PMSE performance of the Stein-rule and positive-part Stein-rule estimators in a regression model with or without proxy variables

Namba, Akio

Ohtani, Kazuhiro

(Citation)
Statistics & Probability Letters,76(9):898-906

(Issue Date)
2006-05

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90000339

PMSE Performance of the Stein-Rule and Positive-Part Stein-rule Estimators in a Regression Model with or without Proxy Variables*

Akio Namba and Kazuhiro Ohtani

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

Abstract

Consider a linear regression model with some relevant regressors are unobservable. In such a situation, we estimate the model by using the proxy variables as regressors or by simply omitting the relevant regressors. In this paper, we derive the explicit formula of the predictive mean squared error (PMSE) of the Stein-rule (SR) estimator and the positive-part Stein-rule (PSR) estimator for the regression coefficients when the proxy variables are used. We examine the effect of using the proxy variables on the risk performances of the SR and PSR estimators. It is shown analytically that the PSR estimator dominates the SR estimator even when the proxy variables are used. Also, our numerical results show that using the proxy variables is preferable to omitting the relevant regressors.

*This research is partially supported by the Grants-in-Aid for 21st Century COE program.

1 Introduction

In the context of linear regression, it is well known that the ordinary least squares (OLS) estimator is the best unbiased estimator (BLUE). However, as shown in Stein (1956) and James and Stein (1961), the Stein-rule (SR) estimator dominates the OLS estimator in terms of predictive mean squared error (PMSE) when the model is specified correctly. Moreover, Baranchik (1970) showed that the SR estimator is further dominated by the positive-part Stein-rule (PSR) estimator when the specified model is correct.

When the data of some independent variables are not available, a researcher may estimate a model by omitting unobservable variables. Several authors investigated the sampling properties of the SR or PSR estimators when relevant regressors are omitted. Some examples are Mittelhammer (1984), Ohtani (1993, 1998) and Namba (2000, 2003). 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 PSR estimators of the misspecified model. Also, Namba (2002) showed exactly that the PSR estimator dominates the SR estimator even when the relevant regressors are omitted.

Also, if the proxy variables for unobservable variables are available, a researcher may use them even when the relevant regressors are unobservable. Many authors investigated the sampling properties of statistics in a regression model when the proxy variables are used. Some examples are McCallum (1972), Wickens (1972), Frost (1979), Ohtani (1981), Ohtani and Hasegawa (1993), and Trenkler and Stahlecker (1996).

Therefore, when the proxy variables are available, the question whether we should use the proxy variables or we should omit the relevant regressors arises naturally. However, such comparison about the SR and PSR estimators has not been made so far. Thus, we derive the explicit formula for the PMSE's of the SR and PSR estimators when the proxy variables are used, and examine their sampling properties. In the next section, we introduce the model and estimators. In section 3, an explicit formula for the PMSE's of the SR and PSR estimators is derived. Using this formula, we execute numerical evaluations to investigate the sampling properties of the estimators in section 4.

2 Model and 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 the matrix of the proxy variables X_2^* is available though X_2 is unobservable. Thus, we can easily specify the following two models. One is the model such that the unobservable regressors X_2 is simply omitted:

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

where $u \sim N(X_2\beta_2, \sigma^2 I_n)$. And the other is the model with the proxy variable X_2^* in place of X_2:

$$y = X_1\beta_1 + X_2^*\beta_2^* + u^*, \quad u^* = X_2\beta_2 - X_2^*\beta_2^* + \epsilon, \tag{3}$$

where $u^* \sim N(X_2\beta_2 - X_2^*\beta_2^*, \sigma^2 I_n)$. Also, We assume that X_1, $[X_1, X_2]$ and $X = [X_1, X_2^*]$ are of full column rank.

Thus, the ordinary least squares (OLS) estimator for $\beta = [\beta_1', \beta_2']'$ based on (1) is

$$b = S^{-1}X'y, \tag{4}$$

where $S = X'X$.

In the context of linear regression, the Stein-rule (SR) estimator based on (1) is

$$b_{\text{SR}} = \left(1 - a\frac{e'e}{b'Sb}\right) b, \tag{5}$$

where $0 \leqslant a \leqslant 2(k-2)/(\nu+2)$, $k = k_1 + k_2$, $\nu = n - k_1 - k_2$ and $e = y - Xb$. As shown in James and Stein (1961), the SR estimator given in (5) dominates the OLS estimator given in (4) in terms of PMSE when the model is specified correctly. Though the SR estimator given in (5) dominates the OLS estimator given in (4), Baranchik (1970) showed that the SR estimator is further dominated by the positive-part Stein-rule (PSR) estimator defined as

$$b_{\text{PSR}} = \max\left[0, 1 - a\frac{e'e}{b'Sb}\right] b. \tag{6}$$

Also, the OLS, SR and PSR estimators for β_1 based on (2) are respectively defined as,

$$b_{\text{O}} = S_{11}^{-1} X_1' y, \tag{7}$$

$$b_{\text{SRO}} = \left(1 - a_1 \frac{e_1' e_1}{b_1' S_{11} b_1}\right) b_1, \tag{8}$$

$$b_{\text{PSRO}} = \max\left[0, 1 - a_1 \frac{e_1' e_1}{b_1' S_{11} b_1}\right] b_1, \tag{9}$$

where $S_{11} = X_1' X_1$, $0 \leqslant a_1 \leqslant 2(k_1 - 2)/(\nu_1 + 2)$, $\nu_1 = n - k_1$, and $e_1 = y - X_1 b_1$. Namba (2002) showed that the PSR estimator given in (9) dominates the SR estimator given in (8) in terms of PMSE even when the matrix of relevant regressors X_2 is omitted in the specified model.

Similarly, we can consider the estimators for $\beta^* = [\beta_1', \beta_2^{*\prime}]'$ based on (3):

$$b^* = S^{*-1} X^{*\prime} y, \tag{10}$$

$$b_{\text{SRP}} = \left(1 - a \frac{e^{*\prime} e^*}{b^{*\prime} S^* b^*}\right) b^*, \tag{11}$$

$$b_{\text{PSRP}} = \max\left[0, 1 - a \frac{e^{*\prime} e^*}{b^{*\prime} S^* b^*}\right] b^*, \tag{12}$$

where $S^* = X^{*\prime} X^*$, $X^* = [X_1, X_2^*]$, $0 \leqslant a \leqslant 2(k-2)/(\nu+2)$ and $e^* = y - X^* b^*$. However, the sampling properties of these estimators have not been examined so far. Thus, we derive the explicit formulae for the PMSE of b_{SRP} and b_{PSRP} in the next section.

3 PMSE of the estimators

To derive the formulae for the PMSE's of estimators, we consider the following pre-test estimator:

$$\widehat{\beta}^*(\tau) = I(F^* \geqslant \tau) \left(1 - a \frac{e^{*\prime} e^*}{b^{*\prime} S^* b^*}\right) b^*, \tag{13}$$

where $I(A)$ is an indicator function such that $I(A) = 1$ if an event A occurs and $I(A) = 0$ otherwise, $F^* = (b^{*\prime} S^* b^*/k)/(e^{*\prime} e^*/\nu)$ is the test statistic for the null hypothesis $H_0 : \beta^* = 0$ against the alternative $H_1 : \beta^* \neq 0$ based on (3), and τ is the critical value of the pre-test. $\widehat{\beta}^*(\tau)$ reduces to b_{SRP} given in (11) when $\tau = 0$, and it reduces to b_{PSRP} when $\tau = a\nu/k$.

The PMSE's of the SR and PSR estimators given in (8) and (9) are respectively defined as

$$\text{PMSE}[b_{\text{SRO}}] = E[(X_1 b_{\text{SRO}} - X\beta)'(X_1 b_{\text{SRO}} - X\beta)], \tag{14}$$

$$\text{PMSE}[b_{\text{PSRO}}] = E[(X_1 b_{\text{PSRO}} - X\beta)'(X_1 b_{\text{PSRO}} - X\beta)]. \tag{15}$$

Namba (2002) derived the explicit formulae for these PMSE's.

Similarly, the PMSE of $\widehat{\beta}^*(\tau)$ is defined as

$$\begin{aligned}
\text{PMSE}[\widehat{\beta}^*(\tau)] &= E[(X^*\widehat{\beta}^*(\tau) - X\beta)'(X^*\widehat{\beta}^*(\tau) - X\beta)] \\
&= E\left[I(F^* \geqslant \tau)\left(1 - a\frac{e^{*\prime}e^*}{b^{*\prime}Sb^*}\right)^2 b^{*\prime}Sb^*\right] \\
&\quad -2E\left[I(F^* \geqslant \tau)\left(1 - a\frac{e^{*\prime}e^*}{b^{*\prime}Sb^*}\right)\beta' X' X^* b^*\right] + \beta' S\beta \\
&= E[I(F^* \geqslant \tau)b^{*\prime}Sb^*] - 2aE[I(F^* \geqslant \tau)e^{*\prime}e^*] \\
&\quad +a^2 E\left[I(F^* \geqslant \tau)\frac{(e^{*\prime}e^*)^2}{b^{*\prime}Sb^*}\right] - 2E[I(F^* \geqslant \tau)\beta' X' X^* b^*] \\
&\quad +2aE\left[I(F^* \geqslant \tau)\frac{e^{*\prime}e^*}{b^{*\prime}Sb^*}\beta' X' X^* b^*\right] + \beta' S\beta \qquad (16)
\end{aligned}$$

Thus, performing some manipulations, we have

$$\begin{aligned}
\text{PMSE}[\widehat{\beta}^*(\tau)]/\sigma^2 &= H(1, 0; \tau) - 2aH(0, 1; \tau) + a^2 H(-1, 2; \tau) \\
&\quad -2J(0, 0; \tau) + 2aJ(-1, 1; \tau) + \lambda_1^* + \lambda_2^*, \qquad (17)
\end{aligned}$$

where

$$H(p, q; \tau) = E\left[I\left(\frac{\nu v_1}{k v_2} \geqslant \tau\right) v_1^p v_2^q\right], \qquad (18)$$

$$J(p, q; \tau) = E\left[I\left(\frac{\nu v_1}{k v_2} \geqslant \tau\right) v_1^p v_2^q \frac{\beta' X' X^* b^*}{\sigma^2}\right], \qquad (19)$$

$v_1 = b^{*\prime}S^*b^*/\sigma^2$, $v_2 = e^{*\prime}e^*/\sigma^2$, $\lambda_1^* = \beta' X' X^* S^{*-1} X^{*\prime} X\beta/\sigma^2$, $\lambda_2^* = \beta' X' M^* X' \beta/\sigma^2$, and $M^* = I - X^* S^{*-1} X^{*\prime}$.

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

$$H(p, q; \tau) = 2^{p+q} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_i(\lambda_1^*) w_j(\lambda_2^*) G_{ij}(p, q; \tau), \qquad (20)$$

$$J(p, q; \tau) = \lambda_1^* 2^{p+q} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} w_i(\lambda_1^*) w_j(\lambda_2^*) G_{i+1,j}(p, q; \tau), \qquad (21)$$

where

$$G_{i,j}(p, q; \tau) = \frac{\Gamma(k/2 + p + i)\Gamma(\nu/2 + q + j)}{\Gamma(k/2 + i)\Gamma(\nu/2 + j)}[1 - I_{\tau^*}(k/2 + p + i, \nu/2 + q + j)], \qquad (22)$$

$I_x(\cdot, \cdot)$ is the incomplete beta function ratio defined as

$$I_x(a_1, a_2) = [B(a_1, a_2)]^{-1} \int_0^x t^{a_1 - 1}(1 - t)^{a_2 - 1} dt,$$

$\tau^* = k\tau/(k\tau + \nu)$ and $w_i(\lambda) = \exp(-\lambda/2)(\lambda/2)^i/i!$. It is easy to show that the infinite series in (20) and (21) converge absolutely. Substituting (20) and (21) into (17), we obtain the explicit formula for the PMSE of $\widehat{\beta}^*(\tau)$. Letting $\tau = 0$ and $\tau = a\nu/k$, we have the PMSE's of the SR and PSR estimators when the proxy variables are used. Though the formulae for the PMSE's of the SR and PSR estimators when the proxy variables are used are similar to those obtained in Namba (2002), they differ in their derivations. (The details are shown in Appendix.)

Differentiating the PMSE of $\widehat{\beta}^*(\tau)$ with respect to τ, and performing some manipulations, we obtain

$$\begin{aligned}\frac{\partial \mathrm{PMSE}[\widehat{\beta}^*(\tau)/\sigma^2]}{\partial \tau} &= -2\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_i(\lambda_1^*)w_j(\lambda_2^*)\frac{\Gamma((\nu+k)/2+i+j+1)}{\Gamma(k/2+i)\Gamma(\nu/2+j)} \\ &\quad \times \frac{k^{k/2+i-1}\nu^{\nu/2+j}\tau^{k/2+i-2}}{(k\tau+\nu)^{(\nu+k)/2+i+j+1}}(k\tau - a\nu)\left[(k\tau - a\nu) - \lambda_1^*\frac{k\tau}{k/2+i}\right], \qquad (23)\end{aligned}$$

since the infinite series in (20) and (21) are absolutely convergent. Thus, the PMSE of $\widehat{\beta}^*(\tau)$ is a monotone decreasing function on $\tau \in [0, a\nu/k]$. Since $\widehat{\beta}^*(\tau)$ reduces to the SR estimator when $\tau = 0$, and reduces to the PSR estimator when $\tau = a\nu/k$, we have the following theorem.

Theorem 1 *The PSR estimator given in (11) dominates the SR estimator given in (12) in terms of PMSE even when the proxy variables are used in place of unobservable variables.*

Since further theoretical analysis is difficult, we execute numerical evaluations in the next section.

4 Numerical analysis

In this section, we compare the PMSE's of the estimators with or without the proxy variables numerically. As discussed in Ohtani and Hasegawa (1993), we have

$$\lambda_2^* = F_{\beta_2}[1 - \{\beta_2' X_2' M_1 X_2^* (X_2^{*\prime} M_1 X_2^*)^{-1} X_2^{*\prime} M_1 X_2 \beta_2 / \beta_2' X_2' M_1 X_2 \beta_2\}], \qquad (24)$$

where $F_{\beta_2} = \beta_2 X_2' M_1 X_2 \beta_2/\sigma^2$, and $M_1 = I - X_1 S_{11}^{-1} X_1'$. F_{β_2} may be interpreted as the magnitude of importance of X_2, since it is the noncentrality parameter which appears in a test for null hypothesis $\mathrm{H}_0 : \beta_2 = 0$ when X_2 is observed. When F_{β_2} is close to zero, using the proxy variables may be rather irrelevant since X_2 is not a significant variables, and vice versa.

If $k_2 = 1$, the part in the braces in (24) reduces to

$$(X_2' M_1 X_2^*)^2 / [(X_2' M_1 X_2)(X_2^{*\prime} M_1 X_2^*)]. \tag{25}$$

Since this is the partial correlation coefficient between X_2 and X_2^* given X_1, the part in the braces in (24) may be interpreted as the magnitude of richness of proxy variables even when $k_2 \geqslant 2$. Denoting the part in the braces in (24)as $r_{22.1}^{*2}$, we have

$$\lambda_2^* = F_{\beta_2}(1 - r_{r22.1}^{*2}). \tag{26}$$

When $r_{22.1}^{*2}$ is close to unity, and especially when $k_2 = 1$, the proxy variable is very rich and vice versa.

Also, as shown in Namba (2002), the PMSE's of the estimators given in (8) and (9) depends on the values of $\lambda_1 = \beta' X' X_1 S_{11}^{-1} X_1' X \beta / \sigma^2 = F_\beta R_1^2$ and $\lambda_2 = \beta' X' M_1 X \beta = F_{\beta_2} = F_\beta (1 - R_1^2)$, 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. It is easy to show that $\lambda_1 + \lambda_2 = F_\beta$, and $\lambda_1^* + \lambda_2^* = F_\beta$.

Thus, in the numerical evaluation, we use the following parameter values: $k_1 = 3, 4, 5, 7$, $k_2 = 1$, $n = 20, 30, 40$, $\lambda_2/F_\beta = 1 - R_1^2 = 0, 0.1, 0.3, 0.5, 0.7, 0{,}9$, $\lambda_2^*/F_\beta = (1 - R_1^2)(1 - r_{22.1}^{*2}) = 0$, 0.1 0.3, 0.5, 0.7, 0,9, and $F_\beta =$ various values. When $\lambda_2/F_\beta = 1 - R_1^2$ is small, the magnitude of model misspecification is regarded as small. Also, λ_2^*/F_β gets much smaller than λ_2/F_β as $r_{r22.1}^{*2}$ gets closer to unity (i.e., the proxy variable gets richer). To compare the PMSE's of the estimators, we evaluated the values of relative PMSE defined as $PMSE[\bar{\beta}]/PMSE[b^*]$, where $\bar{\beta}$ is any estimator. Thus, the estimator $\bar{\beta}$ has smaller PMSE than b^* when the value of relative PMSE is smaller than unity. The numerical evaluations were executed on a personal computer using the FORTRAN code. The double infinite series in $H(p, q; \tau)$ and $J(p, q; \tau)$ were judged to converge when the increment of series got smaller than 10^{-14}. As for the PMSE's of b_{SRO} and b_{PSRO}, we used the formulae derived by Namba (2002). Since the results for the cases of $k = 6$, 8, and $n = 30$ are qualitatively typical, we do not show the results for the other cases.

Table 1 shows the PMSE's of the estimators when the relevant regressor is omitted. As shown in Namba (2002), we see that the PSR estimator b_{PSRO} dominates the SR estimator b_{SRO}

in terms of PMSE. Though the SR and PSR estimators have slightly larger PMSE than the OLS estimator with the proxy variable (i.e., b^*) for large values of F_β and λ_2/F_β, the difference is very small.

From Table 2, as shown in Theorem 1, we actually see that the SR estimator b_{SRP} is dominated by the PSR estimator b_{PSRP} even when the proxy variable is used. As F_β and λ_2^*/F_β get larger, the PMSE's of the SR and PSR estimators get larger. Similar to Table 1, though the SR and PSR estimators have larger PMSE than the OLS estimator for large values of F_β and λ_2/F_β, the difference is very small. Also, comparing Table 1 with Table 2, we see that the PMSE of the PSR estimator with the proxy variable (i.e., b_{PSRP}) is slightly larger than that of the PSR estimator with the omitted regressor (i.e., b_{PSRO}) for $F_\beta = 0$. However, noting that $\lambda_2^* \leqslant \lambda_2$, b_{PSRP} has smaller PMSE than b_{PSRO} over a wide region of parameter space. Thus, it seems that the best choice among the estimators considered in this paper is the PSR estimator with proxy variables.

Appendix

Here, we derive the explicit formulae for $H(p,q;\tau)$ and $J(p,q;\tau)$. First, we derive the formula for $H(p,q;\tau)$. Since $b^* \sim N(S^{*-1}X^{*\prime}X\beta, \sigma^2 S^{*-1})$, $v_1 = b^{*\prime}S^*b^*/\sigma^2 \sim \chi_k'^2(\lambda_1^*)$, where $\lambda_1^* \beta' X' X^* S^{*-1} X^{*\prime} X\beta/\sigma^2$ and $\chi_k'^2(\lambda)$ is a noncentral chi-square distribution with k degrees of freedom and noncentrality parameter λ. Also, since $e^* \sim N(M^* X\beta, \sigma^2 I_n)$ where $M^* = I - X^* S^{*-1} X^{*\prime}$, $v_2 = e^{*\prime}e^*/\sigma^2 \sim \chi_\nu'^2(\lambda_2^*)$ where $\lambda_2^* = \beta' X' M^* X' \beta/\sigma^2$. Moreover, v_1 and v_2 are mutually independent.

Using v_1 and v_2, $H(p,q;\tau)$ is expressed as,

$$H(p,q;\tau) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} K_{ij} \iint_{\nu v_1/k v_2 \geqslant \tau} v_1^{k/2+p+i-1} v_2^{\nu/2+q+j-1} \exp[-(v_1+v_2)/2] dv_1\, dv_2, \tag{27}$$

where

$$K_{ij} = \frac{w_i(\lambda_1^*) w_j(\lambda_2^*)}{2^{(\nu+k)/2+i+j}\Gamma(k/2+i)\Gamma(\nu/2+j)}. \tag{28}$$

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

Making use of the change of variables, $v_3 = v_1/v_2$ and $v_2 = v_4$, the integral in (28) reduces to

$$\int_0^\infty\int_{k\tau/\nu}^\infty v_3^{k/2+p+i-1} v_4^{(\nu+k)/2+p+q+i+j-1} \exp[-v_4(v_3+1)/2] dv_3\, dv_4. \tag{29}$$

Again making use of the change of variable, $v_5 = v_4(v_3+1)/2$ (29) reduces to

$$2^{(\nu+k)/2+p+q+i+j}\Gamma((\nu+k)/2+p+q+i+j)\int_{k\tau/\nu}^\infty \frac{v_3^{k/2+p+i-1}}{(v_3+1)^{(\nu+k)/2+p+q+i+j}} dv_3. \tag{30}$$

Finally, making use of the change of variable, $t = v_3/(v_3+1)$, and performing some manipulations, we have (20) in the text.

Next, we derive the formula for $J(p,q;\tau)$. Differentiating $H(p,q;\tau)$ given in (20) with respect to β, we have

$$\begin{aligned}
\frac{\partial H(p,q;\tau)}{\partial\beta} &= 2^{p+q}\sum_{i=1}^\infty\sum_{j=0}^\infty\left[\frac{\partial w_i(\lambda_1^*)}{\partial\beta}w_j(\lambda_2^*) + w_i(\lambda_1^*)\frac{\partial w_j(\lambda_2^*)}{\partial\beta}\right]G_{ij}(p,q;\tau)\\
&= 2^{p+q}\sum_{i=1}^\infty\sum_{j=0}^\infty\left[\left\{-\frac{1}{2}w_i(\lambda_1^*)+\frac{1}{2}w_{i-1}(\lambda_1^*)\right\}w_j(\lambda_2^*)\frac{2X'X^*S^{*-1}X^{*\prime}X\beta}{\sigma^2}\right.\\
&\qquad\left.+\left\{-\frac{1}{2}w_j(\lambda_2^*)+\frac{1}{2}w_{j-1}(\lambda_2^*)\right\}w_i(\lambda_1^*)\frac{2X'M^*X\beta}{\sigma^2}\right]G_{ij}(p,q;\tau)\\
&= \frac{X'X^*S^{*-1}X^{*\prime}X\beta}{\sigma^2}2^{p+q}\sum_{i=0}^\infty\sum_{j=0}^\infty w_i(\lambda_1^*)w_j(\lambda_2^*)G_{i+1,j}(p,q;\tau)\\
&\quad -\frac{X'X^*S^{*-1}X^{*\prime}X\beta + X'M^*X\beta}{\sigma^2}H(p,q;\tau)\\
&\quad +\frac{X'M^*X\beta}{\sigma^2}2^{p+q}\sum_{i=0}^\infty\sum_{j=0}^\infty w_i(\lambda_1^*)w_j(\lambda_2^*)G_{i,j+1}(p,q;\tau),
\end{aligned} \tag{31}$$

where we may define $w_{-1}(\lambda) = 0$.

Using b^* and v_2, $H(p,q;\tau)$ can be expressed as

$$H(p,q;\tau) = \iint_R \left(\frac{b^{*\prime}S^*b^*}{\sigma^2}\right)^p v_2^q f_{b^*}(b^*) f_{v_2}(v_2) db^*\, dv_2, \tag{32}$$

where

$$f_{b^*}(b^*) = \frac{1}{(2\pi)^{k/2}|\sigma^2 S^{*-1}|^{1/2}}\exp\left[-\frac{(b^*-S^{*-1}X^{*\prime}X\beta)S^*(b^*-S^{*-1}X^{*\prime}X\beta)'}{2\sigma^2}\right], \tag{33}$$

$$f_{v_2}(v_2) = \exp\left(-\frac{\lambda_2^*}{2}\right)\sum_{j=0}^\infty \frac{(\lambda_2^*/2)^j v_2^{\nu/2+j-1}\exp(-v_2/2)}{j!2^{\nu/2+j}\Gamma(\nu/2+j)}, \tag{34}$$

and R is the region such that $\nu b^{*\prime}S^*b^*/(kv_2\sigma^2) \geqslant \tau$. Noting that

$$\frac{\partial f_{b^*}(b^*)}{\partial\beta} = f_{b^*}(b^*)\frac{X'X^*b^* - X'X^*S^{*-1}X^{*\prime}X\beta}{2\sigma^2}. \tag{35}$$

$$\frac{\partial f_{v_2}(v_2)}{\partial\beta} = \left[-f_{v_2}(v_2) + \exp\left(-\frac{\lambda_2^*}{2}\right)\sum_{j=0}^\infty\frac{(\lambda_2^*/2)^j v_2^{\nu/2+j}\exp(-v_2/2)}{j!2^{\nu/2+j+1}\Gamma(\nu/2+j+1)}\right]\frac{X'M^*X\beta}{\sigma^2}, \tag{36}$$

and differentiating (34) with respect to β we obtain

$$\begin{aligned}
\frac{\partial H(p,q;\tau)}{\partial\beta} &= \iint_R \frac{X'X^*b^* - X'X^*S^{*-1}X^{*\prime}X\beta}{\sigma^2}\left(\frac{b^{*\prime}Sb^*}{\sigma^2}\right)^p v_2^q f_{b^*}(b^*)f_{v_2}(v_2)db^*\,dv_2 \\
&\quad -\frac{X'M^*X\beta}{\sigma^2}\iint_R \left(\frac{b^{*\prime}Sb^*}{\sigma^2}\right)^p v_2^q f_{b^*}(b^*)f_{v_2}(v_2)db^*\,dv_2 \\
&\quad +\frac{X'M^*X\beta}{\sigma^2}\iint_R \left(\frac{b^{*\prime}Sb^*}{\sigma^2}\right)^p v_2^q f_{b^*}(b^*) \\
&\qquad\qquad \times\exp\left(-\frac{\lambda_2^*}{2}\right)\sum_{j=0}^{\infty}\frac{(\lambda_2^*/2)^j v_2^{\nu/2+j}\exp(-v_2/2)}{j!2^{\nu/2+j+1}\Gamma(\nu/2+j+1)}db^*\,dv_2 \\
&= E\left[I(F^*\geqslant\tau)v_1^p v_2^q\frac{X'X^*b^*}{\sigma^2}\right] - \frac{X'X^*S^{*-1}X^{*\prime}X\beta + X'M^*X\beta}{\sigma^2}H(p,q;\tau) \\
&\quad +\frac{X'M^*X\beta}{\sigma^2}K_{i,j+1}\iint_R v_1^{k/2+i-1}v_2^{\nu/2+q+j}\exp\left[-\frac{v_1+v_2}{2}\right]dv1\,dv_2 \\
&= E\left[I(F^*\geqslant\tau)v_1^p v_2^q\frac{X'X^*b^*}{\sigma^2}\right] - \frac{X'X^*S^{*-1}X^{*\prime}X\beta + X'M^*X\beta}{\sigma^2}H(p,q;\tau) \\
&\quad +\frac{X'M^*X\beta}{\sigma^2}2^{p+q}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}w_i(\lambda_1^*)w_j(\lambda_2^*)G_{i,j+1}(p,q;\tau),
\end{aligned} \tag{37}$$

where the last equality is obtained by the calculations similar to the ones used to derive $H(p,q;\tau)$. Equating (31) and (37), and multiplying β' from the left, we obtain $J(p,q;\tau)$ given in (21) in the text. Note that we differentiate $H(p,q;\tau)$ with respect to β in (31) and (37) in order to obtain the formula for $J(p,q;\tau)$. When the relevant regressors are omitted, Namba (2002) differentiated a similar term with respect to β_1 in order to obtain the formulae for the PMSE's of the SR and PSR estimators since $\partial\lambda_2/\partial\beta_1 = 0$, where $\lambda_2 = \beta' X' M_1 X\beta$. However, mathematical derivation differs a great deal when the proxy variables are used. We can not obtain the formulae for them by such calculations since $\partial\lambda_2^*/\partial\beta_1 \neq 0$. Thus, (31) and (37) have some additional terms compared to the similar equations in Namba (2002).

Acknowledgement

The authors are grateful to an anonymous referee for his comments and suggestions.

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.

Frost, P. A. (1979) Proxy variables and specification bias. *The Review of Economics and Statistics* 61, 323–325.

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.

McCallum, B. T. (1972) Relative asymptotic bias from errors of omission and measurement. *Econometrica* 40, 757–758.

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

Namba, A. (2002) PMSE performance of the biased estimators in a linear regression model when relevant regressors are omitted. *Econometric Theory* 18, 1086–1098.

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.

Ohtani, K. (1981) On the use of a proxy variable in prediction: An MSE comparison. *The Review of Economics and Statistics* 63, 627–629.

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.

Ohtani, K. & H. Hasegawa (1993) On small sample properties of R^2 in a linear regression model with multivariate t errors and proxy variables. *Econometric Theory* 9, 504–515.

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.

Trenkler, G. & P. Stahlecker (1996) Dropping variables versus use of proxy variables in linear regression. *Journal of Statistical Planning and Inference* 50, 65–75.

Wickens, M. R. (1972) A note on the use of proxy variables. *Econometrica* 40, 759–761.

Table 1: PMSE's for $k = 6$ and $n = 30$ when the relevant regressor is omitted.

		λ_2/F_β					
Estimator	F_β	0.0	0.1	0.3	0.5	0.7	0.9
b_{SRO}	0.0	0.3704	0.3704	0.3846	0.3704	0.3704	0.3704
	1.0	0.4511	0.4531	0.4768	0.4588	0.4603	0.4609
	2.0	0.5126	0.5176	0.5514	0.5300	0.5319	0.5312
	4.0	0.5972	0.6095	0.6624	0.6358	0.6380	0.6335
	6.0	0.6505	0.6700	0.7389	0.7091	0.7119	0.7049
	8.0	0.6859	0.7119	0.7933	0.7617	0.7657	0.7578
	10.0	0.7105	0.7424	0.8330	0.8006	0.8061	0.7988
	15.0	0.7475	0.7913	0.8949	0.8627	0.8723	0.8702
	20.0	0.7676	0.8209	0.9290	0.8983	0.9111	0.9166
	25.0	0.7801	0.8412	0.9496	0.9207	0.9358	0.9494
	30.0	0.7886	0.8564	0.9631	0.9360	0.9526	0.9737
	40.0	0.7995	0.8780	0.9791	0.9554	0.9736	1.0072
	50.0	0.8061	0.8932	0.9879	0.9670	0.9858	1.0290
	100.0	0.8196	0.9326	1.0025	0.9900	1.0089	1.0740
	150.0	0.8241	0.9504	1.0061	0.9976	1.0160	1.0879
b_{PSRO}	0.0	0.2714	0.2714	0.2714	0.2714	0.2714	0.2714
	1.0	0.3712	0.3718	0.3721	0.3713	0.3697	0.3670
	2.0	0.4497	0.4521	0.4542	0.4529	0.4484	0.4410
	4.0	0.5602	0.5687	0.5773	0.5762	0.5663	0.5483
	6.0	0.6297	0.6456	0.6624	0.6631	0.6497	0.6226
	8.0	0.6746	0.6977	0.7226	0.7262	0.7113	0.6774
	10.0	0.7045	0.7342	0.7663	0.7733	0.7584	0.7196
	15.0	0.7463	0.7894	0.8341	0.8485	0.8373	0.7926
	20.0	0.7674	0.8205	0.8715	0.8908	0.8851	0.8396
	25.0	0.7801	0.8411	0.8949	0.9168	0.9163	0.8728
	30.0	0.7886	0.8563	0.9110	0.9339	0.9379	0.8976
	40.0	0.7995	0.8780	0.9319	0.9548	0.9650	0.9324
	50.0	0.8061	0.8932	0.9450	0.9668	0.9807	0.9559
	100.0	0.8196	0.9326	0.9733	0.9900	1.0084	1.0119
	150.0	0.8241	0.9504	0.9835	0.9976	1.0159	1.0348

Table 2: PMSE's for $k = 6$ and $n = 30$ when the proxy variable is used.

		λ_2^*/F_β					
Estimator	F_β	0.0	0.1	0.3	0.5	0.7	0.9
b_{SRP}	0.0	0.3846	0.3846	0.3846	0.3846	0.3846	0.3846
	1.0	0.4755	0.4761	0.4768	0.4767	0.4759	0.4744
	2.0	0.5472	0.5492	0.5514	0.5512	0.5488	0.5444
	4.0	0.6507	0.6563	0.6624	0.6626	0.6575	0.6473
	6.0	0.7197	0.7289	0.7389	0.7401	0.7337	0.7199
	8.0	0.7678	0.7801	0.7933	0.7958	0.7896	0.7746
	10.0	0.8027	0.8174	0.8330	0.8369	0.8318	0.8174
	15.0	0.8578	0.8762	0.8949	0.9018	0.9013	0.8940
	20.0	0.8892	0.9094	0.9290	0.9379	0.9423	0.9456
	25.0	0.9094	0.9303	0.9496	0.9598	0.9684	0.9832
	30.0	0.9234	0.9445	0.9631	0.9740	0.9861	1.0121
	40.0	0.9415	0.9622	0.9791	0.9908	1.0077	1.0537
	50.0	0.9527	0.9726	0.9879	1.0000	1.0200	1.0823
	100.0	0.9759	0.9916	1.0025	1.0152	1.0414	1.1477
	150.0	0.9838	0.9967	1.0061	1.0190	1.0471	1.1707
b_{PSRP}	0.0	0.2753	0.2753	0.2753	0.2753	0.2753	0.2753
	1.0	0.3663	0.3674	0.3690	0.3694	0.3689	0.3675
	2.0	0.4361	0.4398	0.4446	0.4458	0.4439	0.4389
	4.0	0.5306	0.5420	0.5562	0.5604	0.5557	0.5426
	6.0	0.5863	0.6066	0.6316	0.6403	0.6345	0.6146
	8.0	0.6196	0.6484	0.6839	0.6976	0.6925	0.6679
	10.0	0.6400	0.6766	0.7210	0.7398	0.7366	0.7090
	15.0	0.6644	0.7169	0.7768	0.8058	0.8099	0.7806
	20.0	0.6742	0.7389	0.8067	0.8417	0.8537	0.8270
	25.0	0.6792	0.7539	0.8251	0.8631	0.8819	0.8600
	30.0	0.6823	0.7656	0.8379	0.8768	0.9011	0.8848
	40.0	0.6860	0.7834	0.8549	0.8931	0.9245	0.9200
	50.0	0.6880	0.7969	0.8661	0.9022	0.9376	0.9439
	100.0	0.6918	0.8362	0.8918	0.9200	0.9584	1.0021
	150.0	0.6930	0.8559	0.9017	0.9259	0.9631	1.0266