\phi$ given in (7) when $c = 0$ and includes the SR, PSR and Baranchik's (1970) estimators as special cases.

Let h be a $k \times 1$ vector with known elements. If h' is the ith row vector of $S^{-1/2}$, the estimator $h'\tilde{\gamma}_\phi(c)$ is the ith element of the pre-test shrinkage estimator for β in the original model. Since the elements of h are known, we assume that $h'h = 1$ without loss of generality. Then the MSE of $h'\tilde{\gamma}_\phi(c)$ is

$$\begin{aligned}\mathrm{MSE}[h'\tilde{\gamma}_\phi(c)] &= E[(h'\tilde{\gamma}_\phi(c) - h'\gamma)^2] \\ &= E[I(F \geqslant c)\,(1 - \phi(F))^2\,(h'\hat{\gamma})^2] \\ &\quad -2h'\gamma E[I(F \geqslant c)\,(1 - \phi(F))\,h'\hat{\gamma}] + (h'\gamma)^2.\end{aligned} \tag{10}$$

If we define functions $H(p, q; c)$ and $J(p, q; c)$ as

$$H(p, q; c) = E[I(F \geqslant c)(1 - \phi(F))^p(h'\hat{\gamma})^{2q}], \tag{11}$$

$$J(p, q; c) = E[I(F \geqslant c)(1 - \phi(F))^p(h'\hat{\gamma})^{2q}(h'\hat{\gamma})], \tag{12}$$

then the MSE of $h'\tilde{\gamma}_\phi(c)$ is written as

$$\mathrm{MSE}[h'\tilde{\gamma}_\phi(c)] = H(2, 1; c) - 2h'\gamma J(1, 0; c) + (h'\gamma)^2. \tag{13}$$

As shown in Appendix, the explicit formulae of $H(p, q; c)$ and $J(p, q; c)$ are

$$H(p, q; c) = (2\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)G_{ij}(p, q; c), \tag{14}$$

$$J(p, q; c) = h'\gamma(2\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)G_{i+1,j}(p, q; c), \tag{15}$$

where

$$
\begin{aligned}
G_{ij}(p,q;c) = & \frac{\Gamma((\nu+k)/2+q+i+j)\Gamma(1/2+q+i)}{\Gamma(k/2+q+i+j)\Gamma(1/2+i)\Gamma(\nu/2)} \\
& \times \int_{c^*}^{1} \left(1-\phi\left(\frac{t}{1-t}\right)\right)^p t^{k/2+q+i+j-1}(1-t)^{\nu/2-1}dt,
\end{aligned} \tag{16}
$$

$w_i(\lambda) = \exp(-\lambda/2)(\lambda/2)^i/i!$, $\lambda_1 = (h'\gamma)^2/\sigma^2$, $\lambda_2 = \gamma'(I_k - hh')\gamma/\sigma^2$ and $c^* = c/(1+c)$.

Hereafter, we assume that both $H(2,1;c)$ and $J(1,0;c)$ are absolutely convergent. For example, both $H(2,1;c)$ and $J(1,0;c)$ are absolutely convergent if $|\phi(\cdot)| < \infty$. Also, they converge absolutely for $k \geqslant 3$ if $|\phi(\frac{t}{1-t})/t| < \infty$ on $t \in (0,1)$.

Assuming that $\phi(\cdot)$ is continuous and differentiating (16) with respect to c, we have:

$$
\begin{aligned}
\frac{\partial G_{ij}(p,q;c)}{\partial c} = & -\frac{\Gamma((\nu+k)/2+q+i+j)\Gamma(1/2+q+i)}{\Gamma(k/2+q+i+j)\Gamma(1/2+i)\Gamma(\nu/2)} \\
& \times \frac{c^{k/2+q+i+j-1}}{(1+c)^{(\nu+k)/2+q+i+j}} (1-\phi(c))^p .
\end{aligned} \tag{17}
$$

Using (17) and performing some manipulations, we have:

$$
\begin{aligned}
\frac{1}{2\sigma^2}\frac{\partial \mathrm{MSE}[h'\tilde{\gamma}_\phi(c)]}{\partial c} = & -\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{\Gamma((\nu+k)/2+i+j+1)\Gamma(1/2+i+1)}{\Gamma(k/2+i+j+1)\Gamma(1/2+i)\Gamma(\nu/2)} \\
& \times \frac{c^{k/2+i+j}}{(1+c)^{(\nu+k)/2+i+j+1}}(1-\phi(c))\left[(1-\phi(c)) - \frac{\lambda_1}{1/2+j}\right],
\end{aligned} \tag{18}
$$

when $\phi(c)$ is continuous. Thus, the MSE of $h'\tilde{\gamma}_\phi(c)$ is a decreasing function of c when $\phi(c) > 1$.

Assume that $\phi(\cdot)$ is a continuous function such that $\phi(c) > 1$ if $c \leqslant c^{**}$ and $\phi(c) \leqslant 1$ if $c > c^{**}$. Then, the MSE of $h'\tilde{\gamma}_\phi(c)$ is monotonically decreasing on $c \in [0, c^{**}]$. Also, $\tilde{\gamma}_\phi(c)$ reduces to $\hat{\gamma}_\phi$ given in (7) when $c = 0$ and to $\hat{\gamma}_\phi^+$ given in (8) when $c = c^{**}$, respectively. Thus, we obtain the following lemma.

Lemma 1 *When our concern is to estimate each individual regression coefficient, the pre-test shrinkage estimator* $h'\tilde{\gamma}_\phi(c)$ *with* $0 < c \leqslant c^{**}$ *dominates the shrinkage estimator* $h'\hat{\gamma}_\phi$ *in terms of MSE if*

- $\phi(\cdot)$ *is a continuous function such that* $\phi(c) > 1$ *if* $0 < c \leqslant c^{**}$ *for some* c^{**} *and* $\phi(c) \leqslant 1$ *otherwise.*

In particular, the pre-test shrinkage estimator with c^{**} *has the smallest MSE in the class of the estimators with* $0 \leqslant c \leqslant c^{**}$.

Since the pre-test shrinkage estimator with $c = c^{**}$ reduces to the positive-part shrinkage estimator given in (8) when the condition in Lemma 1 is satisfied, we obtain the following theorem.

Theorem 1 *When our concern is to estimate each individual regression coefficient, the positive-part shrinkage estimator* $h'\hat{\gamma}_\phi^+$ *dominates the shrinkage estimator given in* $h'\hat{\gamma}_\phi$ *in terms of MSE if*

- $\phi(\cdot)$ *is a continuous function such that* $\phi(c) \geqslant 1$ *if* $0 < c \leqslant c^{**}$ *for some* c^{**} *and* $\phi(c) < 1$ *otherwise.*

Next, we extend Theorem 1. Since $\widetilde{\gamma}_\phi(c)$ reduces to $\widehat{\gamma}_\phi$ when $c = 0$, we have:

$$\mathrm{MSE}[h'\widehat{\gamma}_\phi] = H(2,1;0) - 2h'\gamma J(1,0;0) + (h'\gamma)^2. \tag{19}$$

Substituting (14) and (15) into (19) and performing some manipulations, we have:

$$\begin{aligned}
&\frac{\mathrm{MSE}[h'\widehat{\gamma}_\phi]}{\sigma^2} \\
&= 2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{\Gamma((\nu+k/2)+i+j+1)\Gamma(1/2+i+1)}{\Gamma(k/2+i+j+1)\Gamma(1/2+i)\Gamma(\nu/2)} \\
&\quad \times \int_0^1 t^{k/2+i+j}(1-t)^{\nu/2-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)\left[\left(1-\phi\left(\frac{t}{1-t}\right)\right)-\frac{\lambda_1}{1/2+i}\right]dt + \lambda_1 \\
&= 2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{\Gamma((\nu+k/2)+i+j+1)\Gamma(1/2+i+1)}{\Gamma(k/2+i+j+1)\Gamma(1/2+i)\Gamma(\nu/2)} \\
&\quad \times \left(\int_{R_1}+\int_{R_2}\right) t^{k/2+i+j}(1-t)^{\nu/2-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)\left[\left(1-\phi\left(\frac{t}{1-t}\right)\right)-\frac{\lambda_1}{1/2+i}\right]dt \\
&\quad +\lambda_1,
\end{aligned} \tag{20}$$

where R_1 is the region such that $\{t|1-\phi(\frac{t}{1-t}) > 0 \text{ and } 0 \leqslant t \leqslant 1\}$, R_2 is the region such that $\{t|1-\phi(\frac{t}{1-t}) \leqslant 0$ and $0 \leqslant t \leqslant 1\}$, and $(\int_{R_1}+\int_{R_2})f(t)dt$ denotes $\int_{R_1} f(t)dt + \int_{R_2} f(t)dt$.

Also, replacing $(1-\phi(\cdot))$ in (20) by $\max[0, 1-\phi(\cdot)]$, we obtain the MSE of the positive-part shrinkage estimator:

$$
\begin{aligned}
&\frac{\text{MSE}[h'\widehat{\gamma}_\phi^+]}{\sigma^2} \\
&\quad = 2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{\Gamma((\nu+k/2)+i+j+1)\Gamma(1/2+i+1)}{\Gamma(k/2+i+j+1)\Gamma(1/2+i)\Gamma(\nu/2)} \\
&\qquad \times \int_{R_1} t^{k/2+i+j}(1-t)^{\nu/2-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)\left[\left(1-\phi\left(\frac{t}{1-t}\right)\right)-\frac{\lambda_1}{1/2+i}\right]dt \\
&\qquad +\lambda_1. \qquad\qquad (21)
\end{aligned}
$$

Subtracting (21) from (20), we have:

$$
\begin{aligned}
&\frac{\text{MSE}[h'\widehat{\gamma}_\phi]-\text{MSE}[h'\widehat{\gamma}_\phi^+]}{\sigma^2} \\
&\quad = 2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{\Gamma((\nu+k/2)+i+j+1)\Gamma(1/2+i+1)}{\Gamma(k/2+i+j+1)\Gamma(1/2+i)\Gamma(\nu/2)} \\
&\qquad \times \int_{R_2} t^{k/2+i+j}(1-t)^{\nu/2-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)\left[\left(1-\phi\left(\frac{t}{1-t}\right)\right)-\frac{\lambda_1}{1/2+i}\right]dt \geqslant 0 \qquad (22)
\end{aligned}
$$

because R_2 is the region such that $\{t|1-\phi(\frac{t}{1-t})\leqslant 0 \text{ and } 0\leqslant t\leqslant 1\}$. Thus, we obtain the following theorem.

Theorem 2 *When our concern is to estimate each individual regression coefficient, the shrinkage estimator is dominated by its positive-part variant in terms of MSE, if*

- $1-\phi(F)<0$ *for some region of* $F\in[0,\infty)$.

$1-\phi(F)\nless 0$ for any region of $F\in[0,\infty)$ implies that the shrinkage estimator $\widehat{\gamma}_\phi$ coincides with the positive-part shrinkage estimator $\widehat{\gamma}_\phi^+$. Therefore, we need not consider a positive-part estimator. Thus, Theorem 2 indicates that the MSE of the shrinkage estimator can be improved whenever there exists a positive part variant.

Appendix

In this appendix, we derive the formulae for $H(p,q;c)$ and $J(p,q;c)$. First, we derive the formula for $H(p,q;c)$. Let $u_1=(h'\widehat{\gamma})^2/\sigma^2$, $u_2=\widehat{\gamma}'(I_k-hh')\widehat{\gamma}/\sigma^2$ and $u_3=e'e/\sigma^2$. Then, $u_1\sim\chi_1'^2(\lambda_1)$ and $u_2\sim\chi_{k-1}'^2(\lambda_2)$, where $\chi_f'^2(\lambda)$ is the noncentral chi-square distribution with f degrees of freedom and noncentrality parameter λ, $\lambda_1=(h'\gamma)^2/\sigma^2$ and $\lambda_2=\gamma'(I_k-hh')\gamma/\sigma^2$. Further, u_3 is distributed as the chi-square distribution with $\nu=n-k$ degrees of freedom, and u_1, u_2 and u_3 are mutually independent.

Using u_1, u_2 and u_3, $H(p,q;c)$ is expressed as

$$
\begin{aligned}
&H(p,q;c) \\
&\quad = (\sigma^2)^q\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}K_{ij}\iiint_R\left(1-\phi\left(\frac{u_1+u_2}{u_3}\right)\right)^q \\
&\qquad \times u_1^{1/2+q+i-1}u_2^{(k-1)/2+j-1}u_3^{\nu/2-1}\exp[-(u_1+u_2+u_3)/2]du_1\,du_2\,du_3, \qquad (23)
\end{aligned}
$$

where

$$K_{ij} = \frac{w_i(\lambda_1)w_j(\lambda_2)}{2^{(\nu+k)/2+i+j}\Gamma(1/2+i)\Gamma((k-1)/2+j)\Gamma(\nu/2)}, \tag{24}$$

$w_i(\lambda) = \exp(-\lambda/2)(\lambda/2)^i/i!$, and R is the region such that $(u_1+u_2)/u_3 \geqslant c$.

Making use of the change of variables, $v_1 = (u_1+u_2)/u_3$, $v_2 = u_1u_3/(u_1+u_2)$ and $v_3 = u_3$, (23) reduces to

$$\begin{aligned}(\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} K_{ij} \int_0^{\infty}\int_0^{v_3}\int_c^{\infty} (1-\phi(v_1))^p\, v_1^{k/2+q+i+j-1} v_2^{1/2+q+i-1} v_3^{\nu/2} (v_3-v_2)^{(k-1)/2+j-1} \\ \times \exp[-v_3(v_1+1)/2] dv_1\, dv_2\, dv_3.\end{aligned} \tag{25}$$

Again, making use of the change of variable, $z_1 = v_2/v_3$, (25) reduces to

$$\begin{aligned}(\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} K_{ij} \frac{\Gamma(1/2+q+i)\Gamma((k-1)/2+j)}{\Gamma(k/2+q+i+j)} \\ \times \int_0^{\infty}\int_c^{\infty} (1-\phi(v_1))^p\, v_1^{k/2+q+i+j-1} v_3^{(\nu+k)/2+q+i+j-1} \exp[-v_3(v_1+1)/2] dv_1\, dv_3.\end{aligned} \tag{26}$$

Further, making use of the change of variable, $z_2 = v_3(v_1+1)/2$, (26) reduces to

$$\begin{aligned}(\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} K_{ij} 2^{(\nu+k)/2+q+i+j} \frac{\Gamma(1/2+q+i)\Gamma((k-1)/2+j)\Gamma((\nu+k)/2+q+i+j)}{\Gamma(k/2+q+i+j)} \\ \times \int_c^{\infty} (1-\phi(v_1))^p \frac{v_1^{k/2+q+i+j-1}}{(1+v_1)^{(\nu+k)/2+q+i+j}} dv_1.\end{aligned} \tag{27}$$

Finally, making use of the change of variable, $t = v_1/(1+v_1)$. we obtain (14) in the text.

Next, we derive the formula for $J(p,q;c)$. Differentiating $H(p,q;c)$ given in (14) with respect to γ, we have

$$\begin{aligned}\frac{\partial H(p,q;c)}{\partial \gamma} &= (2\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\left[\frac{\partial w_i(\lambda_1)}{\partial\gamma} w_j(\lambda_2) + w_i(\lambda_1)\frac{\partial w_j(\lambda_2)}{\partial\gamma}\right] G_{ij}(p,q;c) \\ &= -\frac{hh'\gamma}{\sigma^2}(2\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)G_{ij}(p,q;c) \\ &\quad +\frac{hh'\gamma}{\sigma^2}(2\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)G_{i+1,j}(p,q;c) \\ &\quad -\frac{(I_k-hh')\gamma}{\sigma^2}(2\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)G_{ij}(p,q;c) \\ &\quad +\frac{(I_k-hh')\gamma}{\sigma^2}(2\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)G_{i,j+1}(p,q;c),\end{aligned} \tag{28}$$

where we define $w_{-1}(\lambda_1) = w_{-1}(\lambda_2) = 0$. Since $h'h = 1$, we obtain

$$h'\frac{\partial H(p,q;c)}{\partial\gamma} = -\frac{h'\gamma}{\sigma^2}H(p,q;c) + \frac{h'\gamma}{\sigma^2}(2\sigma^2)^q \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)G_{i+1,j}(p,q;c). \tag{29}$$

Expressing $H(p,q;c)$ by $\widehat{\gamma}$ and $e'e$, we have

$$H(p,q;c) = \iint_{F \geqslant c} (1-\phi(F))^p \, (h'\widehat{\gamma})^{2q} f_N(\widehat{\gamma}) f_e(e'e) d\widehat{\gamma}\, d(e'e), \tag{30}$$

where $F = (\widehat{\gamma}'\widehat{\gamma})/(e'e)$, $f(e'e)$ is the density function of $e'e$, and

$$f_N(\widehat{\gamma}) = \frac{1}{(2\pi)^{k/2}\sigma^k} \exp\left[-\frac{(\widehat{\gamma}-\gamma)'(\widehat{\gamma}-\gamma)}{2\sigma^2}\right], \tag{31}$$

is the density function of $\widehat{\gamma}$.

Differentiating $H(p,q;c)$ given in (30) with respect to γ, and multiplying h' from the left, we obtain

$$h'\frac{\partial H(p,q;c)}{\partial \gamma} = -\frac{h'\gamma}{\sigma^2} H(p,q;c) + \frac{1}{\sigma^2} E\left[I(F \geqslant c)\,(1-\phi(F))^p \, (h'\widehat{\gamma})^{2q} h'\widehat{\gamma}\right]. \tag{32}$$

Equating (29) and (32), we obtain (15) in the text.

References

Baranchik, A. J. (1970) A family of minimax estimators of the mean of a multivariate normal distribution. *The Annals of Mathematical Statistics* 41, 642–645.

Efron, B. & C. Morris (1973) Combining possibly related estimation problems (with discussion). *Journal of the Royal Statistical Society, Series B, Methodological* 35, 379–421.

James, W. & C. Stein (1961) Estimation with quadratic loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, , Volume 1, pp. 361–379. University of California Press, Berkeley.

Judge, G. G. & T. A. Yancey (1986) *Improved Methods of Inference in Econometrics*, North-Holland/Elsevier.

Namba, A. (2003) Dominance of the positive-part shrinkage estimator in a regression model when relevant regressors are omitted. *Statistics & Probability Letters* 63, 375–385.

Nickerson, D. M. (1988) Dominance of the positive-part version of the James-Stein estimator. *Statistics & Probability Letters* 7, 97–103.

Ohtani, K. & H. Kozumi (1996) The exact general formulae for the moments and the mse dominance of the stein-rule and positive-part stein-rule estimators. *Journal of Econometrics* 74, 273–87.

Rao, C. R. & N. Shinozaki (1978) Precision of individual estimators in simultaneous estimation of parameters. *Biometrika* 65, 23–30.

Stein, C. (1956) Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, , Volume 1, pp. 197–206. University of California Press, Berkeley.