ELSEVIER

Journal of Statistical Planning and Inference 50 (1996) 21–35

journal of statistical planning and inference

Risk of a homoscedasticity pre-test estimator of the regression scale under LINEX loss

Judith A. Giles*, David E.A. Giles

Department of Economics, University of Victoria, P.O. Box 3050, Victoria, B.C., Canada, V8W 3P5

Received 28 December 1992; revised 30 March 1993

Abstract

In this paper we consider the risk of an estimator of the error variance after a pre-test for homoscedasticity of the variances in the two-sample heteroscedastic linear regression model. This particular pre-test problem has been well investigated but always under the restrictive assumption of a squared error loss function. We consider an asymmetric loss function — the LINEX loss function — and derive the exact risks of various estimators of the error variance.

AMS Subject Classification: Primary 62J05, 62F11; Secondary 62P20

Keywords: Heteroscedasticity; Preliminary test; Asymmetric loss function; Risk under LINEX loss; Risk under squared error loss

1. Introduction and model framework

We consider a regression model which uses two samples with T_1 and T_2 observations:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{1}$$

or $y = X\beta + u$. y_i is a $(T_i \times i)$ vector of observations on the dependent variable, X_i is a $(T_i \times k_i)$ full-rank non-stochastic matrix of explanatory variables, β_i is a $(k_i \times 1)$ vector of coefficients and u_i is a $(T_i \times 1)$ vector of disturbance terms, $i = 1, 2$. We assume that

$$u \sim N\left(0, \begin{bmatrix} \sigma_1^2 I_{T_1} & 0 \\ 0 & \sigma_2^2 I_{T_2} \end{bmatrix}\right).$$

We also suppose that we are interested in estimating σ_1^2 but we are uncertain of the equality of the error variances and whether the samples should be pooled or not from

* Corresponding author.

SSDI 0378-3758(95)00043-7

an estimation efficiency viewpoint. The usual procedure is to undertake a preliminary test of:

$$\mathrm{H_0}\colon \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad \mathrm{H_A}\colon \sigma_1^2 < \sigma_2^2$$

or equivalently

$$\mathrm{H_0}\colon \psi = 1 \quad \text{vs} \quad \mathrm{H_A}\colon \psi < 1, \tag{2}$$

where $\psi = \sigma_1^2/\sigma_2^2$, and we have assumed a one-sided alternative hypothesis for simplicity. The usual test statistic for (2) is

$$J = \frac{v_1(y_2 - X_2 b_2)'(y_2 - X_2 b_2)}{v_2(y_1 - X_1 b_1)'(y_1 - X_1 b_1)} = \frac{v_1 u_2' M_2 u_2}{v_2 u_1' M_1 u_1} = \frac{s_2^2}{s_1^2}, \tag{3}$$

where $v_i = T_i - k_i$; $M_i = I_{T_i} - X_i(X_i'X_i)^{-1}X_i'$; $b_i = (X_i'X_i)^{-1}X_i'y_i$; $s_i^2 = (y_i - X_i b_i)'(y_i - X_i b_i)/v_i$; $i = 1, 2$. It is straightforward to show that $f(J) = \psi^{-1} f(F_{v_2, v_1})$ where F_{v_2, v_1} is a central F variate with v_2 and v_1 degrees of freedom. The testing strategy is to use the so-called 'always-pool' estimator of σ_1^2, s_{A}^2, if we cannot reject $\mathrm{H_0}$:

$$s_{\mathrm{A}}^2 = (v_1 s_1^2 + v_2 s_2^2)/(v_1 + v_2); \tag{4}$$

but to use the 'never-pool' estimator, s_{N}^2, if we reject $\mathrm{H_0}$:

$$s_{\mathrm{N}}^2 = s_1^2. \tag{5}$$

So, the estimator actually reported is the pre-test estimator:

$$s_{\mathrm{P}}^2 = \begin{cases} s_{\mathrm{N}}^2 & \text{if } J > c, \\ s_{\mathrm{A}}^2 & \text{if } J \leqslant c, \end{cases} \tag{6}$$

where c is the critical value of the test associated with an α% significance level.

The sampling properties of s_{N}^2, s_{A}^2 and s_{P}^2 have been examined in the literature (see, for example, Bancroft, 1944; Toyoda and Wallace, 1975; Ohtani and Toyoda, 1978; Bancroft and Han, 1983; Giles, 1992; Giles and Giles, 1993b) for a survey of this literature) assuming a quadratic loss function.[1] This is a symmetric loss function which implies that under- and over-estimation are equally penalised. However, we may believe that under-estimation of the scale parameter has greater consequences than over-estimation, as under-estimating the error variance in a regression model will lead to calculated t-statistics which make the regressors appear to be more 'significant' than is warranted. A conservative researcher may prefer to err in the opposite direction, which suggests that we should consider the properties of the estimators using an asymmetric loss function which penalises under-estimation more heavily than over-estimation. One such commonly suggested loss function is the LINEX loss function, initially proposed by Varian (1975). When estimating a

[1] The estimation of the coefficient vector under the assumption that $\beta_1 = \beta_2$ after the pre-test of $\mathrm{H_0}$ is considered by, for example, Taylor (1977, 1978) and Greenberg (1980).

parameter θ by $\hat{\theta}$ this loss function is given by:

$$L(\hat{\theta}, \theta) = b(\exp[a(\hat{\theta} - \theta)/\theta] - a(\hat{\theta} - \theta)/\theta - 1), \tag{7}$$

where $a \neq 0$, and $b > 0$. In our investigation we assume (without loss of generality) that $b = 1$. The sign of the shape parameter a reflects the direction of asymmetry — we set $a > 0$ ($a < 0$) if over-estimation is more (less) serious than under-estimation. The magnitude of a reflects the degree of asymmetry. For small values of $|a|$, $L(\hat{\theta}, \theta) \simeq ba^2(\hat{\theta} - \theta)/(2\theta^2)$ which is proportional to a squared error loss. So, the LINEX loss function can be regarded as a generalization of the squared error loss function allowing for asymmetry. In this paper we are particularly interested in choices of $a < 0$.

Various authors have used this form of loss function in a number of studies including Zellner (1986), Rojo (1987), Sadooghi-Alvandi and Nematollahi (1989), Kuo and Dey (1990), Parsian (1990a, b), Sadooghi-Alvandi (1990), Srivastava and Rao (1992), Basu and Ebrahimi (1991), Giles and Giles (1993a), Parsian and Sanjari Farsipour (1992), Parsian et al. (1992), and Sadooghi-Alvandi and Parsian (1992). In particular, Giles and Giles (1993a) consider the estimation of the scale parameter after a pre-test for exact linear restrictions on the regression model's coefficients. They find that the known quadratic risk properties of the pre-test estimator need not be robust to this alternative choice of loss function.

In the next section we derive the risks of s_N^2, s_A^2 and s_P^2 under LINEX loss. We follow in Section 3 with a discussion of some numerical evaluations of the risk functions and Section 4 contains some conclusions.

2. Risk under LINEX loss

We define the (relative) risk of an estimator s_*^2 of σ_1^2 as $R(s_*^2) = E[L(s_*^2, \sigma_1^2)]/\sigma_1^4$. Then, using the LINEX loss function in (7) with $b = 1$ we have:[2]

Theorem 1.

$$R(s_N^2) = e^{-a}(v_1/(v_1 - 2a))^{v_1/2} - 1 \tag{8}$$

$$R(s_A^2) = \frac{e^{-a}\psi^{v_2/2}(v_1 + v_2)^{(v_1+v_2)/2}}{(v_1 + v_2 - 2a)^{v_1/2}(\psi(v_1 + v_2) - 2a)^{v_2/2}} - \frac{av_2(1 - \psi)}{\psi(v_1 + v_2)} - 1 \tag{9}$$

$$R(s_P^2) = R(s_N^2) + e^{-a}\left\{\sum_{i=1}^{\infty} \frac{2^i}{i!} L'_{1i}L_{2i} - \sum_{i=1}^{\infty} \frac{(2a/v_1)^i}{i!} \frac{\Gamma\left(\frac{v_1 + 2i}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)} Q_{0(2i)}\right\} - \frac{av_2(Q_{20} - \psi Q_{02})}{\psi(v_1 + v_2)}, \tag{10}$$

[2] For $a < 0$ these risks are well defined. Some obvious constraints are required for $a > 0$.

where

$$Q_{mn} = \Pr.[F_{(v_2+m,\, v_1+n)} \leqslant (v_2(v_1+n)c\psi)/(v_1(v_2+m))], \quad m, n = 0, 1, \ldots,$$

L'_{1i} *is a* $(1 \times (i+1))$ *vector equal to the* $(i+1)$*th row of Pascal's Triangle;* L_{2i} *is an* $((i+1) \times 1)$ *vector with elements for* $j = 0, 1, \ldots, i$

$$\ell_j = (a/(v_1+v_2))^i \psi^{j-i} \frac{\Gamma\left(\frac{v_1+2j}{2}\right)\Gamma\left(\frac{v_2+2(i-j)}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} Q_{(2(i-j))(2j)}.$$

Proof. See the appendix.

Remarks. (i) As $\alpha \to 0$, $c \to \infty$ and $Q_{mn} \to 1$: we never reject H_0. Then, repeatedly using the Binomial Theorem, $R(s_P^2) \to R(s_A^2)$. Conversely, as $\alpha \to 1, c \to 0$ and $Q_{mn} \to 0$: we always reject H_0. Then, we can show that $R(s_P^2) \to R(s_N^2)$.

(ii) Using the infinite series expansion of the exponential function, it can be shown that (8)–(10) collapse to their quadratic loss counterparts (scaled by $a^2/2$) if a is sufficiently small so that third-order and higher-order terms are negligible. The quadratic loss functions are given by:

$$R_Q(s_N^2) = 2/v_1, \tag{11}$$

$$R_Q(s_A^2) = (v_2^2(\psi - 1)^2 + 2(v_1\psi^2 + v_2))/(\psi^2(v_1+v_2)^2), \tag{12}$$

$$\begin{aligned} R_Q(s_P^2) = (&\psi^2[2(v_1+v_2)^2 - v_2(v_1+2)(2v_1+v_2)Q_{04} + 2v_1v_2(v_1+v_2)Q_{02}] \\ &+ 2v_1v_2\psi[v_1Q_{22} - (v_1+v_2)Q_{20}] + v_1v_2(v_2+2)Q_{40})/ \\ &(v_1\psi^2(v_1+v_2)^2). \end{aligned} \tag{13}$$

We note that these expressions are not identical to those given by, for example, Toyoda and Wallace (1975). The differences arise because we consider risk relative to σ_1^4 while Toyoda and Wallace, for example, define risk to be relative to σ_2^4. We have used the former definition as this results in $R(s_N^2)$ being independent of ψ and risk diagrams that have many characteristics which are similar to those which arise when the pre-test is of exact linear restrictions on the coefficient vector.

(iii) $\lim_{\psi \to 0}(R(s_P^2)) = R(s_N^2)$ while $R(s_A^2) \to \infty$ as $\psi \to 0$. Intuitively, pre-testing leads us to follow the correct strategy of rejecting H_0 when it is in fact very false.

(iv) If H_0 is true ($\psi = 1$) then

$$\begin{aligned} R(s_A^2 \mid \psi = 1)/R(s_N^2) &= (1 - 2a/v_1)^{v_1/2}(1 - 2a/(v_1+v_2))^{-(v_1+v_2)/2} \\ &= [(1 - 2a/v_1)^{v_1}(1 - 2a(v_1+v_2))^{-(v_1+v_2)}]^{1/2}. \end{aligned} \tag{14}$$

Now, using the binomial expansion, we have:

$$(1 - 2a/v_1)^{v_1} = 1 - 2a + 2a^2(v_1 - 1)/v_1 - 4a^3(v_1 - 1)(v_1 - 2)/(3v_1^2) + \cdots$$
$$= 1 - 2a + T_1 + T_2 + \cdots$$

and

$$(1 - 2a/(v_1 + v_2))^{-(v_1+v_2)}$$
$$= 1 - 2a + 2a^2(v_1 + v_2 - 1)/(v_1 + v_2) - 4a^3(v_1 + v_2 - 1)(v_1 + v_2 - 2)/(3(v_1 + v_2)^2) + \cdots$$
$$= 1 - 2a + S_1 + S_2 + \cdots$$

with $S_i > T_i$, $i = 1, 2$ if $a < 0$ (which is the case of interest here). So, $(1 - 2a/v_1)^{v_1}(1 - 2a/(v_1 + v_2))^{-(v_1+v_2)} < 1$ and likewise expression (14) is less than unity and, irrespective of the value of a (<0), imposing valid prior information produces a gain in risk over simply ignoring the second sample. Accordingly, there is a region of the ψ-space over which s_A^2 dominates s_N^2 and also a region for which the converse occurs.

(v) Bancroft (1944) and Toyoda and Wallace (1975) show that under quadratic loss there always exists a range of ψ values over which pre-testing is the preferred strategy. They find that there is a family of pre-test estimators with $c \in (0, 2)$ which strictly dominate the never-pool estimator and dominate the always-pool estimator for a wide range of ψ: s_A^2 has smaller risk than this family of pre-test estimators only around the neighbourhood of the null hypothesis. Ohtani and Toyoda (1978) prove that the pre-test estimator with $c = 1$ strictly dominates all other members with $c \in (0, 2)$. So, under quadratic loss, there exists a ψ-range over which it is preferable to pre-test and to use $c = 1$.

A question of interest is whether this result carries over in general to the LINEX family of loss functions. This can be answered by considering the first and second derivatives of $R(s_P^2)$ with respect to c. Theorem 2 derives the values of c for which $\partial R(s_P^2)/\partial c = 0$.

Theorem 2. *$\partial R(s_P^2)/\partial c = 0$ when $c = 0$, ∞ and 1.*

Proof. See the appendix.

It is straightforward to show that as $c \to 0$ or $c \to \infty$, $\partial^2 R(s_P^2)/\partial c^2 = 0$, so that these two critical values result in points of inflexion of the risk function. In the case of quadratic loss it can also be shown that $\partial^2 R(s_P^2)/\partial c^2 > 0$ when $c = 1$ so that this choice of critical value always results in a minimum of the risk function. However, this cannot be shown in general under a LINEX loss function: it is not possible to sign $\partial^2 R(s_P^2)/\partial c^2$ for all values of a. When the degree of asymmetry is sufficiently large (depending on the values of the other arguments) this second derivative can be negative so that in these cases the choice of $c = 1$ results in a maximum of the pre-test risk function. We now turn to the numerical evaluations of $R(s_A^2)$, $R(s_N^2)$ and $R(s_P^2)$ which will illustrate this last result.

3. Numerical evaluations of the risk functions

We have numerically evaluated the risk functions for various values of α, v_1, v_2, ψ and a (<0). In particular, we consider $v_1 = 6, 10, 20, 30$; $v_2 = 6, 10, 20, 30$; $\alpha = 0, 0.01, 0.05, 0.75, 1$ and that value corresponding to $c = 1$; $a = -0.5, -2.0, -5.0$ and $\psi \in (0, 1]$. The full details of the results relating to all of these cases are available on request.

We have computed the *exact* risk functions on a VAXstation 4000 using a FORTRAN program which incorporates Davies' (1980) algorithm to evaluate the central F probabilities and various other algorithms from Press et al. (1986). The infinite series in (10) converge rapidly with a convergence tolerance of 10^{-6}. We also obtained the corresponding risks under a quadratic loss using equations (11)–(13) though for comparability with the LINEX results we scaled the quadratic results by $a^2/2$.

Unfortunately, we found that our algorithm failed in some cases for high degrees of asymmetry, and in these cases we undertook a Monte Carlo experiment with 5000 replications using the SHAZAM econometrics package (SHAZAM, 1993) on a VAXstation 4000. For the Monte Carlo experiment we assumed $\sigma_1^2 = 1$ so that $\sigma_2^2 = 1/\psi$, and we generated approximate χ^2 random variables to obtain s_1^2 and s_2^2 (this was also undertaken in SHAZAM using the normal random number generator proposed by Brent (1974)). For this particular problem it is not necessary to assign values to the regressors or to the coefficients. Where possible we compared the risks generated from the Monte Carlo experiment with those from the exact evaluations. These comparisons suggested that 5000 replications were sufficient to replicate the exact results to at least three decimal places.

Typical LINEX risk results are illustrated in Figs. 1–3 for $v_1 = 16$ and $v_2 = 8$ with $a = -0.5, -2.0$ and -5.0, respectively. The loss function when $a = -0.5$ exhibits relatively little asymmetry and so, qualitatively, Fig. 1 is similar to that which would be observed under a quadratic loss function. The features discussed in the previous section are clearly evident. In particular, it is preferable to pool the samples when the null hypothesis is true, and the never-pool estimator is strictly dominated by a family of pre-test estimators. The estimator which uses a critical value of unity has the smallest risk of this family.

As discussed in Section 2, *ceteris paribus* there exists a degree of asymmetry such that this latter feature does not occur. Fig. 3, with $a = -5.0$ clearly illustrates such an example. Then we find that the pre-test estimator which uses $c = 1$ has the highest risk of all the considered estimators around the region of the null hypothesis — this estimator no longer strictly dominates the never-pool estimator.

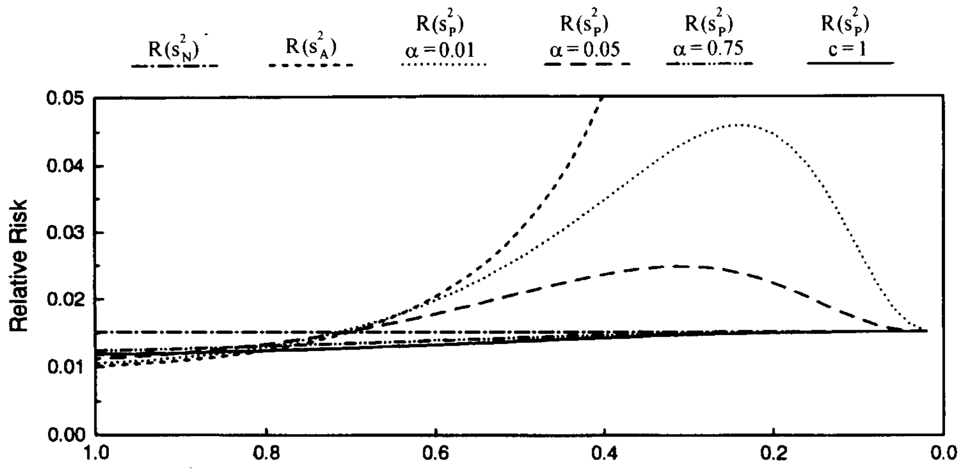

Fig. 1. Relative risk functions $a = -0.5$, $v_1 = 16$, $v_2 = 8$.

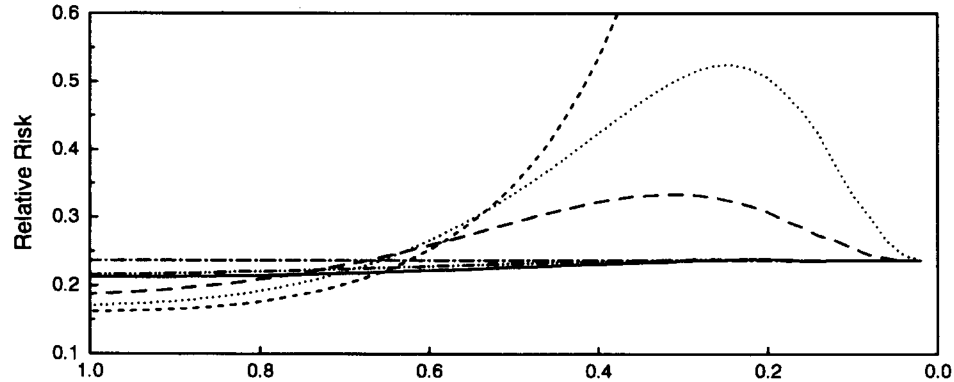

Fig. 2. Relative risk functions $a = -2.0$, $v_1 = 16$, $v_2 = 8$.

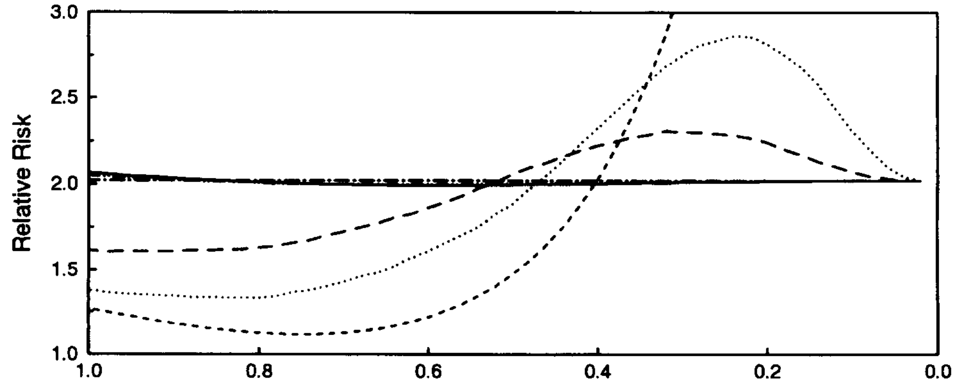

Fig. 3. Relative risk functions $a = -5.0$, $v_1 = 16$, $v_2 = 8$.

Our results also show that the range of ψ over which we prefer the always-pool estimator increases as the degree of asymmetry increases. This suggests that when considering an asymmetric loss function with a reasonable belief that the null hypothesis is true, it is generally preferable to pool the samples without testing.

4. Concluding remarks

In this paper we have extended some well known risk results, associated with pre-testing for variance homogeneity in regression prior to pooling sub-sample information, to the realistic situation where the underlying loss structure is asymmetric. In particular, risk under quadratic loss is generalised to risk under LINEX loss, with under-estimation of the regression scale being penalised more heavily than over-estimation.

This generalisation of the loss structure produces at least two results which differ in an important way from their quadratic loss counterparts. First, in the latter case, pre-testing with a critical value of unity is always preferred to ignoring the prior information. This does not hold if a sufficiently asymmetric LINEX loss is adopted.

Second, the range of ψ over which we prefer the always pool estimator increases as the degree of loss asymmetry increases. This suggests that it may be preferable to use this estimator, rather than undertake a pre-test, if we have a sufficiently asymmetric LINEX loss function. This contrasts with the typical advice under quadratic loss, which is to pre-test using a critical value of unity.

Much remains to be done to determine the sensitivity of established results in the pre-test literature to departures from the usual assumption of quadratic loss. Our results and those of Giles and Giles (1993a) and Giles (1993) suggest that there is less robustness to asymmetric departures than to symmetric ones.

Appendix

Proof of Theorem 1.

$$R(s_{\mathrm{N}}^2) = E(\exp(a(s_{\mathrm{N}}^2 - \sigma_1^2)/\sigma_1^2) - a(s_{\mathrm{N}}^2 - \sigma_1^2)/\sigma_1^2 - 1$$
$$= E(\exp(az_1/v_1 - a)) - E(az_1/v_1 - a) - 1, \tag{A.1}$$

where $z_1 = u_1' M_1 u_1/\sigma_1^2 \sim \chi_{v_1}^2$. Then $E(az_1/v_1 - a) = 0$, as s_{N}^2 is an unbiased estimator of σ_1^2. Further,

$$E(\exp(az_1/v_1 - a)) = \int_0^\infty \mathrm{e}^{az_1/v_1 - a} f(z_1)\,\mathrm{d}z_1,$$

where

$$f(z_1) = (2^{v_1/2}\Gamma(v_1/2))^{-1} z_1^{v_1/2-1}\mathrm{e}^{-z_1/2}.$$

So,

$$E(\exp(az_1/v_1 - a)) = (2^{v_1/2}\Gamma(v_1/2))^{-1}\mathrm{e}^{-a}\int_0^\infty \mathrm{e}^{-z_1(-a/v_1+v_2)}z_1^{v_1/2-1}\,\mathrm{d}z_1.$$

Using the change of variable $t = z_1(-a/v_1 + 1/2)$ we have[3]

$$E(\exp(az_1/v_1 - a)) = \frac{\mathrm{e}^{-a}}{\Gamma\left(\frac{v_1}{2}\right)2^{v_1/2}}\frac{(2v_1)^{v_1/2}}{(v_1-2a)^{v_1/2}}\int_0^\infty \mathrm{e}^{-t}t^{v_1/2-1}\,\mathrm{d}t$$

$$= \frac{\mathrm{e}^{-1}v_1^{v_1/2}}{(v_1-2a)^{v_1/2}}.$$

Substituting these results into (A.1), $R(s_{\mathrm{N}}^2)$ follows.

$$R(s_{\mathrm{N}}^2) = E[\exp(a(s_{\mathrm{A}}^2-\sigma_1^2)/\sigma_1^2 - a(s_{\mathrm{A}}^2-\sigma_1^2)/\sigma_1^2 - 1]$$

$$= E\{\exp(a(w_1+w_2/\psi)/(v_1+v_2)-1) - a((w_1+w_2/\psi)/(v_1+v_2)-1)-1\}, \tag{A.2}$$

where

$$w_i = e^{*\prime}M_i^*e^*/\sigma_2^2,\quad i=1,2;\qquad e^* = [e_1'/\sqrt{\psi}\ e_2'])' \sim N(0,\sigma_2^2 I_T),$$

$$T = T_1+T_2;\qquad \underset{(T\times T)}{M_1^*} = \begin{bmatrix} M_1 & 0\\ 0 & 0\end{bmatrix};\qquad \underset{(T\times T)}{M_2^*} = \begin{bmatrix} 0 & 0\\ 0 & M_2\end{bmatrix}.$$

It is straightforward to show that $w_i \sim \chi_{v_i}^2$ and that w_1 and w_2 are independent. So,

$$E(w_1+w_2/\psi)/(v_1+v_2) - 1 = v_2(1-\psi)/(\psi(v_1+v_2)) \tag{A.3}$$

and

$$E[\exp(a(w_1+w_2/\psi)/(v_1+v_2)-1)]$$

$$= \mathrm{e}^{-a}\int_0^\infty\int_0^\infty \mathrm{e}^{a(w_1+w_2/\psi)/(v_1+v_2)}f(w_1)f(w_2)\,\mathrm{d}w_1\,\mathrm{d}w_2$$

$$= \mathrm{e}^{-a}\left[\int_0^\infty \frac{1}{2^{v_1/2}\Gamma\left(\frac{v_1}{2}\right)}w_1^{v_1/2-1}\mathrm{e}^{-w_1/2+aw_1/(v_1+v_2)}\,\mathrm{d}w_1\right]$$

$$\times\left[\int_0^\infty \frac{1}{2^{v_2/2}\Gamma\left(\frac{v_2}{2}\right)}w_2^{v_2/2-1}\mathrm{e}^{-w_2/2+aw_2/(\psi(v_1+v_2))}\,\mathrm{d}w_2\right].$$

[3] When $a < 0$ (the case of particular interst here) we require no restrictions on its value for the risk functions to be positive. However, certain restrictions are required if $a > 0$ and depending on whether v_1 and/or v_2 are odd or even. As these restrictions are fairly obvious we do not specify them explicitly in this appendix.

Using the change in variables

$$t_1 = w_1[1/2 - a/(v_1 + v_2)]$$

and

$$t_2 = w_2[(\psi(v_1 + v_2) - 2a)/(2\psi(v_1 + v_2))]$$

we have

$$E[\exp(a(w_1 + w_2/\psi)/(v_1 + v_2) - 1)]$$

$$= e^{-a}\left[\frac{v_1 + v_2}{v_1 + v_2 - 2a}\right]^{v_1/2}\left[\frac{\psi(v_1 + v_2)}{\psi(v_1 + v_2) - 2a}\right]^{v_2/2}. \tag{A.4}$$

Substituting (A.3) and (A.4) into (A.2) yields $R(s_A^2)$.

Finally, for the pre-test estimator, s_P^2, we have

$$s_P^2 = s_N^2 + (s_A^2 - s_N^2)I_{[0,c]}(J)$$

$$= \sigma_2^2[\psi(v_1 + v_2)w_1 + (v_1 w_2 - \psi v_2 w_1)I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))]/(v_1(v_1 + v_2))$$

and

$$R(s_P^2) = E\{\exp[a(s_P^2 - \sigma_1^2)/\sigma_1^2] - a(s_P^2 - \sigma_1^2)/\sigma_1^2 - 1\}$$

$$= E\{\exp[a(s_P^2 - \psi\sigma_2^2)/(\psi\sigma_2^2)] - a(s_P^2 - \psi\sigma_2^2)/(\psi\sigma_2^2) - 1\}. \tag{A.5}$$

Now,

$$E[(s_P^2 - \psi\sigma_2^2)/(\psi\sigma_2^2)] = E[\psi(v_1 + v_2)w_1 + (v_1 w_2 - \psi v_2 w_1) \times I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))]/(\psi v_1(v_1 + v_2)) - 1 \tag{A.6}$$

and repeatedly using Lemma 1 of Clarke et al. (1987) (A.6) is

$$E[(s_P^2 - \psi\sigma_2^2)/(\psi\sigma_2^2)] = v_2(Q_{20} - \psi Q_{02})/(\psi(v_1 + v_2)) \tag{A.7}$$

where

$$Q_{mn} = \Pr[(F_{(v_2+m,\, v_1+n)} \leqslant (v_2(v_1 + n)c\psi)/(v_1(v_2 + m)]$$

$m, n = 0, 1, \ldots$

Further,

$$E[\exp(a(s_P^2 - \psi\sigma_2^2)/(\psi\sigma_2^2))]$$

$$= e^{-a}E\{\exp[a[\psi(v_1 + v_2)w_1 + (v_1 w_2 - \psi v_2 w_1)I_{[0,c\psi]}(v_1 w_2/(v_2 w_1)))/(\psi v_1(v_1 + v_2))]\}$$

$$= e^{-a}E\{\exp[b_0 w_1 + (b_2 w_2 - b_1 w_1)I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))]\}$$

$$= e^{-a}E(\exp(Q)), \tag{A.8}$$

where

$$b_0 = a/v_1; \qquad b_1 = av_2/(v_1(v_1+v_2)); \qquad b_2 = a/(\psi(v_1+v_2));$$

$$Q = \exp[b_0 w_1 + (b_2 w_2 - b_1 w_1) I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))]$$
$$= \exp[b_0 w_1 - b_0 w_1 I_{[0,c\psi]}(v_1 w_2/(v_2 w_1)) + b_2 w_2 I_{[0,c\psi]} \times (v_1 w_2/(v_2 w_1)) + (b_0 - b_1) w_1 I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))]. \tag{A.9}$$

Now,

$$E(\exp(Q)) = 1 + E(Q) + \frac{E(Q)^2}{2!} + \frac{E(Q)^3}{3!} + \cdots \tag{A.10}$$

and using Lemma 1 of Clarke (1990) we have

$$E(w_1^{r_1} I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))) = \frac{2^{r_1}\Gamma\left(\frac{v_1+2r_1}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)} Q_{v_2, v_1+2r_1},$$

$$E(w_2^{r_2} I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))) = \frac{2^{r_2}\Gamma\left(\frac{v_2+2r_2}{2}\right)}{\Gamma\left(\frac{v_2}{2}\right)} Q_{v_2+2r_2, v_1},$$

$$E(w_1^{r_1} w_2^{r_2} I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))) = 2^{r_1+r_2}\frac{\Gamma\left(\frac{v_1+2r_1}{2}\right)\Gamma\left(\frac{v_2+2r_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} Q_{v_2+2r_2, v_1+2r_1},$$

where r_1 and r_2 are any real values such that $r_1 > (-v_1/2)$ and $r_2 > (-v_2/2)$.

Using these expressions repeatedly along with definition (A.9) in (A.10) we have

$$E[\exp(Q)]$$
$$= 1 + 2\left\{b_0 \Gamma\left(\frac{v_1+2)}{2}\right)(1 - Q_{02})\Big/\Gamma\left(\frac{v_1}{2}\right) + b_2\Gamma\left(\frac{v_2+2}{2}\right)Q_{20}\Big/\Gamma\left(\frac{v_2}{2}\right) + (b_0 - b_1)\Gamma\left(\frac{v_1+2}{2}\right)Q_{02}\Big/\Gamma\left(\frac{v_1}{2}\right)\right\}$$
$$+ (2^2/2!)\left\{b_0^2\Gamma\left(\frac{v_1+4}{2}\right)(1 - Q_{04})\Big/\Gamma\left(\frac{v_1}{2}\right) + b_2^2\Gamma\left(\frac{v_2+4}{2}\right)Q_{40}\Big/\Gamma\left(\frac{v_2}{2}\right)\right.$$

$$+ 2b_2(b_0 - b_1)\Gamma\left(\frac{v_1+2}{2}\right)\Gamma\left(\frac{v_2+2}{2}\right)Q_{22}\bigg/\left(\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)\right)$$

$$+ (b_0 - b_1)^2\,\Gamma\left(\frac{v_1+4}{2}\right)Q_{04}\bigg/\Gamma\left(\frac{v_1}{2}\right)\Bigg\}$$

$$+ (2^2/3!)\Bigg\{b_0^3\,\Gamma\left(\frac{v_1+6}{2}\right)(1 - Q_{06})\bigg/\Gamma\left(\frac{v_1}{2}\right) + b_2^3\,\Gamma\left(\frac{v_2+6}{2}\right)Q_{60}\bigg/\Gamma\left(\frac{v_2}{2}\right)$$

$$+ 3b_2^2(b_0 - b_1)\Gamma\left(\frac{v_1+2}{2}\right)\Gamma\left(\frac{v_2+4}{2}\right)Q_{42}\bigg/\left(\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)\right)$$

$$+ 3b_2(b_0 - b_1)^2\,\Gamma\left(\frac{v_1+4}{2}\right)\Gamma\left(\frac{v_2+2}{2}\right)Q_{24}\bigg/\left(\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)\right)$$

$$+ (b_0 - b_1)^3\,\Gamma\left(\frac{v_1+6}{2}\right)Q_{06}\bigg/\Gamma\left(\frac{v_1}{2}\right)\Bigg\} + \cdots$$

$$= \sum_{i=0}^{\infty}\frac{(2b_0)^i}{i!}\,\frac{\Gamma\left(\frac{(v_1+2i}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)} - \sum_{i=1}^{\infty}\frac{(2b_0)^i}{i!}\,\frac{\Gamma\left(\frac{v_1+2i}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)}Q_{0(2i)} + \sum_{i=1}^{\infty}\frac{2^i}{i!}L'_{1i}L_{2i},$$

where L'_{1i} is a $(1\times(i+1))$ vector equal to the $(i+1)$th row of Pascal's Triangle; L_{2i} is an $((i+1)\times 1)$ vector with elements

$$\ell_j = b_2^{i-j}(b_0 - b_1)^j\,\Gamma\left(\frac{v_1+2j}{2}\right)\Gamma\left(\frac{v_2+2(i-j)}{2}\right)Q_{(2(i-j))(2j)}\bigg/\left(\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)\right)$$

$$= (a/(v_1+v_2))^i\,\psi^{j-i}\,\Gamma\left(\frac{v_1+2j}{2}\right)\Gamma\left(\frac{v_2+2(i-j)}{2}\right)Q_{(2(i-j))(2j)}\bigg/\left(\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)\right),$$

$$j = 0, 1, \ldots, i.$$

Finally, if $|2a/v_1| < 1$, which is not restrictive in practice, then

$$\sum_{i=1}^{\infty}\frac{(2b_0)^i}{i!}\,\frac{\Gamma\left(\frac{v_1+2i}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)} = \left(\frac{v_1}{v_1-2a}\right)^{v_1/2}.$$

So,

$$E(\exp(Q)) = \left(\frac{v_1}{v_1-2a}\right)^{v_1/2} - \sum_{i=1}^{\infty}\frac{(2a/v_1)}{i!}\,\frac{\Gamma\left(\frac{v_1+2i}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)}Q_{0(2i)} + \sum_{i=1}^{\infty}\frac{2^i}{i!}L'_{1i}L_{2i}. \qquad \text{(A.11)}$$

Substituting (A.11) into (A.8) and then this expression and (A.7) into (A.5) yields the desired result. □

Proof of Theorem 2. Using the infinite series expansion of the exponential function we write

$$
\begin{aligned}
R(s_{\mathrm{P}}^2) &= E\left[\frac{a^2}{2!}((s_{\mathrm{P}}^2 - \sigma_1^2)/\sigma_1^2)^2 + \frac{a^3}{3!}((s_{\mathrm{P}}^2 - \sigma_1^2)/\sigma_1^2)^3 + \cdots\right] \\
&= E\left[\frac{a^2}{2!}([\psi(v_1 + v_2)w_1 - \psi v_1(v_1 + v_2)]\right. \\
&\qquad \times (v_1 w_2 - \psi v_2 w_1) I_{[0,c\psi]}(v_1 w_2/(v_2 w_1)))^2/(\psi v_1(v_1 + v_2))^2 \\
&\qquad + \frac{a^3}{3!}([\psi(v_1 + v_2)w_1 - \psi v_1(v_1 + v_2)] \\
&\qquad \left. \times (v_1 w_2 - \psi v_2 w_1) I_{[0,c\psi]}(v_1 w_2/(v_2 w_1)))^3/(\psi v_1(v_1 + v_2))^3 + \cdots\right] \\
&= E[\Delta^* + (v_1 w_2 - \psi v_2 w_1)\Phi I_{[0,c\psi]}(v_1 w_2/(v_2 w_1))] \qquad \text{(A.12)}
\end{aligned}
$$

where

$$
\Delta^* = \frac{a^2}{2!}\Delta^2 + \frac{a^3}{3!}\Delta^3 + \cdots; \qquad \Delta = \psi(v_1 + v_2)(w_1 - v_1);
$$

$$
\begin{aligned}
\Phi = {} & \frac{a^2}{2!}(2\Delta + (v_1 w_2 - \psi v_2 w_1)) + \frac{a^3}{3!}((v_1 w_2 - \psi v_2 w_1)^2 \\
& + 3\Delta^2 + 3\Delta(v_1 w_2 - \psi v_2 w_1)) + \cdots
\end{aligned}
$$

Now, $I_{[0,c\psi]}(v_1 w_2/(v_2 w_1)) = I_{[0,x]}(w_2)$ where $x = c\psi v_2 w_1/v_1$, and as w_1 and w_2 are independently distributed quadratic forms it follows that (A.12) can be written as

$$
\begin{aligned}
R(s_{\mathrm{P}}^2) &= E_{w_1}\{\Delta^* + E_{w_2}[(v_1 w_2 - \psi v_2 w_1)\Phi I_{[0,x]}(w_2)]\} \\
&= E_{w_1}\left\{\Delta^* + \int_0^x (v_1 w_2 - \psi v_2 w_1)\Phi f(w_2)\,\mathrm{d}w_2\right\},
\end{aligned}
$$

where $f(\cdot)$ is the density function of a $\chi^2_{v_2}$ variate. Then,

$$
\begin{aligned}
\frac{\partial R(s_p^2)}{\partial c} &= E_{w_1}\left\{\frac{\partial x}{\partial c} \times \frac{\partial}{\partial x}\int_0^x (v_1 w_2 - \psi v_2 w_1 \Phi f(w_2)\,\mathrm{d}w_2\right\} \\
&= E_{w_1}\{[(\psi w_1 v_2)^2/v_1](c-1) f(c\psi v_2 w_1/v_1)\Phi^*\}, \qquad \text{(A.13)}
\end{aligned}
$$

where Φ^* is Φ evaluated at $w_2 = c\psi v_2 w_1/v_1$. (A.13) will clearly be zero when $c = 0$, ∞ and 1. □

Acknowledgements

The authors acknowledge the financial support provided by University of Canterbury Research Grants #1770901, 1787147 and 1894183. We are also grateful to Mathew Cunneen and Jason Wong for providing research assistance with some of the numerical evaluations, to Kazuhiro Ohtani for many useful discussions on non-quadratic loss functions, and to Viren Srivastava and two anonymous referees for helpful suggestions and comments.

References

Bancroft, T.A. (1944). On biases in estimation due to the use of preliminary tests of significance. *Ann. Math. Statist.* **15**, 190–204.

Bancroft, T.A. and C.-P. Han (1983). A note on pooling variances. *J. Amer. Statist. Assoc.* **78**, 981–983.

Brent, R.P. (1974). A Gaussian pseudo-random number generator. *Commun. ACM* **17**, 704–706.

Clarke, J.A. (1990). Preliminary-test estimation of the standard error of estimate in linear regression. *Economics Lett.* **34**, 27–32.

Clarke, J.A., D.E.A. Giles and T.D. Wallace (1987). Estimating the error variance in regression after a preliminary test of restrictions on the coefficients. *J. Econometrics* **34**, 293–304.

Davies, R.B. (1980). The distribution of a linear combination of χ^2 random variables (Algorithm AS 155). *Appl. Statist.* **29**, 323–333.

Giles, D.E.A. (1993). Pre-test estimation in regression under absolute error loss. *Economics Letters* **41**, 339–343.

Giles, J.A. (1992). Estimation of the error variance after a preliminary-test of homogeneity in a regression model with spherically symmetric disturbances. *J. Econometrics* **53**, 345–361.

Giles, J.A. and D.E.A. Giles (1993a). Preliminary-test estimation of the regression scale parameter when the loss function is asymmetric. *Commun. Statist.* **22**, 1709–1734.

Giles, J.A. and D.E.A. Giles (1993b). Pre-test estimation and testing in econometrics: Recent developments. *J. Economic Surveys* **7**, 145–197.

Greenberg, E. (1980). Finite sample moments of a preliminary-test estimator in the case of possible heteroscedasticity. *Econometrica* **48**, 1805–1813.

Kuo, L. and D.K. Dey (1990). On the admissibility of the linear estimators of the Poisson mean using LINEX loss functions. *Statist. Decisions* **8**, 201–210.

Ohtani, K. and T. Toyoda (1978). Minimax regret critical values for a preliminary test in pooling variance. *J. Japan Statist. Soc.* **17**, 15–20.

Parsian, A. (1990a). On the admissibility of an estimator of a normal mean vector under a LINEX loss function. *Ann. Institute Math. Statist.* **42**, 657–669.

Parsian, A. (1990b). Bayes estimation using a LINEX loss function. *J. Sci. IROI* **1**, 305–307.

Parsian, A., N. Sanjari Farsipour and N. Nematollahi (1992). On the minimaxity of Pitman type estimator under a LINEX loss function. Mimeograph, Department of Statistics, Shiraz University, Shiraz, Iran.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling (1986). *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, New York.

Rojo, J. (1987). On the admissibility of $c\bar{X} + d$ with respect to the LINEX loss function. *Commun. Statist.* **A16**, 3745–3748.

Sadooghi-Alvandi, S.M. (1990). Estimation of the parameter of a Poisson distribution using a LINEX loss function. *Australian J. Statist.* **32**, 393–398.

Sadooghi-Alvandi, S.M. and N. Nematollahi (1989). A note on the admissibility of $c\bar{X} + d$ relative to the LINEX loss function. *Comm. Statist.* **A18**, 1871–1873.

Sadooghi-Alvandi, S.M. and A. Parsian (1992). Estimation of the binomial parameter n using a LINEX loss function. *Commun. Statist.* **A18**, forthcoming.

SHAZAM (1993). *SHAZAM Econometrics Computer Program: User's Reference Manual Version 7.0*, McGraw-Hill, New York.

Srivastava, V.K. and B.B. Rao (1992). Estimation of disturbance variance in linear regression models under asymmetric criterion. *J. Quantitative Economics* **8**, 341–345.

Taylor, W.E. (1977). Small sample properties of a class of two stage Aitken estimators. *Econometrica* **45**, 497–508.

Taylor, W.E. (1978). The heteroscedastic linear model: Exact finite sample results. *Econometrica* **46**, 663–675.

Toyoda, T. and T.D. Wallace (1975). Estimation of variance after a preliminary test of homogeneity and optimal levels of significance for the pre-test. *J. Econometrics* **3**, 395–404.

Varian, H.R. (1975). A Bayesian approach to real estate assessment. In: S.E. Fienberg and A. Zellner, Eds., *Studies in Bayesian Econometrics and Statistics In Honor of Leonard J. Savage*, North-Holland, Amsterdam, 195–208.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *J. Amer. Statist. Assoc.* **81**, 446–451.