The Annals of Statistics
1996, Vol. 24, No. 3, 931–951

CHOICE OF HIERARCHICAL PRIORS: ADMISSIBILITY IN ESTIMATION OF NORMAL MEANS[1]

BY JAMES O. BERGER AND WILLIAM E. STRAWDERMAN

Purdue University and Rutgers University

In hierarchical Bayesian modeling of normal means, it is common to complete the prior specification by choosing a constant prior density for unmodeled hyperparameters (e.g., variances and highest-level means). This common practice often results in an inadequate overall prior, inadequate in the sense that estimators resulting from its use can be inadmissible under quadratic loss. In this paper, hierarchical priors for normal means are categorized in terms of admissibility and inadmissibility of resulting estimators for a quite general scenario. The Jeffreys prior for the hypervariance and a shrinkage prior for the hypermeans are recommended as admissible alternatives. Incidental to this analysis is presentation of the conditions under which the (generally improper) priors result in proper posteriors.

1. Introduction.

1.1. *The problem.* Use of hierarchical Bayesian models in statistical practice is extensive, yet very little is known about the comparative performance of priors for hyperparameters. As a simple example, suppose $X_i \sim \mathcal{N}(\theta_i, \sigma^2)$ and $\theta_i \sim \mathcal{N}(\beta, \sigma_\pi^2)$ for $i = 1, \ldots, p$ with σ^2 known, and estimation of the θ_i is the goal. Standard choices of the hyperparameter prior are $\pi_1(\beta, \sigma_\pi^2) = 1$ and $\pi_2(\beta, \sigma_\pi^2) = 1/(\sigma^2 + \sigma_\pi^2)$, yet little is known about their comparative performance or properties. The reason typically given for ignoring such questions is that the hyperparameter prior is thought to have less effect on the answers than other features of the model [cf. Goel (1983) and Good (1983)]. While this is true, the effect of the hyperparameter prior can nevertheless be substantial, especially when p is small (or, more generally, when the ratio of p to the number of hyperparameters is small), and the almost standardized use of these models suggests that it is time to address these questions.

The study of this issue is not merely of theoretical interest, because it will be seen that standard choices of the hyperparameter prior, such as π_1 or π_2 above, can lead to inadmissible estimators in terms of mean-squared error. Alternatives will be proposed, both "noninformative" and "robust subjective," that are recommended for general use. Conditions are also given under which the overall posterior is proper or improper, a concern that Bayesians often face in practice when using improper priors. (Note that we focus, here, on

---

December 1993; revised January 1996.
[1]Research supported by NSF Grants DMS-89-23071, DMS-90-23172 and DMS-93-03556.
AMS 1991 *subject classifications*. Primary 62F15; secondary 62C15, 62C20, 62J07.
Key words and phrases. Normal hierarchical models, hyperparameters, inadmissibility, mean-squared error, shrinkage estimation.

improper priors because proper priors automatically result in admissible statistical procedures. Also, we fully encourage use of proper priors based on careful subjective elicitation, but have observed that this is rarely done in practice for hyperparameters.)

We will evaluate hyperparameter priors using the criterion of frequentist admissibility. In particular, we consider the case $\mathbf{X} = (X_1, \ldots, X_p)^t \sim \mathcal{N}_p\ (\boldsymbol{\theta}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ being a known positive definite matrix, and we assume that the goal is to estimate $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^t$ by an estimate $\boldsymbol{\delta}(\mathbf{x}) = (\delta_1(\mathbf{x}), \ldots, \delta_p(\mathbf{x}))^t$, under loss

$$L(\boldsymbol{\theta}, \boldsymbol{\delta}) = \| \boldsymbol{\theta} - \boldsymbol{\delta} \|^2 = \sum_{i=1}^{p} (\theta_i - \delta_i)^2.$$

[The results would be identical for any quadratic loss $(\boldsymbol{\theta} - \boldsymbol{\delta})^t \mathbf{Q} (\boldsymbol{\theta} - \boldsymbol{\delta})$, with $\mathbf{Q}$ positive definite.] The frequentist risk function of $\boldsymbol{\delta}(\mathbf{X})$, or mean-squared error, is

$$R(\boldsymbol{\theta}, \boldsymbol{\delta}) = E_{\boldsymbol{\theta}} L(\boldsymbol{\theta}, \boldsymbol{\delta}(\mathbf{X})) = E_{\boldsymbol{\theta}} \| \boldsymbol{\theta} - \boldsymbol{\delta}(\mathbf{X}) \|^2,$$

where $E_{\boldsymbol{\theta}}$ denotes expectation w.r.t. $\mathbf{X}$. Our goal is to ascertain which hierarchical priors result in admissible Bayes estimators, that is, admissible posterior means. [As usual, an estimator $\boldsymbol{\delta}(\mathbf{x})$ is said to be inadmissible if there exists a $\boldsymbol{\delta}^*(\mathbf{x})$ with $R(\boldsymbol{\theta}, \boldsymbol{\delta}^*) \le R(\boldsymbol{\theta}, \boldsymbol{\delta})$, with strict inequality for some $\boldsymbol{\theta}$; if no such $\boldsymbol{\delta}^*$ exists, then $\boldsymbol{\delta}$ is admissible.]

The stated problem is of clear interest to a frequentist because of the inadmissibility of $\boldsymbol{\delta}_0(\mathbf{X}) = \mathbf{X}$ when $p \ge 3$ [the Stein phenomenon; see Stein (1956)] and the fact that the most successful improved shrinkage estimators have been developed as hierarchical Bayes estimators [see Berger (1980, 1985) and Berger and Robert (1990) for discussion]. It is less clear that Bayesians should be interested in these results. It has been frequently argued, however [cf. Berger and Bernardo (1992a)] that comparison and choice of noninformative priors must involve some type of frequentist computation and that consideration of admissibility of resulting estimators is an often enlightening approach. Indeed, it has been suggested that prior distributions which are "on the boundary of admissibility" are particularly attractive noninformative priors [see Berger (1984) and the references therein]. Finding such is one of our goals.

1.2. *The hierarchical model.* We specialize to the most commonly utilized hierarchical normal model in which the prior distribution for $\boldsymbol{\theta}$ is given in two stages:

$$\boldsymbol{\theta} | \boldsymbol{\beta}, \sigma_\pi^2 \sim \mathcal{N}_p(\mathbf{y}\boldsymbol{\beta}, \sigma_\pi^2 \mathbf{I}), \qquad (\boldsymbol{\beta}, \sigma_\pi^2) \sim \pi_1(\sigma_\pi^2)\pi_2(\boldsymbol{\beta}).$$

Here $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^t$ and σ_π^2 are the unknown "hyperparameters," and $\mathbf{y}$ is a $(p \times k)$ matrix of known covariates which has rank k. [See Morris (1983) and Berger (1985) for background and discussion of this model.] The $p \times p$ identity matrix $\mathbf{I}$ could be replaced by any known positive definite matrix without affecting the results. Analogous results could be derived for second stage priors which incorporate dependence between $\boldsymbol{\beta}$ and σ_π^2, but such are rarely utilized in practice.

Quite general forms of $\pi_1(\sigma_\pi^2)$ will be considered. Indeed, we assume only the following condition.

CONDITION 1. (i) For $c > 0$, $\int_0^c \pi_1(\sigma_\pi^2)\, d\sigma_\pi^2 < \infty$.
(ii) For some constant $C > 0$, $\pi_1(\sigma_\pi^2) \sim C/(\sigma_\pi^2)^a$ as $\sigma_\pi^2 \to \infty$.

Typical values of a that are considered are $a = 0$ (constant prior), $a = 1/2$ (Carl Morris, personal communication), $a = 1$ [see (1.3)] and $a = 3/2$ [see Berger and Deely (1988)].

Choice of $\pi_2(\boldsymbol{\beta})$ is more involved. The following three possibilities will be analyzed.

CASE 1. $\pi_2(\boldsymbol{\beta}) = 1$. This is, by far, the most common choice made in practice.

CASE 2. $\pi_2(\boldsymbol{\beta})$ is $\mathscr{N}_k(\boldsymbol{\beta}^0, A)$, where $\boldsymbol{\beta}^0$ and $\mathbf{A}$ are subjectively specified. [See Berger (1985) for discussion.]

CASE 3. $\pi_2(\boldsymbol{\beta})$ is itself given in two stages:

$$\boldsymbol{\beta}|\lambda \sim \mathscr{N}_k(\boldsymbol{\beta}^0, \lambda\mathbf{A}), \qquad \lambda \sim \pi_3(\lambda), \tag{1.1}$$

where $\boldsymbol{\beta}^0$ and $\mathbf{A}$ (positive definite) are again specified and π_3 satisfies the following condition.

CONDITION 2. (i) For $c > 0$, $\int_0^c \pi_3(\lambda)\, d\lambda < \infty$.
(ii) For some constant $C > 0$, $\pi_3(\lambda) \sim C\lambda^{-b}$ as $\lambda \to \infty$.

The Case 3 prior is of interest for a number of reasons. First, we will see that the commonly employed $\pi_2(\boldsymbol{\beta}) = 1$ frequently results in inadmissible estimators (especially if $k \geq 3$) and hence that alternatives are required. Subjective choices, as in Case 2, are fine, but automatic hyperpriors are also needed. Such are available in Case 3. For instance, if $\boldsymbol{\beta}^0 = \mathbf{0}, \mathbf{A} = \mathbf{I}$ and $\pi_3(\lambda) = \lambda^{-b}$, where $(1 - k/2) < b < 1$, then

$$\pi_2(\boldsymbol{\beta}) \propto \| \boldsymbol{\beta} \|^{-(2b+k-2)} . \tag{1.2}$$

This is recognizable as a standard "shrinkage" prior for $\boldsymbol{\beta}$; indeed, for $b = 0$ this was suggested by Baranchik (1964), although only for the original $\boldsymbol{\theta}$. Priors for $\boldsymbol{\beta}$ must typically have at least this amount of shrinkage in the tails to achieve admissibility.

A second motivation for Case 3 is that it can allow use of subjective information about $\boldsymbol{\beta}$ (through choice of $\boldsymbol{\beta}^0$ and $\mathbf{A}$), yet utilizes this information in a very robust fashion. For instance, if $\pi_3(\lambda)$ is chosen to be inverse gamma

$((b-1)/2, 2/(b-1))$, then $\pi_2(\boldsymbol{\beta})$ is $\mathscr{T}_k(b-1, \boldsymbol{\beta}^0, A)$. The considerable robustness of $\mathscr{T}$ priors here was established in Angers (1987); see also Berger (1985) and Berger and Robert (1990).

A final motivation for consideration of the priors in Case 3 is computational. Such priors are very amenable to computation via Gibbs sampling. For a practical illustration in which $\pi_2(\boldsymbol{\beta}) = \| \boldsymbol{\beta} \|^{-(k-2)}$ [i.e., $\pi_3(\lambda) = 1$] is used, see Andrews, Berger and Smith (1993).

1.3. *Preview.* In Section 2 we present needed technical preliminaries, which include the conditions under which the posterior is proper or improper. Section 3 considers the simpler Cases 1 and 2 from Section 1.2; essentially complete characterizations of admissibility and inadmissibility are possible in these cases. Section 4 considers the much more difficult Case 3 and presents a variety of partial results that cover the estimators of major interest.

It is worthwhile to briefly summarize the methodological recommendations that arise from this work. First, the common noninformative prior $\pi_1(\sigma_\pi^2) = 1$ is typically inadmissible, while priors which behave like C/σ_π^2 for large σ_π^2 are typically admissible. In the case $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$, the Jeffreys (and reference) prior [see Box and Tiao (1973) and Berger and Bernardo (1992b)] is

$$\pi_1(\sigma_\pi^2) = 1/(\sigma^2 + \sigma_\pi^2), \tag{1.3}$$

which is of this admissible form. Hence use of (1.3) or its nonexchangeable analogues [see Berger and Deely (1988) and Ye (1993)] is recommended.

When $k = 1$ or 2, use of $\pi_2(\boldsymbol{\beta}) = 1$ is fine, but it is typically inadmissible if $k \geq 3$. We then recommend either using the default $\pi_2(\boldsymbol{\beta}) = \| \boldsymbol{\beta} \|^{-(k-2)}$ [corresponding to $\pi_3(\lambda) = 1$], which is typically admissible [see Andrews, Berger and Smith (1993) for a practical example], or using a robust subjective Bayes choice as discussed in the previous section. Note that expressions for the resulting estimators are given in Section 2 and require at most two-dimensional integration.

As a final methodological recommendation, we suggest following the "spirit" of these results in more complicated normal hierarchical or random effects models; avoid using constant priors for variances or covariance matrices, or for groups of mean parameters of dimension greater than 2. Instead try adaptions of the above recommended priors. Rigorous verification of these recommendations would be very difficult, but the results in this paper, together with our practical experience, suggest that they are very reasonable.

1.4. *Background.* Many references to normal hierarchical Bayesian analysis can be found in Berger (1985), Berger and Robert (1990) and Ghosh (1992). Results on admissibility or inadmissibility for particular cases include Strawderman (1971, 1973), Berger (1976), Hill (1977) and Zheng (1982).

The most comprehensive technique for proving admissibility and inadmissibility in these models is that of Brown [(1971), Section 6], based on properties of the marginal density

$$m(\mathbf{x}) = \int \frac{\exp\{-(\mathbf{x}-\boldsymbol{\theta})^t \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\theta})/2\}}{(2\pi)^{p/2}(\det \boldsymbol{\Sigma})^{1/2}} \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta} \tag{1.4}$$

corresponding to the (improper) prior $\pi(\boldsymbol{\theta})$. We will utilize only the following two results of Brown's. For these results, define

$$\overline{m}(r) = \int m(\mathbf{x})\, d\varphi(\mathbf{x}), \tag{1.5}$$

$$\underline{m}(r) = \int (1/m(\mathbf{x}))\, d\varphi(\mathbf{x}), \tag{1.6}$$

where $\varphi(\cdot)$ is the uniform probability measure on the surface of the sphere of radius r. Also, suppose $\boldsymbol{\delta}(\mathbf{x})$ is the posterior mean corresponding to the prior $\pi(\boldsymbol{\theta})$.

RESULT 1 (Admissibility). If $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ is uniformly bounded and

$$\int_c^\infty (r^{p-1}\overline{m}(r))^{-1}\, dr = \infty \tag{1.7}$$

for some $c > 0$, then $\boldsymbol{\delta}$ is admissible.

RESULT 2 (Inadmissibility). If

$$\int_c^\infty r^{1-p}\underline{m}(r)\, dr < \infty \tag{1.8}$$

for some $c > 0$, then $\boldsymbol{\delta}$ is inadmissible.

In computing expressions such as (1.5) and (1.6), the following well-known result will be very useful.

RESULT 3 [cf. Kelker (1970)]. If $\mathbf{X}$ has a uniform distribution on the surface of the sphere of radius r and $\hat{\boldsymbol{\beta}} = (\mathbf{y}^t\mathbf{y})^{-1}\mathbf{y}^t\mathbf{X}$, then $W = \| \mathbf{X} - \mathbf{y}\hat{\boldsymbol{\beta}} \|^2 / \| \mathbf{X} \|^2$ has a beta $((p-k)/2, k/2)$ distribution.

2. Preliminaries.

2.1. *Basic expressions.* We first record expressions for $m(\mathbf{x})$ and $\boldsymbol{\delta}(\mathbf{x})$. These are simple modifications of the expressions in Section 4.6.3 of Berger (1985).

For the Case 3 scenario of Section 1.2,

$$\begin{aligned} m(\mathbf{x}) = C \int_0^\infty \int_0^\infty &\left[\frac{\exp\{-(1/2)[\| \mathbf{x} - \mathbf{y}\tilde{\boldsymbol{\beta}} \|_*^2 + \| \tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0 \|_{**}^2]\}}{[\det \mathbf{W}]^{-1/2}[\det(\lambda \mathbf{y}^t\mathbf{W}\mathbf{y}\mathbf{A} + \mathbf{I})]^{1/2}} \right] \\ &\times \pi_1(\sigma_\pi^2)\pi_3(\lambda)\, d\sigma_\pi^2\, d\lambda, \end{aligned} \tag{2.1}$$

where $\mathbf{W} = (\boldsymbol{\Sigma} + \sigma_\pi^2\mathbf{I})^{-1}$, $\tilde{\boldsymbol{\beta}} = (\mathbf{y}^t\mathbf{W}\mathbf{y})^{-1}\mathbf{y}^t\mathbf{W}\mathbf{x}$,

$$\begin{aligned} \| \mathbf{x} - \mathbf{y}\tilde{\boldsymbol{\beta}} \|_*^2 &= (\mathbf{x} - \mathbf{y}\tilde{\boldsymbol{\beta}})^t\mathbf{W}(\mathbf{x} - \mathbf{y}\tilde{\boldsymbol{\beta}}), \\ \| \tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0 \|_{**}^2 &= (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0)^t([\mathbf{y}^t\mathbf{W}\mathbf{y}]^{-1} + \lambda\mathbf{A})^{-1}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) \end{aligned} \tag{2.2}$$

and

$$\boldsymbol{\delta}(\mathbf{x}) = \text{posterior mean} = E^{\pi(\sigma_\pi^2, \lambda|\mathbf{x})}[\boldsymbol{\mu}^*(\mathbf{x}, \lambda, \sigma_\pi^2)], \tag{2.3}$$

where $\pi(\sigma_\pi^2, \lambda|\mathbf{x})$ is proportional to the integrand in (2.1) and

$$\boldsymbol{\mu}^*(\mathbf{x}, \lambda, \sigma_\pi^2) = \mathbf{x} - \boldsymbol{\Sigma}\mathbf{W}(\mathbf{x} - \mathbf{y}\tilde{\boldsymbol{\beta}}) - \boldsymbol{\Sigma}\mathbf{W}\mathbf{y}(\lambda\mathbf{A}\mathbf{y}^t\mathbf{W}\mathbf{y} + \mathbf{I})^{-1}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^0).$$

Above, and henceforth, we use C and c to denote generic constants. An analogous expression for the posterior covariance matrix can be found in Berger (1985).

For Case 2 of Section 1.2, these expressions remain valid with $\lambda = 1$ and the integral over λ removed. The formulas for Case 1 are

$$m(\mathbf{x}) = C\int_0^\infty \frac{\exp\{-(1/2)\parallel \mathbf{x} - \mathbf{y}\tilde{\boldsymbol{\beta}} \parallel_*^2\}}{[\det \mathbf{W}]^{-1/2}[\det(\mathbf{y}^t\mathbf{W}\mathbf{y})]^{1/2}}\pi_1(\sigma_\pi^2)\,d\sigma_\pi^2, \tag{2.4}$$

$$\boldsymbol{\delta}(\mathbf{x}) = E^{\pi(\sigma_\pi^2|\mathbf{x})}[\mathbf{x} - \boldsymbol{\Sigma}\mathbf{W}(\mathbf{x} - \mathbf{y}\tilde{\boldsymbol{\beta}})], \tag{2.5}$$

where $\pi(\sigma_\pi^2|\mathbf{x})$ is proportional to the integrand of (2.4).

2.2. *Boundedness of* $m(\mathbf{x})$ *and* $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$. We need to clarify when the integrals in (2.1) and (2.4) exist and when $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ is uniformly bounded. Note that these imply that the posterior is proper and that the posterior mean exists—important practical issues in their own right. Special cases of these results were given in Berger and Robert (1990).

LEMMA 1. *In the Case* 3 *scenario*, $m(\mathbf{x})$ *and* $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ *are uniformly bounded if*

$$a > 1 - \frac{p}{2}, \qquad b > 1 - \frac{k}{2} \quad \text{and} \quad \frac{p}{2} + a + b - 2 > 0. \tag{2.6}$$

PROOF. Note that, since $\mathbf{y}$ has rank k and $\boldsymbol{\Sigma}$ and $\mathbf{A}$ are positive definite,

$$C(1+\sigma_\pi^2)^{p/2} \le [\det \mathbf{W}]^{-1/2} \le \mathbf{C}'(1+\sigma_\pi^2)^{p/2}, \tag{2.7}$$

$$C\left(1 + \frac{\lambda}{(1+\sigma_\pi^2)}\right)^{k/2} \le [\det(\lambda\mathbf{y}^t\mathbf{W}\mathbf{y}\mathbf{A} + \mathbf{I})]^{1/2} \le \mathbf{C}'\left(1 + \frac{\lambda}{(1+\sigma_\pi^2)}\right)^{k/2}. \tag{2.8}$$

Hence

$$m(\mathbf{x}) \le C\int_0^\infty\int_0^\infty \frac{1}{(1+\sigma_\pi^2)^{p/2}(1+\lambda/(1+\sigma_\pi^2))^{k/2}}\pi_1(\sigma_\pi^2)\pi_3(\lambda)\,d\sigma_\pi^2\,d\lambda.$$

Break this integral up into integrals over $(0, c)\times(0, c)$, $(0, c)\times(c, \infty)$, $(c, \infty)\times(0, c)$ and $(c, \infty)\times(c, \infty)$, where c is chosen so that $\pi_1(\sigma_\pi^2) \le C/\sigma_\pi^{2a}$ for $\sigma_\pi^2 > c$ and $\pi_2(\lambda) \le C/\lambda^b$ for $\lambda > c$ (see Conditions 1 and 2); call these four integrals I_1, I_2, I_3 and I_4, respectively. Note that

$$I_4 \le C\int_c^\infty\int_c^\infty \frac{1}{(1+\sigma_\pi^2)^{a+p/2}(1+\lambda/(1+\sigma_\pi^2))^{k/2}\lambda^b}\,d\sigma_\pi^2\,d\lambda,$$

where recall that we use C as a generic positive constant. Making the change of variables $\mu = (1+\sigma_\pi^2)$ and $v = \lambda/(1+\sigma_\pi^2)$ yields

$$\begin{aligned}
I_4 &\le C\int_0^\infty \int_{\max\{c',c/v\}}^\infty \frac{1}{\mu^{(a+b-1+p/2)}(1+v)^{k/2}v^b}\,d\mu\,dv \\
&= C\int_0^\infty \left[\max\{c', \frac{c}{v}\}\right]^{(2-a-b-p/2)}(1+v)^{-k/2}v^{-b}\,dv \\
&= C\int_0^c (1+v)^{-k/2}v^{(a-2+p/2)}\,dv + C\int_c^\infty (1+v)^{-k/2}v^{-b}\,dv,
\end{aligned}$$

which are finite under the given conditions.

Bounding I_1, I_2 and I_3 is easy, using also Condition 1(i) and Condition 2(i) in Section 1.2. We thus have that $m(\mathbf{x})$ is uniformly bounded under the given conditions.

To show that $\| \boldsymbol{\delta}(\mathbf{x}) - \mathbf{x} \|$ is uniformly bounded, it suffices to show that

$$\boldsymbol{\gamma}(\mathbf{z}) = \frac{\int \boldsymbol{\Sigma}\mathbf{W}\mathbf{z}(\det\mathbf{W})^{-1/2}\exp\{-(1/2)\mathbf{z}^t\mathbf{W}\mathbf{z}\}\pi_1(\sigma_\pi^2)\,d\sigma_\pi^2}{\int(\det\mathbf{W})^{-1/2}\exp\{-(1/2)\mathbf{z}^t\mathbf{W}\mathbf{z}\}\pi_1(\sigma_\pi^2)\,d\sigma_\pi^2} \tag{2.9}$$

is uniformly bounded as a function of $\mathbf{z}$; this suffices because it is straightforward to show, as in Section 2.3.1 of Berger and Robert (1990), that

$$\mathbf{x} - \boldsymbol{\delta}(\mathbf{x}) = E^{\pi_2(\boldsymbol{\beta}|\mathbf{x})}[\boldsymbol{\gamma}(\mathbf{x} - \mathbf{y}\boldsymbol{\beta})],$$

and if the integrand is uniformly bounded for all $(\mathbf{x} - \mathbf{y}\boldsymbol{\beta})$, then so is the expectation.

Define c_2 and c_1 as the maximum and minimum characteristic roots of $\boldsymbol{\Sigma}$, respectively, and [using Condition 1(ii)] choose c_1 so that, for $\sigma_\pi^2 > c_1$, $C/\sigma_\pi^{2a} \le \pi(\sigma_\pi^2) \le C'/\sigma_\pi^{2a}$. Finally, let $c_4 = \max\{1, 2(c_1+c_3)\}$.

For $\| \mathbf{z} \| \le c_4$, it is trivial that $\boldsymbol{\gamma}(\mathbf{x})$ is uniformly bounded (since $\mathbf{W} \le \boldsymbol{\Sigma}^{-1}$). Hence we need only consider the case $\| \mathbf{z} \| > c_4$.

We first find an upper bound on the numerator of (2.9). Clearly,

$$\|\text{numerator}\| \le C\int_0^\infty \frac{\| \mathbf{z} \|}{(c_3+\sigma_\pi^2)}\frac{\exp\{-(1/2)\| \mathbf{z} \|^2/(c_2+\sigma_\pi^2)\}}{(c_3+\sigma_\pi^2)^{p/2}}\pi_1(\sigma_\pi^2)\,d\sigma_\pi^2.$$

Break this up into integrals over $(0, c_1)$ and (c_1, ∞), to be called I_1 and I_2, respectively. Using Condition 1(i), it is clear that

$$\begin{aligned}
I_1 &\le C\int_0^{c_1} \frac{\|\mathbf{z}\|}{c_3}\frac{\exp\{-(1/2)\|\mathbf{z}\|^2/(c_2+c_1)\}}{c_3^{p/2}}\pi_1(\sigma_\pi^2)\,d\sigma_\pi^2 \\
&\le C\|z\|\exp\left\{\frac{-\|\mathbf{z}\|^2}{c_2+c_1}\right\}.
\end{aligned} \tag{2.10}$$

Next, again recalling that we use C as a generic constant and making the change of variables $v = \|\mathbf{z}\|^2/2(c_2 + \sigma_\pi^2)$,

$$
\begin{aligned}
I_2 &\le C \int_{c_1}^{\infty} \|\mathbf{z}\|(c_3 + \sigma_\pi^2)^{-(a+1+p/2)} \exp\{-\tfrac{1}{2}\|\mathbf{z}\|^2/(c_2 + \sigma_\pi^2)\}\, d\sigma_\pi^2 \\
&\le C \int_{c_1}^{\infty} \|\mathbf{z}\|(c_2 + \sigma_\pi^2)^{-(a+1+p/2)} \exp\{-\tfrac{1}{2}\|\mathbf{z}\|^2/(c_2 + \sigma_\pi^2)\}\, d\sigma_\pi^2 \\
&\le C\|\mathbf{z}\|^{-(2a+p-1)} \int_0^{\|\mathbf{z}\|^2/2(c_1+c_2)} v^{(a+p/2-1)} e^{-v}\, dv \\
&\le C\|\mathbf{z}\|^{-(2a+p-1)},
\end{aligned}
\tag{2.11}
$$

the last step using $a > 1 - p/2$. Combining (2.10) and (2.11) yields

$$
\|\text{numerator}\| \le C\|\mathbf{z}\| \exp\{-\|\mathbf{z}\|^2/(c_2 + c_1)\} + C\|\mathbf{z}\|^{-(2a+p-1)}. \tag{2.12}
$$

Next, we find a lower bound on the denominator of (2.9). Clearly, making the change of variables $v = \|\mathbf{z}\|^2/2(c_3 + \sigma_\pi^2)$,

$$
\begin{aligned}
\text{denominator} &> C \int_{c_1}^{\infty} \frac{\exp\{-(1/2)\|\mathbf{z}\|^2/(c_3 + \sigma_\pi^2)\}}{(c_2 + \sigma_\pi^2)^{(a+p/2)}}\, d\sigma_\pi^2 \\
&= C \int_0^{\|\mathbf{z}\|^2/2(c_1+c_3)} \left(c_2 - c_3 + \frac{\|\mathbf{z}\|^2}{2v}\right)^{-(a+p/2)} e^{-v} \frac{\|\mathbf{z}\|^2}{2v^2}\, dv \\
&> C\|\mathbf{z}\|^{-(2a+p-2)} \int_0^1 v^{(a-2+p/2)} e^{-v}\, dv,
\end{aligned}
$$

since $\|\mathbf{z}\|^2 > 2(c_1+c_3)$. Since $a > 1-p/2$, we can conclude that the denominator is greater than

$$
C\|\mathbf{z}\|^{-(2a+p-2)}. \tag{2.13}
$$

Combining (2.12) and (2.13) yields

$$
\|\boldsymbol{\gamma}(\mathbf{z})\| \le C\|\mathbf{z}\|^{(2a+p-1)} \exp\{-\|\mathbf{z}\|^2/(c_2 + c_1)\} + C\|\mathbf{z}\|^{-1},
$$

which is clearly uniformly bounded (recall that $\|\mathbf{z}\| > c_4$). This completes the proof. □

LEMMA 2. *In the Case 2 scenario, $m(\mathbf{x})$ and $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ are uniformly bounded if $a > 1 - p/2$.*

PROOF. Equations (2.7) and (2.8) remain valid with $\lambda = 1$, so that

$$
\begin{aligned}
m(\mathbf{x}) &\le C \int_0^{\infty} \frac{1}{(1 + \sigma_\pi^2)^{p/2}(1 + 1/(1 + \sigma_\pi^2))^{k/2}} \pi_1(\sigma_\pi^2)\, d\sigma_\pi^2 \\
&\le C \int_0^{c} \pi_1(\sigma_\pi^2)\, d\sigma_\pi^2 + C \int_c^{\infty} (\sigma_\pi^2)^{-(a+p/2)}\, d\sigma_\pi^2,
\end{aligned}
$$

which is bounded since $a > 1 - p/2$ [and using Condition 1(i)]. The proof that $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ is uniformly bounded is identical to that in Lemma 1. □

LEMMA 3. *In the Case 1 scenario, $m(\mathbf{x})$ and $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ are uniformly bounded if $a > 1 - (p-k)/2$.*

PROOF. Applying (2.7) and $C(1+\sigma_\pi^2)^{-k/2} \le [\det(\mathbf{y}^t\mathbf{W}\mathbf{y})]^{1/2} \le C'(1+\sigma_\pi^2)^{-k/2}$ to (2.4) yields

$$m(\mathbf{x}) \le C\int_0^\infty \frac{1}{(1+\sigma_\pi^2)^{p/2}(1+\sigma_\pi^2)^{-k/2}}\pi_1(\sigma_\pi^2)\,d\sigma_\pi^2$$
$$\le C\int_0^c \pi_1(\sigma_\pi^2)\,d\sigma_\pi^2 + C\int_c^\infty (\sigma_\pi^2)^{-(a+(p-k)/2)}\,d\sigma_\pi^2,$$

which is finite. The proof that $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ is uniformly bounded is identical to that in Lemma 1. □

3. Results when $\pi_2(\boldsymbol{\beta}) = 1$.

3.1. *Summary.* For this most common choice of $\pi_2(\boldsymbol{\beta})$, admissibility of $\boldsymbol{\delta}(\mathbf{x})$ depends only on p, k and a. Table 1 summarizes the conclusions. For $p = 2$, note that if $a \le \frac{1}{2}$ when $k = 1$, or if $k \ge 2$, then $m(\mathbf{x})$ [and hence $\boldsymbol{\delta}(\mathbf{x})$] is not defined. Hence Table 1 provides a complete resolution of admissibility and inadmissibility when $\pi_2(\boldsymbol{\beta}) = 1$, except when $k \ge 3$ and $a \ge k/2$. We think that $\boldsymbol{\delta}(\mathbf{x})$ is actually inadmissible even in this case, since, in the exchangeable situation or for fixed σ_π^2, a direct shrinkage argument shows $\boldsymbol{\delta}(\mathbf{x})$ to be inadmissible. Since $a \ge k/2$ does not correspond to any previously proposed estimators, we did not attempt to extend the argument to this case.

Our recommendation in Section 1.3 was to use the prior (1.3), which is of the general form in Condition 1 with $a = 1$. Based on Table 1, when $k = 1$ and $p \ge 3$ one could argue instead for the choice $a = 1/2$, this being on the "boundary of admissibility." Indeed, Carl Morris (personal communication) has suggested the choice $\pi_1(\sigma_\pi^2) = 1/\sigma_\pi$, which does satisfy Condition 1 with $a = 1/2$. We find this proposal to be quite reasonable; our recommendation to always use $a = 1$ was mainly based on a desire for simplicity.

TABLE 1
Admissibility and inadmissibility of the Bayes estimator when $\pi_2(\boldsymbol{\beta}) = 1$

	k		
p	1	2	≥ 3
2	$a > \frac{1}{2} \Rightarrow$ admissible	δ not defined	δ not defined
	$a \le \frac{1}{2} \Rightarrow \delta$ not defined		
≥ 3	$a \ge \frac{1}{2} \Rightarrow$ admissible	$a \ge 1 \Rightarrow$ admissible	$a < k/2 \Rightarrow$ inadmissible
	$a < \frac{1}{2} \Rightarrow$ inadmissible	$a < 1 \Rightarrow$ inadmissible	$a \ge k/2$ not analyzed

3.2. *Inadmissibility and admissibility of* $\boldsymbol{\delta}(\mathbf{x})$.

THEOREM 1. *Suppose* $\pi_2(\boldsymbol{\beta}) = 1$ *and* $\pi_1(\sigma_\pi^2)$ *satisfies Condition* 1 *with* $a > 1 - (p-k)/2$. *Then* $\boldsymbol{\delta}(\mathbf{x})$ *is admissible or inadmissible as indicated in Table* 1.

PROOF. By Lemma 3, $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ is uniformly bounded. Hence to establish admissibility or inadmissibility of $\boldsymbol{\delta}(\mathbf{x})$ we need only verify (1.7) or (1.8), respectively.

Note first that

$$\frac{C\|\mathbf{x} - \mathbf{y}^t\boldsymbol{\beta}\|^2}{(1+\sigma_\pi^2)} \le (\mathbf{x} - \mathbf{y}^t\boldsymbol{\beta})^t(\boldsymbol{\Sigma} + \sigma_\pi^2\mathbf{I})^{-1}(\mathbf{x} - \mathbf{y}^t\boldsymbol{\beta}) \le \frac{C'\|\mathbf{x} - \mathbf{y}^t\boldsymbol{\beta}\|^2}{(1+\sigma_\pi^2)}, \tag{3.1}$$

which, together with (2.7) and (2.8), can be directly used to verify that $m(\mathbf{x})$ can be bounded above and below (for appropriate choices of C and C') by

$$m^*(\mathbf{x}) = C\int_0^\infty (1+\sigma_\pi^2)^{-(p-k)/2} \exp\left\{\frac{-C'\|\mathbf{x} - \mathbf{y}\hat{\boldsymbol{\beta}}\|^2}{(1+\sigma_\pi^2)}\right\}\pi_1(\sigma_\pi^2)\, d\sigma_\pi^2,$$

where $\hat{\boldsymbol{\beta}} = (\mathbf{y}^t\mathbf{y})^{-1}\mathbf{y}^t\mathbf{x}$. Defining

$$r^2 = \|\mathbf{x}\|^2, \qquad w = \|\mathbf{x} - \mathbf{y}\hat{\boldsymbol{\beta}}\|^2/\|\mathbf{x}\|^2, \tag{3.2}$$

this can be rewritten as

$$\begin{aligned} m^*(r, w) &= C\int_0^\infty (1+\sigma_\pi^2)^{-(p-k)/2} \exp\{-C'r^2w/(1+\sigma_\pi^2)\}\pi_1(\sigma_\pi^2)\, d\sigma_\pi^2 \\ &\equiv m_1^*(r, w) + m_2^*(r, w), \end{aligned} \tag{3.3}$$

where m_1^* is the integral on $(0, c)$ and m_2^* is the integral on (c, ∞).

Using Condition 1(i), it is clear that m_1^* can be bounded above and below by

$$m_1^{**} = C\exp\{-C'r^2w\}. \tag{3.4}$$

Using Condition 1(ii), one can show that m_2^* can be bounded above and below by

$$\begin{aligned} m_2^{**} &= C\int_c^\infty (1+\sigma_\pi^2)^{-(a+(p-k)/2)} \exp\{-C'r^2w/(1+\sigma_\pi^2)\}\, d\sigma_\pi^2 \\ &= C\int_0^{c'r^2w} (r^2w)^{-(a-1+(p-k)/2)} u^{(a-2+(p-k)/2)} e^{-u}\, du, \end{aligned} \tag{3.5}$$

making the change of variables $u = C'r^2w/(1+\sigma_\pi^2)$.

Proof of inadmissibility. From (1.6), Result 3 and the above

$$m(\mathbf{x}) \geq m^*(r, w) \geq m_1^{**}(r, w) + m_2^{**}(r, w).$$

It is clear that

$$\underline{m}(r) \leq \int_0^1 \frac{g(w)}{m^*(r, w)}\, dw \leq \int_0^{c/r^2} \frac{g(w)}{m_1^{**}(r, w)}\, dw + \int_{c/r^2}^1 \frac{g(w)}{m_2^{**}(r, w)}\, dw, \tag{3.6}$$

where

$$g(w) = Cw^{((p-k)/2-1)}(1-w)^{(k/2-1)}. \tag{3.7}$$

Now, using (3.4),

$$\int_0^{c/r^2} \frac{g(w)}{m_1^{**}(r, w)}\, dw \leq C \int_0^{c/r^2} \exp(C' r^2 w) g(w)\, dw \leq C. \tag{3.8}$$

Next, using (3.5),

$$\int_{c/r^2}^1 \frac{g(w)}{m_2^{**}(r, w)}\, dw \leq C \int_{c/r^2}^1 \frac{r^{(2a-2+p-k)} w^{(a+p-k-2)}(1-w)^{(k/2-1)}}{\int_0^{c' r^2 w} e^{-u} u^{(a-2+(p-k)/2)}\, du}\, dw.$$

However, $c' r^2 w \geq c'c$ over the domain of integration of w, so the integral in the denominator is bounded below by a constant. Hence,

$$\int_{c/r^2}^1 \frac{g(w)}{m_2^{**}(r, w)}\, dw \leq Cr^{(2a-2+p-k)} \int_{c/r^2}^1 w^{(a+p-k-2)}(1-w)^{(k/2-1)}\, dw \leq Cr^{(2a-2+p-k)}, \tag{3.9}$$

since $a + p - k - 2 > (1 - (p-k)/2) + p - k - 2 = (p-k)/2 - 1 > -1$.

Combining (3.6), (3.8) and (3.9), we obtain

$$\underline{m}(r) \leq C + Cr^{(2a-2+p-k)}.$$

Hence, the integral in (1.8) becomes

$$\int_c^\infty r^{1-p} \underline{m}(r)\, dr \leq C \int_c^\infty r^{1-p}\, dr + C \int_c^\infty r^{2a-1-k}\, dr,$$

which is finite if $p \geq 3$ and $a < k/2$. This completes the inadmissibility part of the proof.

Proof of admissibility. From (1.5), Result 3 and the above $m(\mathbf{x}) \leq m^*(r, w) \leq m_1^{**}(r, w) + m_2^{**}(r, w)$, it is clear that

$$\overline{m}(r) \leq \int_0^1 m_1^{**}(r, w) g(w)\, dw + \int_0^1 m_2^{**}(r, w) g(w)\, dw. \tag{3.10}$$

Using (3.4) and (3.7), the first integral can be bounded by

$$\begin{aligned}\int_0^1 m_1^{**}(r, w) g(w)\, dw &= C \int_0^1 \exp(-C' r^2 w) w^{((p-k)/2-1)} (1-w)^{(k/2-1)}\, dw \\ &\leq C \int_0^c \exp(-C' r^2 w) w^{((p-k)/2-1)}\, dw \\ &\quad + C \exp(-C' r^2 c) \int_c^1 g(w)\, dw \\ &\leq C r^{-(p-k)} + C \exp(-C' r^2).\end{aligned} \tag{3.11}$$

Using (3.5), the second integral in (3.10) is

$$\begin{aligned}&\int_0^1 m_2^{**}(r, w) g(w)\, dw \\ &\quad = \frac{C}{r^{(p-k+2a-2)}} \int_0^1 \frac{(1-w)^{(k/2-1)}}{w^a} \int_0^{c' r^2 w} e^{-u} u^{(a-2+(p-k)/2)}\, du\, dw.\end{aligned}$$

Break the integral over w into integrals I_1 and I_2 over $(0, c/r^2)$ and $(c/r^2, 1)$, respectively. Dropping e^{-u} and then $(1-w)^{(k/2-1)}$ in I_1 results in the bound

$$\begin{aligned}I_1 &\leq \frac{C}{r^{(p-k+2a-2)}} \int_0^{c/r^2} \frac{(1-w)^{(k/2-1)}}{w^a} (r^2 w)^{(a-1+(p-k)/2)}\, dw \\ &\leq C \int_0^{c/r^2} w^{((p-k)/2-1)}\, dw \\ &= C r^{-(p-k)}.\end{aligned} \tag{3.12}$$

Bounding the integral over u in I_2 by $\Gamma(a - 1 + (p-k)/2)$, we obtain

$$\begin{aligned}I_2 &\leq \frac{C}{r^{(p-k+2a-2)}} \int_{c/r^2}^1 w^{-a} (1-w)^{(k/2-1)}\, dw \\ &\leq \frac{C}{r^{(p-k+2a-2)}} \left[\int_{c/r^2}^{c'} w^{-a}\, dw + \int_{c'}^1 (1-w)^{(k/2-1)}\, dw \right] \\ &= \begin{cases} C r^{-(p-k+2a-2)}, & \text{if } a < 1, \\ C (\log r) r^{-(p-k)}, & \text{if } a = 1, \\ C r^{-(p-k)}, & \text{if } a > 1. \end{cases}\end{aligned} \tag{3.13}$$

It is clear that (3.11) and (3.12) are of equal or smaller order than (3.13), so that we can conclude that $\overline{m}(r)$ has an upper bound of order (3.13).

Finally, we verify (1.7) for each range of a and k. If $a < 1$, then

$$\int_c^\infty (r^{p-1} \overline{m}(r))^{-1}\, dr \geq \int_c^\infty r^{(2a-1-k)}\, dr = \infty$$

if $a \geq k/2$. This clearly applies only when $k = 1$.

If $a \geq 1$, then

$$\int_c^\infty (r^{p-1}\overline{m}(r))^{-1}\, dr \geq \int_c^\infty (r^{k-1} \log r)^{-1}\, dr = \infty$$

if $k \leq 2$. This completes the verification of the admissibility portions of Table 1. □

4. Results for hierarchical $\pi_2(\boldsymbol{\beta})$.

4.1. *Summary.* The following theorems summarize the results we establish in this section concerning admissibility and inadmissibility of the Bayes rule $\boldsymbol{\delta}(\mathbf{x})$ for the priors defined in Case 2 and Case 3 of Section 1.2.

THEOREM 2. *Suppose* $\pi_2(\boldsymbol{\beta})$ *is* $\mathscr{N}_k(\boldsymbol{\beta}^0, \mathbf{A})$, *with* $\mathbf{A}$ *positive definite, and* $\pi_1(\sigma_\pi^2)$ *satisfies Condition* 1 *with* $a > 1 - p/2$. *Then* $\boldsymbol{\delta}(\mathbf{x})$ *is inadmissible if* $a < 0$ *and admissible if* $a \geq 0$.

THEOREM 3. *Suppose* $\pi_2(\boldsymbol{\beta})$ *has the hierarchical structure defined in Case* 3 *of Section* 1.2, *with* $\pi_1(\sigma_\pi^2)$ *and* $\pi_3(\lambda)$ *satisfying Conditions* 1 *and* 2 *with* $a > 1 - p/2$, $b > 1 - k/2$, *and* $a + b > 2 - p/2$. *Then* $\boldsymbol{\delta}(\mathbf{x})$ *is inadmissible if* $a < 0$ *or* $a + b < 1$ *and is admissible if* a *and* b *are as in Table* 2.

Note that Theorem 3 leaves certain situations unsettled, primarily the case $\{0 \leq a < 1,\ a + b \geq 1\}$. A finer argument would be needed to resolve this situation. The theorem does cover the following cases of major interest:

1. If $k \geq 3$ and $b = 0$ [satisfied by the recommended "shrinkage prior" $\pi_2(\boldsymbol{\beta}) = ||\boldsymbol{\beta}||^{-(k-2)}$], then $\boldsymbol{\delta}(\mathbf{x})$ is admissible if $a \geq 1$ and inadmissible if $a < 1$ (since then $a + b < 1$).
2. If $a = 1$ [corresponding to the recommended priors for $\pi_1(\sigma_\pi^2)$], then $\boldsymbol{\delta}(\mathbf{x})$ is always admissible when $p \geq 3$ and $b \geq 0$ [and when $m(\mathbf{x})$ is defined, i.e., $b > 1 - k/2$]. The requirement $a > 1$ for $k = 1$ and $p = 2$ is a bit annoying (and probably not necessary), but for $k = 1$ and $p = 2$ the choice $\pi_2(\boldsymbol{\beta}) = 1$ will be made most of the time.

TABLE 2
Sufficient conditions for admissibility of $\boldsymbol{\delta}(\mathbf{x})$

	k		
p	**1**	**2**	**≥ 3**
2	$b \geq \frac{1}{2}$ and $a > 1$	—	—
≥ 3	$b > \frac{1}{2}$ and $a \geq 1$	$b > 0$ and $a \geq 1$	$b \geq 0$ and $a \geq 1$

4.2. *Proofs of Theorems* 2 *and* 3.

PROOF OF THEOREM 3. Without loss of generality, we can set $\boldsymbol{\beta}^0 = \mathbf{0}$ in the proof. By Lemma 1, $(\boldsymbol{\delta}(\mathbf{x}) - \mathbf{x})$ is uniformly bounded. Hence to establish admissibility of $\boldsymbol{\delta}(\mathbf{x})$ we need only verify (1.7). To establish inadmissibility it suffices to show (1.8).

Using (2.7) and (3.1), it is clear that $m(\mathbf{x})$ can be bounded above and below by

$$\begin{aligned} m^*(\mathbf{x}) &= C\int_0^\infty \int (1+\sigma_\pi^2)^{-p/2} \exp\left\{\frac{-C'||\mathbf{x}-\mathbf{y}\boldsymbol{\beta}||^2}{(1+\sigma_\pi^2)}\right\}\pi_2(\boldsymbol{\beta})\pi_1(\sigma_\pi^2)\,d\boldsymbol{\beta}\,d\sigma_\pi^2 \\ &= C\int_0^\infty\int_0^\infty\left[\frac{\exp\{-C'[||\mathbf{x}-\mathbf{y}\hat{\boldsymbol{\beta}}||^2/(1+\sigma_\pi^2)+||\hat{\boldsymbol{\beta}}||_{**}^2]\}}{(1+\sigma_\pi^2)^{p/2}[\det(\lambda(1+\sigma_\pi^2)^{-1}\mathbf{y}^t\mathbf{y}\mathbf{A}+\mathbf{I})]^{1/2}}\right] \\ &\qquad\times \pi_1(\sigma_\pi^2)\pi_3(\lambda)\,d\sigma_\pi^2\,d\lambda, \end{aligned}$$

where $\hat{\boldsymbol{\beta}} = (\mathbf{y}^t\mathbf{y})^{-1}\mathbf{y}^t\mathbf{x}$ and $||\cdot||_{**}^2$ is defined in (2.2) with $\mathbf{W} = (1+\sigma_\pi^2)^{-1}\mathbf{I}$ and $\boldsymbol{\beta}^0 = \mathbf{0}$. Next, utilizing (2.8) and the bounds

$$\frac{C||\mathbf{y}\hat{\boldsymbol{\beta}}||^2}{(1+\sigma_\pi^2+\lambda)} \le ||\hat{\boldsymbol{\beta}}||_{**}^2 \le \frac{C'||\mathbf{y}\hat{\boldsymbol{\beta}}||^2}{(1+\sigma_\pi^2+\lambda)},$$

$m^*(\mathbf{x})$ can be bounded above and below by

$$\begin{aligned} m^{**}(\mathbf{x}) &= C\int_0^\infty\int_0^\infty\left[\frac{\exp\left\{(-C'||\mathbf{x}-\mathbf{y}\hat{\boldsymbol{\beta}}||^2/(1+\sigma_\pi^2))-(C''||\mathbf{y}\hat{\boldsymbol{\beta}}||^2/(1+\sigma_\pi^2+\lambda))\right\}}{(1+\sigma_\pi^2)^{p/2}(1+\lambda/(1+\sigma_\pi^2))^{k/2}}\right] \\ &\qquad\times \pi_1(\sigma_\pi^2)\pi_3(\lambda)\,d\sigma_\pi^2\,d\lambda. \end{aligned}$$

Finally, defining

$$r^2 = ||\mathbf{x}||^2 \quad\text{and}\quad w = \frac{||\mathbf{x}-\mathbf{y}\hat{\boldsymbol{\beta}}||^2}{||\mathbf{x}||^2} = 1 - \frac{||\mathbf{y}\hat{\boldsymbol{\beta}}||^2}{||\mathbf{x}||^2}$$

and replacing the C' and C'' by their minimum or maximum, one obtains upper and lower bounds for $m^{**}(\mathbf{x})$ [and hence for $m(\mathbf{x})$] of the form

$$\begin{aligned} &m^{***}(r,w) \\ &\quad= C\int_0^\infty\int_0^\infty\left[\frac{\exp\{-C'r^2(1+\sigma_\pi^2)^{-1}[w+(1-w)/(1+\lambda/(1+\sigma_\pi^2))]\}}{(1+\sigma_\pi^2)^{p/2}(1+\lambda/(1+\sigma_\pi^2))^{k/2}}\right] \\ &\qquad\times \pi_1(\sigma_\pi^2)\pi_3(\lambda)\,d\sigma_\pi^2\,d\lambda. \end{aligned} \tag{4.1}$$

Proof of inadmissibility. Using Conditions 1 and 2,

$$\begin{aligned}
&m^{***}(r,w)\\
&\quad\geq C\int_c^\infty\int_c^\infty \frac{\exp\{-C'r^2(1+\sigma_\pi^2)^{-1}[w+(1-w)/(1+\lambda/(1+\sigma_\pi^2))]\}}{(1+\sigma_\pi^2)^{(a+p/2)}(1+\lambda/(1+\sigma_\pi^2))^{k/2}\lambda^b}\,d\lambda\,d\sigma_\pi^2\\
&\quad\geq C\int_c^\infty \frac{\exp\{-C'r^2(1+\sigma_\pi^2)^{-1}\}}{(1+\sigma_\pi^2)^{(a+p/2)}}\left(\int_c^\infty \frac{1}{(1+\lambda/(1+\sigma_\pi^2))^{k/2}\lambda^b}\,d\lambda\right)d\sigma_\pi^2\\
&\quad\geq \begin{cases} C\displaystyle\int_c^\infty \frac{\exp\{-C'r^2(1+\sigma_\pi^2)^{-1}\}}{(1+\sigma_\pi^2)^{(a+p/2)}-(1-b)^+)}\,d\sigma_\pi^2, & \text{if } b\neq 1,\\ C\displaystyle\int_c^\infty \frac{\exp\{-C'r^2(1+\sigma_\pi^2)^{-1}\}}{(1+\sigma_\pi^2)^{(a+p/2)}(\log(1+\sigma_\pi^2))^{-1}}\,d\sigma_\pi^2, & \text{if } b=1,\end{cases}
\end{aligned}$$

where $+$ denotes the positive part. For $b\neq 1$, making the change of variables $u=c'r^2/(1+\sigma_\pi^2)$ yields

$$\begin{aligned}
m^{***}(r,w) &\geq C\int_0^{cr^2}\frac{\exp\{-u\}u^{(a+p/2-2-(1-b)^+)}}{r^{(p+2a-2-2(1-b)^+)}}\,du\\
&\geq Cr^{-(p+2a-2-2(1-b)^+)}
\end{aligned}$$

for $r>c'$. Applying this to the integral in (1.8) yields

$$\int_c^\infty r^{1-p}\underline{m}(r)\,dr\leq C\int_c^\infty r^{(1-p+p+2a-2-2(1-b)^+)}\,dr,$$

which is finite if $a-(1-b)^+<0$. For $b=1$ the same change of variables leads to $\int_c^\infty r^{1-p}\underline{m}(r)\,dr\leq C\int_c^\infty r^{(-1+2a)}(\log\ r)^{-1}\,dr$, which is finite if $a<0$. Hence (1.8) is satisfied and $\boldsymbol{\delta}(\mathbf{x})$ is inadmissible if $a-(1-b)^+<0$, which happens if either $a<0$ or $a+b<1$, completing the proof.

Proof of admissibility. We break the integral in (4.1) into the four regions $(0,c)\times(0,c)$, $(0,c)\times(c,\infty)$, $(c,\infty)\times(0,c)$ and $(c,\infty)\times(c,\infty)$; call the resulting integrals I_1, I_2, I_3 and I_4, respectively. Clearly, using Conditions 1(i) and 2(i),

$$I_1\leq C\exp\left\{-\frac{C'r^2}{(1+c)}\left[w+\frac{(1-w)}{(1+c)}\right]\right\}\leq C\exp\{-C^*r^2\}. \tag{4.2}$$

Similarly,

$$I_2\leq C\exp\left\{-\frac{C'r^2w}{(1+c)}\right\}\int_c^\infty\frac{\exp\{-C'r^2(1-w)/(1+c+\lambda)\}}{(1+c+\lambda)^{b+k/2}}\,d\lambda, \tag{4.3}$$

where we have used $1\leq(1+c+\lambda)/(1+\lambda)\leq 1+c$. Changing variables to $v=C'r^2(1-w)/(1+c+\lambda)$ results in

$$\begin{aligned}
I_2 &\leq \frac{C\exp\{-C'r^2w\}}{[r^2(1-w)]^{(k/2+b-1)}}\int_0^{c^*r^2(1-w)} e^{-v}v^{(k/2+b-2)}\,dv\\
&\leq \frac{C\exp\{-C'r^2w\}}{[r^2(1-w)]^{(k/2+b-1)}}\min\left\{\Gamma\left(\frac{k}{2}+b-1\right),\frac{[c^*r^2(1-w)]^{(k/2+b-1)}}{(k/2+b-1)}\right\},
\end{aligned}$$

where the second term of "min" arises from bounding $\exp\{-v\}$ by 1. Recalling that w has a beta $((p-k)/2,\ k/2)$ density $g(w)$, it follows that

$$\begin{aligned} I_2^*(r) &\equiv \int_0^1 I_2(r,w)g(w)\,dw \\ &\le C\int_0^{(1-c/r^2)} \frac{\exp\{-C'r^2w\}}{[r^2(1-w)]^{(k/2+b-1)}}\, g(w)\,dw \\ &\quad + C\int_{(1-c/r^2)}^1 \exp\{-C'r^2w\}g(w)\,dw \\ &\le C\int_0^{(1-c/r^2)} \frac{\exp\{-C'r^2w\}w^{((p-k)/2-1)}}{r^{(k+2b-2)}(1-w)^b}\,dw + C\exp\{-C^*r^2\} \\ &\le C\int_0^1 \frac{\exp\{-C'r^2w\}w^{((p-k)/2-1)}}{r^{(k-2)}}\,dw + C\exp\{-C^*r^2\}, \end{aligned}$$

where we have used $r^2(1-w) > c$ on the domain of the next to last integral. Making the change of variables $v = C'r^2w$, yields

$$I_2^*(r) \le C(r^{-(p-2)} + \exp\{-C^*r^2\}). \tag{4.4}$$

To deal with I_3, use Condition 1(ii) and then Condition 2(i) to yield

$$\begin{aligned} I_3 &\le C\int_c^\infty \int_0^c \left[\frac{\exp\{-C'r^2(1+\sigma_\pi^2)^{-1}[w+(1-w)/(1+\lambda/(1+\sigma_\pi^2))]\}}{(1+\sigma_\pi^2)^{(p/2+a)}}\right] \\ &\qquad\qquad \times \pi_2(\lambda)\,d\lambda\,d\sigma_\pi^2 \\ &\le C\int_c^\infty \frac{\exp\{-C^*r^2/(1+\sigma_\pi^2)\}}{(1+\sigma_\pi^2)^{(p/2+a)}}\,d\sigma_\pi^2 \\ &\le Cr^{-(p+2a-2)}. \end{aligned} \tag{4.5}$$

Finally, we must deal with I_4 which, using Conditions 1(ii) and 2(ii), can be bounded by

$$I_4 \le C\int_c^\infty \int_c^\infty \frac{\exp\{-C'r^2(1+\sigma_\pi^2)^{-1}[w+(1-w)/(1+\lambda/(1+\sigma_\pi^2))]\}}{(1+\sigma_\pi^2)^{(p/2+a)}(1+\lambda/(1+\sigma_\pi^2))^{k/2}(1+\lambda)^b}\,d\lambda\,d\sigma_\pi^2.$$

Making the change of variables $v = \lambda/(1+\sigma_\pi^2)$ for the inner integral yields

$$\begin{aligned} I_4 &\le C\int_c^\infty \int_{c/(1+\sigma_\pi^2)}^\infty \frac{\exp\{-C'r^2(1+\sigma_\pi^2)^{-1}[w+(1-w)/(1+v)]\}}{(1+\sigma_\pi^2)^{(p/2+a-1)}(1+v)^{k/2}(1+v(1+\sigma_\pi^2))^b}\,dv\,d\sigma_\pi^2 \\ &\le C\int_c^\infty \int_0^\infty \frac{\exp\{-C'r^2(1+wv)/[(1+\sigma_\pi^2)(1+v)]\}}{(1+\sigma_\pi^2)^{(p/2+a-1)}(1+v)^{(k/2+b)}}\,dv\,d\sigma_\pi^2. \end{aligned}$$

Reversing the order of integration, making the change of variables (for σ_π^2)

$$u = C'r^2(1+wv)/[(1+\sigma_\pi^2)(1+v)]$$

and defining $\psi(r, v, w) = r^2(1+wv)/(1+v)$,

$$
\begin{aligned}
I_4 &\le C\int_0^\infty [\psi(r,v,w)]^{-(p/2+a-2)}(1+v)^{-(k/2+b)} \\
&\qquad \times \left(\int_0^{c^*\psi(r,v,w)} \exp\{-u\}u^{(p/2+a-3)}\,du\right)dv \\
&\le C\int_0^\infty [\psi(r,v,w)]^{-(p/2+a-2)}(1+v)^{-(k/2+b)} \qquad (4.6)\\
&\qquad \times \min\left\{\Gamma\left(\frac{p}{2}+a-2\right), C'[\psi(r,v,w)]^{(p/2+a-2)}\right\}dv \\
&\le C\int_\Omega (1+v)^{-(k/2+b)}\,dv + C\int_{\Omega^c}[\psi(r,v,w)]^{-(p/2+a-2)}(1+v)^{-(k/2+b)}\,dv,
\end{aligned}
$$

where

$$\Omega = \{v\colon \psi(r,v,w) \le c^* \equiv (\Gamma(p/2+a-2)/c')^{1/(p/2+a-2)}\}.$$

Note that we need $a > 2 - p/2$ for the above to be valid; this is satisfied by the conditions in Table 2.

Denoting the two integrals in (4.6) by $I_1^*(r,w)$ and $I_2^*(r,w)$, we have

$$\int_0^1 I_4 g(w)\,dw \le \int_0^1 I_1^*(r,w)g(w)\,dw + \int_0^1 I_2^*(r,w)g(w)\,dw. \qquad (4.7)$$

Now Ω can be rewritten

$$\Omega = \{(v,w)\colon\ 0 < w < c^*/r^2,\ \ v > (r^2-c^*)/(c^*-r^2w)\},$$

so that

$$
\begin{aligned}
&\int_0^1 I_1^*(r,w)g(w)\,dw \\
&\quad \le C\int_0^{c^*/r^2}\left(\int_{(r^2-c^*)/(c^2-r^2w)}^\infty (1+v)^{-(k/2+b)}\,dv\right) \\
&\qquad \times w^{((p-k)/2-1)}(1-w)^{(p/2-1)}\,dw \\
&\quad \le C\int_0^{c^*/r^2}\left(\frac{c^*/r^2-w}{1-w}\right)^{(k/2+b-1)} w^{((p-k)/2-1)}(1-w)^{(k/2-1)}\,dw \\
&\quad \le C\int_0^{c^*/r^2}\left(\frac{c^*}{r^2}-w\right)^{(k/2+b-1)} w^{((p-k)/2-1)}\,dw,
\end{aligned}
$$

since $b \ge 0$ under all cases of Table 2. Making the change of variables $u = r^2w/c^*$, it follows directly that

$$\int_0^1 I_1^*(r,w)g(w)\,dw \le Cr^{-(p-2)}. \qquad (4.8)$$

Next, we consider

$$\begin{aligned}\int_0^1 I_2^*(r,w)g(w)\,dw &= C\iint_{\Omega^c}\left[r^2\left(\frac{1+wv}{1+v}\right)\right]^{-(p/2+a-2)} \\ &\quad\times(1+v)^{-(k/2+b)}w^{((p-k)/2-1)}(1-w)^{(k/2-1)}\,dv\,dw \\ &\le \frac{C}{r^{(p+2a-4)}}\iint_{\Omega^c}\left(\frac{1+wv}{1+v}\right)^{-(p/2+a-2)}(1+v)^{-(k/2+b)}w^{((p-k)/2-1)}\,dv\,dw.\end{aligned} \tag{4.9}$$

Define $\Omega^* = \Omega^c \cap \{(v,w)\colon\ (1+wv)/(1+v) < c' \le 1\}$, and break up (4.9) into integrals over Ω^* and $(\Omega^c - \Omega^*)$. The integral over $(\Omega^c - \Omega^*)$ is trivially bounded since $(1+wv)/(1+v) \ge c'$ and the remainder of the integrand clearly has a finite integral. Hence

$$\begin{aligned}\int_0^1 I_2^*(r,w)g(w)\,dw &\le \frac{C}{r^{(p+2a-4)}}\left(1+\int_{\Omega^*}\left(\frac{1+wv}{1+v}\right)^{-(p/2+a-2)}(1+v)^{-(k/2+b)}w^{((p-k)/2-1)}\,dv\,dw\right) \\ &\le \frac{c}{r^{(p+2a-4)}}\left(1+\int_{\Omega^*}\frac{(1+v)^{((p-k)/2+a-b-2)}w^{((p-k)/2-1)}}{(w(1+v)+1-c')^{(p/2+a-2)}}dvdw\right),\end{aligned}$$

where we used

$$1+wv = w(1+v)+1-w \ge w(1+v)+1-c'$$

on Ω^*. Next, change variables from v to $u = w(1+v)$ and observe that Ω^* transforms to

$$\begin{aligned}&\left\{(u,w)\colon\ \frac{c^*}{r^2} < w\left[1+\frac{(1-w)}{u}\right] < c'\right\} \\ &\quad\subset \left\{(u,w)\colon \frac{c^*}{r^2}\left(\frac{u}{u+1}\right) < w < c'\left(\frac{u}{u+1-c'}\right)\right\}.\end{aligned}$$

Thus we obtain

$$\begin{aligned}\int_0^1 I_2^*(r,w)g(w)\,dw \le \frac{c}{r^{(p+2a-4)}}\left(1+\int_0^\infty \frac{u^{((p-k)/2+a-b-2)}}{(u+1-c')^{(p/2+a-2)}}\left(\int_{c^*u/[r^2(u+1)]}^{c'u/(u+1-c')} w^{b-a}\,dw\right)du.\right.\end{aligned}$$

Computations now have to be done separately for the three cases $b-a<-1$, $b-a=-1$ and $b-a>-1$. The analysis is straightforward and yields

$$\int_0^1 I_2^*(r,w)g(w)\,dw \le \frac{C\log r}{r^{(p+2a-4)}} + \frac{C}{r^{(p+2b-2)}}. \tag{4.10}$$

Combining this with (4.7) and (4.8), we conclude that

$$\int_0^1 I_4 g(w)\,dw \le C\left(\frac{1}{r^{(p-2)}} + \frac{\log r}{r^{(p+2a-4)}} + \frac{1}{r^{(p+2b-2)}}\right). \tag{4.11}$$

To complete the proof, we conclude from (4.2), (4.4), (4.5) and (4.11) that

$$\begin{aligned}\overline{m}(r) &\le \int_0^1 (I_1 + I_2 + I_3 + I_4) g(w)\,dw \\ &\le C\left[e^{-C^* r^2} + \frac{1}{r^{(p-2)}} + \frac{\log r}{r^{(p+2a-4)}} + \frac{1}{r^{(p+2b-2)}}\right].\end{aligned}$$

Hence (1.7) becomes

$$\int_c^\infty (r^{p-1}\overline{m}(r))^{-1}\,dr \ge C\int_c^\infty (r + (\log r) r^{(3-2a)} + r^{(1-2b)})^{-1}\,dr,$$

which is clearly infinite under the conditions of Table 2. This completes the proof of admissibility. □

PROOF OF THEOREM 2. Upper and lower bounds on $m(\mathbf{x})$ are given by (4.1) with $\lambda = 1$ and the integral over λ removed, that is by

$$\begin{aligned} &m^{***}(r, w) \\ &\quad = C\int_0^\infty \frac{\exp\{-C' r^2 (1+\sigma_\pi^2)^{-1}[w + (1-w)/(1 + 1/(1+\sigma_\pi^2))]\}}{(1+\sigma_\pi^2)^{p/2}(1 + 1/(1+\sigma_\pi^2))^{k/2}} \\ &\qquad \times \pi_1(\sigma_\pi^2)\,d\sigma_\pi^2. \end{aligned} \tag{4.12}$$

Proof of inadmissibility. Using Condition 1(ii),

$$\begin{aligned} m^{***}(r, w) &\ge C\int_c^\infty \frac{\exp\{-C' r^2 (1+\sigma_\pi^2)^{-1}\}}{(1+\sigma_\pi^2)^{(p/2+a)}}\,d\sigma_\pi^2 \\ &\ge C r^{-(p+2a-2)}. \end{aligned}$$

Hence

$$\int_c^\infty r^{1-p}\underline{m}(r)\,dr \le C\int_c^\infty r^{(2a-1)}\,dr,$$

which is finite if $a < 0$. By (1.8), $\boldsymbol{\delta}(\mathbf{x})$ is then inadmissible.

Proof of admissibility. Clearly

$$m^{***}(r, w) \le C\int_0^\infty \frac{\exp\{-C^* r^2 (1+\sigma_\pi^2)^{-1}\}}{(1+\sigma_\pi^2)^{p/2}}\,\pi_2(\sigma_\pi^2)\,d\sigma_\pi^2.$$

Breaking this integral up into integrals over $(0, c)$ and (c, ∞) and using Condition 1 yields, after a change of variables in the second integral,

$$m^{***}(r, w) \le C\exp\{-C' r^2\} + C r^{-(p+2a-2)}.$$

Hence

$$\int_C^\infty (r^{p-1}\overline{m}(r))^{-1}\,dr \ge C\int_c^\infty r^{(2a-1)}\,dr,$$

which is infinite if $a \ge 0$. By (1.7), $\boldsymbol{\delta}(\mathbf{x})$ is thus admissible. □

Acknowledgments. We thank Gauri Datta and two referees for helpful comments.

REFERENCES

ANDREWS, R., BERGER, J. and SMITH, M. (1993). Bayesian estimation of fuel economy potential due to technology improvements. *Case Studies in Bayesian Statistics. Lecture Notes in Statist.* **83**. Springer, New York.

ANGERS, J. F. (1987). Development of robust Bayes estimators for a multivariate normal mean. Ph.D. dissertation, Dept. Statistics, Purdue Univ.

BARANCHIK, A. (1964). Multiple regression and estimation of the mean of a multivariate normal distribution. Technical Report 51, Dept. Statistics, Stanford Univ.

BERGER, J. (1976). Admissible minimax estimation of a multivariate normal mean under arbitrary quadratic loss. *Ann. Statist.* **4** 223–226.

BERGER, J. (1980). A robust generalized Bayes estimator and confidence region for a multivariate normal mean. *Ann. Statist.* **8** 716–761.

BERGER, J. (1984). The robust Bayesian viewpoint. In *Robustness in Bayesian Statistics* (J. Kadane, ed.). North-Holland, Amsterdam.

BERGER, J. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer, New York.

BERGER, J. and BERNARDO, J. (1992a). On the development of the reference prior method. In *Bayesian Statistics* (J. M. Bernardo, J. O. Berger A. P. Dawid and A. F. M. Smith, eds.) **4** 35–60. Oxford Univ. Press.

BERGER, J. and BERNARDO, J. (1992b). Reference priors in a variance components problem. *Bayesian Analysis in Statistics and Econometrics*. *Lecture Notes in Statist.* **75** 177–194.

BERGER, J. and DEELY, J. (1988). A Bayesian approach to ranking and selection of related means with alternatives to AOV methodology. *J. Amer. Statist. Assoc.* **83** 364–373.

BERGER, J. and ROBERT, C. (1990). Subjective hierarchical Bayes estimation of a multivariate normal mean: on the frequentist interface. *Ann. Statist.* **18** 617–651.

BOX, G. E. P. and TIAO, G. C. (1973). *Bayesian Inference in Statistical Analysis.* Addison-Wesley, Reading, MA.

BROWN, L. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855–903.

GHOSH, M. (1992). Hierarchical and empirical Bayes multivariate estimation. In *Current Issues in Statistical Inference*. IMS, Hayward, CA.

GOEL, P. K. (1983). Information measures and Bayesian hierarchical models. *J. Amer. Statist. Assoc.* **78** 408–410.

GOOD, I. J. (1983). The robustness of a hierarchical model for multinomials and contingency tables. In *Scientific Inference, Data Analysis, and Robustness* (G. E. P. Box, T. Leonard and C. F. Wu, eds.). Academic Press, New York.

HILL, B. (1977). Exact and approximate Bayesian solutions for inference about variance components and multivariate inadmissibility. In *New Development in the Application of Bayesian Methods* (A. Aykac and C. Brumet, eds.). North-Holland, Amsterdam.

JEFFREYS, H. (1961). *Theory of Probability*. Oxford Univ. Press.

KELKER, D. (1970). Distribution theory of spherical distributions and a location-scale parameter. *Sankhyā Ser. A* **32** 419–430.

MORRIS, C. (1983). Parametric empirical Bayes inference: theory and applications (with discussion). *J. Amer. Statist. Assoc.* **78** 47–65.

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 197–206. Univ. California Press, Berkeley.

STRAWDERMAN, W. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.* **42** 385–388.

STRAWDERMAN, W. (1973). Proper Bayes minimax estimators of the multivariate normal mean for the case of common unknown variances. *Ann. Statist.* **1** 1189–1194.

YE, K. (1993). Sensitivity study of the reference priors in a random effect model. Technical report, Dept. Statistics, Virginia Polytechnic Institute and State Univ., Blacksburg.

ZHENG, Z. (1982). A class of generalized Bayes minimax estimators. In *Statistical Decision Theory and Related Topics* (S. S. Gupta and J. Berger, eds.) **3**. Academic Press, New York.

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907-1399
E-MAIL: berger@stat.purdue.edu

DEPARTMENT OF STATISTICS
RUTGERS UNIVERITY
NEW BRUNSWICK, NEW JERSEY 08903