MSE Performance of a Homogeneous Pre-test estimator consisting of a family of MMSE estimators

Ohtani, Kazuhiro

(Citation)
Kobe University Economic Review,57:1-12

(Issue Date)
2011

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/81003823

(URL)
https://hdl.handle.net/20.500.14094/81003823

MSE PERFORMANCE OF A HOMOGENEOUS PRE-TEST ESTIMATOR CONSISTING OF A FAMILY OF MMSE ESTIMATORS

by **KAZUHIRO OHTANI***

In this paper, we consider a homogeneous pre-test (HO-PT) estimator consisiting of the minimum mean squared error estimator and the adjusted minimum mean squared error estimator. The exact formula for the mean squared error (MSE) of the HO-PT estimator is derived. Since the exact formula for the MSE of the HO-PT estimator is complicated, we examine the MSE performance of the HO-PT estimator by numerical evaluations. Our numerical results show that if the number of regression coefficients is larger than or equal to 3, and the significance level of a pre-test is larger than or equal to 0.25, the HO-PT estimator dominates the OLS estimator. Also, we propose a criterion for choosing the significance level of the pre-test, and we show the significance levels selected under the criterion.

1. Introduction

In the context of regression, the so-called Stein-rule (SR) estimator proposed by Stein (1956) and James and Stein (1961) dominates the ordinary least squares (OLS) estimator in terms of mean squared error (MSE) if the number of regression coefficients is larger than or equal to 3. [The MSE used in this paper is called a weak mean squared error in Wallace (1972).] However, the SR estimator is further dominated by the positive-part Stein-rule (PSR) estimator proposed by Baranchik (1970).

There exist several families of shrinkage estimators. As one of the shrinkage estimators for regression coefficients, Theil (1971) proposed the minimum mean squared error (MMSE) estimator. However, since the MMSE estimator proposed by Theil (1971) includes unknown parameters, Farebrother (1975) proposed an operational variant of the MMSE estimator. [Hereafter we simply call the MMSE estimator proposed by Farebrother (1975) the MMSE estimator.] Ohtani (1996a) derived the exact formula of the MSE of the MMSE estimator, and showed by numerical evaluations that the MMSE estimator dominates the ordinary least squares (OLS) estimator in terms of MSE if the number of regression coefficients is larger than or equal to 3. Further, Ohtani (1996b) considered an adjustment in the degrees of freedom of the MMSE estimator, and showed by numerical evaluations that the adjusted MMSE (AMMSE) estimator can have a smaller MSE than the SR, PSR and MMSE estimators in a wide region of the noncentrality parameter, while the MSE of the AMMSE estimator is slightly larger than that of the OLS estimator when the number of regression coefficients is smaller than or equal to 4 and the value of the noncentrality parameter is considerably large. [See also Ohtani (2000).]

Although the SR and AMMSE estimators belong to different families of shrinkage estimators,

* The author is grateful for JSPS (Japan Society for the Promotion of Science) for partial financial support (Grant-in-Aid for Scientific Research).

Ohtani (1999) combined these two estimators and constructed a heterogenous pre-test (HPT) estimator. Ohtani (1999) showed that if the number of regression coefficients is larger than or equal to 3, the HPT estimator dominates the SR estimator in terms of MSE. In this paper, we consider a pre-test estimator consisiting of the MMSE and AMMSE estimators. Since the MMSE and AMMSE estimators belong to the same family of MMSE estimators, we call the above pre-test estimator the homogeneous pre-test (HO-PT) estimator.

In section 2 the model and estimators are presented, and in section 3 the exact formula for the MSE of the HO-PT estimator is derived. Since the exact formula for the MSE of the HO-PT estimator is complicated, we examine the MSE performance of the HO-PT estimator by numerical evaluations in section 4. Our numerical results show that if the number of regression coefficients is larger than or equal to 3 and the significance level of the pre-test is larger than or equal to 0.25, the HO-PT estimator dominates the OLS estimator. Also, we propose a criterion for choosing the significance level of a pre-test, and show the significance levels selected under the criterion.

2. Model and estimators

We consider a linear regression model,

$$y = X\beta + \epsilon, \tag{1}$$

where y is an $n \times 1$ vector of observations on a dependent variable, X is an $n \times k$ matrix of observations of independent variables, β is a $k \times 1$ vector of coefficients, and ϵ is an $n \times 1$ vector of error terms. We assume that X is nonstochastic and of full column rank, and ϵ is distributed as $N(0,\ \sigma^2 I_n)$, where I_n is an $n \times n$ identity matrix.

The ordinary least squares (OLS) estimator of β is

$$b = S^{-1}X'y, \tag{2}$$

where $S = X'X$. The minimum mean squared error (MMSE) estimator and the adjusted MMSE (AMMSE) estimator are respectively

$$b_M = \Big(\frac{b'Sb}{b'Sb + e'e/\nu}\Big)b, \tag{3}$$

$$b_{AM} = \Big(\frac{b'Sb/k}{b'Sb/k + e'e/\nu}\Big)b, \tag{4}$$

where $e = y - Xb$ and $\nu = n - k$.

As is shown in Ohtani (1996b, 2000), in the case of $3 \le k \le 5$, the AMMSE estimator has a much smaller MSE than the MMSE estimator when the value of a noncentrality parameter is close to zero, while the MMSE estimator has smaller MSE than the AMMSE estimator when the value of a noncentrality parameter is large. Thus, especially when $3 \leq k \leq 5$, we may consider the following homogeneous pre-test (HO-PT) estimator:

$$\widehat{\beta}_c = I(F < c)\, b_{AM} + I(F \geq c)\, b_M, \tag{5}$$

where $F = (b'Sb/k)/(e'e/\nu)$ is a test statistic for the null hypothesis $H_0 : \beta = 0$ against the alternative hypothesis $H_1 : \beta \neq 0$, $I(A)$ is an indicator function such that $I(A) = 1$ if an event A occurs and $I(A) = 0$ otherwise, and c is a critical value of the pre-test.

3. MSE of the HO-PT estimator

In this paper we use the MSE defined by

$$MSE(\widehat{\beta}_c) = E[(\widehat{\beta}_c - \beta)S(\widehat{\beta}_c - \beta)]. \tag{6}$$

This MSE is called the weak MSE in Wallace (1972).

Substituting (3) and (4) in (5), and further substituting (5) into (6), we obtain

$$\begin{aligned} MSE(\widehat{\beta}_c) = E\Big[I(F < c)\left(\frac{b'Sb}{b'Sb + (k/\nu)e'e}\right)^2 b'Sb\Big] \\ + \; E\Big[I(F \geq c)\left(\frac{b'Sb}{b'Sb + (1/\nu)e'e}\right)^2 b'Sb\Big] \\ - \; 2\Big\{E\Big[I(F < c)\left(\frac{b'Sb}{b'Sb + (k/\nu)e'e}\right)\beta'Sb\Big] \\ + \; E\Big[I(F \geq c)\left(\frac{b'Sb}{b'Sb + (1/\nu)e'e}\right)\beta'Sb\Big]\Big\} \\ + \; \beta'S\beta. \end{aligned} \tag{7}$$

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

$$H(p, q, \theta; c) = E\Big[I(F < c)\left(\frac{b'Sb}{b'Sb + \theta\, e'e}\right)^p (b'Sb)^q\Big], \tag{8}$$

$$J(p, q, \theta; c) = E\Big[I(F < c)\left(\frac{b'Sb}{b'Sb + \theta\, e'e}\right)^p (b'Sb)^q\, (\beta'Sb)\Big], \tag{9}$$

where $\theta = k/\nu$ or $\theta = 1/\nu$, then the MSE of $\widehat{\beta}_c$ is written as

$$\begin{aligned} MSE(\widehat{\beta}_c) = H(2, 1, k/\nu; c) + H(2, 1, 1/\nu; \infty) - H(2, 1, 1/\nu; c) \\ - \; 2\Big\{J(1, 0, k/\nu; c) + J(1, 0, 1/\nu; \infty) - J(1, 0, 1/\nu; c)\Big\} \\ + \; \beta'S\beta. \end{aligned} \tag{10}$$

As is shown in the Appendix, the explicit formulas of $H(p,q,\theta;c)$ and $J(p,q,\theta;c)$ are given by

$$H(p,q,\theta;c) = (2\,\sigma^2)^q \sum_{i=0}^{\infty} w_i(\lambda)\, G_i(p,q,\theta;c), \tag{11}$$

$$J(p,q,\theta;c) = (\beta' S\beta)\,(2\,\sigma^2)^q \sum_{i=0}^{\infty} w_i(\lambda)\, G_{i+1}(p,q,\theta;c), \tag{12}$$

where

$$G_i(p,q,\theta;c) = \frac{\Gamma((\nu+k)/2+q+i)}{\Gamma(k/2+i)\,\Gamma(\nu/2)} \times \int_0^{kc/(\nu+kc)} \left[\frac{t}{t+\theta(1-t)}\right]^p t^{k/2+q+i-1}\,(1-t)^{\nu/2-1}\,dt, \tag{13}$$

$w_i(\lambda) = exp(-\lambda/2)\,(\lambda/2)^i / \,i\,!$ and $\lambda = \beta' S\beta/\sigma^2$. Substituting (11) and (12) in (10), the MSE of the HO-PT estimator normalized by σ^2 is finally written as

$$\begin{aligned} MSE(\widehat{\beta}_c)/\sigma^2 = {} & 2\Big[\sum_{i=0}^{\infty} w_i(\lambda)\, G_i(2,1,k/\nu;c) \\ & + \sum_{i=0}^{\infty} w_i(\lambda)\, G_i(2,1,1/\nu;\infty) \\ & - \sum_{i=0}^{\infty} w_i(\lambda)\, G_i(2,1,1/\nu;c)\Big] \\ & -2\,\lambda\Big[\sum_{i=0}^{\infty} w_i(\lambda)\, G_{i+1}(1,0,k/\nu;c) \\ & + \sum_{i=0}^{\infty} w_i(\lambda)\, G_{i+1}(1,0,1/\nu;\infty) \\ & - \sum_{i=0}^{\infty} w_i(\lambda)\, G_{i+1}(1,0,1/\nu;c)\Big] \\ & + \lambda. \end{aligned} \tag{14}$$

Since the MSE of the HO-PT estimator is very complicated, it is difficult to examine the MSE performance analytically. Thus, we examine the MSE performance of the HO-PT estimator by numerical evaluations in the next section.

4. Numerical analysis

The parameter values used in the numerical evaluations are : k = 3, 4, 5, 8; n = 20, 30, 40; λ = various values; α (significance level of the pre-test) = 0.0, 0.01, 0.05, 0.10, 0.25, 0.50, 1.0. When $\alpha = 0.0$ (i.e., $c = \infty$), the HO-PT estimator reduces to the AMMSE estimator. Also, when $\alpha = 1.0$ (i.e., $c = 0$), the HO-PT estimator reduces to the MMSE estimator. The numerical evaluations were excuted on a personal computer, using FORTRAN code. In evaluating the integral in $G_i(p, q, \theta; c)$ given in (13), we used Simpson's 3/8 rule with 500 equal subdivisions. The infinite series in $H(p, q, \theta; c)$ and $J(p, q, \theta; c)$ were judged to converge when the increment of the series got smaller than 10^{-12}.

Since the results for $n = 30$ are typical, we discuss the results for $n = 30$ in detail. Tables 1 to 4 show the numerical results for $k = 3, 4, 5, 8$ and $n = 30$. The values shown in the tables are those of the relative MSE of the HO-PT estimator to the OLS estimator (i.e., $MSE(\widehat{\beta}_c)/MSE(b)$). Thus, if the value shown in the tables is smaller than unity, the HO-PT estimator has smaller MSE than the OLS estimator.

From the tables, we see that when $k \geq 3$ and $\alpha = 1.0$, the HO-PT estimator (MMSE estimator) dominates the OLS estimator. Also, when $\alpha = 0.0$ and $k \leq 4$, the HO-PT estimator (AMMSE estimator) does not dominate the OLS estimator. However, when $\alpha = 0.0$ and $k \geq 5$, the HO-PT estimator (AMMSE estimator) dominates the OLS estimator. These results coincide with those in Ohtani (1996b, 2000).

We see from Table 1 that when $k = 3$, the HO-PT estimator with $\alpha \leq 0.1$ does not dominate the OLS estimator since some relative MSE's for $\lambda \geq 10$ are larger than unity. However, the HO-PT estimator with $\alpha \geq 0.25$ dominates the OLS estimator since the relative MSE's are uniformly smaller than unity. Comparing Tables 1 to 3, we see that as the value of k increases from 3 to 5, the range of λ where the relative MSE's are larger than unity decreases. In particular, when $k = 5$, only the HO-PT estimator with $\alpha = 0.01$ does not dominate the OLS estimator.

The relative MSE of the AMMSE estimator (i.e., $\alpha = 0$) is much smaller than that of the MMSE estimator (i.e., $\alpha = 1.0$) around $\lambda = 0$. Although the relative MSE around $\lambda = 0$ increases as the value of α increases from zero to one, the relative MSE of the HO-PT estimator with $0.01 \leq \alpha \leq 0.5$ is still much smaller than that of the MMSE estimator (i.e., $\alpha = 1.0$) around $\lambda = 0$. When we choose the value of α, one criterion may be that the HO-PT estimator dominate the OLS estimator and simultaneously have the relative MSE as small as possible around $\lambda = 0$. When $k \geq 5$, the AMMSE estimator dominates the OLS estimator. Thus, $\alpha = 0$ is selected under this criterion when $k \geq 5$. Although we do not show all the numerical results, our numerical results show that if we choose such significance levels from our numerical results for $k = 3$ and $k = 4$, the selected values of α are as follows: $\alpha = 0.25$ for $k = 3$ and $n = 20, 30, 40$; $\alpha = 0.1$ for $k = 4$ and $n = 20, 30$; $\alpha = 0.0$ for $k = 4$ and $n = 40$.

In sum, our numerical results show that if the number of regression coefficients is larger than or equal to 3 and the significance level of the pre-test is larger than or equal to 0.25, the HO-PT estimator dominates the OLS estimator. Also, we proposed a criterion for choosing the

Table 1

Relative MSE's of the HO-PT estimators for $k = 3$ and $n = 30$.

	AMMSE	significance level					MMSE
λ	0.00	.01	.05	.10	.25	.50	1.0
.00	.3599	.3664	.3920	.4216	.4959	.5802	.6316
.10	.3748	.3821	.4093	.4396	.5135	.5942	.6412
.30	.4035	.4125	.4426	.4743	.5470	.6207	.6593
.50	.4308	.4415	.4745	.5074	.5784	.6452	.6764
.75	.4630	.4760	.5124	.5464	.6149	.6734	.6963
1.00	.4933	.5086	.5482	.5829	.6485	.6991	.7148
1.25	.5217	.5395	.5820	.6171	.6793	.7224	.7319
1.50	.5484	.5689	.6139	.6492	.7077	.7436	.7477
2.00	.5972	.6230	.6724	.7071	.7575	.7805	.7761
2.50	.6403	.6718	.7243	.7575	.7992	.8110	.8006
3.00	.6786	.7158	.7702	.8011	.8340	.8363	.8218
3.50	.7125	.7555	.8107	.8387	.8628	.8573	.8402
4.00	.7427	.7914	.8464	.8709	.8867	.8747	.8562
5.00	.7935	.8532	.9047	.9216	.9223	.9013	.8823
10.00	.9306	1.0227	1.0272	1.0117	.9788	.9558	.9473
15.00	.9799	1.0619	1.0281	1.0064	.9817	.9717	.9694
20.00	.9996	1.0503	1.0096	.9947	.9828	.9796	.9791
25.00	1.0081	1.0283	.9974	.9898	.9852	.9844	.9843
30.00	1.0119	1.0110	.9924	.9892	.9876	.9874	.9874
50.00	1.0137	.9939	.9931	.9930	.9930	.9930	.9930
75.00	1.0115	.9955	.9955	.9955	.9955	.9955	.9955
100.00	1.0095	.9967	.9967	.9967	.9967	.9967	.9967
125.00	1.0080	.9974	.9974	.9974	.9974	.9974	.9974
150.00	1.0069	.9979	.9979	.9979	.9979	.9979	.9979

Table 2

Relative MSE's of the HO-PT estimators for $k = 4$ and $n = 30$.

	AMMSE	significance level					MMSE
λ	0.00	.01	.05	.10	.25	.50	1.0
.00	.3400	.3469	.3746	.4073	.4928	.5990	.6829
.10	.3515	.3590	.3882	.4217	.5074	.6107	.6892
.30	.3736	.3826	.4147	.4499	.5354	.6328	.7014
.50	.3949	.4054	.4403	.4769	.5620	.6534	.7130
.75	.4203	.4327	.4711	.5092	.5932	.6773	.7266
1.00	.4444	.4589	.5005	.5400	.6224	.6992	.7393
1.25	.4673	.4840	.5287	.5692	.6495	.7193	.7513
1.50	.4890	.5080	.5557	.5969	.6749	.7378	.7625
2.00	.5293	.5532	.6061	.6481	.7203	.7702	.7828
2.50	.5658	.5948	.6522	.6941	.7595	.7974	.8008
3.00	.5989	.6333	.6941	.7352	.7931	.8203	.8167
3.50	.6289	.6688	.7323	.7717	.8219	.8395	.8308
4.00	.6562	.7017	.7669	.8041	.8464	.8557	.8433
5.00	.7036	.7604	.8263	.8579	.8847	.8808	.8645
10.00	.8475	.9484	.9811	.9788	.9571	.9349	.9239
15.00	.9125	1.0215	1.0091	.9911	.9655	.9526	.9490
20.00	.9454	1.0344	1.0004	.9837	.9680	.9630	.9620
25.00	.9637	1.0234	.9895	.9790	.9717	.9700	.9698
30.00	.9747	1.0085	.9837	.9784	.9755	.9750	.9750
50.00	.9921	.9872	.9853	.9852	.9851	.9851	.9851
75.00	.9976	.9902	.9901	.9901	.9901	.9901	.9901
100.00	.9993	.9926	.9926	.9926	.9926	.9926	.9926
125.00	1.0000	.9941	.9941	.9941	.9941	.9941	.9941
150.00	1.0003	.9951	.9951	.9951	.9951	.9951	.9951

Table 3

Relative MSE's of the HO-PT estimators for $k = 5$ and $n = 30$.

	AMMSE	significance level				MMSE
λ	0.00	.01	.05	.10	.25	.50	1.0
.00	.3273	.3343	.3628	.3969	.4890	.6100	.7227
.10	.3366	.3442	.3740	.4091	.5015	.6202	.7273
.30	.3547	.3635	.3960	.4327	.5258	.6396	.7361
.50	.3721	.3822	.4173	.4556	.5490	.6579	.7445
.75	.3931	.4048	.4432	.4833	.5766	.6792	.7545
1.00	.4131	.4267	.4682	.5098	.6026	.6989	.7639
1.25	.4323	.4477	.4924	.5353	.6271	.7172	.7728
1.50	.4507	.4681	.5157	.5598	.6501	.7341	.7811
2.00	.4852	.5067	.5599	.6056	.6923	.7640	.7966
2.50	.5168	.5429	.6010	.6477	.7295	.7896	.8104
3.00	.5459	.5767	.6392	.6861	.7621	.8114	.8229
3.50	.5727	.6085	.6747	.7210	.7907	.8300	.8341
4.00	.5974	.6383	.7075	.7528	.8157	.8459	.8442
5.00	.6413	.6928	.7657	.8074	.8562	.8709	.8617
10.00	.7846	.8844	.9391	.9503	.9436	.9262	.9145
15.00	.8580	.9798	.9911	.9804	.9582	.9442	.9395
20.00	.8995	1.0146	.9947	.9791	.9616	.9549	.9535
25.00	.9248	1.0175	.9871	.9749	.9652	.9627	.9623
30.00	.9414	1.0085	.9811	.9737	.9693	.9684	.9683
50.00	.9716	.9847	.9811	.9808	.9807	.9807	.9807
75.00	.9841	.9871	.9871	.9870	.9870	.9870	.9870
100.00	.9894	.9903	.9903	.9903	.9903	.9903	.9903
125.00	.9922	.9922	.9922	.9922	.9922	.9922	.9922
150.00	.9939	.9935	.9935	.9935	.9935	.9935	.9935

Table 4

Relative MSE's of the HO-PT estimators for $k = 8$ and $n = 30$.

	AMMSE	significance level					MMSE
λ	0.00	.01	.05	.10	.25	.50	1.0
.00	.3077	.3144	.3427	.3779	.4780	.6231	.8008
.10	.3137	.3207	.3500	.3860	.4869	.6308	.8030
.30	.3253	.3331	.3644	.4019	.5044	.6457	.8072
.50	.3367	.3452	.3786	.4176	.5213	.6599	.8113
.75	.3505	.3601	.3960	.4368	.5419	.6768	.8162
1.00	.3638	.3746	.4131	.4555	.5617	.6928	.8209
1.25	.3768	.3888	.4298	.4738	.5807	.7079	.8254
1.50	.3894	.4026	.4461	.4917	.5990	.7221	.8298
2.00	.4135	.4294	.4779	.5261	.6336	.7482	.8379
2.50	.4362	.4549	.5083	.5587	.6655	.7713	.8454
3.00	.4576	.4794	.5374	.5897	.6947	.7917	.8523
3.50	.4778	.5029	.5653	.6190	.7216	.8098	.8587
4.00	.4968	.5254	.5920	.6467	.7462	.8258	.8647
5.00	.5320	.5680	.6419	.6975	.7890	.8523	.8754
10.00	.6615	.7396	.8272	.8693	.9099	.9181	.9119
15.00	.7421	.8578	.9259	.9437	.9472	.9386	.9326
20.00	.7954	.9331	.9683	.9681	.9571	.9485	.9456
25.00	.8324	.9748	.9811	.9730	.9608	.9556	.9545
30.00	.8592	.9933	.9819	.9726	.9638	.9613	.9609
50.00	.9170	.9885	.9774	.9758	.9752	.9751	.9751
75.00	.9468	.9839	.9829	.9829	.9829	.9829	.9829
100.00	.9613	.9870	.9870	.9870	.9870	.9870	.9870
125.00	.9697	.9895	.9895	.9895	.9895	.9895	.9895
150.00	.9752	.9912	.9912	.9912	.9912	.9912	.9912

significance level of the pre-test in the cases of $k = 3$ and $k = 4$. If we ues the significance levels selected under the criterion, then the HO-PT estimator dominates the OLS estimator and the maximum gain in MSE around $\lambda = 0$ can be obtained.

Appendix

First, we derive the explicit formula of $H(p, q, \theta; c)$ given in equation (11). We define u_1 and u_2 as $u_1 = b'Sb/\sigma^2$ and $u_2 = e'e/\sigma^2$. Since b is distributed as $N(\beta, \sigma^2 S^{-1})$, u_1 is distributed as the noncentral chi-square distribution with k degrees of freedom and noncentrality parameter $\lambda = \beta' S\beta/\sigma^2$. Also, u_2 is distributed as the chi-square distribution with $\nu = n - k$ degrees of freedom. Using u_1 and u_2, $H(p, q, \theta; c)$ is written as

$$\begin{aligned} H(p, q, \theta; c) &= E\Big[I(F < c) \Big(\frac{b'Sb}{b'Sb + \theta\, e'e} \Big)^p (b'Sb)^q \Big] \\ &= E\Big[I(u_1/u_2 < kc/\nu) \Big(\frac{u_1}{u_1 + \theta u_2} \Big)^p (\sigma^2\, u_1)^q \Big] \\ &= \sigma^{2q} \int\!\!\int_{u_1/u_2 < kc/\nu} \Big(\frac{u_1}{u_1 + \theta u_2} \Big)^p u_1^q\, f_1(u_1)\ f_2(u_2)\ du_1\, du_2, \end{aligned} \tag{15}$$

where

$$f_1(u_1) = \sum_{i=0}^{\infty} exp(-\lambda/2) \frac{(\lambda/2)^i}{i!} \frac{1}{2^{k/2+i}\,\Gamma(k/2+i)} u_1^{k/2+i-1} exp(-u_1/2), \tag{16}$$

$$f_2(u_2) = \frac{1}{2^{\nu/2}\,\Gamma(\nu/2)} u_2^{\nu/2-1} exp(-u_2/2). \tag{17}$$

Substituting (16) and (17) into (15) and making use of the change of variables, $\tau_1 = u_1/u_2$ and $\tau_2 = u_2$, (15) reduces to

$$\sigma^{2q} \sum_{i=0}^{\infty} K_i \int_0^{kc/\nu} \int_0^{\infty} \frac{\tau_1^{k/2+p+q+i-1}\ \tau_2^{(\nu+k)/2+q+i-1}}{(\theta + \tau_1)^p} exp[-(1+\tau_1)\tau_2/2]\, d\tau_2\, d\tau_1, \tag{18}$$

where

$$K_i = w_i(\lambda) \frac{1}{2^{(\nu+k)/2+i}\,\Gamma(k/2+i)\,\Gamma(\nu/2)}, \tag{19}$$

$$w_i(\lambda) = exp(-\lambda/2) \frac{(\lambda/2)^i}{i!}. \tag{20}$$

Again, making use of the change of variable, $z = (1 + \tau_1)\tau_2/2$, (18) reduces to

$$\sigma^{2q}\sum_{i=0}^{\infty}K_i\,2^{(\nu+k)/2+q+i}\int_0^{\infty}z^{(\nu+k)/2+q+i-1}\,exp(-z)\,dz$$

$$\times\int_0^{kc/\nu}\frac{\tau_1^{k/2+p+q+i-1}}{(\theta+\tau_1)^p\,(1+\tau_1)^{(\nu+k)/2+q+i}}\,d\tau_1$$

$$=\ \sigma^{2q}\sum_{i=0}^{\infty}K_i\,2^{(\nu+k)/2+q+i}\,\Gamma((\nu+k)/2+q+i)$$

$$\times\int_0^{kc/\nu}\frac{\tau_1^{k/2+p+q+i-1}}{(\theta+\tau_1)^p\,(1+\tau_1)^{(\nu+k)/2+q+i}}\,d\tau_1. \tag{21}$$

Finally, substituting (19) into (21) and making use of the change of variable, $t=\tau_1/(1+\tau_1)$, we obtain (11) in the text.

Next, we derive the explicit formula of $J(p,q,\theta;c)$ given in equation (12). Noting that $\lambda=\beta' S\beta/\sigma^2$, we differentiate $H(p,q,\theta;c)$ with respect to β :

$$\frac{\partial H(p,q,\theta;c)}{\partial\beta}=(2\sigma^2)^q\sum_{i=0}^{\infty}\Big[\frac{\partial w_i(\lambda)}{\partial\beta}\Big]\,G_i(p,q,\theta;c)$$

$$=-\Big(\frac{S\beta}{\sigma^2}\Big)\,(2\sigma^2)^q\sum_{i=0}^{\infty}w_i(\lambda)\,G_i(p,q,\theta;c)$$

$$+\ \Big(\frac{S\beta}{\sigma^2}\Big)\,(2\sigma^2)^q\sum_{i=0}^{\infty}w_{i-1}(\lambda)\,G_i(p,q,\theta;c), \tag{22}$$

where we define $w_{-1}(\lambda)=0$. Putting $j=i-1$ in the second term in (22), (22) reduces to

$$-\Big(\frac{S\beta}{\sigma^2}\Big)\,H(p,q,\theta;c)+\Big(\frac{S\beta}{\sigma^2}\Big)\,(2\sigma^2)^q\sum_{j=0}^{\infty}w_j(\lambda)G_{j+1}(p,q,\theta;c). \tag{23}$$

As an alternative expression of $H(p,q,\theta;c)$, we express $H(p,q,\theta;c)$ by b and $e'e$:

$$H(p,q,\theta;c)=\int\int_{F<c}\Big(\frac{b'Sb}{b'Sb+\theta e'e}\Big)^p\,(b'Sb)^q\,f_1(b)\,f_2(e'e)\,db\,de'e, \tag{24}$$

where $F=(b'Sb/k)/(e'e/\nu)$,

$$f_1(b_1)=\frac{1}{(2\pi)^{k/2}\,|\sigma^2\,S^{-1}|^{1/2}}\,exp\Big[-\frac{(b-\beta)'S(b-\beta)}{2\sigma^2}\Big], \tag{25}$$

and $f_2(e'e)$ is a density function of $e'e$.

Diffrentiating (24) with respect to β , we obtain

$$\frac{\partial H(p,q,\theta;c)}{\partial \beta}$$

$$= \int\int_{F<c} \left(\frac{b'Sb}{b'Sb+\theta e'e}\right)^p (b'Sb)^q \left(\frac{Sb-S\beta}{\sigma^2}\right) f_1(b)\, f_2(e'e)\, db\, de'e$$

$$= \frac{1}{\sigma^2} \int\int_{F<c} \left(\frac{b'Sb}{b'Sb+\theta e'e}\right)^p (b'Sb)^q\, (Sb)\, f_1(b)\, f_2(e'e)\, db\, de'e$$

$$-\frac{S\beta}{\sigma^2} \int\int_{F<c} \left(\frac{b'Sb}{b'Sb+\theta e'e}\right)^p (b'Sb)^q\, f_1(b)\, f_2(e'e)\, db\, de'e$$

$$= \frac{1}{\sigma^2} E\Big[I(F<c)\left(\frac{b'Sb}{b'Sb+\theta e'e}\right)^p (b'Sb)^q\, (Sb)\Big]$$

$$-\frac{S\beta}{\sigma^2} H(p,q,\theta;c). \tag{26}$$

Equating (23) and (26), we obtain (12) in the text.

REFERENCES

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

Farebrother, R.W. (1975), "The minimum mean square error linear estimator and ridge regression," *Technometrics*, 17, 127-128.

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

Ohtani, K. (1996a), "Exact small sample properties of an operational variant of the minimum mean squared error estimator," *Communications in Statistics-Theory and Methods*, 25, 1223-1231.

Ohtani, K. (1996b), "On an adjustment of degrees of freedom in the minimum mean squared error estimator,' *Communications in Statistics-Theory and Methods*, 25, 3049-3058.

Ohtani, K. (1999), "MSE performance of a heterogeneous pre-test estimator," *Statistics and Probability Letters*, 41, 65-71.

Ohtani, K. (2000), *Shrinkage Estimation of a Linear Regression Model in Econometrics*, (Nova Science Publishers, New York).

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* 1, (Berkeley, University of California Press). 197-206.

Theil, H. (1971), *Principles of Econometrics*, (New York, John Wiley).

Wallace, T.D. (1972), Weaker Criteria and Tests for Linear Restrictions in Regression," *Econometrica*, 40, 689 - 698