PMSE dominance of the positive-part shrinkage estimator in a regression model when relevant regressors are omitted

Namba, Akio

(Citation)
Statistics & Probability Letters,63(4):375-385

(Issue Date)
2003-07

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90000341

PMSE Dominance of the Positive-Part Shrinkage Estimator in a Regression Model when Relevant Regressors are Omitted

Akio Namba*

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

Abstract

In this paper we consider a regression model with omitted relevant regressors and a general family of shrinkage estimators of regression coefficients. We derive the formula for the predictive mean squared error (PMSE) of the estimators. It is shown analytically that the positive-part shrinkage estimator dominates the ordinary shrinkage estimator even when there are omitted relevant regressors. Also, as an example, our result is applied to the double k-class estimator.

Key Words: Positive-part estimator, predictive mean squared error, Dominance

AMS classification: 62J07, 62C99.

*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 problem of estimating the mean vector of a multivariate normal distribution, Stein (1956) and James and Stein (1961) showed that the maximum likelihood (ML) estimator is dominated by the Stein-rule (SR) estimator in terms of mean squared error (MSE). The SR estimator can be applied in a linear regression model. In the context of linear regression, the SR estimator dominates the ordinary least squares (OLS) estimator in terms of predictive mean squared error (PMSE) if the model is specified correctly. Since the findings of Stein (1956), lots of estimators have been proposed and their sampling properties have been examined. In particular, Baranchik (1964) showed that the SR estimator is further dominated by the positive-part Stein-rule (PSR) estimator. Also, Baranchik (1970) proposed a family of estimators which dominate the OLS estimator in terms of PMSE. These estimators are called the shrinkage estimators since they are obtained by shrinking the OLS estimator toward the origin. In general, it is expected that the PMSE of the shrinkage estimator can be improved by using its positive-part variant if the model is specified correctly. Nickerson (1988) considered a class of estimators and showed that the estimators are dominated by their positive-part variants.

The shrinkage estimators are biased even when there are no omitted relevant regressors. However, in most of practical situations, it is hard to determine which regressors should be included in the model. Thus, the researcher may exclude relevant regressors mistakenly. If there are omitted relevant regressors, even the OLS estimator is not unbiased. In such situations, there may be a strong incentive to use the shrinkage estimators from the viewpoint of PMSE. However, there are few researches on the sampling properties of shrinkage estimators when relevant regressors are omitted in the specified model. Some exceptions are Mittelhammer (1984), Ohtani (1993, 1998) and Namba (2000). Mittelhammer (1984) showed that the SR estimator no longer dominates the OLS estimator when relevant regressors are omitted. Ohtani (1993) derived the formulae for the PMSE's of the SR and the PSR estimator of the misspecified model. In particular, Namba (2000) showed exactly that the PSR estimator dominates the SR estimator even when the relevant regressors are omitted. Thus, it may be expected that the shrinkage estimators are dominated by their positive-part variants when there are omitted relevant regressors.

Thus, in this paper, we consider a general class of shrinkage estimators of regression coefficients when relevant regressors are omitted. The plan of the paper is as follows. In section 2 the model and the estimators are presented. In section 3 we derive the PMSE of the general class of shrinkage estimators. We propose the pre-test shrinkage estimator which dominates the shrinkage estimator. Moreover, it is shown analytically that the shrinkage estimators are dominated by their positive-part variants whether there are omitted relevant regressors or not. In section 4, as an example, our result is applied to the double k-class estimator proposed by Ullah and Ullah (1978).

2 Model and the Estimators

Consider a linear regression model,

$$y = X_1\beta_1 + X_2\beta_2 + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I_n), \tag{1}$$

where y is an $n \times 1$ vector of observations on a dependent variable, X_1 and X_2 are $n \times k_1$ and $n \times k_2$ matrices of observations on nonstochastic independent variables, β_1 and β_2 are $k_1 \times 1$ and $k_2 \times 1$ vectors of regression coefficients, and ϵ is an $n \times 1$ vector of normal error terms. We assume that X_1 and $[X_1, X_2]$ are of full column rank.

Suppose that the matrix of regressors X_2 is omitted mistakenly and the model is specified as

$$y = X_1\beta_1 + \eta, \quad \eta = X_2\beta_2 + \epsilon. \tag{2}$$

Then, based on the misspecified model, the ordinary least squares (OLS) estimator of β_1 is

$$b_1 = S_{11}^{-1} X_1' y, \tag{3}$$

where $S_{11} = X_1' X_1$.

Also, based on the misspecified model, the Stein-rule estimator proposed by Stein (1956) and James and Stein (1961) is

$$b_{S1} = \left(1 - \frac{a e_1' e_1}{b_1' S_{11} b_1}\right) b_1, \tag{4}$$

where $e_1 = y - X_1 b_1$, $\nu_1 = n - k_1$ and a is a constant such that $0 \leq a \leq 2(k_1 - 2)/(\nu_1 + 2)$. Here, we define the predictive mean squared error (PMSE) as

$$PMSE[\bar{\beta}_1] = E[(X_1\bar{\beta}_1 - X\beta)'(X_1\bar{\beta}_1 - X\beta)], \tag{5}$$

where $X = [X_1, X_2]$, $\beta' = [\beta_1', \beta_2']$ and $\bar{\beta}_1$ is any estimator of β_1. The meaning and some discussions of the PMSE are given in Mittelhammer (1984) and Ohtani (1993). When there are no omitted relevant regressors (i.e., $X = X_1$), the SR estimator dominates the OLS estimator in terms of PMSE for $k_1 \geq 3$. Also, the PMSE of the SR estimator is minimized when $a = (k_1 - 2)/(\nu_1 + 2)$ if the model is specified correctly.

Baranchik (1970) proposed a family of estimators:

$$b_{B1} = \left[1 - \frac{r(F)}{F}\right] b_1, \tag{6}$$

where $F = (b_1' S_{11} b_1)/(e_1' e_1)$. If we define $F_1 = \nu_1 F/k_1$, F_1 is the test statistic for the null hypothesis $H_0 : \beta_1 = 0$ against the alternative $H_1 : \beta_1 \neq 0$ based on the misspecified model. This estimator dominates the OLS estimator if

(i). $r(\cdot)$ is monotone, nondecreasing,

(ii). $0 \leq r(\cdot) \leq 2(k_1 - 2)/(\nu_1 + 2)$, and

(iii). relevant regressors are not omitted.

Also, as is shown in Baranchik (1964), when there are no omitted regressors, the SR estimator is further dominated by the positive-part Stein-rule estimator (PSR) defined as

$$b_{PS1} = \max\left[0, 1 - \frac{a e_1' e_1}{b_1' S_{11} b_1}\right] b_1 \tag{7}$$

Moreover, Namba (2000) showed that the PSR estimator dominates the SR estimator in terms of PMSE even when there are omitted relevant regressors. Thus, it is expected that there may be some conditions such that the shrinkage estimators are dominated by their positive-part variant even when relevant regressors are omitted.

Thus, in this paper, we consider the following general class of shrinkage estimators:

$$\widehat{\beta}_1 = (1 - \phi(F)) b_1, \tag{8}$$

and its positive-part variant:

$$\widehat{\beta}_1^{+} = \max[0, 1 - \phi(F)] b_1, \tag{9}$$

where $\phi(F)$ is any real value function of F. We assume that the second moments of (8) and (9) exist since our analysis is based on PMSE. In general, $\phi(\cdot)$ is positive and continuous. Hereafter, we call these estimators the shrinkage estimator and the positive-part shrinkage estimator respectively. Most of the estimators proposed so far are included in (8).

In the next section it is shown that the positive-part shrinkage estimator dominates the shrinkage estimator in terms of PMSE even when there are omitted relevant regressors.

3 PMSE of the estimator

In this section, we consider the following pre-test shrinkage estimators:

$$\tilde{\beta}_1 = I(F \geq c)(1 - \phi(F))b_1, \tag{10}$$

where $I(A)$ is an indicator function such that $I(A) = 1$ if an event A occurs and $I(A) = 0$ otherwise. The estimator given in (10) includes the SR, PSR and Baranchik's (1970) estimators as special cases. Also, (10) reduces to (8) when $c = 0$.

The PMSE of $\tilde{\beta}_1$ is defined as

$$\begin{aligned} PMSE[\tilde{\beta}_1] &= E[(X_1\tilde{\beta}_1 - X\beta)'(X_1\tilde{\beta}_1 - X\beta)] \\ &= E[\tilde{\beta}_1' S_{11} \tilde{\beta}_1] - 2E[\beta' X' X_1 \tilde{\beta}_1] + \beta' S \beta, \end{aligned} \tag{11}$$

where $S = X'X$.

Denoting $u_1 = b_1' S_{11} b_1/\sigma^2$ and $u_2 = e_1' e_1/\sigma^2$, (11) reduces to

$$\begin{aligned} PMSE[\tilde{\beta}_1] = E&\left[I\left(\frac{u_1}{u_2} \geq c\right)\left(1 - \phi\left(\frac{u_1}{u_2}\right)\right)^2 u_1 \sigma^2 \right] \\ &-2E\left[I\left(\frac{u_1}{u_2} \geq c\right)\left(1 - \phi\left(\frac{u_1}{u_2}\right)\right) \beta' X' X_1 b_1 \right] + \beta' S \beta. \end{aligned} \tag{12}$$

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

$$H(p, q; c) = E\left[I\left(\frac{u_1}{u_2} \geq c\right)\left(1 - \phi\left(\frac{u_1}{u_2}\right)\right)^p u_1^q \right], \tag{13}$$

$$J(p, q; c) = E\left[I\left(\frac{u_1}{u_2} \geq c\right)\left(1 - \phi\left(\frac{u_1}{u_2}\right)\right)^p u_1^q (\beta' X' X_1 b_1/\sigma^2) \right], \tag{14}$$

we obtain

$$PMSE[\tilde{\beta}_1]/\sigma^2 = H(2, 1; c) - 2J(1, 0; c) + \lambda_1 + \lambda_2, \tag{15}$$

where $\lambda_1 = \beta' X' X_1 S_{11}^{-1} X_1' X \beta/\sigma^2$, $\lambda_2 = \beta' X' M_1 X \beta/\sigma^2$, $M_1 = I_n - X_1 S_{11}^{-1} X_1'$ and $\lambda_1 + \lambda_2 = \beta' S \beta/\sigma^2$.

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

$$H(p, q; c) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_i(\lambda_1) w_j(\lambda_2) G_{ij}(p, q; c), \tag{16}$$

$$J(p, q; c) = \lambda_1 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_i(\lambda_1) w_j(\lambda_2) G_{i+1,j}(p, q; c), \tag{17}$$

where $w_i(\lambda) = \exp(-\lambda/2)(\lambda/2)^i/i!$ and

$$G_{ij}(p,q;c) = \frac{2^q\Gamma((k_1+\nu_1)/2+q+i+j)}{\Gamma(k_1/2+i)\Gamma(\nu_1/2+j)} \times \int_{c^*}^1 t^{k_1/2+q+i-1}(1-t)^{\nu_1/2+j-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)^p dt, \quad (18)$$

where $c^* = c/(1+c)$.

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 $k_1 \geq 3$ if $|\phi(\frac{t}{1-t})/t| < \infty$ on $t \in (0,1)$.

Substituting (16) and (17) in (15), we obtain the explicit formula for the PMSE of $\tilde{\beta}_1$:

$$\begin{aligned} PMSE[\tilde{\beta}_1]/\sigma^2 = &\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{2\Gamma((k_1+\nu_1)/2+i+j+1)}{\Gamma(k_1/2+i)\Gamma(\nu_1/2+j)} \\ &\times \int_{c^*}^1 t^{k_1/2+i}(1-t)^{\nu_1/2+j-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)^2 dt \\ &-2\lambda_1\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{\Gamma((k_1+\nu_1)/2+i+j+1)}{\Gamma(k_1/2+i+1)\Gamma(\nu_1/2+j)} \\ &\times \int_{c^*}^1 t^{k_1/2+i}(1-t)^{\nu_1/2+j-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right) dt + \lambda_1 + \lambda_2. \end{aligned} \quad (19)$$

Supposing that $\phi(\cdot)$ is continuous and differentiating (19) with respect to c and performing some manipulations we obtain

$$\begin{aligned} \frac{\partial PMSE[\tilde{\beta}_1]/\sigma^2}{\partial c} = &-\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{2\Gamma((k_1+\nu_1)/2+i+j+1)}{\Gamma(k_1/2+i)\Gamma(\nu_1/2+j)}\frac{c^{k_1/2+i}}{(1+c)^{(k_1+\nu_1)/2+i+j}} \\ &\times(1-\phi(c))\left[(1-\phi(c)) - \frac{\lambda_1}{k_1/2+i}\right]. \end{aligned} \quad (20)$$

From (20), the PMSE of $\tilde{\beta}_1$ is monotone decreasing function of c when $\phi(c) \geq 1$.

Assume that $\phi(\cdot)$ is a continuous function such that $\phi(c) \geq 1$ if $c \leq c^{**}$ and $\phi(c) < 1$ if $c > c^{**}$. Then, the PMSE of $\tilde{\beta}_1$ is monotonically decreasing on $c \in [0, c^{**}]$. Also, $\tilde{\beta}_1$ reduces to the shrinkage estimator given in (8) when $c = 0$ and to the positive-part shrinkage estimator given in (9) when $c = c^{**}$. Thus, we obtain the following lemma.

Lemma 1 *The pre-test shrinkage estimator with* $0 < c \leq c^{**}$ *given in (10) dominates the shrinkage estimator given in (8) in terms of PMSE whether there are omitted relevant regressors or not if*

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

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

Since the pre-test shrinkage estimator with $c = c^{**}$ reduces to the positive-part shrinkage estimator given in (9) when the condition in Lemma 1 is satisfied, we obtain the following theorem.

Theorem 1 *The positive-part shrinkage estimator given in (9) dominates the shrinkage estimator given in (8) in terms of PMSE whether there are omitted relevant regressors or not if*

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

Next, we extend Theorem 1. From (15), the PMSE of the shrinkage estimator given in (8) is

$$
\begin{aligned}
&PMSE[\widehat{\beta}_1]/\sigma^2 \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{2\Gamma((k_1+\nu_1)/2+i+j+1)}{\Gamma(k_1/2+i)\Gamma(\nu_1/2+j)} \\
&\quad \times \int_0^1 t^{k_1/2+i}(1-t)^{\nu_1/2+j-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)\left[\left(1-\phi\left(\frac{t}{1-t}\right)\right)-\frac{\lambda_1}{k_1/2+i}\right]dt \\
&\qquad +\lambda_1+\lambda_2 \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{2\Gamma((k_1+\nu_1)/2+i+j+1)}{\Gamma(k_1/2+i)\Gamma(\nu_1/2+j)} \\
&\quad \times \left(\int_{R_1}+\int_{R_2}\right) t^{k_1/2+i}(1-t)^{\nu_1/2+j-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)\left[\left(1-\phi\left(\frac{t}{1-t}\right)\right)-\frac{\lambda_1}{k_1/2+i}\right]dt \\
&\qquad +\lambda_1+\lambda_2, \qquad (21)
\end{aligned}
$$

where R_1 is the region such that $\{t|1-\phi(\frac{t}{1-t})>0 \text{ and } 0\le t\le 1\}$, R_2 is the region such that $\{t|1-\phi(\frac{t}{1-t})\le 0 \text{ and } 0\le t\le 1\}$, 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 $1-\phi(\cdot)$ by $\max[0,1-\phi(\cdot)]$, we obtain the PMSE of the positive-part shrinkage estimator given in (9):

$$
\begin{aligned}
&PMSE[\widehat{\beta}_1^+]/\sigma^2 \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{2\Gamma((k_1+\nu_1)/2+i+j+1)}{\Gamma(k_1/2+i)\Gamma(\nu_1/2+j)} \\
&\quad \times \int_{R_1} t^{k_1/2+i}(1-t)^{\nu_1/2+j-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)\left[\left(1-\phi\left(\frac{t}{1-t}\right)\right)-\frac{\lambda_1}{k_1/2+i}\right]dt \\
&\qquad +\lambda_1+\lambda_2. \qquad (22)
\end{aligned}
$$

Subtracting (22) from (21), we obtain

$$
\frac{PMSE(\widehat{\beta}_1)-PMSE(\widehat{\beta}_1^+)}{\sigma^2}
$$

$$
\begin{aligned}
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1)w_j(\lambda_2)\frac{2\Gamma((k_1+\nu_1)/2+i+j+1)}{\Gamma(k_1/2+i)\Gamma(\nu_1/2+j)} \\
&\quad \times \int_{R_2} t^{k_1/2+i}(1-t)^{\nu_1/2+j-1}\left(1-\phi\left(\frac{t}{1-t}\right)\right)\left[\left(1-\phi\left(\frac{t}{1-t}\right)\right)-\frac{\lambda_1}{k_1/2+i}\right]dt \\
&\geq 0, \qquad (23)
\end{aligned}
$$

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

Theorem 2 *The positive-part shrinkage estimator given in (9) dominates the shrinkage estimator given in (8) in terms of PMSE whether there are omitted relevant regressors or not if*

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

Theorem 2 implies that using the positive-part shrinkage estimator is preferable to using the shrinkage estimator in terms of PMSE as long as $1-\phi(\cdot)$ can be negative.

In the next section we apply these theorems to the double k-class estimator proposed by Ullah and Ullah (1978).

4 Example

In this section, we consider the double k-class (KK) estimator proposed by Ullah and Ullah (1978):

$$
\begin{aligned}
b_{KK1} &= \left(1-\frac{k_{1d}e_1'e_1}{y'y-k_{2d}e_1'e_1}\right)b_1 \\
&= \left(1-\frac{k_{1d}}{F+(1-k_{2d})}\right)b_1 \\
&= (1-\phi_{KK}(F))\,b_1 \qquad (24)
\end{aligned}
$$

where k_{1d} and k_{2d} are constants chosen appropriately and $\phi_{KK}(F)=k_{1d}/(F+1-k_{2d})$.

When relevant regressors are not omitted, the double k-class estimator dominates the OLS estimator when $0\leq k_{1d}\leq 2(k_1-2)/(\nu_1+2)$ and $1-k_{2d}>0$ since Baranchik's (1970) conditions are satisfied. Also, when $k_{2d}<1$ and $k_{1d}>1-k_{2d}$, $\phi_{KK}(F)\geq 1$ for $F\leq k_{1d}+k_{2d}-1$ and $\phi(F)<1$ otherwise. Thus, when $0\leq k_{1d}\leq 2(k_{1d}-2)/(\nu_1+2)$ and $k_{2d}>1-k_{1d}$, the positive-part double k-class (PKK) estimator defined as

$$
b_{PKK1} = \max\left[0, 1-\frac{k_{1d}e_1'e_1}{y'y-k_{2d}e_1'e_1}\right]b_1 \qquad (25)
$$

dominates the KK estimator given in (24) in terms of PMSE whether there are relevant regressors or not. Also, the PKK estimator dominates the OLS estimator if the model is specified correctly.

To examine the PMSE performance of the KK estimator and the PKK estimator, we executed the numerical evaluation. As is shown in Ohtani (1993), the noncentrality parameters are expressed as $\lambda_1 = F_\beta R_1^2$ and $\lambda_2 = F_\beta(1 - R_1^2)$, where $F_\beta = \beta' S\beta/\sigma^2$ and $R_1^2 = \beta' X' X_1 S_{11}^{-1} X_1' X\beta/\beta' S\beta$. F_β is the noncentrality parameter which appeared in the test for the null hypothesis that all the regression coefficients are zeros in the correctly specified model. Also, R_1^2 is interpreted as the coefficient of determination in regression of $X\beta$ on X_1. Thus, if R_1^2 is close to unity, the magnitude of model misspecification is regarded as small, and vice versa.

The parameter values used in the numerical evaluations were $k_1 = 3$, 5, 8, $n = 20$, 30, 40, $R_1^2 = 0.1$, 0.3, 0.5, 0.7, 0.9, 1.0 and $F_\beta =$ various values. Also, we use $k_{1d} = (k_1 - 2)/(\nu_1 + 2)$ and $k_{2d} = 1 - k_{1d}/2$. In this case, the condition $0 \le k_{1d} \le 2(k_{1d} - 2)/(\nu_1 + 2)$ and $k_{2d} > 1 - k_{1d}$ is satisfied. To compare the PMSE's of the estimators, we evaluated the values of relative PMSE defined as $PMSE[\bar{\beta}_1]/PMSE[b_1]$, where $\bar{\beta}_1$ is any estimator of β_1. Thus, the estimator $\bar{\beta}_1$ has smaller PMSE than the OLS estimator when the value of relative PMSE is smaller than unity. The numerical evaluations were executed on a personal computer using the FORTRAN code. In evaluating the integral in $G_{ij}(p, q; c)$ given in (18), we used Simpson's rule with 200 equal subdivisions. The double infinite series in $H(p, q; c)$ and $J(p, q; c)$ were judged to converge when the increment of series got smaller than 10^{-12}. Since the results for the cases of $k_1 = 5$, and $n = 20$ are qualitatively typical, we do not show the results for the other cases.

The result for $k_1 = 5$ and $n = 20$ is shown in Table 1. As is shown in Theorem 1, the PKK estimator dominates the KK estimator. Though the PMSE of the PKK estimator is slightly larger than unity for some values of λ_1 and λ_2, it is much smaller than unity over a wide region of parameter space. Thus, there is a incentive to use the PKK estimator in terms of PMSE.

5 Appendix

In this Appendix, we derive the formulae for $H(p, q; c)$ and $J(p, q; c)$. First, we derive the formula for $H(p, q; c)$. The OLS estimator b_1 is distributed as $N(\beta_1 + S_{11}^{-1} S_{12}\beta_2, \sigma^2 S_{11}^{-1}) = N(S_{11}^{-1} X_1' X\beta, \sigma^2 S_{11}^{-1})$, where $X = [X_1, X_2]$, $\beta' = [\beta_1', \beta_2']$ and $S_{12} = X_1' X_2$. Thus, $u_1 = b_1' S_{11} b_1/\sigma^2$ is distributed as the non-central chi-square distribution with k_1 degrees of freedom and noncentrality parameter $\lambda_1 = \beta' X' X_1 S_{11}^{-1} X_1' X\beta/\sigma^2$. Also, the residual vector e_1 is distributed as $N(M_1 X\beta, \sigma^2 M_1)$, where $M_1 = (I_n - X_1 S_{11}^{-1} X_1')$. Thus, $u_2 = e_1' e_1/\sigma^2$ is distributed as the non-central chi-square distribution with ν_1 degrees of freedom and noncentrality parameter $\lambda_2 = \beta' X' M_1 X\beta/\sigma^2$. Furthermore, u_1 and u_2 are mutually independent.

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

$$H(p,q;c) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} K_{ij} \iint_R u_1^{k_1/2+q+i-1} u_2^{\nu_1/2+j-1} \left(1-\phi\left(\frac{u_1}{u_2}\right)\right)^p \times \exp[-(u_1+u_2)/2] du_2\, du_1, \tag{26}$$

where

$$K_{ij} = \frac{w_i(\lambda_1)w_j(\lambda_2)}{2^{(k_1+\nu_1)/2+i+j}\Gamma(k_1/2+i)\Gamma(\nu_1/2+j)}, \tag{27}$$

$w_i(\lambda) = \exp(-\lambda/2)(\lambda/2)^i/i!$, and R is the region such that $\{u_1, u_2 | u_1/u_2 \geq c\}$.

Making use of the change of variables, $v_1 = u_1/u_2$ and $v_2 = u_2$, (26) reduces to:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} K_{ij} \int_c^{\infty}\int_0^{\infty} v_1^{k_1/2+q+i-1} v_2^{(k_1+\nu_1)/2+q+i+j-1}(1-\phi(v_1))^p \times \exp[-v_2(1+v_1)/2] dv_2\, dv_1. \tag{28}$$

Again, making use of the change of variable, $z = (1+v_1)v_2/2$, (28) reduces to:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} K_{ij} 2^{(k_1+\nu_1)/2+q+i+j}\Gamma((k_1+\nu_1)/2+q+i+j) \times \int_c^{\infty} \frac{v_1^{k_1/2+q+i-1}}{(1+v_1)^{(k_1+\nu_1)/2+q+i+j}}(1-\phi(v_1))^p dv_1. \tag{29}$$

Finally, making use of the change of variable, $t = v_1/(1+v_1)$, and performing some manipulations, we obtain (16) in the text.

Next, we derive the formula for $J(p,q;c)$. Noting that $\partial\lambda_1/\partial\beta_1 = 2(S_{11}\beta_1 + S_{12}\beta_2)/\sigma^2$ and $\partial\lambda_2/\partial\beta_1 = 0$, and differentiating $H(p,q;c)$ given in (16) with respect to β_1, we obtain:

$$\begin{aligned}
\frac{\partial H(p,q;c)}{\partial\beta_1}
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\left[\frac{\partial w_i(\lambda_1)}{\partial\beta_1}\right] w_j(\lambda_2) G_{ij}(p,q;c) \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\left[\frac{S_{11}\beta_1+S_{12}\beta_2}{\sigma^2}\right][-w_i(\lambda_1)+w_{i-1}(\lambda_1)] w_j(\lambda_2) G_{ij}(p,q;c) \\
&= -\left[\frac{S_{11}\beta_1+S_{12}\beta_2}{\sigma^2}\right] H(p,q;c) \\
&\quad + \left[\frac{S_{11}\beta_1+S_{12}\beta_2}{\sigma^2}\right]\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1) w_j(\lambda_2) G_{i+1,j}(p,q;c)
\end{aligned} \tag{30}$$

Since $u_1 = b_1' X_{11} b_1/\sigma^2$ and $b_1 \sim N(S_{11}^{-1}X_1'X\beta, \sigma^2 S_{11}^{-1})$, $H(p,q;c)$ can be expressed as

$$H(p,q;c) = \iint_R \left(1-\phi\left(\frac{u_1}{u_2}\right)\right)^p u_1^q f_N(b_1) f_2(u_2) du_2\, db_1, \tag{31}$$

where $f_2(u_2)$ is the density function of u_2 and

$$f_N(b_1) = \frac{1}{(2\pi)^{k_1/2}|\sigma^2 S_{11}^{-1}|^{1/2}} \exp\left[-\frac{(b_1 - S_{11}^{-1}X_1'X\beta)'S_{11}(b_1 - S_{11}^{-1}X_1'X\beta)}{2\sigma^2}\right]. \tag{32}$$

Differentiating (31) with respect to β_1, we obtain:

$$\begin{aligned}\frac{\partial H(p,q;c)}{\partial \beta_1} &= \iint_R \left(1 - \phi\left(\frac{u_1}{u_2}\right)\right)^p u_1^q f_N(b_1) f_2(u_2) \left[\frac{S_{11}b_1 - S_{11}\beta_1 - S_{12}\beta_2}{\sigma^2}\right] du_2\, db_1 \\ &= \frac{1}{\sigma^2} E\left[I\left(\frac{u_1}{u_2} \geq c\right)\left(1 - \phi\left(\frac{u_1}{u_2}\right)\right)^p u_1^q S_{11} b_1\right] \\ &\quad - \left[\frac{S_{11}\beta_1 + S_{12}\beta_2}{\sigma^2}\right] H(p,q;c). \end{aligned} \tag{33}$$

Equating (30) and (33), and multiplying $\beta' X' X_1 S_{11}^{-1}$ from the left, we obtain (17) in the text.

Acknowledgement

The author is grateful to Kazuhiro Ohtani for his very helpful comments and suggestions.

References

Baranchik, A.J. (1964), Multiple regression and estimation of the mean of a multivariate normal distribution, Stanford University Technical Report No. 51.

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

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

Mittelhammer R.C. (1984), Restricted least squares, pre-test, OLS and Stein rule estimators: Risk comparison under model misspecification, *Journal of Econometrics*, 25, 151-164.

Namba, A. (2000), PMSE performance of the biased estimators in a linear regression model when relevant regressors are omitted, *Econometric Theory*, forthcoming.

Nickerson, D.M. (1989), Dominance of the positive-part version of the James-Stein estimator, *Statistics and Probability Letters*, 7, 97-103.

Ohtani, K. (1993), A comparison of the Stein-rule and positive-part Stein-rule estimators in a misspecified linear regression model, *Econometric Theory*, 9, 668-679.

Ohtani, K. (1998), MSE performance of the minimum mean squared error estimators in a linear regression model when relevant regressors are omitted, *Journal of Statistical Computation and Simulation*, 61, 61-75.

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

Ullah, A. and S. Ullah (1978), Double k-class estimators of coefficients in linear regression, *Econometrica*, 46, 705-722.

Table 1: Relative PMSE for $k_1 = 5$ and $n = 20$.

Estimator	F_β	R_1^2 0.1	0.3	0.5	0.7	0.9	1.0
PKK	.0	.3848	.3848	.3848	.3848	.3848	.3848
	1.0	.4724	.4781	.4828	.4867	.4895	.4906
	2.0	.5379	.5508	.5609	.5681	.5723	.5732
	4.0	.6303	.6568	.6758	.6873	.6910	.6895
	6.0	.6930	.7300	.7544	.7672	.7683	.7636
	8.0	.7390	.7834	.8102	.8224	.8203	.8126
	10.0	.7745	.8239	.8511	.8616	.8564	.8463
	15.0	.8366	.8917	.9149	.9197	.9094	.8954
	20.0	.8777	.9328	.9497	.9495	.9368	.9211
	25.0	.9075	.9600	.9704	.9666	.9530	.9368
	30.0	.9302	.9790	.9836	.9773	.9636	.9472
	40.0	.9631	1.0032	.9988	.9895	.9763	.9604
	50.0	.9861	1.0176	1.0070	.9961	.9835	.9683
	100.0	1.0438	1.0435	1.0204	1.0068	.9962	.9841
	150.0	1.0692	1.0506	1.0239	1.0095	.9995	.9894
KK	.0	.3980	.3980	.3980	.3980	.3980	.3980
	1.0	.4857	.4913	.4959	.4995	.5021	.5030
	2.0	.5514	.5639	.5735	.5800	.5833	.5838
	4.0	.6445	.6697	.6869	.6965	.6985	.6962
	6.0	.7081	.7425	.7638	.7740	.7730	.7675
	8.0	.7550	.7954	.8181	.8272	.8230	.8147
	10.0	.7915	.8353	.8576	.8648	.8580	.8473
	15.0	.8560	.9015	.9187	.9209	.9097	.8956
	20.0	.8994	.9410	.9518	.9499	.9369	.9212
	25.0	.9313	.9668	.9716	.9667	.9530	.9368
	30.0	.9559	.9845	.9843	.9773	.9636	.9472
	40.0	.9919	1.0067	.9990	.9895	.9763	.9604
	50.0	1.0171	1.0198	1.0071	.9961	.9835	.9683
	100.0	1.0796	1.0437	1.0204	1.0068	.9962	.9841
	150.0	1.1048	1.0506	1.0239	1.0095	.9995	.9894