Bayesian estimation subject to minimaxity of the mean of a multivariate normal distribution in the case of a common unknown variance: a case for Bayesian robustness

by

Anirban DasGupta and Herman Rubin

Purdue University

Technical Report #86-10

Purdue University
Department of Statistics

April 1986

Bayesian estimation subject to minimaxity of the mean of a multivariate normal distribution in the case of a common unknown variance: a case for Bayesian robustness

by

Anirban DasGupta and Herman Rubin

1. Introduction

In multiparameter estimation problems, one addresses the problem of estimating a vector of p parameters $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_p)$ under a certain loss $L(\underset{\sim}{\theta}, \underset{\sim}{\eta}, \underset{\sim}{a})$, where $\underset{\sim}{\eta}$ is some k-dimensional nuisance parameter. Since the pioneering work of Stein (1955), researchers in this general area have proved the existence of the Stein-effect in an incredible number of estimation problems. Typically, the message of all such works is that if the number of dimensions p is sufficiently large, then a classical estimator δ_0 is inadmissible and that there exists a startingly large collection of dominating estimators; one then finds the need to develop a systematic theory for selecting one (or more generally, a sub-class of) estimator(s) for actual use in specific problems. The following restricted risk Bayes problem suggests itself very naturally in this context.

Let $\pi(\underset{\sim}{\theta},\underset{\sim}{\eta})$ denote a prior distribution on the unknown parameters of the problem. Let also $\mathcal{D}_\varepsilon$ denote the class of estimators

$$\mathcal{D}_\varepsilon = \{\delta : \; R(\underset{\sim}{\theta},\underset{\sim}{\eta},\delta) \leq R(\underset{\sim}{\theta},\underset{\sim}{\eta},\delta_0) + \varepsilon \;\; \forall \, \underset{\sim}{\theta}, \underset{\sim}{\eta}\},$$

where ε is a non-negative (typically small) real number. Characterization of estimators in $\mathcal{D}_\varepsilon$ is virtually impossible (even in special cases), but working with simpler sub-classes of $\mathcal{D}_\varepsilon$ is tractable in quite a few problems. Let also $r(\pi,\delta)$ denote the usual integrated Bayes risk of an estimator δ with respect to the prior π. The restricted risk Bayes estimator (if one exists) is defined as the estimator $\delta_{\pi,\varepsilon}$ such that $r(\pi,\delta_{\pi,\varepsilon}) = \inf_{\delta\varepsilon \mathcal{D}_\varepsilon} r(\pi,\delta)$. More generally, one may consider the problem of minimizing the posterior expected loss conditioned on the data. For a thorough discussion of how the restricted risk Bayes rules pertain to robust Bayesian issues, see Berger (1982), Berger and DasGupta (1986), and Larry Brown's discussion on Berger (1983).

Several authors have considered versions of this restricted risk Bayes problem in special cases; see, for example, Hodges and Lehmann (1952), Efron and Morris (1971), Bickel (1980), Casella and Strawderman (1981), Berger (1982), DasGupta (1985), Marazzi (1985), Berger and DasGupta (1986). As is well-known, as stated above, the restricted Bayes problem is exceedingly difficult and no neat mathematical solution seems possible. For this reason, it makes sense to solve a slightly simplified version of this problem (originally suggested in Berger (1982)), a version that is not a sweeping simplification in any of the problems we have thought of, yet is very nicely tractable in many cases. It is well known that in a large number of simultaneous estimation problems, one can find an unbiased estimator $D\delta(\underset{\sim}{X})$ of $\Delta(\underset{\sim}{\theta},\underset{\sim}{\eta}) = R(\underset{\sim}{\theta},\underset{\sim}{\eta},\delta) - R(\underset{\sim}{\theta},\underset{\sim}{\eta},\delta_0)$; clearly,

$\mathcal{D}^*_\varepsilon = \{\delta\colon D\delta(\underset{\sim}{X}) \leq \varepsilon\ \forall\ \underset{\sim}{X}\} \subseteq \mathcal{D}_\varepsilon$; in specific problems, $\mathcal{D}^*_\varepsilon$ may be a proper subset of $\mathcal{D}_\varepsilon$ (see Hwang (1979) for an example). The modified restricted risk Bayes problem is to find $\delta^*_{\pi,\varepsilon}$ such that $r(\pi,\delta^*_{\pi,\varepsilon}) = \inf_{\delta\varepsilon \mathcal{D}^*_\varepsilon} r(\pi,\delta)$. A good amount of evidence has accumulated by now that $\delta^*_{\pi,\varepsilon}$ can be written down in a closed form in quite a few problems, and that in all of these cases it is attractively simple; see Berger (1982), Berger and DasGupta (1986). For reasons of simplicity, we will consider $\varepsilon = 0$ which makes the estimator $\delta^*_{\pi,\varepsilon}$ full minimax; ordinarily, one would expect to trade off large sacrifices in the overall Bayes risk in return for this full robustness. The beauty of the approach lies in the fact that when p is reasonably large, the sacrifices are next to nothing, yet we are assured of full minimaxity of the new estimator. At this stage, we would like to remind the reader that minimizing the posterior expected loss for each $\underset{\sim}{X}$ is a more attractive problem, but in general presents new difficulties. This is because the restricted risk Bayes actions $a(\underset{\sim}{x})$ for different $\underset{\sim}{x}$, when put together to form a strategy $a(\underset{\sim}{X})$, may go outside of the class $\mathcal{D}_\varepsilon^*$ and thus the infimum may not be attained in $\mathcal{D}_\varepsilon^*$. Thus, roughly speaking, $\mathcal{D}_\varepsilon^*$ should be moderately rich to start with, so this catastrophe does not occur. We will see in the subsequent sections that the resultant estimator is within $\mathcal{D}_\varepsilon^*$ in our problems; we can thus completely avoid the hard technical variational arguments, and yet solve the more general posterior restricted risk Bayes problem.

We would consider the problem of estimating the mean vector $\underset{\sim}{\theta}$ of a p dimensional multivariate normal distribution with common unknown variance σ^2. We would consider the normalized quadratic loss

$$L(\underset{\sim}{\theta},\sigma^2,a) = \frac{||\underset{\sim}{\theta} - \underset{\sim}{a}||^2}{\sigma^2} \quad . \tag{1.1}$$

The classical estimator $\delta_0(\underset{\sim}{X}) = \underset{\sim}{X}$ has constant risk p and is known to be inadmissible for $p \geq 3$. We would assume that the unknown variance σ^2 is estimated on the basis of W, where $(m+2)W \sim \sigma^2\chi^2(m)$ for some $m \geq 1$, and is distributed independently of the X_i's (the multiplier (m+2) upfront makes some calculations a little easier).

It was proved in Strawderman (1973) (also see Baranchik (1970) and Efron and Morris (1976)), that if

$$\delta(\underset{\sim}{X},W) = (1 - \frac{(p-2)\tau(F,W)}{F}) \cdot \underset{\sim}{X} \tag{1.2}$$

belongs to $\mathcal{D}_0^* = \{\delta: 0 < \tau \leq 2, \tau(F,W)\downarrow$ in W for each fixed F,

$\tau(F,W)\uparrow$ in F for each fixed W$\}$,

then $R(\underset{\sim}{\theta},\sigma^2,\delta) < R(\underset{\sim}{\theta},\sigma^2,\underset{\sim}{X}) = p \quad \forall\, \underset{\sim}{\theta},\sigma^2$; in the above, $F = \frac{||\underset{\sim}{X}||^2}{W}$. When $\tau(F) = \tau(F,W)$ does not involve W (i.e., δ is orthogonally invariant), the conditions on $\tau(F)$ can be somewhat relaxed (one does not need full monotonicity of $\tau(F)$); see Efron and Morris (1976). For certain kinds of priors $\pi(\underset{\sim}{\theta},\sigma^2)$, the Bayes estimator of $\underset{\sim}{\theta}$ is orthogonally invariant (so that the corresponding $\tau_\pi(F,W)$ is free of W); it will be clear from the proof that the restricted risk Bayes estimate for such priors is the same no matter whether we seek a solution of the problem within the $\mathcal{D}_0^*$ defined above or within the enlarged class of Efron and Morris (1976). One nice thing about this is that their enlarged class is the class of all orthogonally invariant estimators such that an appropriate unbiased estimator of the risk difference is uniformly ≤ 0. When $\tau_\pi(F,W)$ is allowed to depend on both F and W, the class $\mathcal{D}_0^*$ is not necessarily the class of all estimators of the form (1.2) that uniformly dominate $\underset{\sim}{X}$. Thus, we have really granted ourselves one more simplification of Berger's (1982) modified version of the problem; in return, we can handle priors which do not produce orthogonally invariant Bayes estimators. Note also that the classical multiple regression problem is a special case of the set up described above, once one recalls that the least squares estimate $\hat{\underset{\sim}{\beta}}$ and the error sum of squares S^2 form a sufficient statistic in this problem and a linear transformation on $\hat{\underset{\sim}{\beta}}$ makes its dispersion matrix proportional to the identity matrix I. Below we provide a very simple example to give the reader an idea of how the proof will go in all cases considered.

Example.

$$\text{Let } \underset{\sim}{X} \sim N_p(\underset{\sim}{\theta},\sigma^2 I),$$
$$(m+2)W \sim \sigma^2 \chi^2(m).$$

Suppose conditionally given σ^2, θ_i's are i.i.d. $N(0,\sigma^2)$, and σ^2 has some prior density $\pi_0(\sigma^2)$. It is easy to see that under the loss (1.1), the Bayes estimate of θ_i is

$$\delta_{\pi,i}(\underset{\sim}{X},W) = \frac{E(\frac{\theta_i}{\sigma^2}\,|\underset{\sim}{X},W)}{E(\frac{1}{\sigma^2}|\underset{\sim}{X},W)} = \frac{X_i}{2}\,. \tag{1.3}$$

$$\text{Define now } \delta^*_{\pi,0}(\underset{\sim}{X}) = \frac{\underset{\sim}{X}}{2} \quad \text{if } F \le 4(p-2)$$
$$= (1 - \frac{2(p-2)}{F}) \cdot \underset{\sim}{X} \quad \text{if } F > 4(p-2). \tag{1.4}$$

It is clear that $\delta^* \varepsilon \mathcal{D}_0$; also, because of the nature of the loss function, $\delta^*_{\pi,0}$ would minimize the posterior expected loss for all $\underset{\sim}{X},W$ if $||\delta^*_{\pi,0} - \delta_\pi||^2 \le ||\delta - \delta_\pi||^2$ for all $\delta\varepsilon\mathcal{D}^*_0$ (a formal proof is postponed till the next section). Now notice that

$$||\delta^*_{\pi,0} - \delta_\pi||^2 = 0 \le ||\delta - \delta_\pi||^2 \quad \text{if } F \le 4(p-2);\ \text{for } F > 4(p-2),$$

$$||\delta^*_{\pi,0} - \delta_\pi||^2 = (\frac{1}{2} - \frac{2(p-2)}{F})^2 \cdot ||\underset{\sim}{X}||^2$$
$$= \frac{(p-2)^2}{F^2} (\frac{F}{2(p-2)} - 2)^2 \cdot ||\underset{\sim}{X}||^2$$
$$\le \frac{(p-2)^2}{F^2} (\frac{F}{2(p-2)} - \tau(F,W))^2 \cdot ||\underset{\sim}{X}||^2$$
$$= ||\delta - \delta_\pi||^2 \tag{1.5}$$

The first equality follows from definition of $\delta^*_{\pi,0}$; the second line is straight algebra; the inequality is due to the fact that for all $\delta \varepsilon \mathcal{D}^*_0$, $\tau(F,W) \leq 2$ and that $\frac{F}{2(p-2)} \geq 2$ for $F \geq 4(p-2)$; the last equality again follows from definition of $\delta(X,W)$.

In more general cases, almost all the effort goes in showing that the restricted risk Bayes rule belongs to $\mathcal{D}^*_0$; the proof of the fact that it minimizes the conditional expected loss is no more difficult. The rest of the paper emphasizes technical arguments to show what more general priors can be handled. In the last section, we provide analytical formulas for Bayes risks (and the usual RSL's) of the restricted risk Bayes rules; a concrete example is fully worked out and it is shown that when enough degrees of freedom are available to estimate σ^2, the restricted risk Bayes rule does practically as good as the unconstrained Bayes rule.

2. Some background machinery

In this section, we will prove some technical lemmas which would later be used to generalize the results of the example above for much more general priors. Our target will be to show that for a given prior $\pi(\underset{\sim}{\theta},\sigma^2)$, the estimator that minimizes the posterior expected loss for each fixed $\underset{\sim}{X}, W$ is of the form

$$\delta^*_{\pi,0}(\underset{\sim}{X},W) = (1 - \tau_\pi(F,W)) \cdot \underset{\sim}{X} \text{ if } F\tau_\pi(F,W) \leq 2(p-2)$$
$$= \left(1 - \frac{2(p-2)}{F}\right) \cdot \underset{\sim}{X} \text{ if } F\tau_\pi(F,W) > 2(p-2) .$$

Note that $\delta^*_{\pi,0}$ coincides with the unconstrained Bayes rule for $F\tau_\pi(F,W) \leq 2(p-2)$. The proof will include the following two steps:

(i) Showing that $||\delta^*_{\pi,0} - \delta_\pi||^2 \leq ||\delta - \delta_\pi||^2 \; \forall \underset{\sim}{X}, W; \forall \delta \varepsilon \mathcal{D}^*_0$

(ii) Showing that $\delta^*_{\pi,0} \; \varepsilon \; \mathcal{D}^*_0$.

Proving (i) is always easy; in order to prove (ii), it would be necessary to prove that $\tau(F,W) = \frac{F\tau_\pi(F,W)}{2(p-2)}$ satisfies the monotonicity restrictions needed for uniform domination. Typically, the monotone decreasing nature as a function of W for each fixed F causes no problem for the priors we would consider. It is the monotone increasing behavior as a function of F that needs non-trivial proofs. We would assume that the conditional density of $\underset{\sim}{\theta}$ given σ^2 is spherically symmetric and that the density of σ^2 is of the form

$$\pi_0(\sigma^2) = e^{-\frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\beta-1}, \; a \geq 0. \tag{2.1}$$

Restriction to such priors seems necessary for the Bayes estimate to be of the form (1.2). We remark that (2.1) includes the family of conjugate priors on σ^2 as also Jeffrey's non-informative prior. We also remark that the priors on $(\underset{\sim}{\theta},\sigma^2)$ are hierarchical in nature and it is a very common practice to put a conjugate or a non-informative prior at the second stage in hierarchical Bayes problems.

Lemma 1 Let $g(\frac{\Sigma\theta_i^2}{\sigma^2}) \sigma^{-p}$ denote the conditional density of $\underset{\sim}{\theta}$ given σ^2 and let $\pi_0(\sigma^2)$ given by (2.1) above be the marginal density of σ^2.

For $s \geq 0$, define

$$K^*(s) = e^{-\frac{s}{2}} g(s) ,$$

$$K(s) = K^*(s) - \frac{1}{2}\int_s^\infty K^*(t)dt \tag{2.2}$$

Let $$f_1\left(\frac{\Sigma X_j^2}{\sigma^2}\right) = \int \frac{e^{\frac{\Sigma\theta_j x_j}{\sigma^2}} K\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right)}{\sigma^p} d\underset{\sim}{\theta} ,$$

and $$f_2\left(\frac{\Sigma x_j^2}{\sigma^2}\right) = \int \frac{e^{\frac{\Sigma\theta_j x_j}{\sigma^2}} k^*\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) d\underset{\sim}{\theta}}{\sigma^p} \tag{2.3}$$

Then the Bayes estimate of θ_i is given by

$$\delta_{\pi,i}(\underset{\sim}{X},W) = (1 - \tau_\pi(F,W)) \cdot X_i ,$$

where

$$\tau_\pi(F,W) = \frac{\int_0^\infty e^{-\frac{z}{2} - \frac{z}{2F}(n+\frac{2a}{W})} f_1(z)z^{\alpha-3}dz}{\int_0^\infty e^{-\frac{z}{2}} e^{-\frac{z}{2F}(n + \frac{2a}{W})} f_2(z)z^{\alpha-3} dz} , \tag{2.4}$$

$\alpha = \frac{p+m}{2} + \beta$, and $n = m + 2$.

Proof: Since $L(\underset{\sim}{\theta},\sigma^2,\underset{\sim}{a}) = \frac{||\underset{\sim}{\theta}-\underset{\sim}{a}||^2}{\sigma^2}$,

$$\delta_{\pi,i}(\underset{\sim}{X},\underset{\sim}{W}) = \frac{\iint \frac{\theta_i}{\sigma^2} \frac{e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2}}{\sigma^p} g\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) e^{-\frac{(m+2)W}{2\sigma^2}} e^{-\frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-2} d\underset{\sim}{\theta}\, d\sigma^2}{\iint \frac{1}{\sigma^2} \frac{e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2}}{\sigma^p} g\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) e^{-\frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-2} d\underset{\sim}{\theta}\, d\sigma^2}$$

$$= x_i - \frac{\iint \frac{\partial}{\partial\theta_i} e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2} \frac{g\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right)}{\sigma^p} e^{-\frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-2} d\underset{\sim}{\theta}\, d\theta^2}{\iint \frac{1}{\sigma^2} e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2} \frac{g(\Sigma\theta_j^2)}{\sigma^p} e^{-\frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-2} d\underset{\sim}{\theta}\, d\sigma^2}$$

$$= x_i + 2\, \frac{\iint \frac{\theta_i}{\sigma^2} \frac{e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2}}{\sigma^p} g'\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) e^{-\frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-2} d\underset{\sim}{\theta}\, d\sigma^2}{\iint \frac{1}{\sigma^2} \frac{e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2}}{\sigma^p} g\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) e^{-\frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-2} d\underset{\sim}{\theta}\, d\sigma^2}$$

$$= x_i + \frac{\iint \frac{e^{\frac{\Sigma\theta_j x_j}{\sigma^2}}}{\sigma^p} \frac{\partial}{\partial\theta_i} k\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) e^{-\frac{\Sigma x_j^2}{2\sigma^2} - \frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-2} d\underset{\sim}{\theta}\, d\sigma^2}{\iint \frac{e^{\frac{\Sigma\theta_j x_j}{\sigma^2}}}{\sigma^p} k^*\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) e^{-\frac{\Sigma x_j^2}{2\sigma^2} - \frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-1} d\underset{\sim}{\theta}\, d\sigma^2}$$

$$= x_i \left[1 - \frac{\iint \frac{e^{\frac{\Sigma\theta_j x_j}{\sigma^2}}}{\sigma^p} k\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) e^{-\frac{\Sigma x_j^2}{2\sigma^2} - \frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-1} d\underset{\sim}{\theta}\, d\sigma^2}{\iint \frac{e^{\frac{\Sigma\theta_j x_j}{\sigma^2}}}{\sigma^p} k^*\left(\frac{\Sigma\theta_j^2}{\sigma^2}\right) e^{-\frac{\Sigma x_j^2}{2\sigma^2} - \frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha-1} d\underset{\sim}{\theta}\, d\sigma^2} \right]$$

$$= x_i \left[1 - \frac{\int e^{-\frac{\Sigma x_j^2}{2\sigma^2}} \cdot f_1(\frac{\Sigma x_j^2}{\sigma^2}) e^{-\frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} (\frac{1}{\sigma^2})^{\alpha-1} d\sigma^2}{\int e^{-\frac{\Sigma x_j^2}{2\sigma^2}} f_2(\frac{\Sigma x_j^2}{\sigma^2}) e^{-\frac{(m+2)W}{2\sigma^2} - \frac{a}{\sigma^2}} (\frac{1}{\sigma^2})^{\alpha-1} d\sigma^2} \right],$$

which reduces to (2.4) on substituting $\frac{\Sigma x_j^2}{\sigma^2} = z$.

Lemma 2

$$f_1(z) = k_1 z^{-\frac{p-2}{4}} \int_0^\infty I_{\frac{p-2}{2}}(\sqrt{zs}) s^{\frac{p-2}{4}} k(s) ds,$$

and
$$f_2(z) = k_1 z^{-\frac{p-2}{4}} \int_0^\infty I_{\frac{p-2}{2}}(\sqrt{zs}) s^{\frac{p-2}{4}} k^*(s) ds, \qquad (2.5)$$

where k_1 is some absolute constant and $I_\nu(u) = k_2 \cdot u^\nu \int_{-1}^{1} e^{ut}(1-t^2)^{\nu-\frac{1}{2}} dt$ is the Bessel function of order ν.

Proof: We will prove the first formula. The second one follows similarly. By definition,

$$f_1(\Sigma y_j^2) = \int e^{\Sigma\theta_j y_j} k(\Sigma\theta_j^2) d\underset{\sim}{\theta}, \qquad \text{where } y_j = \frac{x_j}{\sigma}.$$

Transforming to Polar coordinates $\eta_1, \eta_2, \ldots, \eta_{p-1}$, r, where η_1 is the angle between $\underset{\sim}{\theta}$ and $\underset{\sim}{y}$, and $r = ||\underset{\sim}{\theta}||$, one has,

$$f_1(\Sigma y_j^2) = k_3 \iint e^{||\underset{\sim}{y}|| r \cos \eta_1} (\sin \eta_1)^{p-2} d\eta_1 r^{p-1} k(r^2) dr. \qquad (2.6)$$

Letting $||\underset{\sim}{y}||^2 = z$, substituting $\cos \eta_1 = t$ and $r^2 = s$, the result now follows by straight algebra.

Lemma 3

$\frac{f_1(z)}{f_2(z)}$ is monotone non-increasing if g is log-convex and monotone non-decreasing if g is log-concave.

Proof: Clearly, in view of Lemma 2, the result would follow if we can show that $\frac{k(s)}{k^*(s)}$ is monotone non-increasing for log-convex g, monotone non-decreasing for log-concave g, and that $p(s|z) = I_{\frac{p-2}{2}}(\sqrt{zs})s^{\frac{p-2}{4}} k^*(s)$ is an MLR family of measures (note $p(s|z)$ is not a probability measure; the corresponding normalized measures are also MLR if $p(s|z)$ is MLR). We first show $\frac{k(s)}{k^*(s)}$ is monotone non-increasing if g is log-convex. First note that,

$$\frac{k(s)}{k^*(s)} = 1 - \frac{1}{2}\frac{\int_s^\infty e^{-\frac{t}{2}} g(t)dt}{e^{-\frac{s}{2}} g(s)} . \tag{2.7}$$

Next, integrating by parts,

$$\frac{1}{2}\int_s^\infty e^{-\frac{t}{2}} g(t)dt = e^{-\frac{s}{2}} g(s) + \int_s^\infty e^{-\frac{t}{2}} g'(t)dt$$

$$\Rightarrow \frac{1}{2} = \frac{e^{-\frac{s}{2}} g(s)}{\int_s^\infty e^{-\frac{t}{2}} g(t)dt} + \frac{\int_s^\infty e^{-\frac{t}{2}} g'(t)dt}{\int_s^\infty e^{-\frac{t}{2}} g(t)dt} . \tag{2.8}$$

Substituting into (2.7) yields

$$\frac{k(s)}{k^*(s)} = 1 - \frac{1}{2}\left\{\frac{1}{2} - \frac{\int_s^\infty e^{-\frac{t}{2}} g'(t)dt}{\int_s^\infty e^{-\frac{t}{2}} g(t)dt}\right\}^{-1} . \tag{2.9}$$

Note now $\frac{g'(t)}{g(t)}$ is increasing if g is log-convex, and $E[\frac{g'(T)}{g(T)}\Big| T \geq s]$ is therefore increasing in s (whatever be the distribution of T). (2.9) now gives that $\frac{k(s)}{k^*(s)}$ is decreasing for log-convex g. The same argument also shows that it is increasing for log-concave g. In order to show $p(s|z)$ has the MLR property, it will be necessary to show that

$$R(s) = \frac{\int_{-1}^{1} e^{\sqrt{z_2}st} (1-t^2)^{\frac{p-3}{2}} dt}{\int_{-1}^{1} e^{\sqrt{z_1}st} (1-t^2)^{\frac{p-3}{2}} dt} \quad \text{is increasing in s for } z_2 > z_1 > 0.$$

Now by straightforward computation,

$$R'(s) \geq 0 \Leftrightarrow \sqrt{z_2}\, \frac{\int_{-1}^{1} e^{\sqrt{z_2}st}\, t(1-t^2)^{\frac{p-3}{2}} dt}{\int_{-1}^{1} e^{\sqrt{z_2}st} (1-t^2)^{\frac{p-3}{2}} dt} \geq \sqrt{z_1}\, \frac{\int_{-1}^{1} e^{\sqrt{z_1}st}\, t(1-t^2)^{\frac{p-3}{2}} dt}{\int_{-1}^{1} e^{\sqrt{z_1}st}\, t(1-t)^{2})^{\frac{p-3}{2}} dt}$$

It's clear that $\int_{-1}^{1} te^{\sqrt{z_2}st} (1-t^2)^{\frac{p-3}{2}} dt > 0 \; \forall s, z_2 > 0.$ (2.10)

Hence (2.10) will follow if

$$\frac{\int_{-1}^{1} e^{\sqrt{z_2}st}\, t(1-t^2)^{\frac{p-3}{2}} dt}{\int_{-1}^{1} e^{\sqrt{z_2}st}\, t(1-t^2)^{\frac{p-3}{2}} dt} - \frac{\int_{-1}^{1} e^{\sqrt{z_1}st}\, t(1-t^2)^{\frac{p-3}{2}} dt}{\int_{-1}^{1} e^{\sqrt{z_1}st}\, t(1-t^2)^{\frac{p-3}{2}} dt} ,$$

which is an immediate consequence of the MLR property of the family of densities

$e^{\sqrt{z}st}\,(1-t^2)^{\frac{p-3}{2}}$.

<u>Lemma 4</u>. $\tau_\pi(F,W)$ is decreasing in F for each fixed W if g is log-convex and increasing in F for each fixed W if g is log-concave.

Proof: Follows from Lemma 3 and the fact that (2.4) implies that $\tau_\pi(F,W) = E[\frac{f_1(Z)}{f_2(Z)}]$ with respect to an MLR family of densities.

<u>Lemma 5</u>. $\tau_\pi(F,W)$ is decreasing in W for each fixed F if g is log-convex or if a = 0 in (2.1).

Proof: Again follows from (2.4) and Lemma 3 using MLR arguments. For a = 0, $\tau_\pi(F,W)$ is independent of W for any g. At this stage we present a brief discussion of what the preceding lemmas tell us. They show that for log-concave g and a = 0, $\tau_\pi(F)$ may not be positive but if it is then $F\tau_\pi(F)$ will be monotone increasing in F as $\tau_\pi(F)$ itself is so. Thus for log-concave g, the effort would be directed towards showing that $\tau_\pi(F) \geq 0$ (note that $\tau_\pi(F) \geq 0$ is required for uniform domination by $\delta^*_{\pi,0}$). On the other hand, for log-convex g, which are often decreasing, whatever be a, $\tau_\pi(F,W)$ is guaranteed positive but is monotone decreasing in F, so that the effort will be directed towards showing that $F\tau_\pi(F,W)$ is monotone increasing in F. For a general log-convex g, it seems extremely difficult to exactly specify some easily verifiable conditions on g for $F\tau_\pi(F,W)$ to be monotone increasing (one needs more than mere log-convexity for this to happen, because $F\tau_\pi(F,W)$ is <u>not</u> monotone increasing in F even for <u>all</u> finite scale mixtures of normal priors, i.e., $g(s) = \sum_{i=1}^{k} \varepsilon_i e^{-\frac{c_i}{2} s}$, which are all log-convex. We will next show that monotonicity of $F\tau_\pi(F,W)$ generally causes no problem if the c_i's and ε_i's are not too different; we will also prove a parallel monotonicity result for certain continuous scale-mixtures

$$g(s) = \int e^{-\frac{s}{2\tau^2}} p(\tau^2)d\tau^2.$$

The technicalities are easier to handle for continuous mixtures than for discrete mixtures. We will then prove similar monotonicity results for certain log-convex g which are not scale mixtures of normals and exhibit a class of log-concave priors for which $\tau_\pi(F) \geq 0$ for every F.

3. Priors which are scale-mixtures of normals

Lemma 6

Let $g(s) = \sum_{i=1}^{k} \varepsilon_i \, c_i^{\frac{p}{2}} e^{-\frac{c_i}{2} s}$, where $0 < \varepsilon_i < 1$, $\Sigma\varepsilon_i = 1$, $c_i > 0$. Then

$$\tau_\pi(F,W) = \frac{\Sigma\varepsilon_i a_i^{\frac{p}{2}+1} (n+a_iF + \frac{2a}{W})^{2-\alpha}}{\Sigma\varepsilon_i a_i^{\frac{p}{2}} (n+a_iF + \frac{2a}{W})^{2-\alpha}} \tag{3.1}$$

where $a_i = \frac{c_i}{1+c_i}$.

Proof: Note that $k^*(s) = \Sigma\varepsilon_i c_i^{\frac{p}{2}} e^{-\frac{s}{2}(1+c_i)}$.

$\therefore$ by straight algebra, $$f_2(u) = \int \frac{e^{\frac{\Sigma\theta_j x_j}{\sigma^2}} k^*(\frac{\Sigma\theta_j^2}{\sigma^2})}{\sigma^p} d\underset{\sim}{\theta}$$

$$= \Sigma\varepsilon_i a_i^{\frac{p}{2}} e^{\frac{u}{2}(1-a_i)}. \tag{3.2}$$

Therefore, using (2.4), the denominator of $\tau_\pi(F,W)$

$$= \Sigma\varepsilon_i a_i^{\frac{p}{2}} \int e^{-z[\frac{n+\frac{2a}{W}}{2F} + \frac{a_i}{2}]} z^{\alpha-3} dz$$

$$= \Gamma(\alpha-2)(2F)^{\alpha-2}\, \Sigma\varepsilon_i a_i^{\frac{p}{2}} (n+\tfrac{2a}{W} + a_i F)^{2-\alpha}. \tag{3.3}$$

Using the definition of k(s), it follows as in (3.2) that

$f_1(u) = \Sigma\varepsilon_i a_i^{\frac{p}{2}+1} e^{\frac{u}{2}(1-a_i)}$, and hence the numerator of (2.4) is equal to

$$\Gamma(\alpha-2)(2F)^{\alpha-2}\, \Sigma\varepsilon_i a_i^{\frac{p}{2}+1} (n+\tfrac{2a}{W} + a_i F)^{2-\alpha} \tag{3.4}$$

(2.4) now gives the result.

Lemma 7. For g(s) as in Lemma 6, $F\tau_\pi(F,W)$ is increasing in F for each fixed W if

$$(\alpha-2)(\frac{a_{max}}{a_{min}} - 1) \leq 1, \tag{3.5}$$

where $a_{max} = \max\limits_{1\leq i\leq k} a_i$ and $a_{min} = \min\limits_{1\leq i\leq k} a_i$.

Proof: $\frac{\partial}{\partial F} F\tau_\pi(F,W)$

$$= \tau_\pi(F,W) + F\cdot\frac{\partial}{\partial F}\tau_\pi(F,W)$$

$$= \frac{\Sigma\varepsilon_i a_i^{\frac{p}{2}+1}(n+a_i F+\frac{2a}{W})^{2-\alpha}}{\Sigma\varepsilon_i a_i^{\frac{p}{2}}(n+a_i F+\frac{2a}{W})^{2-\alpha}}$$

$$+ F(2-\alpha)\left\{\frac{\Sigma\varepsilon_i a_i^{\frac{p}{2}+2}(n+a_iF+\frac{2a}{W})^{1-\alpha}}{\Sigma\varepsilon_i a_i^{\frac{p}{2}}(n+a_iF+\frac{2a}{W})^{2-\alpha}}\right.$$

$$\left. - \frac{\Sigma\varepsilon_i a_i^{\frac{p}{2}+1}(n+a_iF+\frac{2a}{W})^{2-\alpha}\Sigma\varepsilon_i a_i^{\frac{p}{2}+1}(n+a_iF+\frac{2a}{W})^{1-\alpha}}{(\Sigma\varepsilon_i a_i^{\frac{p}{2}}(n+a_iF+\frac{2a}{W})^{2-\alpha})^2}\right\} \geq 0$$

$$\Leftrightarrow \Sigma\varepsilon_i a_i^{\frac{p}{2}+1}(n+a_iF+\frac{2a}{W})^{2-\alpha}\Sigma\varepsilon_i a_i^{\frac{p}{2}}(n+a_iF+\frac{2a}{W})^{2-\alpha}$$

$$+ (\alpha-2)F\left\{\Sigma\varepsilon_i a_i^{\frac{p}{2}+1}(n+a_iF+\frac{2a}{W})^{2-\alpha}\Sigma\varepsilon_i a_i^{\frac{p}{2}+1}(n+a_iF+\frac{2a}{W})^{1-\alpha}\right.$$

$$\left. - \Sigma\varepsilon_i a_i^{\frac{p}{2}+2}(n+a_iF+\frac{2a}{W})^{1-\alpha}\Sigma\varepsilon_i a_i^{\frac{p}{2}}(n+a_iF+\frac{2a}{W})^{2-\alpha}\right\} \geq 0. \qquad (3.6)$$

Define $\varepsilon_i a_i^{\frac{p}{2}+1}(n+a_iF+\frac{2a}{W})^{2-\alpha} = d_i$

$$\varepsilon_i a_i^{\frac{p}{2}}(n+a_iF+\frac{2a}{W})^{2-\alpha} = c_i$$

$$\frac{a_iF}{n+a_iF+\frac{2a}{W}} = \delta_i. \qquad (3.7)$$

Then (3.6) reduces to

$$\Sigma d_i\ \Sigma c_i + (\alpha-2)\{\Sigma d_i\ \Sigma c_i\delta_i - \Sigma d_i\delta_i\ \Sigma c_i\} \geq 0. \qquad (3.8)$$

Now $\Sigma c_i\ \Sigma d_i\delta_i - \Sigma d_i\ \Sigma c_i\delta_i$

$$\leq \{(\max \frac{c_i}{d_i})(\max \frac{d_i}{c_i}) - 1\}\Sigma d_i\ \Sigma c_i\delta_i$$

$$\leq \left(\frac{a_{max}}{a_{min}} - 1\right)\Sigma d_i \; \Sigma c_i \qquad \text{as} \quad \frac{d_i}{c_i} = a_i \quad \text{and} \quad \delta_i \leq 1. \tag{3.9}$$

(3.8) now gives the result.

(3.5) is fairly restrictive; however, it is a general sufficient condition for arbitrary k and mixing constants $\{\varepsilon_i\}$. In specific cases, one will almost invariably be able to accommodate a much larger spectrum of c_i's in the definition of g(s). The following Lemma for a mixture of 2 normals gives evidence for this.

<u>Lemma 8</u>. Let $g(s) = \frac{1}{2} \sum_{i=1}^{2} c_i^{\frac{p}{2}} e^{-\frac{c_i}{2} s}$.

Let $a_i = \frac{c_i}{1+c_i}$, $i = 1,2$. Then $F\tau_\pi(F,W)$ is increasing in F for every fixed W if

$$(\alpha-2)^2\left(\frac{a_{max}}{a_{min}} - 1\right)^4 \leq 16\left(\frac{a_{max}}{a_{min}}\right)^3. \tag{3.10}$$

Proof: Multiplying throughout by F^{p+1} and setting $a_i F = x_i$, (3.6) becomes

$$\Sigma x_i^{\frac{p}{2}+1}\left(n+x_i+\frac{2a}{W}\right)^{2-\alpha} \Sigma x_i^{\frac{p}{2}}\left(n+x_i+\frac{2a}{W}\right)^{2-\alpha}$$

$$+ (\alpha-2)\left\{\Sigma x_i^{\frac{p}{2}+1}\left(n+x_i+\frac{2a}{W}\right)^{2-\alpha} \Sigma x_i^{\frac{p}{2}+1}\left(n+x_i+\frac{2a}{W}\right)^{1-\alpha}\right.$$

$$\left. - \Sigma x_i^{\frac{p}{2}+2}\left(n+x_i+\frac{2a}{W}\right)^{1-\alpha} \Sigma x_i^{\frac{p}{2}}\left(n+x_i+\frac{2a}{W}\right)^{2-\alpha}\right\} \geq 0. \tag{3.11}$$

The quantity within { } in (3.11) is

$$\sum_i \sum_j x_i^{\frac{p}{2}+1} x_j^{\frac{p}{2}+1} (n+x_i+\frac{2a}{W})^{1-\alpha}(n+x_j+\frac{2a}{W})^{2-\alpha}$$

$$- \sum_i \sum_j x_i^{\frac{p}{2}} x_j^{\frac{p}{2}+2} (n+x_i+\frac{2a}{W})^{2-\alpha}(n+x_j+\frac{2a}{W})^{1-\alpha}$$

$$= (n+\frac{2a}{W})\sum_i \sum_j x_i^{\frac{p}{2}} x_j^{\frac{p}{2}+1} (n+x_i+\frac{2a}{W})^{1-\alpha}(n+x_j+\frac{2a}{W})^{1-\alpha}(x_i-x_j)$$

$$= -(n+\frac{2a}{W})x_1^{\frac{p}{2}} x_2^{\frac{p}{2}}(n+x_1+\frac{2a}{W})^{1-\alpha}(n+x_2+\frac{2a}{W})^{1-\alpha}(x_1-x_2)^2. \qquad (3.12)$$

Thus (3.11) implies that it suffices to show

$$x_1^{p+1}(n+x_1+\frac{2a}{W})^{4-2\alpha} + x_2^{p+1}(n+x_2+\frac{2a}{W})^{4-2\alpha} + x_1^{\frac{p}{2}} x_2^{\frac{p}{2}}(n+x_1+\frac{2a}{W})^{2-\alpha}$$

$$(n+x_2+\frac{2a}{W})^{2-\alpha}(x_1+x_2) - (\alpha-2)(n+\frac{2a}{W})x_1^{\frac{p}{2}} x_2^{\frac{p}{2}}(n+x_1+\frac{2a}{W})^{1-\alpha}(n+x_2+\frac{2a}{W})^{1-\alpha}(x_1-x_2)^2 \geq 0 \qquad (3.13)$$

$$\Leftrightarrow \frac{x_1^{\frac{p}{2}+1}}{x_2^{\frac{p}{2}}} (n+x_1+\frac{2a}{W})^{2-\alpha}(n+x_2+\frac{2a}{W})^{\alpha-2} + \frac{x_2^{\frac{p}{2}+1}}{x_1^{\frac{p}{2}}} (n+x_1+\frac{2a}{W})^{\alpha-2}(n+x_2+\frac{2a}{W})^{2-\alpha}$$

$$+ (x_1+x_2) - (\alpha-2)(n+\frac{2a}{W}) \frac{(x_1-x_2)^2}{(n+x_1+\frac{2a}{W})(n+x_2+\frac{2a}{W})} \geq 0 \qquad (3.14)$$

$$= A + B + C - D \quad \text{(say)}. \qquad (3.15)$$

By the concavity of logz on $(0,\infty)$, $\log(\frac{A+B}{2}) \geq \frac{1}{2}\log A + \frac{1}{2}\log B = \log\sqrt{x_1x_2}$

$\Rightarrow A+B \geq 2\sqrt{x_1x_2}$. So is C.

Hence $A + B + C \geq 4\sqrt{x_1x_2}$. (3.16)

On the other hand, assuming without loss of generality that $a_1 = a_{max} \geq a_{min} = a_2$,

$$\frac{(n+\frac{2a}{W})(x_1-x_2)^2}{(n+x_1+\frac{2a}{W})(n+x_2+\frac{2a}{W})} \leq \frac{x_2^2\,(\frac{a_{max}}{a_{min}}-1)^2}{(n+x_1+\frac{2a}{W})} \leq x_2\,\frac{(\frac{a_{max}}{a_{min}}-1)^2}{\frac{a_{max}}{a_{min}}} . \tag{3.17}$$

The lemma now follows from (3.16) and (3.17).

It's clear that Lemma 8 is much stronger than Lemma 7 in the special case treated in Lemma 8. For example, with $p = 5$, $m = 5$, a non-informative prior on σ^2 (i.e., $\beta = 2$), Lemma 7 only allows $\frac{a_{max}}{a_{min}} \leq 1.2$ whereas Lemma 8 allows $\frac{a_{max}}{a_{min}} \leq 3.1$. It's not clear to us right now, however, if Lemma 8 can be generalized to an arbitrary finite mixture with $k > 2$. The next Lemma proves monotonicity of $F\tau_\pi(F,W)$ in F when g is a continuous scale-mixture of normals.

Lemma 9

Let $g(s) = \int e^{-\frac{s}{2\tau^2}} p(\tau^2)d\tau^2$. Let also $f(a)$ denote the density of $z = \frac{1}{1+\tau^2}$. Then $F\tau_\pi(F,W)$ is increasing in F (for fixed W) if $\frac{zf'(z)}{f(z)}$ is decreasing in z.

Proof: As in (2.11), it follows that

$$\tau_\pi(F,W) = \frac{\int_0^1 z^{\frac{p}{2}+1}(n+zF+\frac{2a}{W})^{2-\alpha}f(z)dz}{\int_0^1 z^{\frac{p}{2}}(n+zF+\frac{2a}{W})^{2-\alpha}f(z)dz} . \tag{3.18}$$

Let $0 < F_1 < F_2 < \infty$. We need to show

$$F_2 \cdot \frac{\int_0^1 z^{\frac{p}{2}+1}(n+zF_2+\frac{2a}{W})^{2-\alpha}f(z)dz}{\int_0^1 z^{\frac{p}{2}}(n+zF_2+\frac{2a}{W})^{2-\alpha}f(z)dz} \geq F_1 \cdot \frac{\int_0^1 z^{\frac{p}{2}+1}(n+zF_1+\frac{2a}{W})^{2-\alpha}f(z)dz}{\int_0^1 z^{\frac{p}{2}}(n+zF_1+\frac{2a}{W})^{2-\alpha}f(z)dz}$$

$$\Leftrightarrow \int_0^1\int_0^1 \Big\{(F_2x)^{\frac{p}{2}+1}(F_1y)^{\frac{p}{2}}(n+F_2x+\frac{2a}{W})^{2-\alpha}(n+F_1y+\frac{2a}{W})^{2-\alpha}$$

$$- (F_2x)^{\frac{p}{2}}(F_1y)^{\frac{p}{2}+1}(n+F_2x+\frac{2a}{W})^{2-\alpha}(n+F_1y+\frac{2a}{W})^{2-\alpha}\Big\} f(x)f(y)dxdy \geq 0$$

$$\Leftrightarrow \int_0^1\int_0^1 h(F_2x,\ F_1y)f(x)f(y)dxdy \geq 0,$$

where $h(x,y) = (x-y)(xy)^{\frac{p}{2}}(n+x+\frac{2a}{W})^{2-\alpha}(n+y+\frac{2a}{w})^{2-\alpha}$.

$$\Leftrightarrow \int_0^{F_2}\int_0^{F_1} h(x,y)f(\frac{x}{F_2})f(\frac{y}{F_1})dxdy \geq 0. \tag{3.19}$$

It's clear that $h(x,y) = -h(y,x)$ and $h(x,y) > 0$ for $x > y$. Consequently (3.19) will follow if we can show that

$$\int_0^{F_1}\int_0^{F_1} h(x,y)f(\frac{x}{F_2})f(\frac{y}{F_1})dxdy \geq 0$$

which in turn will follow if it can be proved that the density at (x,y) is greater than the density at (y,x) whenever $x > y$. This is what we prove below. Clearly, it needs to be shown that

$$F_2 > F_1,\ x > y \Rightarrow f(\frac{x}{F_2})f(\frac{y}{F_1}) \geq f(\frac{y}{F_2})f(\frac{x}{F_1})$$

$$\Leftrightarrow \frac{f(\frac{x}{F_2})}{f(\frac{y}{F_2})} \geq \frac{f(\frac{x}{F_1})}{f(\frac{y}{F_1})}\ . \tag{3.20}$$

For fixed x, y (x > y), define $p(a) = \frac{f(ax)}{f(ay)}$, a > 0.

Clearly, $p'(a) = \frac{xf'(ax)}{f(ay)} - \frac{yf'(ay)f(ax)}{f^2(ay)}$

$$= \frac{1}{af^2(ay)} \{axf'(ax)f(ay)-ayf(ax)f'(ay)\}$$

$$= \frac{f(ax)}{af(ay)} \{\frac{axf'(ax)}{f(ax)} - \frac{ayf'(ay)}{f(ay)}\}$$

$$\leq 0 \text{ since } \frac{zf'(z)}{f(z)} \text{ is } \downarrow \text{ in } z.$$

Hence $\frac{f(ax)}{f(ay)}$ decreasing in a if x > y, which proves (3.20). This proves the Lemma.

We now give examples of a few mixture distributions which satisfy the sufficient condition of Lemma 9.

Example 1

Let $p(\tau^2) = e^{-\alpha/\tau^2}(\tau^2)^{1-\beta}$, $\beta > 2$, $\alpha \geq \beta-1$. Then

$f(z) = e^{-\frac{\alpha z}{1-z}}(1-z)^{1-\beta}z^{\beta-3}$. By direct computation,

$\frac{zf'(z)}{f(z)} = -\frac{\alpha z}{(1-z)^2} + \frac{(\beta-1)z}{1-z} + (\beta-3)$, which is decreasing in z for $\alpha \geq \beta-1$.

Note that the family of mixing distributions $p(\tau^2)$ generate the spherically symmetric t and Cauchy priors.

Example 2. Let $p(\tau^2) = (\tau^2)^{n-1}(1+\tau^2)^{-(m+n)}$.

Then $f(z) = z^{m-1}(1-z)^{n-1}$.

Again, by direct computation, $\frac{zf'(z)}{f(z)} = (m-1) - \frac{(n-1)z}{1-z}$, which is decreasing in z for $n \geq 1$.

Again, note that these are the scale mixtures which generate Proper Bayes minimax estimates for the multivariate normal mean with a common unknown variance (see Strawderman (1973)).

In the next section, we will treat certain other priors which cannot be written as scale mixtures of normals. Before proceeding to the next section, we formally state a theorem below which has essentially been proved already.

Theorem 1. Let the condtional density of $\underset{\sim}{\theta}$ given σ^2 be of the form $g(\frac{\Sigma\theta_j^2}{\sigma^2})$ and let σ^2 have a prior of the form given by (2.1). If g is a finite scale mixture of normals satsifying the conditions of either Lemma 7 or Lemma 8, or if g is a continuous scale-mixture satisfying the hypothesis of Lemma 9, then $\delta^*_{\pi,0}$ defined as

$$\delta^*_{\pi,0}(\underset{\sim}{X},W) = (1-\tau_\pi(F,W))\cdot\underset{\sim}{X} \quad \text{if} \quad F\tau_\pi(F,W) \leq 2(p-2)$$

$$= (1 - \frac{2(p-2)}{F})\cdot\underset{\sim}{X} \quad \text{otherwise}$$

minimizes for every $\underset{\sim}{X}$ and W the posterior expected loss within the class $\mathcal{D}^*_0$.

Proof: Follows from Lemmas 5, 6, 7, 8, and 9.

4. Priors which are not scale mixtures of normals

It was proved in Berger (1975) that if $g(\underset{\sim}{\theta})$ is a spherically symmetric prior of the form $g(\Sigma\theta_j^2) = \int e^{-\frac{1}{2\sigma^2}\Sigma\theta_i^2} dF(\sigma^2)$, then g is completely monotonic (i.e., $(-1)^n \frac{d^n}{ds^n} g(s) \geq 0 \ \forall s$), and conversely a completely monotonic prior is some scale mixture of normals as above. Relatively simple examples of functions which are not completely monotonic are $g(s) = s^n e^{-s/2}$, $n \neq 0$. Note that g is log-convex and decreasing if $n < 0$ and log-concave if $n > 0$. In view of the discussion following

Lemma 5, we should try to prove that $\tau_\pi(F,W) \geq 0$ for $n > 0$ and that $F\tau_\pi(F,W)$ is increasing for $n < 0$. It turns out that one needs conditions on n in both cases. First we derive an alternative expression for $\tau_\pi(F,W)$ which comes very handy later on.

Lemma 10. Define $m(\frac{\Sigma x_j^2}{\sigma^2}) = \int \frac{e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j-x_j)^2}}{\sigma^p} g(\frac{\Sigma\theta_j^2}{\sigma^2})d\underset{\sim}{\theta}$ where g is arbitrary subject to existence of all integrals and derivatives considered below. Then $\tau_\pi(F,W)$ can also be expressed as

$$\tau_\pi(F,W) = -2\,\frac{\int_0^\infty m'(zF)e^{-\frac{z}{2}[(m+2)+\frac{2a}{W}]}z^{\alpha-3}dz}{\int_0^\infty m(zF)e^{-\frac{z}{2}[(m+2)+\frac{2a}{W}]}z^{\alpha-3}dz}\,. \tag{4.1}$$

Proof: Familiar calculations and arguments similar to those of Lemma 1 give (4.1).

Lemma 11. Let $g(s) = s^n e^{-\frac{s}{2}}$, where $0 < n \leq \frac{p}{2}$. Let also a = 0 in (2.1). Then, $\tau_\pi(F,W) \equiv \tau_\pi(F) \geq 0$ for all F.

Proof: In view of Lemma 4, if there exists F_0 such that $\tau_\pi(F_0) < 0$, then $\limsup_{F\to 0} \tau_\pi(F) < 0$. We will show that for $n \leq \frac{p}{2}$, $\lim_{F\to 0} \tau_\pi(F)$ exists and is ≥ 0. Using the arguments of Lemma 2, one has, for any $z > 0$,

$$m(z) = e^{-\frac{z}{2}}\cdot\Phi(z), \text{ where}$$

$$\Phi(z) = \text{constant} \times \int_0^\infty \int_{-1}^{1} e^{\sqrt{sz}t}(1-t^2)^{\frac{p-3}{2}}dt\; s^{n+\frac{p-2}{2}}e^{-s}ds$$

$$\therefore m'(z) = -\frac{1}{2}e^{-\frac{z}{2}}\Phi(z)+e^{-\frac{z}{2}}\Phi'(z). \tag{4.2}$$

We would prove that $\lim_{z\to 0} \Phi'(z) = \Phi'(0)$ (say) exists and that $\Phi'(0) - \frac{1}{2}\Phi(0) \le 0$, which in view of (4.1) and an application of the Dominated Convergence Theorem would then imply that $\lim \tau_\pi(F) \ge 0$. For purposes of showing that $\Phi'(0) - \frac{1}{2}\Phi(0) \le 0$, we would ignore the constant in the definition of $\Phi(z)$. Clearly, $\Phi(z) = \int_0^\infty \int_0^1 (e^{\sqrt{sz}t} + e^{-\sqrt{sz}t})(1-t^2)^{\frac{p-3}{2}} dt\, s^{n+\frac{p-2}{2}} e^{-s} ds.$

By the Dominated Convergence Theorem,

$$\Phi(0) = \lim_{z\to 0} \Phi(z) = B(\tfrac{1}{2}, \tfrac{p-1}{2}) \cdot \Gamma(n+ \tfrac{p}{2}). \tag{4.3}$$

Again, by the Dominated Convergence Theorem,

$$\lim_{z\to 0} \Phi'(z) = \int_0^\infty \int_0^1 t^2(1-t^2)^{\frac{p-3}{2}} dt\, e^{-\frac{s}{2}} s^{n+\frac{p}{2}} ds$$

$$= \frac{1}{2} \cdot B(\tfrac{3}{2}, \tfrac{p-1}{2})\, \Gamma(n+ \tfrac{p}{2} +1) \tag{4.4}$$

(4.3) and (4.4) imply that $\Phi'(0) - \frac{1}{2}\Phi(0) \le 0$

$$\text{iff} \quad \frac{1}{2}\frac{\Gamma(\frac{3}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{p+2}{2})} \cdot \Gamma(n+\tfrac{p}{2}+1) - \frac{1}{2}\frac{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})}{\Gamma(\frac{p}{2})}\Gamma(n+ \tfrac{p}{2}) < 0$$

$$\text{iff} \quad n \le \tfrac{p}{2}.$$

<u>Theorem 2</u>. $\delta^*_{\pi,0}(\underset{\sim}{X},W)$ defined as before is the restricted risk Bayes rule within $\mathcal{D}^*_0$ if the conditional density of $\underset{\sim}{\theta}$ given σ^2 is

$g(\frac{\Sigma\theta_j^2}{\sigma^2}) = (\frac{\Sigma\theta_j^2}{\sigma^2})^n e^{-\frac{1}{2\sigma^2}\Sigma\theta_j^2}$ with $0 < n \le \frac{p}{2}$ and σ^2 has density given by (2.1) with $a = 0$.

Proof: Follows from Lemmas 4, 5, and 11.

We now prove that priors of the form $g(s) = s^n e^{-\frac{s}{2}}$ for $n < 0$ can also be handled if n is sufficiently close to zero. The arguments used below do not exactly say how close to zero n has to be, but we think it is possible to specify such a lower bound on n. Also, simply for convenience, we have taken $a = 0$ (in the prior for σ^2) in the following Lemma, although $a > 0$ can also be handled using essentially the same arguments. Also recall that the definition of $\delta^*_{\pi,0}$ really needs monotonicity of $F\tau_\pi(F,W)$ only in the region $F\tau_\pi(F,W) \leq 2(p-2)$.

Lemma 12.

Let $g(s) = s^n e^{-\frac{s}{2}}$, where $n < 0$. Also let $a = 0$ in (2.1). Then there exists an $n_0 > 0$ such that for $n > -n_0$, $F\tau_\pi(F)$ is increasing in F for $F\tau_\pi(F) \leq 2(p-2)$.

Proof: Define for $n \leq 0$,

$$h(n,F) = (F\tau_\pi(F))' = \tau_\pi(F) + F\tau'_\pi(F).$$

Let $C_n = \{F: \ F\tau_\pi(F) \leq 2(p-2)\}$.

We need to show that $\exists n_0 \ni n > -n_0 \Rightarrow h(n,F) \geq 0 \ \forall F \in C_n$.

We would prove this as follows. If the above assertion is false, then we can find a sequence (n,F_n) with $n \to 0$ and $F_n \in C_n$ such that $h(n,F_n) < 0$ for every n. We would prove that F_n (or a subsequence thereof) necessarily converges to F_0, where F_0 may be zero. If $F_0 > 0$, by the joint continuity of h(n,F) in n and F at $n = 0$ and all $F > 0$, we would have a contradiction because $h(0,F) = \frac{1}{2} \ \forall F > 0$. We would also show that $\lim_{n,F\to 0} h(n,F)$ exists and is also $\frac{1}{2}$ so that we would arrive at the same contradiction even if F_0 were to be zero. We now prove these things one at a time.

In order to prove that F_n has a convergent subsequence, we would prove that $C_n \subseteq C_0$ ($\forall n < 0$) and hence $F_n \in C_0$ $\forall n$. Since $C_0 \cup \{0\}$ is a compact set, F_n has a subsequence converging to some $F_0 \in C_0$.

Using (4.1), $\forall n < 0$,

$$\tau_\pi(F) = -2 \frac{\int m_n'(zF) e^{-\frac{z}{2}\cdot(m+2)} z^{\alpha-3} dz}{\int m_n(zF) e^{-\frac{z}{2}(m+2)} z^{\alpha-3} dz}, \tag{4.5}$$

where $m_n(\frac{\Sigma x_j^2}{\sigma^2})$ stands for $\int \frac{1}{\sigma^p} e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2} g(\Sigma\theta_j^2) d\underset{\sim}{\theta}$.

Now
$$\frac{m_n(z)}{m_0(z)} = \frac{\int \frac{1}{\sigma^p} e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2} s^n e^{-\frac{s}{2}} d\underset{\sim}{\theta}}{\int \frac{1}{\sigma^p} e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j - x_j)^2} e^{-\frac{s}{2}} d\underset{\sim}{\theta}} \tag{4.6}$$

where $s = \frac{\Sigma\theta_j^2}{\sigma^2}$.

By using the arguments of Lemmas 2 and 3, $\frac{m_n(z)}{m_0(z)}$ can be expressed as $E(s^n)$, where s (given z) has the density

$$p(s|z) = I_{\frac{p-2}{2}}(\sqrt{sz}) s^{\frac{p-2}{4}} e^{-s}.$$

It was proved in Lemma 3 that $p(s|z)$ is MLR in z; furthermore, s^n is monotone decreasing in s, $\forall n < 0$, which now implies that $\frac{m_n(z)}{m_0(z)}$ is decreasing in z;

$$\therefore \quad \frac{m_n'(z)}{m_n(z)} \leq \frac{m_0'(z)}{m_0(z)} = -\frac{1}{4}$$

$$\Rightarrow \quad \tau_\pi(F) \geq \frac{1}{2} \quad \forall F \quad \text{(using (4.5))}. \tag{4.7}$$

$$\therefore \quad F \in C_n \Rightarrow \frac{F}{2} \leq F\tau_\pi(F) \leq 2(p-2) \Rightarrow F \in C_0.$$

This proves that $C_n \subseteq C_0$, $\forall n < 0$.

It now only remains to show that the joint limit $\lim\limits_{n,F\to 0} h(n,F)$ exists and is equal to $h(0,0) = \frac{1}{2}$. Towards this end, recall that $h(n,F) = \tau_\pi(F) + F\tau_\pi'(F)$ and hence it suffices to show that $\lim\limits_{n,F\to 0} \tau_\pi(F)$ and $\lim\limits_{n,F\ 0} \tau_\pi'(F)$ both exist and equal $\frac{1}{2}$ and 0 respectively.

By definition,

$$\tau_\pi(F) = -2\,\frac{\int m_n'(zF)e^{-\frac{(m+2)z}{2}} z^{\alpha-3}dz}{\int m_n(zF)e^{-\frac{(m+2)z}{2}} z^{\alpha-3}dz} \tag{4.8}$$

$$\therefore \quad \tau_\pi'(F) = -2\left\{\frac{\int m_n''(zF)e^{-(m+2)\frac{z}{2}}z^{\alpha-2}dz}{\int m_n(zF)e^{-(m+2)\frac{z}{2}}z^{\alpha-3}dz} - \frac{\int m_n'(zF)e^{-(m+2)\frac{z}{2}}z^{\alpha-3}dz\int m_n'(zF)e^{-(m+2)\frac{z}{2}}z^{\alpha-2}dz}{(\int m_n(zF)e^{-(m+2)\frac{z}{2}}z^{\alpha-3}dz)^2}\right\} \tag{4.9}$$

It suffices to show that each of the following four integrals

$$I_1(n,F) = \int m_n(zF)e^{-(m+2)\frac{z}{2}}z^{\alpha-3}dz$$

$$I_2(n,F) = \int m_n'(zF)e^{-(m+2)\frac{z}{2}} z^{\alpha-3}dz$$

$$I_3(n,F) = \int m_n''(zF)e^{-(m+2)\frac{z}{2}} z^{\alpha-2}dz$$

$$I_4(n,F) = \int m_n'(zF)e^{-(m+2)\frac{z}{2}} z^{\alpha-2}dz \tag{4.10}$$

are jointly continuous in n and F at (0,0). We would prove this only for $I_1(n,F)$. The proof runs along the same lines for the other integrals. We need to show that $0 < F < \delta_0$ and $-n < \varepsilon \Rightarrow$

$$m_n(zF) < M(z) \text{ such that}$$

$$\int M(z)e^{-(m+2)\frac{z}{2}} z^{\alpha-3}dz < \infty. \tag{4.11}$$

The Dominated Convergence Theorem would then imply that $I_1(n,F)$ is jointly continuous at $(n,F) = (0,0)$. Assume without loss $\delta_0 \leq 1$ and $\varepsilon \leq 1$. By (4.2),

$$m_n(zF) = C \times e^{-\frac{z}{2}} \int_0^\infty \int_0^1 (e^{\sqrt{szF}\,t} + e^{-\sqrt{szF}\,t})(1-t^2)^{\frac{p-3}{2}} e^{-s} s^{n+\frac{p-2}{2}} ds$$

$$\leq 2C\, e^{-\frac{z}{2}} \int_0^\infty \int_0^1 (1-t^2)^{\frac{p-3}{2}} e^{\sqrt{sz}}\, e^{-(1-\delta)s} e^{-\delta s} s^{n+\frac{p-2}{2}} ds \tag{4.12}$$

(where $0 < \delta \leq \frac{1}{2}$ is a fixed number)

$\leq 2C^* e^{-\frac{z}{2}} \int_0^\infty e^{\sqrt{sz}-(1-\delta)z} ds$, since $\int_0^1 (1-t^2)^{\frac{p-3}{2}} dt$ is an absolute constant and $e^{-\delta s} s^{n+\frac{p-2}{2}}$ is uniformly bounded in s and n by an absolute constant. (4.13)

Completing squares, (4.13) gives

$$m_n(zF) \le 2C^* e^{-\frac{z}{2}} \int_0^\infty e^{-(1-\delta)[\sqrt{s} - \frac{\sqrt{z}}{2(1-\delta)}]^2} ds \times e^{\frac{z}{4(1-\delta)}}$$

$$\le K\, e^{-z\frac{(1-2\delta)}{4(1-\delta)}} \tag{4.14}$$

(4.14) clearly implies (4.11).

The joint continuity of I_2, I_3, and I_4 uses similar bounds but $m_n'(zF)$ and $m_n''(zF)$ need to be handled now. No problems arise and we omit the details. This proves the lemma.

Theorem 3. There exists a small enough $n_0 > 0$ such that if the conditional density of $\underset{\sim}{\theta}$ given σ^2 is $g(s) = s^n e^{-\frac{s}{2}}$ with $s = \frac{\Sigma\theta_j^2}{\sigma^2}$ and $0 < -n \le n_0$, and if σ^2 has density given by (2.1) with $a = 0$, then $\delta^*_{\pi,0}$ is the Restricted Risk Bayes rule within $\mathcal{D}^*_0$.

Proof: Follows from Lemmas 5 and 12.

5. Bayes risks, RSL's, and an example.

In this section we explicitly implement the analysis for a $N(\underset{\sim}{0}, I)$ prior at the first stage and an inverted Gamma prior as in (2.1) at the second stage. It is customary to judge the amount of gains in subjective Bayes risk one has to sacrifice in exchange for Bayesian robustness in terms of the Relative Savings Loss (RSL) of Efron and Morris (1971), defined as,

$$RSL(\delta^*_{\pi,0}) = \frac{r(\pi,\delta^*_{\pi,0}) - r(\pi,\delta_\pi)}{r(\pi,\delta_0) - r(\pi,\delta_\pi)} \,. \tag{4.1}$$

See Efron and Morris (1971), Berger (1982), or Berger and DasGupta (1986) for discussions and conditional versions of the RSL. We would develop upper bounds on the RSL for a general spherically symmetric prior at the first stage (recall that low values of the RSL are desirable) and work out these bounds in the normal case (the bounds have been obtained such that the RSL is exactly equal to the upper bound in the special normal case).

Lemma 13. For any $\delta(\underset{\sim}{X},W)$,

$$r(\pi,\delta) - r(\pi,\delta_\pi) = E_m\{||\delta-\delta_\pi||^2 \cdot E(\frac{1}{\sigma^2}\Big|\underset{\sim}{X},W)\},$$

where E_m denotes expectation with respect to the marginal distribution of $\underset{\sim}{X}$ and W.

Proof: Follows from definition of $L(\underset{\sim}{\theta},\sigma^2,\delta)$.

Lemma 14. The joint density of $\underset{\sim}{X}$ and W is given as

$$f(\underset{\sim}{x},W) = \text{constant} \times \frac{W^{\frac{m}{2}-1}}{(WF)^{\alpha-3}} \int_0^\infty m(z) e^{-\frac{z}{2F}(n+\frac{2a}{W})} z^{\alpha-4} dz,$$

where $F = \frac{\Sigma x_j^2}{W}$, $n = m+2$, and $\alpha = \frac{p}{2} + \frac{m}{2} + \beta$.

Proof: By definition,

$$f(\underset{\sim}{x},W) = \iint \frac{1}{\sigma^p} e^{-\frac{1}{2\sigma^2}\Sigma(\theta_j-x_j)^2} \frac{g(\frac{\Sigma\theta_j^2}{\sigma^2})}{\sigma^p} e^{-\frac{(m+2)W}{2\sigma^2}} (\frac{W}{\sigma^2})^{\frac{m}{2}-1} e^{-\frac{a}{\sigma^2}} (\frac{1}{\sigma^2})^{\beta-1} d\underset{\sim}{\theta} d\sigma^2$$

$$= W^{\frac{m}{2}-1} \int m(\frac{\Sigma x_j^2}{\sigma^2}) e^{-\frac{W}{2\sigma^2}[(m+2)+\frac{2a}{W}]} (\frac{1}{\sigma^2})^{\alpha-2} d\sigma^2$$

$$= \text{constant} \times \frac{W^{\frac{m}{2}-1}}{(FW)^{\alpha-3}} \int_0^\infty m(z) e^{-\frac{z}{2F}(n+\frac{2a}{W})} z^{\alpha-4} dz \text{ (substitute } z = \frac{\Sigma x_j^2}{\sigma^2}\text{)}.$$

<u>Lemma 15</u>. The marginal density of F and W is given by

$$f(F,W) = \text{constant} \times \frac{W^{\frac{p}{2}+\frac{m}{2}-1} F^{\frac{p}{2}-1}}{(FW)^{\alpha-3}} \int_0^\infty m(z) e^{-\frac{z}{2F}(n+\frac{2a}{W})} z^{\alpha-4} dz.$$

Proof: Transforming to Polar coordinates gives the marginal density of U and W, where $U = ||\underset{\sim}{X}||$. Now make another transformation to $(F,W) = (\frac{U^2}{W}, W)$ and the result follows.

<u>Lemma 16</u>.

$$E(\frac{1}{\sigma^2}|\underset{\sim}{X},W) = \text{constant} \frac{\int m(z) e^{-\frac{z}{2F}(n+\frac{2a}{W})} z^{\alpha-3} dz}{FW \int m(z) e^{-\frac{z}{2F}(n+\frac{2a}{W})} z^{\alpha-4} dz}.$$

Proof: By definition,

$$E(\frac{1}{\sigma^2} \underset{\sim}{X},W) = \frac{\int m(\frac{\Sigma x_j^2}{\sigma^2}) e^{-\frac{W}{2\sigma^2}(n+\frac{2a}{W})} (\frac{1}{\sigma^2})^{\alpha-1} d\sigma^2}{\int m(\frac{\Sigma x_j^2}{\sigma^2}) e^{-\frac{W}{2\sigma^2}(n+\frac{2a}{W})} (\frac{1}{\sigma^2})^{\alpha-2} d\sigma^2},$$

from which the result follows on substituting $\frac{\Sigma x_j^2}{\sigma^2} = z$.

<u>Lemma 17</u>. Assume $\sup_z \frac{m'(z)}{m(z)} = -M < 0$. Then

$$r(\pi,\delta_0) - r(\pi,\delta_\pi) \geq K \times 4M^2 \cdot (\frac{2}{n})^{\frac{m}{2}} a^{3-\beta} \Gamma\frac{m}{2}\, \Gamma(\beta-3) \int_0^\infty m(z) z^{\frac{p}{2}} dz,$$

where n = m+2, and K is a suitable absolute constant.

Proof: From Lemma 13,

$$r(\pi,\delta_0) - r(\pi,\delta_\pi) = E_m\{\tau_\pi^2(F,W)\cdot||\underset{\sim}{X}||^2\cdot E(\frac{1}{\sigma^2}|\underset{\sim}{X},W)\}$$

$$= 4\times\text{constant} \iint \frac{(\int_0^\infty m'(z)e^{-\frac{z}{2F}(n+\frac{2a}{W})}z^{\alpha-3}dz)^2}{(\int_0^\infty m(z)e^{-\frac{z}{2F}(n+\frac{2a}{W})}z^{\alpha-3}dz)^2} \cdot \frac{\int_0^\infty m(z)e^{-\frac{z}{2F}(n+\frac{2a}{W})}z^{\alpha-3}dz}{\int_0^\infty m(z)e^{-\frac{z}{2F}(n+\frac{2a}{W})}z^{\alpha-4}dz}$$

$$\times W^{2-\beta}F^{2-\frac{m}{2}-\beta}(\int_0^\infty m(z)e^{-\frac{z}{2F}(n+\frac{2a}{W})}z^{\alpha-4}dz)dFdW$$

(using Lemmas 15 and 16)

$$\geq 4M^2\times\text{constant} \iint W^{2-\beta}F^{2-\frac{m}{2}-\beta}(\int_0^\infty m(z)e^{-\frac{z}{2F}(n+\frac{2a}{W})}z^{\alpha-3}dz)dFdW$$

$$= \text{constant}\times 4M^2\cdot\Gamma(\frac{m}{2}+\beta-3)\cdot 2^{\frac{m}{2}+\beta-3}\ \frac{(2a)^{3-\beta}}{n^{\frac{m}{2}}}\cdot B(\frac{m}{2},\ \beta-3)\int_0^\infty m(z)z^{\frac{p}{2}}dz$$

(integrating first with respect to F and W)

$$= \text{constant}\times 4M^2\cdot(\frac{2}{n})^{\frac{m}{2}}\cdot a^{3-\beta}\cdot\Gamma\frac{m}{2}\ \Gamma(\beta-3)\int_0^\infty m(z)z^{\frac{p}{2}}dz.$$

<u>Lemma 18</u>. Assume g(s) is such that $F\tau_\pi(F,W)$ is increasing in F and decreasing in W (for fixed W and fixed F respectively). Then

$$r(\pi,\delta^*_{\pi,0}) - r(\pi,\delta_\pi)$$

$$\leq 4K\Gamma(\beta-3)a^{3-\beta}(\int_{\psi(0)}^\infty F^{-\frac{m}{2}-1}\int_0^\infty \frac{(m'(z)+\frac{p-2}{F}m(z))^2}{m(z)}e^{-\frac{nz}{2F}}z^{\alpha-\beta}dz)dF,$$

where $\psi(0) = \inf_W \sup\{F: F\tau_\pi(F,W) \leq 2(p-2)\}$.

Proof: Using Lemmas 13, 15, and 16, $r(\pi,\delta^*_{\pi,0}) - r(\pi,\delta_\pi)$

$$= \iint \left\{ \tau_\pi^2(F,W) - \frac{4(p-2)\tau_\pi(F,W)}{F} + \frac{4(p-2)^2}{F^2} \right\} \left(\int m(z) e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz \right)$$

$$\times W^{2-\beta} F^{2-\frac{m}{2}-\beta} I_{F\tau_\pi(F,W) > 2(p-2)} dF dW. \qquad (5.2)$$

Define now, for each $W > 0$, $\psi(W) = \sup\{F: F\tau_\pi(F,W) \le 2(p-2)\}$. We claim that $\psi(W)$ is non-decreasing in W. If not, $\exists\ 0 < W_1 < W_2 < \infty$ such that $\psi(W_1) > \psi(W_2)$. By definition of $\psi(W_1)$, $\exists\ F_0 > \psi(W_2) \ni F_0\tau_\pi(F_0,W_1) \le 2(p-2)$ $\Rightarrow F_0\tau_\pi(F_0,W_2) \le F_0\tau_\pi(F_0,W_1) \le 2(p-2)$, which contradicts the definition of $\psi(W_2)$.

Hence, letting $\psi(0) = \inf_{W>0} \psi(W)$, one has, from (5.2) and Lemma 10,

$r(\pi,\delta^*_{\pi,0}) - r(\pi,\delta_\pi)$

$$\le K\cdot\int_0^\infty \int_{\psi(0)}^\infty \left\{ 4\frac{\left(\int_0^\infty -m'(z) e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz\right)^2}{\left(\int_0^\infty m(z) e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz\right)^2} + \frac{8(p-2)}{F} \frac{\int_0^\infty m'(z) e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz}{\int_0^\infty m(z) e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz} \right.$$

$$\left. + \frac{4(p-2)^2}{F^2} \right\} \left(\int_0^\infty m(z) e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz \right) W^{2-\beta} F^{2-\frac{m}{2}-\beta} dF dW$$

$$(5.3)$$

$$\le K\cdot\int_0^\infty \int_{\psi(0)}^\infty \left\{ 4\left(\int_0^\infty \frac{(-m'(z))^2}{m(z)} e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz\right) + \frac{8(p-2)}{F}\left(\int_0^\infty m'(z) e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz\right.\right.$$

$$\left.\left. + \frac{4(p-2)^2}{F^2} \left(\int_0^\infty m(z) e^{-\frac{z}{2F}\left(n+\frac{2a}{W}\right)} z^{\alpha-3} dz\right) \right\} F^{2-\frac{m}{2}-\beta} W^{2-\beta} dF dW \qquad (5.4)$$

(use Schwartz's inequality on the first term)

$$= K\cdot\Gamma(\beta-3)a^{3-\beta}\int_{\psi(0)}^{\infty} F^{\beta-3+2-\frac{m}{2}-\beta}\int_0^{\infty}\left\{4\cdot\frac{(-m'(z))^2}{m(z)} + \frac{8(p-2)}{F}m'(z) + \frac{4(p-2)^2}{F^2}m(z)\right\}$$

$$\times e^{-\frac{nz}{2F}} z^{\alpha-3+3-\beta} dz dF$$

(integrating out W)

$$= 4K\Gamma(\beta-3)a^{3-\beta}\int_{\psi(0)}^{\infty} F^{-\frac{m}{2}-1}\left(\int_0^{\infty}\frac{(m'(z)+\frac{p-2}{F}m(z))^2}{m(z)} e^{-\frac{nz}{2F}} z^{\alpha-\beta}dz\right)dF \qquad (5.5)$$

The following corollaries give expression for $r(\pi,\delta_0) - r(\pi,\delta_\pi)$ and $r(\pi,\delta^*_{\pi,0}) - r(\pi,\delta_\pi)$ in the special case when conditional on σ^2, $\underset{\sim}{\theta} \sim N_p(\underset{\sim}{0},\sigma^2 I)$.

<u>Corollary 1</u>. Let $g(s) = e^{-\frac{s}{2}}$. Then

$$r(\pi,\delta_0)-r(\pi,\delta_\pi) = K\cdot 4^{\frac{p}{2}}\cdot\left(\frac{2}{n}\right)^{\frac{m}{2}} a^{3-\beta}\Gamma\frac{m}{2}\,\Gamma(\beta-3)\Gamma\left(\frac{p}{2}+1\right).$$

Proof: Follows from Lemma 17 on observing that $m(z) = e^{-\frac{z}{4}}$. Note that all upper bounds used in Lemma 17 are equalities in the normal case.

<u>Corollary 2</u>. Let $g(s) = e^{-\frac{s}{2}}$.

Then $r(\pi,\delta^*_{\pi,0}) - r(\pi,\delta_\pi)$

$$= K\cdot 4^{\frac{p}{2}}\cdot\left(\frac{2}{n}\right)^{\frac{m}{2}} a^{3-\beta}\cdot\Gamma(\beta-3)\,\frac{1}{n^2}\,\Gamma\left(\frac{p}{2}+\frac{m}{2}+1\right)\times\Big[n^2B_x\left(\frac{p}{2}+1,\frac{m}{2}\right)-4(p-2)nB_x\left(\frac{p}{2},\frac{m}{2}+1\right)$$

$$+ 4(p-2)^2B_x\left(\frac{p}{2}-1,\frac{m}{2}+2\right)\Big],$$

where $B_x(r,s) = \int_x^1 t^{r-1}(1-t)^{s-1}$ is the incomplete Beta integral, and $x = \frac{2(p-2)}{2(p-2)+n}$.

Proof: In (5.5), using $m(z) = e^{-\frac{z}{4}}$, and using the facts that

$$\int_0^\infty e^{-\frac{z}{4}(1+\frac{2n}{F})} z^{\alpha-\beta} dz = \frac{4^{\frac{p}{2}+\frac{m}{2}+1}\,\Gamma(\frac{p}{2}+\frac{m}{2}+1)\cdot F^{\frac{p}{2}+\frac{m}{2}+1}}{(F+2n)^{\frac{p}{2}+\frac{m}{2}+1}},$$ that $\psi(0) = 4(p-2)$

for $g(s) = e^{-\frac{s}{2}}$, and that for $s > 0$, $\int_{4(p-2)}^\infty \frac{F^{\frac{p}{2}-s}}{(F+2n)^{\frac{p}{2}+\frac{m}{2}+1}} dF$

$= (2n)^{-\frac{m}{2}-s} B_x(\frac{p}{2}-s+1, \frac{m}{2}+s)$, one gets the result. Again note that all lower bounds used in Lemma 18 are exact equalities.

Theorem 4. Let $\theta|\sigma \sim N_p(0,\sigma^2 I)$, and let σ^2 have density given by (2.1). Then the Relative Savings Loss of the restricted risk Bayes rule is given by

$$RSL(\delta^*_\pi,0) = \frac{1}{n^2 B(\frac{p}{2}+1, \frac{m}{2})}\Big[n^2 B_x(\frac{p}{2}+1, \frac{m}{2}) - 4(p-2)nB_x(\frac{p}{2}, \frac{m}{2}+1)$$
$$+4(p-2)^2 B_x(\frac{p}{2}-1, \frac{m}{2}+2)\Big],$$

where $n = m+2$.

Proof: Follows from Corollaries 1 and 2.

It's interesting that for each p and m, the RSL's are independent of the hyperparameters of the prior on σ^2. The RSL's are monotone decreasing in m and p, and for large m, decreases to zero at an exponential rate. This is the assertion of the following proposition.

Proposition. If θ and σ^2 have prior distributions as in Theorem 4 then

$$\lim_{m\to\infty} RSL(\delta^*_{\pi,0}) = 1-\psi_{p+2}(2(p-2)) + \frac{4(p-2)}{p}\{\psi_p(2(p-2))-\psi_{p-2}(2(p-2))\},$$

where $\psi_p(z) = p[\chi^2(p) \le z]$.

Proof: First note that for s > 0,

$$R(m,s) = \frac{B_X(\frac{p}{2}+1-s, \frac{m}{2}+s)}{B(\frac{p}{2}+1-s, \frac{m}{2}+s)} = P[X \ge \frac{2(p-2)}{2(p-1)+m}], \tag{5.6}$$

where $X \sim Beta(\frac{p}{2}+1-s, \frac{m}{2}+s)$. Using the fact that $\frac{\frac{m}{2}+s}{\frac{p}{2}+1-s}\cdot\frac{X}{1-X} \sim F(p+2-2s, m+2s)$,

one has, from (5.6),

$$R(m,s) = P[F \ge \frac{\frac{m}{2}+s}{\frac{m}{2}+1}\cdot\frac{2(p-2)}{(p-2)+4-2s}]. \tag{5.7}$$

As $m\to\infty$, $F \overset{\mathcal{L}}{\to} \frac{\chi^2(p+2-2s)}{p+2-2s}$. Therefore, by Slutsky's Theorem,

$R(m,s) \to P[\chi^2(p+2-2s) \ge 2(p-2)] = 1-\psi_{p+2-2s}(2(p-2))$. The result now follows from Theorem 4 using the fact that $\frac{B(\frac{p}{2}+1-s, \frac{m}{2}+s)}{B(\frac{p}{2}+1, \frac{m}{2})} = \frac{\Gamma(\frac{p}{2}+1-s)\Gamma(\frac{m}{2}+s)}{\Gamma(\frac{p}{2}+1)\ \Gamma(\frac{m}{2})}$.

Theorem 4 is a little surprising; it enables us to calculate the RSL's in the special normal case in closed analytical forms. In what follows, we have provided a table for RSL's of the restricted risk Bayes rule. Non-normal priors are also perfectly manageable because of Lemmas 17 and 18; however, one possibly has to resort to numerical integration in such cases, and the theory of this section gives only upper bounds for the RSL's when the prior is non-normal and not exact RSL's. Finally, we remark that a normal prior at the first stage should not be a cause of concern from robust Bayesian viewpoints, because $\delta^*_{\pi,0}$ is full minimax to start with and this has the ultimate robustness built-in in it. However, one should perhaps also try t, Cauchy, and double-exponential first stage priors; work in this direction is currently under progress.

Table 1: RSL($\delta^*_{\pi,0}$) when $\underset{\sim}{\theta}|\sigma^2 \sim N(0,\sigma^2 I)$

m \ p	3	4	5	6	10	12	15
4	.50000	.36000	.29289	.25364	.18614	.17129	.15718
8	.41377	.26030	.19065	.15171	.08926	.07656	.06494
12	.37858	.22138	.15225	.11470	.05744	.4652	.03687
20	.34756	.18814	.12047	.08496	.03424	.02545	.01810
30	.33101	.17089	.10443	.07037	.02402	.01660	.01071
50	.31724	.15681	.09163	.05899	.01680	.01064	.00607
∞	.29567	.13534	.07269	.04275	.00804	.00400	.00159

The numbers indicate that if one has roughly 20 to 30 degrees of freedom for estimating σ^2, then for $p \geq 6$ the sacrifices in the subjective Bayes risk are nominal, yet we are assured of full robustness via minimaxity. Note however that the rate of decrease to the limiting value for each fixed p is going to be small and it would take quite a large sample for estimating σ^2 before the RLS's are practically the same as the limiting values. Also, interestingly, for each p, the limiting RSL's are the same as the RSL's of the restricted risk Bayes rules for the known variance case (see Berger (1982)). In short, the moral of this paper is that striving for Bayesian robustness, contrary to popular opinion, is not a wild goose chase; at least in most normal problems, it's something that can be achieved in small enough dimensions without any perceivable dent on the subjective Bayes risks. An even more conservative formulation of the problem (now under study) would be to try to limit the RSL's simultaneously under a class of priors and a class of losses. Of substantial practical interest is also the problem when the variances are different. Of necessity, the technicalities are more involved.

Acknowledgments. The first author is grateful to James Berger for valuable discussions, constructive suggestions, and active encouragement throughout the work. Our appreciation also goes to Peter Matthews for several useful suggestions.

Bibliography

Baranchik, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. Ann. Math. Statist. 41, 642-645

Berger, James (1975). Minimax estimation of location vectors for a wide class of densities. Ann. Statist. 3, 1318-1328.

Berger, James (1982). Estimation in continuous exponential families: Bayesian estimation subject to risk restrictions and inadmissibility results. Statistical Decision Theory and Related Topics III (Vol. 1) 109-141, Academic Press, New York.

Berger, James (1983). The robust Bayesian viewpoint. Robustness in Bayesian Statistics, J. Kadane (ed.). North-Holland, Amsterdam.

Berger, James and DasGupta, Anirban (1986). Estimation of multiple gamma scale-parameters: Bayes estimation subject to uniform domination (To appear).

Bickel, P. J. (1980). Minimax estimation of the mean of a normal distribution subject to doing well at a point. Techical Report, University of California, Berkeley.

Casella, George and Strawderman, W. E. (1981). Estimating a bounded normal mean. Ann. Statist. 9, 870-878.

DasGupta, Anirban (1985). Bayes minimax estimation in multiparameter families when the parameter space is restricted to a bounded convex set. Sankhyā, vol. 47, Series A, Part c, 326-332.

Efron, B. and Morris, C. (1971). Limiting the risk of Bayes and empirical Bayes estimators - Part I: The Bayes case. Jour. Amer. Statist. Assoc. 66, 807-815.

Efron, B. and Morris, C. (1976). Families of minimax estimators of the mean of a multivariate normal distribution. Ann. Statist. 4, 11-21.

Hodges, J. L. and Lehmann, E. L. (1952). The use of previous experience in reaching statistical decisions. Ann. Math. Statist. 23, 392-407.

Hwang, J. T. (1979). Improving upon inadmissible estimators in discrete exponential families. Ph.D. Thesis, Department of Statistics, Purdue University, Mimeograph Series #79-14.

Marazzi, A. (1985). On constrained minimization of the Bayes risk for the linear model. Statistics and Decisions Vol. 3 No. 3/4, 277-296.

Strawderman, W. E. (1973). Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances. Ann. Statist. 1, 1189-1194.

Stein, C. (1955). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Proc. 3rd Berkeley Symp. Math. Statist. Prob. 1 197-206. University of California Press, Berkeley.