Estimating the error variance after a pre-test for an inequality restriction on the coefficients

Alan T.K. Wan

Department of Applied Statistics & OR, City University of Hong Kong, Kowloon, Hong Kong

Received 14 June 1994; revised 3 February 1995

Abstract

Estimation of the regression error variance after a preliminary test of an inequality constraint on the coefficient vector is considered. We derive the exact finite sample risk of several inequality restricted and pre-test estimators of $\sigma^2$. These estimators are associated with the maximum likelihood, least squares and minimum mean squared error component estimators. Optimal critical values for the pre-test according to a mini-max regret criterion are numerically calculated. Furthermore, we examine the robustness of the optimal choice of critical values and the risk properties of the estimators of $\sigma^2$ to model mis-specification through the exclusion of relevant regressors.

AMS Classification: Primary 62J05

Keywords: Error variance; Inequality restriction; Mini-max regret; Pre-test

1. Introduction

In the literature on pre-test estimation in the linear regression model, most discussions focus on the estimators of the coefficient or prediction vectors. By way of comparison, the estimation of the regression error variance has received much less attention. As argued in a recent survey paper by Giles and Giles (1993a), this is not surprising as $\sigma^2$ is often regarded as a nuisance parameter when interest centers on $\beta$. However, in practical terms, the estimation of $\sigma^2$ is necessary if the researcher is interested in the analysis of the precision of the estimators of $\beta$, or if hypothesis tests are to be carried out.

When the pre-test in question involves a set of linear equality restrictions on the coefficients, the inferential complications created by pre-testing for the properties of the resulting estimator of $\sigma^2$ are well known. See, for example, Clarke et al. (1987a, b), Ohtani (1988), Giles and Clarke (1989), Giles (1990, 1991a, b), Giles and Lieberman
This literature focuses on the risk properties of pre-test estimators of \( \sigma^2 \), the robustness of these properties to model mis-specification, as well as the choice of an optimal size of the pre-test. Relatively less known, however, are the properties of the estimators of \( \sigma^2 \) that result when a pre-test is undertaken to check the validity of a hypothesis represented by one or more inequality restrictions on the coefficient vector. Wan (1994a), among other things, investigates the properties of the estimator of \( \sigma^2 \) after a pre-test of an inequality restriction on the coefficients when estimation is based on the maximum likelihood method. He also considers the robustness of the properties of this pre-test estimator to mis-specification through the omission of relevant regressors from the model. A related problem of estimating the error variance after a one-sided pre-test of the mean is considered by Ohtani (1991).

The aim of this paper is twofold. First, we extend the analysis of Wan (1994a) on the risk properties of several inequality restricted and pre-test estimators of \( \sigma^2 \) in a model where relevant regressors are unwittingly omitted. These estimators are associated with the maximum likelihood, least squares and minimum mean squared error component estimators. We find that with the latter two component estimators, regardless of the degree of model mis-specification, there exists certain levels of pre-test such that the risk of the pre-test estimator is uniformly smaller than that of the corresponding unrestricted estimator. This contrasts with the results obtained when the method of estimation is based on the maximum likelihood principle, as examined in Wan (1994a). The situation where the model being correctly specified is nested as a special case in our analysis. Second, we consider the choice of an optimal critical value for a pre-test of an inequality restriction when estimating \( \sigma^2 \), and examine the way in which the omission of relevant regressors affects this choice.

2. The model and estimators

Consider the data generating process

\[
y = X\beta + Z\eta + \varepsilon; \quad \varepsilon \sim N(0, \sigma^2 I)
\]  

where \( y \) and \( \varepsilon \) are \( n \times 1 \) vectors; \( X \) and \( Z \) are nonstochastic matrices of full column rank and are \( n \times k \) and \( n \times p \) respectively; \( \beta \) and \( \eta \) are unknown coefficient vectors of dimensions \( k \times 1 \) and \( p \times 1 \) respectively. We assume, however, that the set of regressors \( Z \) is mistakenly omitted from the model. The fitted model is therefore

\[
y = X\beta + \mu; \quad \mu \sim N(Z\eta, \sigma^2 I).
\]  

As the researcher is unaware of the mis-specification in the model, \( E(\mu) \) is incorrectly assumed to be zero. The prior information available to the researcher is represented by

\[
H_0: C'\beta \geq r,
\]  

where \( C' \) is a \( 1 \times k \) vector and \( r \) is a known scalar.
Following Judge and Yancey (1981, 1986), (1)-(3) can be reparameterized as

\[ y = H\theta + B\pi + \varepsilon \]  
\[ y = H\theta + \mu \]  

and

\[ H_0: \theta_1 \geq r_0 \]

respectively, where \( H = XS^{-1/2}Q' \);  
\( B = ZT^{-1/2}V' \);  
\( \theta = QS^{1/2}\beta \);  
\( \pi = VT^{1/2}\eta \);  
\( S = X'X \);  
\( T = Z'Z \);  
\( V'V = I_p \);  
\( Q \) is an orthogonal matrix such that

\[
QS^{-1/2}C(C'S^{-1}C)^{-1}C'S^{-1/2}Q' = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},
\]

\( r_0 \) is a (positive) scalar multiple of \( r \); and \( \theta_1 \) is the first element in \( \theta \).

Under (5), the unrestricted estimator of \( \theta \) is \( \tilde{\theta} = H'y \). The equality restricted estimator, which takes into account the restriction \( \theta_1 = r_0 \), is

\[
\theta^* = (r_0, \tilde{\theta}_{(k-1)})',
\]

where

\[
\tilde{\theta}_{(k-1)} = (0, I_{(k-1)})\tilde{\theta}.
\]

The corresponding unrestricted estimator of \( \sigma^2 \) is \( \hat{\sigma}^2 = \tilde{\sigma}^2/(n + \delta) \) and the equality restricted estimator of \( \sigma^2 \) is \( \sigma^*^2 = e^*e^*/(n + \gamma) \), where \( \tilde{\sigma} \) and \( e^* \) are the vectors of residuals corresponding to the use of \( \tilde{\theta} \) and \( \theta^* \) of \( \theta \) in model (5) respectively. The maximum likelihood (ML) estimators of \( \sigma^2 \) correspond to \( \delta = \gamma = 0 \). The least squares (LS) estimators of \( \sigma^2 \) correspond to \( \delta = -k \) and \( \gamma = -k + 1 \), while the minimum mean squared error (MM) estimators correspond to \( \delta = -k + 2 \) and \( \gamma = -k + 3 \).

If the unrestricted estimator of \( \theta \) satisfies (6), then the restriction is believed to be nonbinding and \( \sigma^2 \) is estimated by \( \hat{\sigma}^2 \), otherwise the restriction is treated as binding and \( \sigma^2 \) is estimated by \( \sigma^*^2 \). This procedure gives rise to the following inequality restricted estimator:

\[
\sigma^{*^2} = \begin{cases} 
\hat{\sigma}^2 & \text{if } \tilde{\theta}_1 \geq r_0 \\
\sigma^*^2 & \text{if } \tilde{\theta}_1 < r_0 
\end{cases}
\]

where \( I_{(-\infty, r_0]}(\tilde{\theta}_1) \) is an indicator function which is 1 if \( \tilde{\theta}_1 \) lies in the subscripted interval and zero otherwise. Using the fact that \( I_{(r_0, \infty)}(\tilde{\theta}_1) = 1 - I_{(-\infty, r_0]}(\tilde{\theta}_1) \) and after some manipulations, we can write

\[
\sigma^{*^2} = \hat{\sigma}^2 + I_{(-\infty, r_0]}(u_1)[(\sigma u_1 - \tau)^2 - \hat{\sigma}^2(\gamma - \delta)]/(n + \gamma),
\]

where \( u_1 = (\tilde{\theta}_1 - \theta_1)/\sigma \) is normally distributed with mean \( \zeta/\sigma \), where \( \zeta = (H'B\pi)_1 \) is the first element of \( H'B\pi \), and \( \tau = r_0 - \theta_1 \).
Often the investigator is uncertain of the validity of the inequality restriction and so may conduct a pre-test of $H_0: \theta_1 \geq r$. As the investigator is unaware of the misspecification in the model, $H_0$ is tested using the statistic $t'' = \sqrt{n(\bar{\theta}_1 - r_0)}\hat{\sigma}^{-1}/\sqrt{n+\delta}$, which has a doubly noncentral $t$ distribution with $v = n - k$ degrees of freedom and noncentrality parameters $\lambda_1^2 = (r - \xi)^2/2\sigma^2$ and $\lambda_2^2 = n'B'(I - HH')B\pi/2\sigma^2$. We reject the null hypothesis if $t'' \leq c$, where $c$ is the size-$\alpha$ critical value for the central $t$ variate with $v$ degrees of freedom, and use $\hat{\sigma}^2$. We do not reject $H_0$ if $t > c$ and use $\sigma^{**2}$. This mechanism gives rise to the inequality pre-test estimator

$$
\hat{\sigma}^2 = \begin{cases} 
\hat{\sigma}^2 & \text{if } t'' \leq c \\
\sigma^{**2} & \text{if } t'' > c
\end{cases} = I_{(-\infty, c)}(t'')\hat{\sigma}^2 + I_{(c, \infty)}(t'')\sigma^{**2}. \quad (9)
$$

Using the properties of indicator functions, recognizing that

$$
I_{(-\infty, c)}(t''_i)I_{(-\infty, \tau/0)}(u_1) = I_{(-\infty, (c' + \delta)/0)}(u_1),
$$

and after some manipulations, we can write

$$
\hat{\sigma}^2 = \sigma^{**2} - I_{(-\infty, (c' + \delta)/0)}(u_1)[((\sigma u_1 - \tau)^2 - \hat{\sigma}^2(\gamma - \delta))/(n + \delta)], \quad (10)
$$

where $c' = c\sqrt{(n + \delta)/v}$.

The sampling properties of the inequality restricted and pre-test estimators of $\beta$ have been examined rather thoroughly in the literature. See, for example, Thomson and Schmidt (1982), Judge and Yancey (1981, 1986), Hasegawa (1989) and Wan (1994a, b, 1995). Wan (1994a) also considers the risks of $\sigma^{**2}$ and $\hat{\sigma}^2$ (under squared error loss) for the case where the component estimators are the maximum likelihood estimators (i.e. $\gamma = \delta = 0$). In the next section, we examine the risks of $\sigma^{**2}$ and $\hat{\sigma}^2$ under the more general case given by expressions (8) and (10).

3. The risk properties of $\sigma^{**2}$ and $\hat{\sigma}^2$

For an estimator $\hat{\sigma}^2$ of $\sigma^2$, the risk under squared error loss is given by $\rho(\hat{\sigma}^2, \sigma^2) = E[(\hat{\sigma}^2 - \sigma^2)^2]/\sigma^2$. Now, from Giles (1990),

$$
\rho(\hat{\sigma}^2, \sigma^2) = [2(v + 4\lambda_2) + (v + 2\lambda_2 - (n + \delta))^2]/(n + \delta)^2 \quad (11)
$$

and

$$
\rho(\sigma^{**2}, \sigma^2) = [2[1 + v + 4(\lambda_1^2 + \lambda_2)] + [1 - k + 2(\lambda_1^2 + \lambda_2) - \gamma]^2]/(n + \gamma^2). \quad (12)
$$

Under the stated assumptions, the risk of $\sigma^{**2}$ and $\hat{\sigma}^2$ are given by the following theorem, the derivation of which is outlined in Appendix A.
Theorem 1. When $\lambda_1 \leqslant 0$,
\[
\rho(\sigma^{**2}, \sigma^2) = \rho(\delta^2, \sigma^2) + \{2\lambda_1^2(n + \delta)^2[2\lambda_1^2 - 2(n + \gamma)] \\
+ 2(v + 2\lambda_2)(n + \delta)[2\lambda_1^2(n + \delta) + (n + \gamma)(\gamma - \delta)] \\
- [v(v + 2) + 4\lambda_2^2 + 4\lambda_2 v + 8\lambda_2] \\
\times (\gamma - \delta)(2n + \gamma + \delta)] P_1 / 2[(n + \gamma)(n + \delta)]^2 \\
+ [8\lambda_1^3 + 4\lambda_1 v + 8\lambda_2 \lambda_2 - 4\lambda_1(n + \gamma)] P_2 / [\Gamma(1/2)(n + \gamma)^2] \\
+ [6\lambda_1^2 + v - (n + \gamma) + 2\lambda_2] P_3 / (n + \gamma)^2 \\
+ 8\lambda_1 P_4 / [\Gamma(1/2)(n + \gamma)] + 3P_5 / (2(n + \gamma)^2)
\] (13)

\[
\rho(\delta^2, \sigma^2) = \rho(\sigma^{**2}, \sigma^2) - 3E_{5,v}(2(n + \gamma)) - 8\lambda_1 E_{4,v} / [\Gamma(1/2)(n + \gamma)^2] \\
+ (n + \gamma - 6\lambda_1^2) \times E_{3,v} / (n + \gamma)^2 \\
+ [4\lambda_1(n + \gamma) - 8\lambda_1^3] E_{2,v} / [\Gamma(1/2)(n + \gamma)^2] \\
+ [2\lambda_1^2(n + \gamma) - \lambda_1^4] E_{1,v} / (n + \gamma)^2 - v E_{3,v+2} / (n + \gamma)^2 \\
- 4v\lambda_1 E_{3,v+2} / [\Gamma(1/2)(n + \gamma)^2] \\
- [v(n + \gamma)(v - \delta) + 2v\lambda_1^2(n + \delta)] E_{1,v+2} / [(n + \delta)(n + \gamma)^2] \\
- 2\lambda_2 E_{3,v+4} / (n + \gamma)^2 - 8\lambda_1 \lambda_2 E_{2,v+4} / [\Gamma(1/2)(n + \gamma)^2] \\
+ [v(v + 2)(\gamma - \delta)(2n + \gamma + \delta) / [2(n + \delta)(n + \delta)^2] \\
- 4\lambda_1^2 \lambda_2 / (n + \gamma)^2 - 2(\gamma - \delta)\lambda_2 / [(n + \delta)(n + \gamma)] E_{1,v+4} \\
+ (\gamma - \delta)(v + 2)2\lambda_2(2n + \gamma + \delta) E_{1,v+6} / [(n + \delta)(n + \gamma)]^2 \\
+ (\gamma - \delta)2\lambda_1^2(2n + \gamma + \delta) E_{1,v+8} / [(n + \delta)(n + \gamma)]^2
\] (14)

Alternatively, when $\lambda_1 > 0$,
\[
\rho(\sigma^{**2}, \sigma^2) = \rho(\delta^2, \sigma^2) + \{2\lambda_1^2(n + \delta)^2[2\lambda_1^2 - 2(n + \gamma)] \\
+ 2(v + 2\lambda_2)(n + \delta)[2\lambda_1^2(n + \delta) + (n + \gamma)(\gamma - \delta)] \\
- [v(v + 2) + 4\lambda_2^2 + 4\lambda_2 v + 8\lambda_2] \\
\times (\gamma - \delta)(2n + \gamma + \delta)] [1 - P_1] / [(n + \gamma)(n + \delta)]^2 \\
+ [8\lambda_1^3 + 4\lambda_1 v + 8\lambda_2 \lambda_2 - 4\lambda_1(n + \gamma)] P_2 / [\Gamma(1/2)(n + \gamma)^2] \\
+ [6\lambda_1^2 + v - (n + \gamma) + 2\lambda_2] (2 - P_3) / (n + \gamma)^2 \\
+ 8\lambda_1 P_4 / [\Gamma(1/2)(n + \gamma)^2] + (6 - 3P_5) / (2(n + \gamma)^2),
\] (15)
\[ \rho(\hat{\sigma}^2, \sigma^2) = \rho(\sigma^{**2}, \sigma^2) - 3[E_{5,v}/2 + G_{5,v}]/(n + \gamma) - 8\lambda_1 E_{4,v}/[\Gamma(\frac{1}{2})(n + \gamma)^2] \\
+ (n - 6\lambda_1^2)[E_{3,v} + 2G_{3,v}]/(n + \gamma)^2 \\
+ [4\lambda_1(n + \gamma) - 8\lambda_1^2]E_{2,v}/[\Gamma(\frac{1}{2})(n + \gamma)^2] \\
+ [2\lambda_1^2(n + \gamma) - \lambda_1^4]E_{1,v} + 2G_{1,v}]/(n + \gamma)^2 \\
- v(E_{3,v+2} + 2G_{3,v+2}]/(n + \gamma)^2 - 4\lambda_1 E_{2,v+2}/[\Gamma(\frac{1}{2})(n + \gamma)^2] \\
- [v(n + \gamma)(\gamma - \delta) + 2\lambda_1^2(n + \delta)] \\
\times [E_{1,v+2} + 2G_{1,v+2}]/[(n + \delta)(n + \gamma)^2] \\
- 2\lambda_2^2[E_{3,v+4} + 2G_{3,v+4}]/(n + \gamma)^2 - 8\lambda_1\lambda_2 E_{2,v+4}/[\Gamma(\frac{1}{2})(n + \gamma)^2] \\
+ \{v(v + 2)(\gamma - \delta)(2n + \gamma + \delta)/[2(n + \delta)(n + \gamma)^2] \\
- 4\lambda_1^2\lambda_2/(n + \gamma)^2 - (\gamma - \delta)\lambda_2/[n + \delta(n + \gamma)]\} \\
\times [E_{1,v+4} + 2G_{1,v+4} + (\gamma - \delta)(v + 2) \times 2\lambda_2(2n + \gamma + \delta)[E_{1,v+6} + 2G_{1,v+6}]/[(n + \delta)(n + \gamma)^2] \\
+ (\gamma - \delta)2\lambda_2^2(2n + \gamma + \delta)[E_{1,v+8} + 2G_{1,v+8}]/[(n + \delta)(n + \gamma)^2].
\]

(16)

where

\[ E_{i,j} = e^{-\lambda_2} \sum_{t=0}^{\infty} \frac{\lambda_2^t}{t!} 2^{i/2 + t} \Gamma(J/2 + t) \]

\[ \times \int_0^\infty P(\chi^2_i \geq (cq_j/\sqrt{v} + \sqrt{2}\lambda_1)^2)(q_j)^{1/2 + t - 1} e^{-q_j^2/2} dq_j \]

\[ G_{i,j} = e^{-\lambda_2} \sum_{t=0}^{\infty} \frac{\lambda_2^t}{t!} 2^{i/2 + t} \Gamma(J/2 + t) \]

\[ \times \int_0^{2\lambda_1^2|q_j|^2} P(\chi^2_i < (cq_j/\sqrt{v} + \sqrt{2}\lambda_1)^2)(q_j)^{1/2 + t - 1} e^{-q_j^2/2} dq_j, \]

\[ P_I = P(\chi^2_I \geq 2\lambda_1^2), \] and \( q_j^2 \) is a noncentral Chi-squared random variable with \( J \) degrees of freedom and \( I = 1, \ldots, 5 \).

Using the Convergence Theorem given in Judge and Yancey (1986, p. 77), we can show that for any given \( \lambda_2, \lambda_1 P(\chi^2_I \geq 2\lambda_1^2) \to 0 \) as \( \lambda_1 \to \pm \infty \). Accordingly, as \( \lambda_1 \to - \infty \), \( \rho(\sigma^{**2}, \sigma^2) \to \rho(\hat{\sigma}^2, \sigma^2) \). Similarly, as \( \lambda_1 \to \infty \), \( \rho(\sigma^{**2}, \sigma^2) \to \rho(\sigma^2, \sigma^2) \). Also, as \( \lambda_1 \to \pm \infty \), \( E_{II} \to 0 \) while \( G_{II} \to 1 \). On the other hand, for any finite \( \lambda_1 \) and nonzero \( c \), \( G_{II} \) and \( E_{II} \), both approach zero as \( \lambda_2 \to \infty \). Furthermore, \( E_{II} \) approaches \( P_I \) while \( G_{II} \) approaches \( 1 - P_I \) as \( c \to 0^+ \). Both \( E_{II} \) and \( G_{II} \) approach zero as \( c \to - \infty \).

Using these results, we can show that \( \rho(\hat{\sigma}^2, \sigma^2) \to \rho(\hat{\sigma}^2, \sigma^2) \) as \( \lambda_1 \to \pm \infty \) (given \( \lambda_2 \)), and that \( \rho(\hat{\sigma}^2, \sigma^2) \to \rho(\sigma^{**2}, \sigma^2) \) as \( \lambda_2 \to \infty \) (given \( \lambda_1 \)). \( \rho(\hat{\sigma}^2, \sigma^2) \) approaches \( \rho(\hat{\sigma}^2, \sigma^2) \)
as \( c \to 0 \), and approaches \( \rho(\sigma^{**2}, \sigma^2) \) as \( c \to -\infty \). These are analogous to the corresponding results when estimating \( E(y) \) (see Wan, 1994a).

Numerical evaluations of the risks of \( \sigma^{**2} \) and \( \hat{\sigma}^2 \) have been carried out for \( n = 20, 30, 40, 50, 80 \), \( k = 2, 5, 10, 15, \ldots, n - 5 \), \( \alpha = 0.01, 0.05, 0.10, 0.25, 0.40 \), \( \lambda_1 \in [-10, 10] \) and various choices of \( \lambda_2 \). They were performed on a VAX 6340 computer using double precision FORTRAN code which incorporates the subroutine GAMMQ given in Press et al. (1986) to evaluate \( P_I \). The subroutine DOIAJF from the NAG (1991) subroutine library is used to calculate the integrals \( E_{II} \) and \( G_{II} \). The infinite series in \( E_{II} \) and \( G_{II} \) converges with a convergence tolerance of \( 10^{-15} \). Some typical results are illustrated in Figs. 1–4.

![Relative risk functions of \( \hat{\sigma}_L^2, \sigma_{ML}^{**2}, \sigma_{ML}^{*2} \) and \( \hat{\sigma}_L^2 \) for \( n = 20, k = 5 \) and \( \lambda_2 = 0 \).](image1)

![Relative risk functions of \( \hat{\sigma}_L^2, \sigma_{ML}^{**2}, \sigma_{ML}^{*2} \) and \( \hat{\sigma}_L^2 \) for \( n = 20, k = 5 \) and \( \lambda_2 = 10 \).](image2)
Fig. 3. Relative risk functions of $\hat{\sigma}_{1s}^2$, $\hat{\sigma}_{2s}^2$, $\hat{\sigma}_{**2}$ and $\hat{\sigma}_{1s}^2$ for $n = 20$, $k = 5$ and $\lambda_2 = 0$.

Fig. 4. Relative risk functions of $\hat{\sigma}_{1s}^2$, $\hat{\sigma}_{2s}^2$, $\hat{\sigma}_{**2}$ and $\hat{\sigma}_{1s}^2$ for $n = 20$, $k = 5$ and $\lambda_2 = 10$.

Figs. 1 and 2 illustrate the results given in Wan (1994a), where $\sigma_{**2}$ and $\hat{\sigma}^2$ are based on the ML component estimators. It is shown there that with a relatively large $\lambda_2$, $\hat{\sigma}_{2s}^2$ can uniformly dominate $\sigma_{**2}$ and $\hat{\sigma}_{ML}^2$. When $\lambda_2$ is relatively small, $\hat{\sigma}_{ML}^2$ can strictly dominate $\sigma_{**2}$, but does not dominate both $\sigma_{**2}$ and $\hat{\sigma}_{ML}^2$ simultaneously.

The situation changes considerably when the component estimators are LS or MM estimators. Regardless of the value of $\lambda_2$, there always exists a family of inequality pre-test estimators which simultaneously dominate all other estimators over certain regions in the parameter space. A sub-class of this family of inequality pre-test estimators also strictly dominates the unrestricted estimator over the entire $\lambda_1$ range. Within this sub-class of pre-test estimators, the estimator with $c = -1$ (for LS) or $c = -\sqrt{v/(v+2)}$ (for MM) has the minimum risk in the region where pre-testing is
the best. It is shown in Appendix A that the minimum risk boundary is achieved by using $\sigma^{**2}$ for $\lambda_1 \in (-\infty, \lambda_1^*)$ and $\hat{\sigma}^2 | c = c^*$ for $\lambda_1 \in (\lambda_1^*, \infty)$, where $c^* = -\sqrt{v(\gamma - \delta)/(n + \delta)}$ (i.e., $c^* = -1$ for LS), or $c^* = -\sqrt{v/(v + 2)}$ (for MM) and $\lambda_1^*$ is that value of $\lambda_1 > 0$ for which $\rho(\sigma^{**2}, \sigma^2) = \rho(\hat{\sigma}^2, \sigma^2 | c = c^*)$. This coincides with the results of Giles (1990, 1991a, b) where the pre-test in question is the form of an equality hypothesis. When $\lambda_2$ is relatively small, the risk of the inequality pre-test estimator with $c = c^*$ is larger than the risk of $\hat{\sigma}^2$ with $c < c^*$ in the region $\lambda_1 < \lambda_1^*$.

Our numerical results also show that, depending on the choice of $c$, the risk gain from using $\sigma^{**2}$ and $\hat{\sigma}^2$ in the region $\lambda_1 < \lambda_1^*$ can be very slight. More strikingly, it is found that when $\lambda_2$ is sufficiently large, there exists certain sizes of the pre-test, including that corresponding to $c^*$, such that the risk of $\hat{\sigma}^2$, attains that of $\sigma^{**2}$ in the region $\lambda_1 < \lambda_1^*$. Given our previous result that the pre-test estimator with $c = c^*$ achieves minimum risk in the rest of the parameter space, the present finding implies that $\hat{\sigma}^2 | c = c^*$ is a strictly dominating estimator when $\lambda_2$ is large. Figs. 3 and 4 illustrate some of these results.

4. Numerical calculations of optimal critical values

Given that $\lambda_1$ is unknown in practice, and that there is no dominating estimator except for the case where $\lambda_2$ is large, we need to ask what choice of critical value will bring the risk of $\hat{\sigma}^2$ as close as possible to the minimum that can be achieved. In this section, we use the mini-max regret criterion (Gin, 1967) to obtain optimal critical values for the pre-test problem discussed in this paper. This criterion has been used by many authors in searching for optimal test size in various other pre-test contexts. See, for example, Brook (1976), Ohtani and Toyoda (1978), Giles and Lieberman (1991), Giles et al. (1993) and Wan (1995). The last paper considers mini-max regret critical value for an inequality restriction when the interest centers on the estimation of $\beta$.

For any given $\lambda_1$ and $c$, regret is defined as the difference between the risk of $\hat{\sigma}^2$ and the minimum risk boundary, i.e.,

$$\text{REG}(\hat{\sigma}^2 | c) = \rho(\hat{\sigma}^2, \sigma^2 | c) - \inf(\rho(\hat{\sigma}^2, \sigma^2 | c)).$$

As discussed in Appendix A, when $\lambda_1$ is small, the minimum risk boundary of $\rho(\hat{\sigma}^2, \sigma^2)$ is given by the risks corresponding to $c = \infty$ when $\lambda_1 \leq \lambda_1^*$, and to $c = c^* = -\sqrt{v(\gamma - \delta)/(n + \delta)}$ (i.e., $c = 0$ for ML, $c = -1$ for LS and $c = -\sqrt{v/(v + 2)}$ for MM) when $\lambda_1 > \lambda_1^*$ respectively. Hence (17) can be written as

$$\text{REG}(\hat{\sigma}^2 | c) = \rho(\hat{\sigma}^2, \sigma^2 | c) - \min(\rho(\sigma^{**2}, \sigma^2), \rho(\hat{\sigma}^2, \sigma^2 | c = c^*).$$

The mini-max regret procedure is to find a critical value such that the maximum regret of not being on the minimum risk boundary is minimized. If we denote $d_L$ and $d_U$ as the maximum regret in the regions $\lambda_1 \leq \lambda_1^*$ and $\lambda_1 > \lambda_1^*$ respectively, then we find that decreasing $c$ from $c^*$ decreases $d_L$, but increases $d_U$. Because of this monotonicity property, the mini-max regret procedure is to find the critical value $c^MX$.
such that \( d^L = d^U \). That is, both regrets are simultaneously minimized. It is readily shown that \( c^{\text{MX}} \) is unique.

We performed numerical computations to calculate \( c^{\text{MX}} \) for \( n = 20, 30, 40, 50, 80, k = 2, 5, 10, 15, \ldots, (n - 5), \lambda_2 = 0, 2, 10, 25, 50 \). Brent’s (1974) algorithm was used to search for the value of \( \lambda^*_2 \). The Golden Section Search Routine given in Press et al. (1986) was used to compute the mini-max regret critical values. These were incorporated into a double precision Fortran program executed on a VAX 6340 computer. Table 1 illustrates these results. It is shown that with the maximum likelihood based pre-test estimator, for any given \( n \), \( |c^{\text{ML}}| \) increases as \( k \) increases, i.e., \( c^{\text{ML}} \) is not

<table>
<thead>
<tr>
<th>( n )</th>
<th>( k )</th>
<th>( \lambda_2 = 0 )</th>
<th>( \lambda_2 = 2 )</th>
<th>( \lambda_2 = 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( c^{\text{ML}} )</td>
<td>( c^{\text{LS}} )</td>
<td>( c^{\text{MM}} )</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>-1.16</td>
<td>-1.13</td>
<td>-1.73</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-2.58</td>
<td>-1.13</td>
<td>-1.17</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>-3.93</td>
<td>-1.13</td>
<td>-1.70</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>-4.78</td>
<td>-1.12</td>
<td>-1.64</td>
</tr>
<tr>
<td>30</td>
<td>2</td>
<td>-1.17</td>
<td>-1.14</td>
<td>-1.75</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-2.62</td>
<td>-1.14</td>
<td>-1.74</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>-4.03</td>
<td>-1.13</td>
<td>-1.73</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>-5.03</td>
<td>-1.13</td>
<td>-1.71</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>-6.33</td>
<td>-1.12</td>
<td>-1.64</td>
</tr>
<tr>
<td>40</td>
<td>2</td>
<td>-1.18</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-2.64</td>
<td>-1.14</td>
<td>-1.75</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>-4.06</td>
<td>-1.14</td>
<td>-1.75</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>-5.09</td>
<td>-1.14</td>
<td>-1.74</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>-6.64</td>
<td>-1.13</td>
<td>-1.71</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>-7.62</td>
<td>-1.12</td>
<td>-1.64</td>
</tr>
<tr>
<td>50</td>
<td>2</td>
<td>-1.18</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-2.65</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>-4.08</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>-5.12</td>
<td>-1.14</td>
<td>-1.75</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>-6.72</td>
<td>-1.14</td>
<td>-1.74</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>-7.94</td>
<td>-1.13</td>
<td>-1.71</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>-9.73</td>
<td>-1.12</td>
<td>-1.64</td>
</tr>
<tr>
<td>80</td>
<td>2</td>
<td>-1.18</td>
<td>-1.14</td>
<td>-1.77</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-2.66</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>-4.11</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>-5.16</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>-6.79</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>-8.09</td>
<td>-1.14</td>
<td>-1.76</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>-9.20</td>
<td>-1.14</td>
<td>-1.75</td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>-10.15</td>
<td>-1.14</td>
<td>-1.74</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>-10.94</td>
<td>-1.13</td>
<td>-1.71</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>-11.24</td>
<td>-1.12</td>
<td>-1.64</td>
</tr>
</tbody>
</table>
invariant to the model's degrees of freedom. This differs from the results that were observed when estimating $E(y)$ as shown in Wan (1995), but is consistent with the results of Giles and Lieberman (1991) who consider the case where the linear restriction is in the form of a strict equality. For any given $k$, $c^\text{MX}_{\text{ML}}$ is roughly constant as $n$ varies. Other things being equal, $c^\text{MX}_{\text{ML}}$ approaches zero as $\lambda_2$ approaches infinity, reflecting the fact that $\hat{\sigma}^2_{\text{ML}}$ strictly dominates when the model is considerably mis-specified.

For the LS and MM cases, our results in Table 1 suggest that when $\lambda_2 = 0$, the mini-max regret critical values for using $\hat{\sigma}^2_{\text{LS}}$ or $\hat{\sigma}^2_{\text{MM}}$ are roughly constant for moderate to high degrees of freedom (approximately $-1.14$ for the case of LS and $-1.76$ for the case of MM). When the possibility of omitted variables is allowed for, the optimal critical value differs only marginally from $c = -1$ in the LS case and from $c = -\sqrt{v/(v+2)}$ in the MM case. Again, this result is not surprising, given that the pre-test estimators corresponding to these critical values are the minimum risk estimators over almost the entire $\lambda_1$ range.

5. Conclusions

Unlike the case when one is estimating the prediction vector or estimating $\sigma^2$ using the ML principle, our results illustrate the advantages of pre-testing over unrestricted estimation when estimating $\sigma^2$ using the least squares or minimum mean squared error criterion. We show that there exists a class of pre-test estimators of the LS and MM families of component estimators which strictly dominate the unrestricted estimators of their respective families. The estimator corresponding to $c^* = -\sqrt{v/(v-\delta)/(n+\delta)}$ also dominates all other estimators that we have considered when the model is sufficiently mis-specified. Therefore, unless there is a strong evidence that both the constraint and the model specification errors are small, naively imposing the restriction without testing is also not recommended.

When the model is correctly specified, the mini-max regret optimal critical value for estimating $\sigma^2$ using the LS principle ($c = -1.14$ (approx.)) roughly coincides with the corresponding mini-max regret optimal critical value tabulated by Wan (1995) for the estimation of $E(y)$ ($c = -1.12$ (approx.)). This suggests that a critical value of around $-1.13$ can be adopted in an applied situation where the risk of both the prediction vector and the error variance can be simultaneously minimized. However, this recommendation again relies heavily on the assumption of a correctly specified model, as Wan (1995) shows that the mini-max regret critical value for estimating $E(y)$ is generally not robust to mis-specification of the model. Little prescription can be offered with regard to an optimal choice of test size for the estimation of both the prediction vector and error variance in the presence of model mis-specification, except to say that the correct specification of the model is of paramount importance. The risk properties of $c^{**}$ and $\hat{\sigma}^2$ under loss functions other than the squared error loss are however yet to be explored.
Appendix A

Proof of Theorem 1. We only outline the proof of Theorem 1. Full details are available upon request from the author.

Using (8) and the definition of risk under squared error loss, the risks of the $\sigma^{**2}$ and $\sigma^2$ can be written as

$$
\rho(\sigma^{**2}, \sigma^2) = \rho(\sigma^2, \sigma^2) + E\{I_{(-\infty, -\sigma)}(u_1)[(\sigma u_1 - \tau)^2 - \sigma^2(\gamma - \delta))/(n + \gamma)] \times [2(\sigma^2 - \sigma^2) + ((\sigma u_1 - \tau)^2 - \sigma^2(\gamma - \delta))/(n + \gamma)]\}/\sigma^4
$$

(A.1)

and

$$
\rho(\sigma^2, \sigma^2) = \rho(\sigma^{**2}, \sigma^2) + E\{I_{(-\infty, -\sigma + \tau)}(u_1)[(\sigma u_1 - \tau)^2 - \sigma^2(\gamma - \delta))/(n + \gamma)] \times [2(\sigma^2 - \sigma^2) + ((\sigma u_1 - \tau)^2 - \sigma^2(\gamma - \delta))/(n + \gamma)]\}/\sigma^4
$$

(A.2)

respectively.

The evaluation of (A.1) involves the evaluations of $E(\sigma^2 I_{(-\infty, -\sigma)}(u_1)u_i^j)$, $i = 2, 4$ and $j = 0, 1, 2, 3, 4$. Given that $u_1$ is distributed independently of $\sigma^2$, $E[\sigma^2 I_{(-\infty, -\sigma)}(u_1)u_i^j]$ can be written as $E(\sigma^2)E[I_{(-\infty, -\sigma)}(u_1)u_i^j]$.

Now, from the properties of Chi-square random variables, $E(\sigma^2) = \sigma^2(v + 2\lambda_2)/(n + \delta)$. Furthermore, $E(\sigma^4) = \sigma^4E(\chi^2_{\lambda_2})/(n + \delta)^2$. Applying Lemma 2 of Judge and Bock (1978, p. 322) or Lemma 1 of Clarke et al. (1987a), we can show that

$$
E(\sigma^2) = \sigma^2(v + 2\lambda_2)/(n + \delta)
$$

(A.3)

Now, $tr(I_v) = v$. $E(\chi^2_{\lambda_2}) = (v + 2\lambda_2)$ and $E[\chi^2_{\lambda_2}] = (v + 2\lambda_2)$. It then follows that

$$
E(\sigma^4) = \sigma^4[v(v + 2 + 2\lambda_2)]/(n + \delta)^2
$$

(A.4)

In evaluating $E[I_{(-\infty, -\sigma)}(u_1)u_i^j]$, $j = 0, 1, 2, 3, 4$, we use Lemma 1 of Wan (1994a). For example, when $j = 3$ and $\lambda_1 \leq 0$, we obtain

$$
E[I_{(-\infty, -\sigma)}(u_1)u_i^3] = \frac{\xi^2}{2\sigma^3}P_1 - \frac{3\xi^2}{\sigma^2\sqrt{2\pi}}P_2 + \frac{3\xi}{2\sigma}P_3 - \frac{\sqrt{2}}{\sqrt{\pi}}P_4.
$$

The expressions for the other cases can be similarly derived by applying Lemma 1 of Wan (1994a). Using these results and after some algebraic manipulations yield (13) and (15) directly.

To evaluate (A.2), we need to evaluate $E[\sigma^2 I_{(-\infty, -\sigma + \tau)}(u_1)u_i^j]$, $i = 0, 2, 4$ and $j = 0, 1, \ldots, 4$. When $i = 0$, this expectation can be evaluated using Lemma 2 of Wan.
(1994a). For example, for the case of $j = 4$, and $\lambda_1 > 0$,
\[
E[I_{(-\infty, (c' + \gamma)\sigma)}(u_1)\hat{\sigma}_1^4] = \frac{\xi^4}{2\sigma^4} E_{1,v} + \frac{\xi^4}{\sigma^4} G_{1,v} - \frac{4\xi^3}{\sigma^4 \sqrt{2\pi}} E_{2,v} + \frac{3\xi^2}{\sigma^4} E_{3,v}
\]
\[
+ \frac{6\xi^2}{\sigma^2} G_{3,v} + \frac{4\xi}{\sigma} \sqrt{\frac{2}{\pi}} E_{4,v} + \frac{3}{2} \frac{E_{5,v} - 3G_{5,v}}{\sigma^2},
\]

Now when $i = 2$, using Theorem 2 in Judge and Bock (1978), we can write
\[
E[I_{(-\infty, (c' + \gamma)\sigma)}(u_1)\hat{\sigma}_1^3] = \frac{\sigma^2}{n + \delta} \left( \sqrt{\frac{1}{v}} E[I_{(-\infty, cq_{1,1}/\sqrt{v} + \gamma\sigma)}(u_1)] \right.
\]
\[
+ 2\lambda_2 E[I_{(-\infty, cq_{1,1}/\sqrt{v} + \gamma\sigma)}(u_1)] \right). \tag{A.5}
\]
Using this result repeatedly, we obtain
\[
E[I_{(-\infty, (c' + \gamma)\sigma)}(u_1)\hat{\sigma}_1^4]
\]
\[
= \frac{\sigma^4}{(n + \delta)^2} \left( \sqrt{\frac{1}{v}} E[I_{(-\infty, cq_{1,1}/\sqrt{v} + \gamma\sigma)}(u_1)] \right.
\]
\[
+ 2\lambda_2 E[I_{(-\infty, cq_{1,1}/\sqrt{v} + \gamma\sigma)}(u_1)] \right)
\]
\[
+ 2\lambda_2 (v + 4) E[I_{(-\infty, cq_{1,1}/\sqrt{v} + \gamma\sigma)}(u_1)] + 2\lambda_2 E[I_{(-\infty, cq_{1,1}/\sqrt{v} + \gamma\sigma)}(u_1)] \right). \tag{A.6}
\]

The expectation terms in (A.5) and (A.6) can be evaluated using Lemma 2 of Wan (1994a). Eqs. (14) and (16) then follow directly.

**Proof that the risk of $\hat{\sigma}^2$ achieves minimums at $c = -\infty$, or $c = -\sqrt{v(\gamma - \delta)/(n + \delta)}$.** Using the fact that $\hat{\sigma}^2 = \sigma^2 q^2_n/(n + \delta)$, where $q^2_n \sim \chi^2_{(v, \sigma)}$, let $\omega = u_1 - \xi/\sigma$, a standard normal variable, and recognizing the fact that $\tau/\sigma = \sqrt{2\lambda_1}$, (A.2) can be rewritten as
\[
\rho(\hat{\sigma}^2, \sigma^2) = \rho(\sigma^{**2}, \sigma^2) - E[I_{(-\infty, cq_{1,1}/\sqrt{v} + \sqrt{2\lambda_1})}(\omega)] \left( (\sigma\omega - \sigma\sqrt{2\lambda_1})^2 - \sigma^2 \times (\gamma - \delta)q^2_n/(n + \delta))\right) / (n + \delta)
\]
\[
+ ((\sigma\omega - \sigma\sqrt{2\lambda_1})^2 - \sigma^2 (\gamma - \delta)q^2_n/(n + \delta)) / (n + \gamma) \right) / \sigma^4. \tag{A.7}
\]
First, let us consider the case when $cq_{1,1}/\sqrt{v} + \sqrt{2\lambda_1} \leq 0$. If we let $d = |cq_{1,1}/\sqrt{v} + \sqrt{2\lambda_1}| = -cq_{1,1}/\sqrt{v} - \sqrt{2\lambda_1} \geq 0$, then (A.7) may be written as
\[
\rho(\hat{\sigma}^2, \sigma^2) = \rho(\sigma^{**2}, \sigma^2)
\]
\[
- E[I_{(-\infty, d)}(\omega)] \left( (\sigma\omega - \sigma\sqrt{2\lambda_1})^2 - \sigma^2 (\gamma - \delta)q^2_n/(n + \delta)) / (n + \gamma) \right)
\]
\[
\times \left[ 2\sigma^2 (q^2_n - (n + \delta)) / (n + \delta) + ((\sigma\omega - \sigma\sqrt{2\lambda_1})^2 - \sigma^2 (\gamma - \delta)q^2_n/(n + \delta)) / (n + \gamma) \right] / \sigma^4. \tag{A.8}
\]
Now from the properties of odd and even functions, we can show that for any given \( q_v \), if \( d \geq 0 \), then

\[
E[I(-\infty, d)(\omega)\omega^j] = \begin{cases} E[I(d, \infty)(\omega)\omega^j] & \text{if } j \text{ is even or zero}, \\ -E[I(d, \infty)(\omega)\omega^j] & \text{if } j \text{ is odd}. \end{cases}
\]  

(A.9)

Hence,

\[
E[I(-\infty, d)(\omega - \sigma \sqrt{2\lambda_1})^2] = E[I(d, \infty)(\sigma\omega) + \sigma \sqrt{2\lambda_1})^2] \]

(A.10)

and

\[
E[I(-\infty, -d)(\omega - \sigma \sqrt{2\lambda_1})^4] = E[I(d, \infty)(\sigma\omega) + \sigma \sqrt{2\lambda_1})^4].
\]

(A.11)

Therefore, (A.7) becomes

\[
\rho(\tilde{\sigma}^2, \sigma^2) = \rho(\sigma^{**2}, \sigma^2) - E[I(d, \infty)(\omega)\omega^j] \\
- \frac{1}{2} \left[ \int_0^\infty f(\psi)\psi^{d/2} \, d\psi - \int_0^d f(\psi)\psi^{d/2} \, d\psi \right].
\]

(A.13)

where \( f(\psi) = \psi^{-1/2}e^{-\psi/2}/\sqrt{2\Gamma(1/2)} \).

It is straightforward to show that for any given \( q_v \),

\[
E[I(d, \infty)(\omega)\omega^j] = \frac{1}{2} \left[ \int_0^\infty f(\psi)\psi^{d/2} \, d\psi - \int_0^d f(\psi)\psi^{d/2} \, d\psi \right].
\]

(A.13)

Hence (A.12) can be written as

\[
\rho(\tilde{\sigma}^2, \sigma^2) = \rho(\sigma^{**2}, \sigma^2) \\
- \frac{1}{2} E_q \left\{ \left[ \int_0^\infty f(\psi) - \int_0^d f(\psi) \right] \left[ (\sigma\sqrt{\psi} + \sqrt{2\lambda_1})^2 - \sigma^2(\gamma - \delta)q_v^2/(n + \delta)/(n + \gamma) \right] \\
+ ((\sigma\sqrt{\psi} + \sigma \sqrt{2\lambda_1})^2 - \sigma^2(\gamma - \delta)q_v^2/(n + \delta)/(n + \gamma)) \right\} \sigma^4
\]

(A.14)

From (A.14), for any given \( q_v \) such that \( cq_v/\sqrt{v} + \sqrt{2\lambda_1} \leq 0 \),

\[
\frac{\partial \rho(\tilde{\sigma}^2, \sigma^2)}{\partial c} = \frac{1}{2} \left\{ \frac{\partial^2}{\partial c} f(d^2)[(\sigma d + \sigma \sqrt{2\lambda_1})^2 - \sigma^2(\gamma - \delta)q_v^2/(n + \delta)/(n + \gamma)] \\
\times [2\sigma^2(q_v^2 - (n + \delta))/(n + \delta) \\
+ ((\sigma d + \sigma \sqrt{2\lambda_1})^2 - \sigma^2(\gamma - \delta)q_v^2/(n + \delta)/(n + \gamma))] \sigma^4 \right\} \sigma^4
\]

(A.15)
Using the fact that \( d = -c q_v / \sqrt{v} - \sqrt{2} \lambda_1 \), which is nonnegative, we can show that for any given \( q_v \),

\[
\frac{\partial \rho(d^2, \sigma^2)}{\partial c} = \{ (c q_v / \sqrt{v} + \sqrt{2} \lambda_1) (q_v / \sqrt{v}) f(d^2) \\
	imes \left[ (\sigma^2 q_v^2 (c^2 / v - (\gamma - \delta)/(n + \delta)) / (n + \gamma) ) \\
	imes [2\sigma^2 (q_v^2 - (n + \delta)) / (n + \delta) \\
+ (\sigma^2 q_v^2 (c^2 / v - (\gamma - \delta)/(n + \delta)) / (n + \gamma) ) / \sigma^4. \right]
\]

(A.16)

Sufficient conditions for (A.16) to be zero are \( c = -\infty \), which implies that \( f(d^2) = 0 \), or \( c^2 / v - (\gamma - \delta)/(n + \delta) = 0 \), i.e. \( c = -\sqrt{v(\gamma - \delta)/(n + \delta)} \). The latter condition reduces to \( c = 0 \) for ML, \( c = -1 \) for LS and \( c = -\sqrt{v/(v + 2)} \) for MM.

Next, consider the case where \( c q_v / \sqrt{v} + \sqrt{2} \lambda_1 > 0 \). For purposes of analysis, we re-defined \( d \) as \( |c q_v / \sqrt{v} + \sqrt{2} \lambda_1| = c q_v / \sqrt{v} + \sqrt{2} \lambda_1 > 0 \). Now, when \( c q_v / \sqrt{v} + \sqrt{2} \lambda_1 > 0 \),

\[
\rho(d^2, \sigma^2) = \rho(\sigma^{**2}, \sigma^2) - E[I_{-\infty, d}(\omega)](\sigma \omega - \sigma \sqrt{2} \lambda_1)^2 \\
- \sigma^2 (\gamma - \delta) q_v^2 / (n + \delta) / (n + \gamma) \\
\times [2\sigma^2 (q_v^2 - (n + \delta)) / (n + \delta) + (\sigma \omega - \sigma \sqrt{2} \lambda_1)^2 \\
- \sigma^2 (\gamma - \delta) q_v^2 / (n + \delta) / (n + \gamma) ) / \sigma^4 \] \)

(A.17)

Using the properties of odd and even functions again, for any given \( q_v \), if \( d \geq 0 \), then

\[
E[I_{-\infty, d}(\omega) \omega^j] = \begin{cases} 
 g_j - E[I_{d, \omega}(\omega) \omega^j] & \text{if } j \text{ is even or zero,} \\
 - E[I_{d, \omega}(\omega) \omega^j] & \text{if } j \text{ is odd,}
\end{cases} \] \)

(A.18)

where \( g_j = 2^{1/2} \Gamma((j + 1)/2 \Gamma(1/2)) \).

Using this result and steps similar to those in the previous case, one can show that for any \( q_v \) such that \( c q_v / \sqrt{v} + \sqrt{2} \lambda_1 > 0 \),

\[
\frac{\partial \rho(d^2, \sigma^2)}{\partial c} = \{ (c q_v / \sqrt{v} + \sqrt{2} \lambda_1) (q_v / \sqrt{v}) f(d^2) \\
	imes \left[ (\sigma^2 q_v^2 (c^2 / v - (\gamma - \delta)/(n + \delta)) / (n + \gamma) ) \\
	imes [2\sigma^2 (q_v^2 - (n + \delta)) / (n + \delta) \\
+ (\sigma^2 q_v^2 (c^2 / v - (\gamma - \delta)/(n + \delta)) / (n + \gamma) ) / \sigma^4. \right]
\]

(A.19)

Consistent with the case where \( c q_v + \sqrt{2} \lambda_1 \leq 0 \), a sufficient condition for (A.19) to be zero is \( c^2 / v - (\gamma - \delta)/(n + \delta) = 0 \), i.e. \( c = -\sqrt{v(\gamma - \delta)/(n + \delta)} \) or \( c = -\infty \). Although it is difficult to analyze the second derivative corresponding to (A.17) and (A.19), it is found numerically the inequality pre-test estimators corresponding to these values of \( c \) form the minimum risk boundary of \( \sigma^2 \). More specifically, when \( \lambda_2 \) is small, the
minimum risk boundary of $\hat{\sigma}^2$ is given by the risk of $\hat{\sigma}^2 | c = -\infty$ when $\lambda_1 \leq \lambda_1^*$, and the risk of $\hat{\sigma}^2 | c = -\sqrt{\nu(\gamma - \delta)/(n + \delta)}$ where $\lambda_1 > \lambda_1^*$. When $\lambda_2$ is large, the minimum risk boundary of $\hat{\sigma}^2$ is given by the risk of $\hat{\sigma}^2 | c = -\sqrt{\nu(\gamma - \delta)/(n + \delta)}$ over the entire range of $\lambda_1$.

Acknowledgements

This paper is based on Chapters 7 and 8 of my doctoral thesis. The assistance given by David Giles and Judith Giles is gratefully acknowledged. Thanks are also due to William Griffiths, George Judge, Kazuhiro Ohtani, Nripesh Podder and the referee for their helpful comments.

References


