Bayesian Analysis (2015) **10**, Number 3, pp. 627–664

# Necessary and Sufficient Conditions for High-Dimensional Posterior Consistency under $g$-Priors

Douglas K. Sparks*, Kshitij Khare†, and Malay Ghosh‡

**Abstract.** We examine *necessary and sufficient* conditions for posterior consistency under $g$-priors, including extensions to hierarchical and empirical Bayesian models. The key features of this article are that we allow the number of regressors to grow at the same rate as the sample size and define posterior consistency under the sup vector norm instead of the more conventional Euclidean norm. We consider in particular the empirical Bayesian model of George and Foster (2000), the hyper-$g$-prior of Liang et al. (2008), and the prior considered by Zellner and Siow (1980).

**Keywords:** empirical Bayes, $g$-prior, hyper-$g$-prior, posterior consistency.

## 1 Introduction

Arnold Zellner made pioneering contributions to the fields of statistics and econometrics. One of his works, the $g$-prior (Zellner, 1986), has become a cornerstone of research in Bayesian statistics. The $g$-prior specifies that a vector of regression coefficients is normally distributed a priori with some mean (typically zero) and covariance matrix equal to a scalar multiple (typically denoted by $g$) of the covariance matrix of the maximum likelihood estimator. These priors are useful for conventional hierarchical and empirical Bayesian analysis (Ghosh et al., 1982) for linear regression models, but their application extends well beyond to variable selection (George and Foster, 2000), Bayesian classification of high-dimensional low–sample size data (Mallick et al., 2005), and many other interesting topics of research. The excellent article of Liang et al. (2008) provides a succinct account of mixtures of $g$-priors for Bayesian variable selection.

One very important but often neglected issue in the selection of priors is to examine the consistency of resulting posteriors in the frequentist sense. We will provide a formal definition in Section 2, but in plain language, this means that as one accumulates more and more samples, the posterior distribution of the parameter under consideration gets closer and closer to its true value, eventually becoming degenerate at this point in the limit. Recently, the notion of posterior consistency has also been considered in nonparametric settings (Barron et al., 1999; Ghosal et al., 2000).

In the $g$-prior model, if the number of regressors $p$ does not vary with $n$, then it can easily be seen that the resulting posterior is inconsistent if $g$ is fixed, but the

*Stanford University, dksparks@stanford.edu
†University of Florida, kdkhare@stat.ufl.edu
‡University of Florida, ghoshm@stat.ufl.edu

 DOI: 10.1214/14-BA893

problem disappears when $g \equiv g_n$ with $g_n \to \infty$. See Section 2 for the details of these results. Now suppose instead that the number of regressors $p \equiv p_n$ increases with $n$ but satisfies $p_n < n$ and $p_n/n \to \alpha$, where $0 \leq \alpha < 1$. This situation represents the so-called "large $p$, large $n$" regime, which has been considered in the context of model selection. Berger et al. (2003) provide scenarios where the Bayes factor is consistent but the Bayesian Information Criterion (BIC) is not, with the explanation that BIC may be a poor approximation to the Bayes factor when $p_n \to \infty$. Moreno et al. (2010) examine consistency of the Bayes factor for nested normal linear models with $p_n \to \infty$, including the case where $p_n$ grows at the same rate as the sample size. Also, Jiang (2007) addressed the variable selection problem when $p_n > n$ and provided convergence rates for the fitted densities in a broad class of generalized linear models.

In the context of parameter estimation as examined here, Ghosal (1999) considered certain types of high-dimensional linear models and provided a valuable contribution by proving not only posterior consistency but also asymptotic normality of the posterior distribution. However, our work differs from Ghosal (1999) in three principal respects. First, and perhaps most fundamentally, the $g$-prior model itself involves an unknown sampling variance $\sigma^2$ with an associated prior (the prior on the regression coefficients is taken to be conditional on $\sigma^2$). Such a structure is not included in the class of models considered in Ghosal (1999). Second, we provide *necessary and sufficient* conditions for posterior consistency in three of the four $g$-prior models we consider. While we readily admit that stronger results such as asymptotic normality are perhaps more useful whenever posterior consistency occurs, our *necessary* conditions demonstrate circumstances in which posterior consistency fails to occur at all, which we believe to be interesting in their own right. Third, our work allows the parameter space for the $p_n$-dimensional vector of regression coefficients to be taken as $\mathbb{R}^{p_n}$, as is natural. This contrasts with Ghosal (1999), which essentially requires the restriction of the parameter space to a sequence of compact sets.

Bontemps (2011) also extended the work of Ghosal (1999) in several ways by permitting the model to be misspecified and the number of regressors to grow proportionally to the sample size, the latter of which is also a feature of our work. However, our work differs from Bontemps (2011), most notably by allowing the consideration of models where the sampling variance $\sigma^2$ is assumed to be unknown. There are also differences in the assumptions. In particular, Bontemps (2011) does not make any assumption analogous to the eigenvalue bounds that we will later impose in (A3). On the other hand, unlike Bontemps (2011), we do not make any assumptions on the asymptotic behavior of the true coefficient vector $\boldsymbol{\beta}_{0n}$. We must also emphasize once again that, unlike Bontemps (2011), we provide conditions that are both *necessary and sufficient* for posterior consistency. This establishes circumstances in which posterior consistency definitively does *not* occur, which can in some cases be rather surprising (see the remarks following Theorem 2, for example). The recent work of Armagan et al. (2013) establishes sufficient conditions for posterior consistency in linear models under shrinkage priors. Again, the most notable difference between the models considered in Armagan et al. (2013) and the $g$-prior based models considered in this paper is that the variance parameter $\sigma^2$ is assumed to be known in Armagan et al. (2013). Lee and Oh (2013) consider a high dimensional Bayesian Principal Components Analysis regression setup with $p_n > n$

and normal priors, and examine posterior consistency (in the $\ell_2$-norm) and convergence rates under appropriate assumptions on the rank of the design matrix.

Other authors have addressed the asymptotic properties of $g$-prior models, but for model selection instead of parameter estimation. Fernandez et al. (2001) provided both theoretical results and simulation-based evidence for the consistency of posterior model probabilities under particular choices for the $g$-prior hyperparameter $g \equiv g_n$. Liang et al. (2008) took a more theoretical approach and proved the consistency of posterior model probabilities under hierarchical and empirical Bayesian $g$-prior models, but only in the case where the dimensionality $p_n$ of the full model is fixed. More recently, Shang and Clayton (2011) provided similar results in the case where $p_n \to \infty$, albeit under a considerable number of assumptions. They also note that these results can be extended to $p_n > n$, the so-called "large $p$, small $n$" regime, when combined with certain dimension reduction approaches. See also the work of Zhang et al. (2009).

Another new feature of our work is that we have established posterior consistency under the sup vector norm $\ell_\infty$ ($||\boldsymbol{x}||_\infty = \max_{1\le i\le p} |x_i|$) rather than the conventional $\ell_2$ ($||\boldsymbol{x}||_2 = [\sum_{i=1}^p x_i^2]^{1/2}$) vector norm. The choice is motivated primarily because the $\ell_\infty$ norm introduces added flexibility to our procedure, since it is weaker than the $\ell_2$ norm (as a vector norm), noting that $||\boldsymbol{x}||_\infty \le ||\boldsymbol{x}||_2$. In particular, for proving consistency when the number of covariates $p_n$ grows with the sample size, the sup norm approach allows $p_n$ to grow at a faster rate than is possible under the $\ell_2$ norm. The simplest yet most convincing fact in this regard is the following. For the linear model $\boldsymbol{Y}_n = \boldsymbol{X}_n\boldsymbol{\beta}_n + \boldsymbol{e}_n$ with i.i.d. Gaussian errors and $\boldsymbol{X}_n^T\boldsymbol{X}_n = n\mathbf{I}_{p_n}$ (orthogonal covariates, $\mathbf{I}_{p_n}$ denotes the identity matrix of dimension $p_n$), the MLE for $\boldsymbol{\beta}_n$ is consistent under the $\ell_2$ vector norm if and only if $p_n = o(n)$. However, the MLE for $\boldsymbol{\beta}_n$ is still consistent under the $\ell_\infty$ norm for any $p_n < n$. See remark immediately following Lemma 1.

As discussed above, if $p_n \to \infty$, it is harder to prove posterior consistency under the $\ell_2$ norm as compared to the $\ell_\infty$ norm. However, in the same vein, it is harder to prove posterior inconsistency under the $\ell_\infty$ norm as compared to the $\ell_2$ norm. In particular, any necessary condition for posterior consistency under the $\ell_\infty$ norm is also a necessary condition for posterior consistency under the $\ell_2$ norm. Hence, this paper also provides novel necessary conditions for posterior consistency under the conventional $\ell_2$ norm (note that assumption (A2) in Section 2 subsumes the case $p_n/n \to 0$).

The outline of the remaining sections is as follows. Section 2 provides necessary and sufficient conditions for posterior consistency for a nonstochastic sequence $\{g_n, n \ge 1\}$. In the process, we demonstrate the posterior consistency or inconsistency of some popular recommendations regarding the choice of $g_n$. Section 3 provides necessary and sufficient conditions for posterior consistency in an empirical Bayesian context in which $g_n$ is estimated from the data. Section 4 provides necessary and sufficient conditions for posterior consistency under the hierarchical hyper-$g$-prior model (Liang et al., 2008). Section 5 considers the celebrated Zellner-Siow prior (Zellner and Siow, 1980) and provides a sufficient (though not necessary) condition for posterior consistency under this model. At the end of each of Sections 2–5, the interpretations and implications of the results are briefly discussed. Some final remarks are made in Section 6. It should be noted that although the key results of Sections 3–5 yield the same condition for posterior consistency, the techniques used to prove these results differ substantially among the three models. Furthermore, the coincidence of the conditions in Theorems 2–4 should not be misconstrued as a suggestion that the same condition would be shared by other hierarchical or empirical Bayesian $g$-prior models. Specifically, this condition is not shared by Theorem 1, yet the non-hierarchical model addressed by Theorem 1 can be considered as a hierarchical model with a sequence of degenerate hyperpriors. Moreover, it should again be noted that the conditions in Theorems 1–3 are both necessary and sufficient, but the condition for the Zellner-Siow $g$-prior model provided in Theorem 4 is merely sufficient, and its necessity or lack thereof is not presently clear.

## 2 Non-Hierarchical Model

Consider the usual linear model $\boldsymbol{Y}_n = \boldsymbol{X}_n\boldsymbol{\beta}_n + \boldsymbol{e}_n$, with response $\boldsymbol{Y}_n = (Y_{n,1}, \ldots, Y_{n,n})^T$, covariates $\boldsymbol{X}_n = (\boldsymbol{x}_{n,1}, \ldots, \boldsymbol{x}_{n,n})^T$, regression coefficients $\boldsymbol{\beta}_n = (\beta_{n,1}, \ldots, \beta_{n,p_n})^T$ and errors $\boldsymbol{e}_n = (e_1, \ldots, e_n)^T$. We now impose the following assumptions:

(A1) The errors are distributed as $\boldsymbol{e}_n \sim N_n(\mathbf{0}_n, \sigma^2\mathbf{I}_n)$. Here $\mathbf{0}_n$ denotes the vector of length $n$ with all zero entries.

(A2) The number of regressors $p_n$ is a nondecreasing sequence with $p_n < n$ and $p_n/n \to \alpha$, where $0 \le \alpha < 1$.

(A3) The eigenvalues $\lambda_{n,1}, \ldots, \lambda_{n,p_n}$ of the matrix $n(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1}$ satisfy $0 < \lambda_{\min} \le \inf_{n,i}\lambda_{n,i} \le \sup_{n,i}\lambda_{n,i} \le \lambda_{\max} < \infty$ for some $\lambda_{\min}$ and $\lambda_{\max}$.

Note that (A3) implies that $\lambda_{\max}^{-1}\boldsymbol{I}_{p_n} \le n^{-1}\boldsymbol{X}_n^T\boldsymbol{X}_n \le \lambda_{\min}^{-1}\boldsymbol{I}_{p_n}$. This assumption is identical to assumption (A2) of Armagan et al. (2013).

The goal in such a model is estimation of $\boldsymbol{\beta}_n$. Minimal sufficiency leads to the reduction $(\hat{\boldsymbol{\beta}}_n, S_n)$, where $\hat{\boldsymbol{\beta}}_n = (\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1}\boldsymbol{X}_n^T\boldsymbol{Y}_n$, the maximum likelihood estimator of $\boldsymbol{\beta}_n$, and $S_n = ||\boldsymbol{Y}_n - \boldsymbol{X}_n\hat{\boldsymbol{\beta}}_n||_2^2$, the error sum of squares. Note that conditional on $\boldsymbol{\beta}_n$ and $\sigma^2$, $\hat{\boldsymbol{\beta}}_n$ and $S_n$ are mutually independent with $\hat{\boldsymbol{\beta}}_n \mid \boldsymbol{\beta}_n, \sigma^2 \sim N_{p_n}(\boldsymbol{\beta}_n, \sigma^2(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1})$ and $S_n \mid \boldsymbol{\beta}_n, \sigma^2 \sim \sigma^2\chi^2_{n-p_n}$.

Now suppose priors are specified as $\boldsymbol{\beta}_n \mid \sigma^2 \sim N_{p_n}(\boldsymbol{\gamma}_n, g\sigma^2(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1})$ (Zellner's $g$-prior) and $\sigma^2 \sim \text{InverseGamma}(a/2, b/2)$, where we permit $a \ge -2$ and $b \ge 0$ to accommodate such improper priors as $\pi(\sigma^2) \propto 1/\sigma^2, 1/\sigma$, or 1. Suppose further that $g \equiv g_n$ is specified as a known sequence of constants. This collection of likelihoods and priors comprises our non-hierarchical $g$-prior model, which we denote by $P_M$. One motivation for the use of such a model is the convenient form of the Bayes estimator under squared error loss,

$$\hat{\boldsymbol{\beta}}_n^{\mathrm{B}} := E_M(\boldsymbol{\beta}_n \mid \hat{\boldsymbol{\beta}}_n, S_n) = \frac{g_n}{g_n+1}\hat{\boldsymbol{\beta}}_n + \frac{1}{g_n+1}\boldsymbol{\gamma}_n,$$

where $\hat{\boldsymbol{\beta}}_n$ denotes the MLE.

We now introduce the formal definition of posterior consistency.

**Definition.** *Let* $\boldsymbol{\beta}_{0n} \in \mathbb{R}^{p_n}$ *for each* $n \geq 1$, *and let* $\sigma_0^2 > 0$. *Now let* $P_0$ *denote the distribution of* $\{(\hat{\boldsymbol{\beta}}_n, S_n), n \geq 1\}$ *under the model* $\boldsymbol{Y}_n = \boldsymbol{X}_n \boldsymbol{\beta}_{0n} + \boldsymbol{e}_n$, *where* $\boldsymbol{e}_n \sim N_n(\boldsymbol{0}_n, \sigma_0^2 I_n)$, *for each* $n \geq 1$. *The sequence of posterior distributions* $P_M(\boldsymbol{\beta}_n \mid \hat{\boldsymbol{\beta}}_n, S_n)$ *is said to be consistent under the* $\ell_\infty$ *norm at* $\{(\boldsymbol{\beta}_{0n}, \sigma_0^2), n \geq 1\}$ *if* $P_M(||\boldsymbol{\beta}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 0$ *a.s.*$(P_0)$ *for every* $\epsilon > 0$.

It should be immediately noted that the type of posterior consistency considered herein is fundamentally different from what could instead be considered in the analysis of Bayesian methodology, that is, convergence of the posterior under the same model $P_M$ under which it is derived. In this case, one is assuming that the prior associated with the model $P_M$ is in some sense "true." However, this approach is perhaps too favorable in that posterior consistency is quite easy to achieve. In fact, in this approach, a quite general result due to Doob (1948) states that posterior consistency occurs on a set of parameter values with probability 1 under the prior associated with $P_M$. Instead, the type of posterior consistency considered herein is fundamentally frequentist in nature, that is, the values $\boldsymbol{\beta}_{0n}$ and $\sigma_0^2$ are considered fixed but unknown.

The frequentist properties of Bayesian methods have been of interest for some time. Even pure frequentists may be interested in originally Bayesian procedures, or limits and approximations thereof, due to considerations such as admissibility and the convenient elimination of nuisance parameters. Indeed, it was shown as early as Laplace (1774) that in simple cases, the posterior distribution and the distribution of the maximum likelihood estimator are comparable for large sample sizes. More sophisticated versions of such results have been developed in more recent times (Bernstein, 1934; Diaconis and Freedman, 1986; Ghosh et al., 1982; LeCam, 1982; von Mises, 1964).

We now provide a lemma establishing strong frequentist consistency of the MLE $\hat{\boldsymbol{\beta}}_n$ in the $\ell_\infty$ norm.

**Lemma 1.** *Let* $\boldsymbol{Z}_n \sim N_{p_n}(\boldsymbol{0}_{p_n}, n^{-1}\boldsymbol{V}_n)$, *where* $p_n < n$, *and where the eigenvalues* $\omega_{n,1}, \ldots, \omega_{n,p_n}$ *of* $\boldsymbol{V}_n$ *satisfy* $\sup_{n,i} \omega_{n,i} = \omega_{\max} < \infty$. *Then* $||\boldsymbol{Z}_n||_\infty \to 0$ *almost surely.*

**Proof** (Proof of Lemma 1)**.** *First note that* $Var(Z_{n,i}) = n^{-1}V_{n,ii} \leq n^{-1}\omega_{\max}$, *and* $n^{1/2}V_{n,ii}^{-1/2}Z_{n,i} \sim N(0,1)$. *Now let* $\epsilon > 0$. *Since*

$$\sum_{n=1}^{\infty} P\left(||\boldsymbol{Z}_n||_\infty > \epsilon\right) = \sum_{n=1}^{\infty} P\left(\max_{1 \leq i \leq p_n} |Z_{n,i}| > \epsilon\right),$$

*it follows that*

$$\begin{aligned} \sum_{n=1}^{\infty} P\left(||\boldsymbol{Z}_n||_\infty > \epsilon\right) &\leq \sum_{n=1}^{\infty} \sum_{i=1}^{p_n} P\left(n^{1/2}V_{n,ii}^{-1/2}\,|Z_{n,i}| > \epsilon\left(n^{-1}V_{n,ii}\right)^{-1/2}\right) \\ &\leq \sum_{n=1}^{\infty} \sum_{i=1}^{p_n} P\left(n^{1/2}V_{n,ii}^{-1/2}\,|Z_{n,i}| > \epsilon\,\omega_{\max}^{-1/2} n^{1/2}\right) \end{aligned}$$

$$\leq \sum_{n=1}^{\infty} \sum_{i=1}^{p_n} \frac{15\omega_{\max}^3}{\epsilon^6 n^3} < \infty$$

*by applying Markov's inequality to* $n^3 V_{n,ii}^{-3} Z_{n,i}^6$. *The result follows from the Borel-Cantelli lemma, noting that* $p_n < n$.

Observe that Lemma 1 under $P_0$ with assumptions (A1)–(A3) and $\boldsymbol{Z}_n = \hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}$ implies that $||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty \to 0$ a.s.$(P_0)$. Thus, the MLE $\hat{\boldsymbol{\beta}}_n$ retains strong frequentist consistency in the $\ell_\infty$ norm even as $p_n$ grows at a rate exactly proportional to $n$. To contrast this with the behavior of the MLE under the conventional $\ell_2$ vector norm, note that we have the upper bound

$$||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_2^2 \leq \frac{\sigma_0^2 \lambda_{\max}}{n} \left(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}\right)^T \frac{1}{\sigma_0^2} \boldsymbol{X}_n^T \boldsymbol{X}_n \left(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}\right)$$

and a similar lower bound with $\lambda_{\max}$ replaced by $\lambda_{\min}$. Since

$$\left(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}\right)^T \frac{1}{\sigma_0^2} \boldsymbol{X}_n^T \boldsymbol{X}_n \left(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}\right) \sim \chi_{p_n}^2,$$

it can be immediately seen that $p_n = o(n)$ is required for strong frequentist consistency of the MLE $\hat{\boldsymbol{\beta}}_n$ under the $\ell_2$ norm.

In a Bayesian analysis, Lemma 1 leads to the following useful lemma, which essentially states that $\boldsymbol{\beta}_{0n}$ may be replaced by $\hat{\boldsymbol{\beta}}_n$ in the definition of posterior consistency.

**Lemma 2.** *In the g-prior model (both hierarchical and non-hierarchical),* $P_M(||\boldsymbol{\beta}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 0$ *a.s.*$(P_0)$ *for every* $\epsilon > 0$ *if and only if* $P_M(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 0$ *a.s.*$(P_0)$ *for every* $\epsilon > 0$.

**Proof** (Proof of Lemma 2)**.** *The triangle inequality implies that*

$$\begin{aligned} P_M & \left(\left|\left|\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > 2\epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) - P_M \left(\left|\left|\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}\right|\right|_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) \\ & \leq P_M \left(||\boldsymbol{\beta}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) \\ & \leq P_M \left(\left|\left|\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon/2 \mid \hat{\boldsymbol{\beta}}_n, S_n\right) + P_M \left(\left|\left|\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}\right|\right|_\infty > \epsilon/2 \mid \hat{\boldsymbol{\beta}}_n, S_n\right). \end{aligned}$$

*When conditioning on* $\hat{\boldsymbol{\beta}}_n$ *and* $S_n$,

$$P_M \left(||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) = I\left(||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon\right),$$

*where* $I(\cdot)$ *denotes the usual indicator function. Lemma 1 implies that* $||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty \to 0$ *a.s.*$(P_0)$*, from which it follows that* $I(||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon) \to 0$ *a.s.*$(P_0)$ *for all* $\epsilon > 0$. *This and the above inequalities immediately yield the result.*

To establish results on posterior consistency or inconsistency in the non-hierarchical $g$-prior model, we first define $T_n := (\hat{\boldsymbol{\beta}}_n - \boldsymbol{\gamma}_n)^T \boldsymbol{X}_n^T \boldsymbol{X}_n (\hat{\boldsymbol{\beta}}_n - \boldsymbol{\gamma}_n)$, so that $T_n/\sigma^2$ is the usual frequentist likelihood ratio test statistic for a test of $H_0 : \boldsymbol{\beta}_n = \boldsymbol{\gamma}_n$ vs. $H_a : \boldsymbol{\beta}_n \neq \boldsymbol{\gamma}_n$ with known variance $\sigma^2$. Then the joint posterior $\pi_n(\boldsymbol{\beta}_n, \sigma^2 \mid \hat{\boldsymbol{\beta}}_n, S_n)$ is given by

$$\pi_n(\boldsymbol{\beta}_n, \sigma^2 \mid \hat{\boldsymbol{\beta}}_n, S_n) \propto \exp\left[-\frac{1}{2}\left(\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^{\mathrm{B}}\right)^T \left(\frac{g_n}{g_n+1}\sigma^2\left(\boldsymbol{X}_n^T\boldsymbol{X}_n\right)^{-1}\right)^{-1}\left(\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^{\mathrm{B}}\right)\right] \times \left(\sigma^2\right)^{-(n+p_n+a)/2} \exp\left[-\frac{1}{2\sigma^2}\left(S_n + b + \frac{T_n}{g_n+1}\right)\right],$$

and integrating out $\boldsymbol{\beta}_n$ from this yields the marginal posterior of $\sigma^2$,

$$\pi_n(\sigma^2 \mid \hat{\boldsymbol{\beta}}_n, S_n) \propto \left(\sigma^2\right)^{-(n+a)/2} \exp\left[-\frac{1}{2\sigma^2}\left(S_n + b + \frac{T_n}{g_n+1}\right)\right],$$

i.e., $\sigma^2 \mid \hat{\boldsymbol{\beta}}_n, S_n \sim \text{InverseGamma}((n+a-2)/2,\ \widetilde{T}_n/2)$, where we define $\widetilde{T}_n := S_n + b + (g_n+1)^{-1}T_n$. For notational convenience, for each $n \geq 1$, define

$$\begin{aligned}
\breve{\lambda}_{0n} &:= \frac{n||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2}{(\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n})^T \boldsymbol{X}_n^T \boldsymbol{X}_n (\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n})}, \\
\theta_{0n} &:= E_0(T_n) = p_n\sigma_0^2 + n\breve{\lambda}_{0n}^{-1}\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\,, \\
\widetilde{\theta}_{0n} &:= E_0(\widetilde{T}_n) = (n-p_n)\sigma_0^2 + b + \frac{1}{g_n+1}\left(p_n\sigma_0^2 + n\breve{\lambda}_{0n}^{-1}\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right),
\end{aligned}$$

and note that $\lambda_{\min} \leq \breve{\lambda}_{0n} \leq \lambda_{\max}$ since $\lambda_{\max}^{-1}\boldsymbol{I}_{p_n} \leq n^{-1}\boldsymbol{X}_n^T\boldsymbol{X}_n \leq \lambda_{\min}^{-1}\boldsymbol{I}_{p_n}$.

The following lemmas establish the behavior of various quantities under $P_0$, and they will be heavily used in proving posterior consistency or inconsistency in both the non-hierarchical and hierarchical $g$-prior models. The proof of each lemma can be found in the Appendix.

**Lemma 3.** $(n-p_n)^{-1}S_n \to \sigma_0^2$ *a.s.*$(P_0)$.

**Lemma 4.** *If* $\alpha > 0$ *or* $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 > 0$, *then* $T_n/\theta_{0n} \to 1$ *a.s.*$(P_0)$.

**Lemma 5.** $\widetilde{T}_n/\widetilde{\theta}_{0n} \to 1$ *a.s.*$(P_0)$.

The following lemmas regarding the normal distribution will be useful in establishing the condition for posterior consistency in the non-hierarchical case. The proofs are provided in the Appendix.

**Lemma 6.** *Let* $\boldsymbol{Z}_n \sim N_{p_n}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$, $\boldsymbol{\Sigma}_n$ *positive definite,* $n \geq 1$. *If* $||\boldsymbol{\mu}_n - \boldsymbol{\xi}_n||_\infty \nrightarrow 0$, *then there exist* $\epsilon > 0$ *and a subsequence* $k_n$ *of* $n$ *such that* $P(||\boldsymbol{Z}_{k_n} - \boldsymbol{\xi}_{k_n}||_\infty > \epsilon) \geq 1/2$ *for all* $n$.

**Lemma 7.** *Let* $Z \sim N(\mu, \tau^2)$. *Then* $P(|Z| \leq \xi) \leq 1 - 2\Phi(-\xi/\tau)$ *for every* $\xi \geq 0$, *where* $\Phi$ *is the standard normal cdf.*

**Lemma 8.** *Let* $\boldsymbol{Z}_n \sim N_{p_n}(\boldsymbol{0}_{p_n}, \boldsymbol{\Sigma}_n)$ *for each* $n \geq 1$, *where* $\boldsymbol{\Sigma}_n$ *has each diagonal entry equal to 1 and eigenvalues* $\omega_{n,1}, \ldots, \omega_{n,p_n}$. *If* $\inf_{n,i}\omega_{n,i} = \omega_{\min}$, *then* $\inf_{n,i} \mathit{Var}(Z_i \mid Z_{i+1}, \ldots, Z_{p_n}) \geq \omega_{\min}$.

Finally, one additional lemma provides a key result about the marginal posterior of $\sigma^2$. Again, the proof is deferred to the Appendix.

**Lemma 9.** *In the non-hierarchical $g$-prior model, the posterior distribution of $\sigma^2$ satisfies $P_M(\widetilde{\theta}_{0n}/2n \le \sigma^2 \le 2\widetilde{\theta}_{0n}/n \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 1$ a.s.$(P_0)$.*

Note that although $g_n$ does not appear explicitly in the result in Lemma 9, the result nevertheless does depend on the choice of $g_n$ since it is involved in the quantity $\widetilde{\theta}_{0n}$.

We now state and prove the necessary and sufficient condition for posterior consistency in the non-hierarchical $g$-prior model.

**Theorem 1.** *In the non-hierarchical $g$-prior model $P_M$, posterior consistency occurs if and only if both $(g_n+1)^{-1}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty \to 0$ and $g_n(g_n+1)^{-2}(\log p_n)n^{-1}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2 \to 0$.*

The proof of this theorem in provided in the Appendix.

## 2.1 Interpretations and Implications

In the same vein as frequentist consistency, posterior consistency can be conceptualized as the idea that the center (not necessarily the mean) of the posterior distribution converges to the true value while the spread (not necessarily the variance) of the posterior distribution converges to zero. In light of this, it is noteworthy that the two conditions in Theorem 1 arise from precisely such considerations. The first condition controls the convergence to zero of the $\ell_\infty$-distance between the posterior's center and the true value $\boldsymbol{\beta}_{0n}$, while the second condition controls the convergence of the posterior's spread to zero. Both conditions are necessary for posterior consistency to hold.

In the simple case where $p_n$ does not increase with $n$, it is typical to fix the prior mean as $\boldsymbol{\gamma}_n = \boldsymbol{\gamma}$ and to assume that $\boldsymbol{\beta}_{0n} = \boldsymbol{\beta}_0$ also does not vary with $n$. In this case it can be immediately seen that although the second condition of Theorem 1 is satisfied, the first condition fails except in the serendipitous case that $\boldsymbol{\gamma} = \boldsymbol{\beta}_0$. Of course, the result is somewhat obvious even without appealing to Theorem 1, since the posterior mean is simply a weighted average of the MLE $\hat{\boldsymbol{\beta}}_n$, which is strongly consistent for $\boldsymbol{\beta}_0$, and the prior mean $\boldsymbol{\gamma}$ with weights $g(g+1)^{-1}$ and $(g+1)^{-1}$. In this case, the situation may be remedied by taking any choice of $g_n$ that tends to infinity. For instance, the unit information prior (Kass and Wasserman, 1995) is equivalent to taking $g_n = n$, while $g_n = \max\{n, p_n^2\}$ has also been recommended (Fernandez et al., 2001). Either choice yields posterior consistency in the fixed-$p$ case.

The result of Theorem 1 becomes more interesting when $p_n \to \infty$. Suppose that $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty = O(1)$, but $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 = O(p_n)$. This can happen, for example, if (a) $\boldsymbol{\gamma}_n = \hat{\boldsymbol{\beta}}_n$, or (b) the entries of $\boldsymbol{\beta}_{0n}$ are uniformly bounded and $\boldsymbol{\gamma}_n = c\hat{\boldsymbol{\beta}}_n$ where $0 \le c < 1$ (follows immediately from Lemma 1). In this case, the first condition is satisfied as long as $g_n \to \infty$, but the second condition imposes the additional requirement that $g_n$ must grow faster than $p_n n^{-1} \log p_n$. The aforementioned choices of $g_n = n$ or $g_n = \max\{n, p_n^2\}$ provide posterior consistency in this case as well.

As another special case, suppose $p_n = O(n)$ exactly, but suppose only a finite number $m > 0$ of components of $\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}$ are nonzero and these $m$ components remain fixed as $n$ grows. This circumstance could arise with the logical choice $\boldsymbol{\gamma}_n = \mathbf{0}_{p_n}$ if only the first few covariates are present in the "true" frequentist model $P_0$, but covariates continue to be added as the sample size increases. Then any $g_n \to \infty$ ensures posterior consistency. This case is admittedly uninteresting in the non-hierarchical model, but we will revisit its behavior later under empirical and hierarchical Bayesian models.

# 3 Empirical Bayesian Model

A popular approach is to avoid specifying $g$ or $g_n$ altogether by the use of an empirical Bayes method (George and Foster, 2000) in which the value of $g$ is estimated from the data. The most common technique is to use the value of $g$ that maximizes its marginal likelihood, restricted to $g \geq 0$. By integrating out $\boldsymbol{\beta}_n$ and $\sigma^2$ from the joint distribution of $\hat{\boldsymbol{\beta}}_n, S_n, \boldsymbol{\beta}_n, \sigma^2$, the marginal likelihood of $g$ is found to be

$$L(g; \hat{\boldsymbol{\beta}}_n, S_n) \propto (g+1)^{(n-p_n+a-2)/2} \left[(g+1)(S_n+b) + T_n\right]^{-(n+a-2)/2},$$

for which the maximizing value of $g$ subject to $g \geq 0$ is

$$\hat{g}_n^{\mathrm{EB}} := \max\left\{0, \left(\frac{n-p_n+a-2}{S_n+b}\right)\left(\frac{T_n}{p_n}\right) - 1\right\}.$$

We first provide a lemma (proven in the Appendix) that addresses the behavior of $\hat{g}_n^{\mathrm{EB}}$.

**Lemma 10.** *If* $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 > 0$, *then* $\liminf_{n\to\infty} \hat{g}_n^{EB} > 0$ *a.s.*$(P_0)$.

Since $\hat{g}_n^{\mathrm{EB}}$ is simply a function of $(\hat{\boldsymbol{\beta}}_n, S_n)$, the empirical Bayes posterior is identical to the simple non-hierarchical Bayes posterior, but with the data-dependent quantity $\hat{g}_n^{\mathrm{EB}}$ in place of $g_n$. Thus, while Theorem 1 would allow us to immediately state a necessary and sufficient condition for posterior consistency in terms of $\hat{g}_n^{\mathrm{EB}}$, an alternative condition not involving data-dependent quantities would be preferable. The following result gives precisely such a condition and establishes its necessity and sufficiency.

**Theorem 2.** *In the empirical Bayes $g$-prior model, posterior consistency occurs if and only if either $\alpha = 0$ or there does not exist a subsequence $k_n$ of $n$ and a constant $A > 0$ such that $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to A$ and $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \nrightarrow 0$.*

The proof of this theorem is provided in the Appendix.

## 3.1 Interpretations and Implications

It should be noted that there is no immediately obvious remedy for inconsistency in an empirical Bayesian $g$-prior model due to the failure of the conditions in Theorem 2. For any particular non-hierarchical $g$-prior model, Theorem 1 implies that there always exists a choice of $g_n$ growing sufficiently fast to ensure posterior consistency (although the choice may depend on $\boldsymbol{\beta}_{0n}$). However, such options are not available in the empirical Bayes approach, since $g$ is selected via a specified function of the data.

Another salient consequence of Theorem 2 is that if $p_n = o(n)$, then the empirical Bayes model exhibits posterior consistency for all values of $\boldsymbol{\gamma}_n$ and $\boldsymbol{\beta}_{0n}$. However, if $p_n = O(n)$ exactly, then the situation is not as simple. For example, if $\boldsymbol{\gamma}_n = \mathbf{0}_{p_n}$ for every $n$ and $\lim_{n\to\infty} \|\boldsymbol{\beta}_{0n}\|_2^2 = \infty$, then $\|\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}\|_2^2$ converges to $\infty$ for every subsequence $k_n$, which implies that posterior consistency occurs. Similarly, if $\boldsymbol{\gamma}_n = \hat{\boldsymbol{\beta}}_n$ for every $n$, then by Lemma 1, $\|\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}\|_\infty$ converges to zero for every subsequence $k_n$, which implies that posterior consistency occurs. On the other hand, suppose that only a fixed number $p^\star > 0$ of components of $\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}$ are nonzero and these $p^\star$ components remain fixed as $n$ grows. Then clearly both $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty$ and $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2$ converge to constants, so the condition of Theorem 2 fails, and the posterior is inconsistent.

This behavior is perhaps somewhat surprising. If the prior mean $\boldsymbol{\gamma}_n$ is imagined as a guess for the true $\boldsymbol{\beta}_{0n}$, then one might speculate that posterior inconsistency would only occur when the guess is quite bad, i.e., when $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2$ or $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty$ grows too quickly. However, in the empirical Bayesian setting, Theorem 2 shows that this is not the case. Intuitively, the reason is that if we allow the data to determine the value of $g$, then a prior mean $\boldsymbol{\gamma}_n$ that is "too close" to $\boldsymbol{\beta}_{0n}$ (in the $\ell_2$ sense) may cause the data to choose $g$ values that tend to a finite constant, rather than to infinity, which leads to posterior inconsistency. An open question regarding this behavior is whether this interesting behavior is in some way dependent on the Gaussian tails imposed by the $g$-prior model. However, the derivation of a similar condition for a hierarchical $g$-prior model considered later in Theorem 3 casts doubt on this possibility, since the hierarchical model simply corresponds to some marginal prior with heavier tails.

## 4 Hyper-$g$-Prior Hierarchical Model

An alternative approach to the specification of $g$ is a hierarchical model in which $g$ is considered a hyperparameter and is given a hyperprior $\pi_n(g)$. Under this model, the joint posterior is given by

$$\begin{aligned}
&\pi_n(\boldsymbol{\beta}_n, \sigma^2, g \mid \hat{\boldsymbol{\beta}}_n, S_n) \\
&\propto \exp\left[-\frac{1}{2}\left[\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^{\mathrm{B}}(g)\right]^T \left(\frac{g}{g+1}\sigma^2\left(\boldsymbol{X}_n^T\boldsymbol{X}_n\right)^{-1}\right)^{-1}\left[\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^{\mathrm{B}}(g)\right]\right] \\
&\qquad \times \left(\sigma^2\right)^{-(n+p_n+a)/2} \exp\left[-\frac{1}{2\sigma^2}\left(S_n + b + \frac{T_n}{g+1}\right)\right] g^{-p_n/2}\,\pi_n(g),
\end{aligned}$$

where $\tilde{\boldsymbol{\beta}}_n^{\mathrm{B}}(g) := E(\boldsymbol{\beta}_n \mid g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n) = g(g+1)^{-1}\hat{\boldsymbol{\beta}}_n + (g+1)^{-1}\boldsymbol{\gamma}_n$. Integrating out $\boldsymbol{\beta}_n$ and subsequently $\sigma^2$ yields the marginal posteriors

$$\pi_n(\sigma^2, g \mid \hat{\boldsymbol{\beta}}_n, S_n) \propto \left(\sigma^2\right)^{-(n+a)/2} \exp\left[-\frac{1}{2\sigma^2}\left(S_n + b + \frac{T_n}{g+1}\right)\right](g+1)^{-p_n/2}\,\pi_n(g), \tag{1}$$

$$\pi_n(g \mid \hat{\boldsymbol{\beta}}_n, S_n) \propto (g+1)^{-p_n/2} \left( S_n + b + \frac{T_n}{g+1} \right)^{-(n+a-2)/2} \pi_n(g). \tag{2}$$

The following technical lemma, which is proven in the Appendix, establishes a relationship between posterior consistency in the hierarchical $g$-prior model and the convergence of a particular sequence of posterior probabilities. Note that the lemma makes no assumptions on the particular form of the hyperprior $\pi_n(g)$.

**Lemma 11.** *In a hierarchical g-prior model, suppose that* $n^{-3} T_n^2 E_M[g^2(g+1)^{-4} \mid \hat{\boldsymbol{\beta}}_n, S_n] \to 0$ *a.s.*$(P_0)$. *Then posterior consistency occurs if and only if* $P_M[(g+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n] \to 0$ *a.s.*$(P_0)$ *for every* $\epsilon > 0$.

The form of the marginal posterior of $g$ in (2) suggests that a convenient choice of hyperprior is $\pi_n(g) \propto (g+1)^{-c/2}$ for some constant $c$, called the hyper-$g$-prior (Liang et al., 2008). This prior is proper for $c > 2$, and there exists an argument (Liang et al., 2008) for taking $2 < c \leq 4$, but we instead permit $c$ to take any real value in the present analysis. The hyper-$g$-prior yields the posterior

$$\begin{aligned} \pi_n(g \mid \hat{\boldsymbol{\beta}}_n, S_n) &\propto (g+1)^{-(p_n+c)/2} \left( S_n + b + \frac{T_n}{g+1} \right)^{-(n+a-2)/2} \\ &\propto (g+1)^{(n-p_n+a-c-2)/2} \big[ (g+1)(S_n+b) + T_n \big]^{-(n+a-2)/2}. \end{aligned} \tag{3}$$

It will also be useful to define the transformation

$$u := \frac{(g+1)(S_n+b)}{(g+1)(S_n+b)+T_n}, \qquad W_n := \frac{S_n+b}{S_n+b+T_n}, \tag{4}$$

so that $g \geq 0$ if and only if $u \geq W_n$. The next lemma asserts that Lemma 11 applies with this choice of hyperprior. The proof can be found in the Appendix.

**Lemma 12.** *With the hyper-g-prior,* $n^{-3} T_n^2 E_M[g^2(g+1)^{-4} \mid \hat{\boldsymbol{\beta}}_n, S_n] \to 0$ *a.s.*$(P_0)$.

To examine the behavior of the posterior probabilities in Lemma 11 under the hyper-$g$-prior, we begin by using the posterior in (3) to write

$$\begin{aligned} &P_M\left[ \left( \frac{1}{g+1} \right) ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n \right] \\ &\quad = P_M\left[ g < \frac{1}{\epsilon} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty - 1 \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n \right] \\ &\quad = \frac{\displaystyle\int_0^{q_n(\epsilon)} (g+1)^{(n-p_n+a-c-2)/2} \big[ (g+1)(S_n+b)+T_n \big]^{-(n+a-2)/2} \, dg}{\displaystyle\int_0^{\infty} (g+1)^{(n-p_n+a-c-2)/2} \big[ (g+1)(S_n+b)+T_n \big]^{-(n+a-2)/2} \, dg}, \end{aligned}$$

where we define $q_n(\epsilon) := \max\{0, \epsilon^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty - 1\}$. Now define

$$\widetilde{L}_n(\epsilon) := \frac{\epsilon^{-1} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty (S_n+b)}{\epsilon^{-1} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty (S_n+b) + T_n}, \qquad L_n(\epsilon) := \max\left\{ W_n, \widetilde{L}_n(\epsilon) \right\},$$

and apply the transformation in (4) to obtain

$$P_M\left[\left(\frac{1}{g+1}\right)||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty>\epsilon\;\middle|\;\hat{\boldsymbol{\beta}}_n,S_n\right] = \frac{\displaystyle\int_{W_n}^{L_n(\epsilon)} u^{(n-p_n+a-c-2)/2}(1-u)^{(p_n+c-4)/2}\,du}{\displaystyle\int_{W_n}^{1} u^{(n-p_n+a-c-2)/2}(1-u)^{(p_n+c-4)/2}\,du} = \frac{P_M[W_n<U_n<L_n(\epsilon)\mid\hat{\boldsymbol{\beta}}_n,S_n]}{P_M(U_n>W_n\mid\hat{\boldsymbol{\beta}}_n,S_n)}, \tag{5}$$

where $U_n\sim\text{Beta}((n-p_n+a-c)/2,\ (p_n+c-2)/2)$ and is independent of $\hat{\boldsymbol{\beta}}_n$ and $S_n$ under $P_M$. Note that by the properties of the beta distribution, $U_n\to 1-\alpha$ a.s.$(P_0)$, and $P_M(U_n>W_n\mid\hat{\boldsymbol{\beta}}_n,S_n)>0$ for all $n$ a.s.$(P_0)$ since $W_n<1$ for all $n$ a.s.$(P_0)$. We now introduce several technical results regarding these quantities that will be useful in proving the main theorem. The proofs are deferred to the Appendix.

**Lemma 13.** *If* $\liminf_{n\to\infty}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\ge\delta$ *for some* $\delta>0$*, then* $\limsup_{n\to\infty}W_n\le(1-\alpha)\lambda_{\max}\sigma_0^2/(\delta+\lambda_{\max}\sigma_0^2)<1-\alpha$ *a.s.*$(P_0)$.

**Lemma 14.** *If* $||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\to\infty$*, then (i)* $W_n\to 0$ *a.s.*$(P_0)$*, and also (ii)* $L_n(\epsilon)\to 0$ *a.s.*$(P_0)$ *for every* $\epsilon>0$.

**Lemma 15.** *If* $\liminf_{n\to\infty}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty>0$ *and* $||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\to A$*, where* $0<A<\infty$*, then (i) for every* $\epsilon>0$*, there exists* $L^\star(\epsilon)<1$ *such that* $\limsup_{n\to\infty}L_n(\epsilon)\le L^\star(\epsilon)$ *a.s.*$(P_0)$*, and (ii) for every* $\zeta<1$*, there exists* $\epsilon_\zeta>0$ *such that*

$$\liminf_{n\to\infty}L_n\left(\epsilon_\zeta\right)>\zeta\ \ a.s.(P_0).$$

To prove our main result, we will also need the following lemma, which provides a simple result about beta random variables, the proof of which is in the Appendix.

**Lemma 16.** *Let* $Z_n\sim$ *Beta*$(a_n,b_n)$ *for* $n\ge 1$*, where* $a_n/n\to 1-\alpha$ *and* $b_n/n\to\alpha$*, with* $0\le\alpha<1$*. Then* $P(1-\alpha-\epsilon\le Z_n\le 1-\alpha+\epsilon)\to 1$ *for every* $\epsilon>0$.

We may now state and prove the main result, a necessary and sufficient condition for posterior consistency in the hyper-$g$-prior hierarchical model. Interestingly, this condition is identical to the one given in Theorem 2 for the empirical Bayesian model.

**Theorem 3.** *In the g-prior model with the hyper-g-prior, posterior consistency occurs if and only if either* $\alpha=0$ *or there does not exist a subsequence* $k_n$ *of* $n$ *and a constant* $A>0$ *such that* $||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_2^2\to A$ *and* $||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_\infty\nrightarrow 0$.

The proof of this theorem is provided in the Appendix.

### 4.1 Interpretations and Implications

It should not be entirely surprising that the empirical Bayesian and hyper-$g$-prior hierarchical models share the same necessary and sufficient condition for posterior consistency. Indeed, the choice $c = 0$ yields the Uniform$(0, \infty)$ hyperprior on $g$, and in this case the marginal posterior and likelihood of $g$ coincide. More generally, we should expect an adequately well-behaved hierarchical model to exhibit broadly similar behavior to the empirical Bayesian model, since both models essentially permit the data to determine the value of $g$.

## 5 Zellner-Siow Hierarchical Model

Another popular choice for the hyperprior $\pi_n(g)$ is $g \sim \text{InverseGamma}(1/2, n/2)$, called the Zellner-Siow hyperprior (Zellner and Siow, 1980). The motivation behind this choice is clearest when $\boldsymbol{X}_n^T \boldsymbol{X}_n = n\boldsymbol{I}_{p_n}$, in which case it leads to marginal Cauchy priors for each component of $\boldsymbol{\beta}_n$. In this section, we will provide a sufficient condition for posterior consistency with the Zellner-Siow hyperprior. It still remains an open problem to determine if the condition is also necessary.

For general $\boldsymbol{X}_n^T \boldsymbol{X}_n$, the Zellner-Siow hyperprior yields the posterior

$$\begin{aligned}\pi_n(g \mid \hat{\boldsymbol{\beta}}_n, S_n) &\propto (g+1)^{-(p_n)/2} \left( S_n + b + \frac{T_n}{g+1} \right)^{-(n+a-2)/2} g^{-3/2} \exp\left( -\frac{n}{2g} \right) \\ &\propto (g+1)^{(n-p_n+a-2)/2} \big[ (g+1)(S_n+b) + T_n \big]^{-(n+a-2)/2} g^{-3/2} \\ &\quad \times \exp\left( -\frac{n}{2g} \right). \end{aligned} \tag{6}$$

We begin with a lemma showing that Lemma 11 applies in this model. The proof is deferred to the Appendix.

**Lemma 17.** *With the Zellner-Siow hyperprior,* $n^{-3}\, T_n^2\, E_M[g^2(g+1)^{-4} \mid \hat{\boldsymbol{\beta}}_n, S_n] \to 0$ *a.s.*$(P_0)$.

Now consider the form of the posterior probabilities in Lemma 11 under this hyperprior. By once again making the transformation in (4), we may write

$$\begin{aligned} &P_M\left[ \left( \frac{1}{g+1} \right) ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n \right] \\ &= \frac{\displaystyle\int_{W_n}^{L_n(\epsilon)} \frac{u^{(n-p_n+a-2)/2}(1-u)^{(p_n-4)/2}}{\{[u-W_n]\,/\,[W_n(1-u)]\}^{3/2}} \exp\left[ -\frac{nW_n(1-u)}{2(u-W_n)} \right] du}{\displaystyle\int_{W_n}^{1} \frac{u^{(n-p_n+a-2)/2}(1-u)^{(p_n-4)/2}}{\{[u-W_n]\,/\,[W_n(1-u)]\}^{3/2}} \exp\left[ -\frac{nW_n(1-u)}{2(u-W_n)} \right] du} \end{aligned}$$

$$= \frac{\displaystyle\int_{W_n}^{L_n(\epsilon)} f_n(u) \left[\frac{u-W_n}{1-u}\right]^{-3/2} \exp\left[-\frac{nW_n(1-u)}{2(u-W_n)}\right] du}{\displaystyle\int_{W_n}^{1} f_n(u) \left[\frac{u-W_n}{1-u}\right]^{-3/2} \exp\left[-\frac{nW_n(1-u)}{2(u-W_n)}\right] du}, \tag{7}$$

where $f_n$ is the density of a $\text{Beta}[(n-p_n+a)/2, (p_n-2)/2]$ random variable with respect to Lebesgue measure. The following lemma (proven in the Appendix) addresses the lower tail probabilities of such a sequence.

**Lemma 18.** *Let $Z_n \sim \text{Beta}(a_n, b_n)$ for $n \geq 1$, where $a_n/n \to 1-\alpha$ and $b_n/n \to \alpha$, with $0 \leq \alpha < 1$, and let $\xi \geq 0$. Then (i) $P(Z_n \leq \xi) \leq 4^n \xi^{n(1-\alpha)}$ for all sufficiently large $n$ if $\alpha > 0$, and (ii) $P(Z_n \leq \xi) \leq \xi^{n/2}$ for all sufficiently large $n$ if $\alpha = 0$.*

Note that the bound provided by Lemma 18 in the case where $0 < \alpha < 1$ is only useful if $\xi^{1-\alpha} < 1/4$. Now let $Q_n(\epsilon)$ and $R_n$ denote the numerator and denominator, respectively, of (7). The following lemmas establish some results regarding these quantities that will effectively provide the proof of the main theorem. Their proofs are provided in the Appendix.

**Lemma 19.** *If $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 > 0$, then there exists a finite constant $K$ such that $R_n \geq \exp(-nK)$ for all sufficiently large $n$ a.s.$(P_0)$.*

**Lemma 20.** *If $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to \infty$, then there exists a sequence of constants $\kappa_n(\epsilon) \to \infty$ such that $Q_n(\epsilon) \leq \exp[-n\kappa_n(\epsilon)]$ for all sufficiently large $n$ a.s.$(P_0)$.*

**Lemma 21.** *If $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to A > 0$, $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty > 0$, and $\alpha = 0$, then $Q_n(\epsilon)/R_n \to 0$ a.s.$(P_0)$ for every $\epsilon > 0$.*

We may now state the main theorem, which establishes the same sufficient condition for posterior consistency under the Zellner-Siow hyperprior as for the conjugate hyperprior and empirical Bayes models of the previous sections. However, unlike Theorems 2 and 3, it does not establish the necessity of the condition, which remains an open question.

**Theorem 4.** *In the $g$-prior model with the Zellner-Siow hyperprior, posterior consistency occurs if either $\alpha = 0$ or there does not exist a subsequence $k_n$ of $n$ and a constant $A > 0$ such that $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to A$ and $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \nrightarrow 0$.*

The proof of this theorem is provided in the Appendix.

## 5.1 Interpretations and Implications

Since the same condition is sufficient for posterior consistency under both the hyper-$g$-prior and Zellner-Siow hierarchical models, one might wonder if this condition is sufficient for posterior consistency under every hierarchical model. However, the falsehood of such a claim is made clear by the observation that the non-hierarchical model, for which the sufficient condition differs, is simply a special case of the hierarchical model in which the hyperprior $\pi_n$ is specified to be degenerate at $g_n$. In actuality, the posterior consistency or inconsistency of hierarchical models with other hyperpriors on $g$ remains a topic for future consideration.

# 6 Summary

We have derived conditions for posterior consistency under $g$-priors by defining posterior consistency under the $\ell_\infty$ vector norm, which allows useful results to be obtained even when the number of parameters $p \equiv p_n$ grows in proportion to the sample size $n$. Using this definition, we have obtained conditions for posterior consistency under a variety of $g$-prior models. First, we have obtained a necessary and sufficient condition for posterior consistency in the non-hierarchical model in which $g \equiv g_n$ is specified as a series of constants. Additionally, we have derived a necessary and sufficient condition for posterior consistency under both the empirical Bayesian $g$-prior model (George and Foster, 2000) and the hyper-$g$-prior model (Liang et al., 2008). Interestingly, we have found that the condition is the same for both models, and we have illustrated that the necessity of the condition proves posterior *inconsistency* in a somewhat surprising scenario. Finally, we have shown that this same condition is sufficient for posterior consistency in the Zellner-Siow $g$-prior model (Zellner and Siow, 1980), but the condition's necessity or lack thereof remains an open question for future consideration.

# Appendix: Proofs

**Proof** (Proof of Lemma 3)**.** *Under $P_0$, the expectation and fourth central moment of $S_n$ are $E_0(S_n) = (n-p_n)\sigma_0^2$ and $(\mu_4)_0(S_n) = 12(n-p_n)(n-p_n+4)\sigma_0^8$. Let $\epsilon > 0$. Then*

$$\sum_{n=1}^{\infty} P_0\left(\left|\frac{S_n}{n-p_n} - \sigma_0^2\right| > \epsilon\right) \leq \frac{12\sigma_0^8}{\epsilon^4}\sum_{n=1}^{\infty}\frac{n-p_n+4}{(n-p_n)^3} < \infty,$$

*so $(n-p_n)^{-1}S_n \to \sigma_0^2$ a.s.$(P_0)$ by the Borel-Cantelli lemma.*

**Proof** (Proof of Lemma 4)**.** *Note that under $P_0$, $T_n/\sigma_0^2$ has a noncentral chi-square distribution with $p_n$ degrees of freedom and noncentrality parameter $\frac{1}{2}n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2$. Then the fourth central moment of $T_n$ under $P_0$ is*

$$\begin{aligned}
&(\mu_4)_0(T_n) \\
&:= E_0\left[(T_n - \theta_{0n})^4\right] = E_0\left\{\left[T_n - \left(p_n\sigma_0^2 + n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right)\right]^4\right\} \\
&= 12\sigma_0^4\left(p_n\sigma_0^2 + 2n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right)^2 + 48\sigma_0^6\left(p_n\sigma_0^2 + 4n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right) \\
&\leq 12\sigma_0^4\left(2p_n\sigma_0^2 + 2n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right)^2 + 48\sigma_0^6\left(4p_n\sigma_0^2 + 4n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right) \\
&= 48\sigma_0^4\theta_{0n}^2 + 192\sigma_0^6\theta_{0n}. \qquad (8)
\end{aligned}$$

*Define $\delta := \liminf_{n\to\infty}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2$. Observe that if $\alpha > 0$, then $\theta_{0n} \geq p_n\sigma_0^2 > \alpha n\sigma_0^2/2$ for all sufficiently large $n$, and so $\theta_{0n}^{-1} = O(n^{-1})$. If $\delta > 0$, then $\theta_{0n} \geq$*

$n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 > n\lambda_{\max}^{-1}\delta/2$ *for all sufficiently large* $n$, *and so* $\theta_{0n}^{-1} = O(n^{-1})$. *Either way,* $\theta_{0n}^{-1} = O(n^{-1})$, *so the fourth central moment of* $T_n/\theta_{0n}$ *under* $P_0$ *is*

$$(\mu_4)_0\left(\frac{T_n}{\theta_{0n}}\right) \leq \frac{48\sigma_0^4}{\theta_{0n}^2} + \frac{192\sigma_0^6}{\theta_{0n}^3} = O(n^{-2}).$$

*Then for any* $\epsilon > 0$,

$$\sum_{n=1}^{\infty} P_0\left(\left|\frac{T_n}{\theta_{0n}} - 1\right| > \epsilon\right) \leq \sum_{n=1}^{\infty} \frac{1}{\epsilon^4}(\mu_4)_0\left(\frac{T_n}{\theta_{0n}}\right) < \infty,$$

*which implies that* $T_n/\theta_{0n} \to 1$ *a.s.*$(P_0)$ *by the Borel-Cantelli lemma.*

**Proof** (Proof of Lemma 5)**.** *It follows from (8) that the fourth central moment of* $\widetilde{T}_n$ *under* $P_0$ *is*

$$\begin{aligned}
(\mu_4)_0\left(\widetilde{T}_n\right) &:= E_0\left[\left(\widetilde{T}_n - \widetilde{\theta}_{0n}\right)^4\right] \\
&= E_0\left\{\left[S_n - (n-p_n)\sigma_0^2 + \frac{T_n}{g_n+1} - \frac{p_n\sigma_0^2 + n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2}{g_n+1}\right]^4\right\} \\
&\leq 8E_0\left\{\left[S_n - E_0(S_n)\right]^4\right\} + 8E_0\left\{\left[\frac{T_n}{g_n+1} - E_0\left(\frac{T_n}{g_n+1}\right)\right]^4\right\} \\
&= 96(n-p_n)(n-p_n+4)\sigma_0^8 + \frac{12\sigma_0^4}{(g_n+1)^4}\left(p_n\sigma_0^2 + 2n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right)^2 \\
&\quad + \frac{48\sigma_0^6}{(g_n+1)^4}\left(p_n\sigma_0^2 + 4n\breve{\lambda}_{0n}^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right) \\
&\leq 96(n-p_n+4)^2\sigma_0^8 + 48\sigma_0^4\widetilde{\theta}_{0n}^2 + 192\sigma_0^6\widetilde{\theta}_{0n}.
\end{aligned}$$

*Since* $\widetilde{\theta}_{0n} \geq (n-p_n)\sigma_0^2$, *the fourth central moment of* $\widetilde{T}_n/\widetilde{\theta}_{0n}$ *under* $P_0$ *is*

$$\begin{aligned}
&(\mu_4)_0\left(\frac{\widetilde{T}_n}{\widetilde{\theta}_{0n}}\right) \\
&\leq \frac{96(n-p_n+4)^2\sigma_0^8}{\widetilde{\theta}_{0n}^4} + \frac{48\sigma_0^4}{\widetilde{\theta}_{0n}^2} + \frac{192\sigma_0^6}{\widetilde{\theta}_{0n}^3} \leq \frac{96(n-p_n+4)^2}{(n-p_n)^4} + \frac{48}{(n-p_n)^2} + \frac{192}{(n-p_n)^3},
\end{aligned}$$

*which is* $O(n^{-2})$. *Then for any* $\epsilon > 0$,

$$\sum_{n=1}^{\infty} P_0\left(\left|\frac{\widetilde{T}_n}{\widetilde{\theta}_{0n}} - 1\right| > \epsilon\right) \leq \sum_{n=1}^{\infty} \frac{1}{\epsilon^4}(\mu_4)_0\left(\frac{\widetilde{T}_n}{\widetilde{\theta}_{0n}}\right) < \infty,$$

*which implies that* $\widetilde{T}_n/\widetilde{\theta}_{0n} \to 1$ *a.s.*$(P_0)$ *by the Borel-Cantelli lemma.*

**Proof** (Proof of Lemma 6)**.** *Assume* $||\boldsymbol{\mu}_n - \boldsymbol{\xi}_n||_\infty \nrightarrow 0$. *Then there exists a subsequence* $k_n$ *of* $n$ *and a* $\delta > 0$ *such that* $||\boldsymbol{\mu}_{k_n} - \boldsymbol{\xi}_{k_n}||_\infty > \delta$ *for all* $n$. *There also exists an* $i_n$, $1 \leq i_n \leq p_n$, *such that* $|\mu_{k_n,i_n} - \theta_{k_n,i_n}| = ||\boldsymbol{\mu}_{k_n} - \boldsymbol{\xi}_{k_n}||_\infty > \delta$ *for all* $n$. *Then either*

$\mu_{k_n,i_n} < \theta_{k_n,i_n} - \delta$ *(Case 1) or* $\mu_{k_n,i_n} > \theta_{k_n,i_n} + \delta$ *(Case 2). Now let* $0 < \epsilon < \delta$*, and note that* $P\left(||\boldsymbol{Z}_{k_n} - \boldsymbol{\xi}_{k_n}||_\infty \leq \epsilon\right) \leq P\left(\xi_{k_n,i_n} - \delta \leq Z_{k_n,i_n} \leq \xi_{k_n,i_n} + \delta\right)$*. Recall that* $\boldsymbol{\Sigma}_{k_n}$ *is assumed positive definite. Then in Case 1,*

$$P\left(\xi_{k_n,i_n} - \delta \leq Z_{k_n,i_n} \leq \xi_{k_n,i_n} + \delta\right) \leq P\left(\mu_{k_n,i_n} \leq Z_{k_n,i_n}\right) = 1/2,$$

*while in Case 2,*

$$P\left(\xi_{k_n,i_n} - \delta \leq Z_{k_n,i_n} \leq \xi_{k_n,i_n} + \delta\right) \leq P\left(Z_{k_n,i_n} \leq \mu_{k_n,i_n}\right) = 1/2.$$

*Either way,* $P(||\boldsymbol{Z}_{k_n} - \boldsymbol{\xi}_{k_n}||_\infty > \epsilon) \geq 1/2$ *for all* $n$.

**Proof** (Proof of Lemma 7)**.** *Note that for any* $t > 0$, $\Phi(z+t) - \Phi(z-t)$ *is maximized at* $z = 0$*. Hence,*

$$\begin{aligned} P(|Z| \leq \xi) = P(-\xi \leq Z \leq \xi) &= P\left(\frac{-\xi - \mu}{\tau} \leq \frac{Z - \mu}{\tau} \leq \frac{\xi - \mu}{\tau}\right) \\ &= \Phi\left(\frac{\xi}{\tau} - \frac{\mu}{\tau}\right) - \Phi\left(-\frac{\xi}{\tau} - \frac{\mu}{\tau}\right) \\ &\leq \Phi\left(\frac{\xi}{\tau}\right) - \Phi\left(-\frac{\xi}{\tau}\right), \end{aligned}$$

*from which it immediately follows that* $P(|Z| \leq \xi) \leq 1 - 2\Phi(-\xi/\tau)$.

**Proof** (Proof of Lemma 8)**.** *For each* $i = 1, \ldots, p_n$*, partition* $\boldsymbol{\Sigma}_n$ *as*

$$\boldsymbol{\Sigma}_n = \begin{bmatrix} \boldsymbol{\Sigma}_{n,i,11} & \boldsymbol{\Sigma}_{n,i,1i} & \boldsymbol{\Sigma}_{n,i,12} \\ \boldsymbol{\Sigma}_{n,i,1i}^T & \Sigma_{n,ii} & \boldsymbol{\Sigma}_{n,i,2i} \\ \boldsymbol{\Sigma}_{n,i,12}^T & \boldsymbol{\Sigma}_{n,i,2i}^T & \boldsymbol{\Sigma}_{n,i,22} \end{bmatrix},$$

*where the submatrices* $\boldsymbol{\Sigma}_{n,i,11}$ *and* $\boldsymbol{\Sigma}_{n,i,22}$ *along the diagonal have dimension* $(i-1) \times (i-1)$ *and* $(p_n - i) \times (p_n - i)$*, respectively. Then define* $\widetilde{\Sigma}_{n,i} := Var(Z_i \mid Z_{i+1}, \ldots, Z_{p_n})$*, so that* $\widetilde{\Sigma}_{n,i} = \Sigma_{n,ii} - \boldsymbol{\Sigma}_{n,i,2i}\boldsymbol{\Sigma}_{n,i,22}^{-1}\boldsymbol{\Sigma}_{n,i,2i}^T$*. Note that* $\widetilde{\Sigma}_{n,i}^{-1}$ *is the first diagonal entry of*

$$\begin{bmatrix} \Sigma_{n,ii} & \boldsymbol{\Sigma}_{n,i,2i} \\ \boldsymbol{\Sigma}_{n,i,2i}^T & \boldsymbol{\Sigma}_{n,i,22} \end{bmatrix}^{-1},$$

*which has eigenvalues bounded above by* $\omega_{\min}^{-1}$ *since the eigenvalues of a principal submatrix are bounded below by the smallest eigenvalue of the full matrix. Hence* $\widetilde{\Sigma}_{n,i}^{-1} \leq \omega_{\min}^{-1}$*, and the result immediately follows.*

**Proof** (Proof of Lemma 9)**.** *Recall that* $\widetilde{T}_n/\widetilde{\theta}_{0n} \to 1$ *a.s.*$(P_0)$ *by Lemma 5. Then for all sufficiently large* $n$*,*

$$\begin{aligned} &P_M\left(\frac{\widetilde{\theta}_{0n}}{2n} \leq \sigma^2 \leq \frac{2\widetilde{\theta}_{0n}}{n} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) \\ \geq\, &P_M\left(\frac{3\widetilde{T}_n}{4(n+a-4)} \leq \sigma^2 \leq \frac{5\widetilde{T}_n}{4(n+a-4)} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) \;\; a.s.(P_0) \end{aligned}$$

$$
\begin{aligned}
&= P_M\left(\left|\sigma^2 - E_M\left(\sigma^2 \mid \hat{\boldsymbol{\beta}}_n, S_n\right)\right| \leq \frac{\widetilde{T}_n}{4(n+a-4)} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\geq 1 - \left(\frac{4(n+a-4)}{\widetilde{T}_n}\right)^2 \left(\frac{2\widetilde{T}_n^2}{(n+a-4)^2(n+a-6)}\right) \\
&= 1 - \frac{32}{n+a-6} \to 1,
\end{aligned}
$$

*where the last inequality is a consequence of Chebyshev's inequality, for which we note that* $\mathit{Var}_M(\sigma^2 \mid \hat{\boldsymbol{\beta}}_n, S_n) = 2(n+a-4)^{-2}(n+a-6)^{-1}\widetilde{T}_n^2$.

**Proof** (Proof of Theorem 1)**.** *By Lemma 2, we may replace* $\boldsymbol{\beta}_{0n}$ *with* $\hat{\boldsymbol{\beta}}_n$ *in the definition of posterior consistency. We will now consider four cases.*

*Case 1: Suppose* $(g_n+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty \nrightarrow 0$. *Then since* $||\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n||_\infty \geq ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty - ||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty$ *and* $||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty \to 0$ *a.s.*$(P_0)$ *by Lemma 1, it follows that* $(g_n+1)^{-1}||\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n||_\infty \nrightarrow 0$ *a.s.*$(P_0)$. *Now observe that under* $P_M$,

$$
\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n \mid \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n \sim N_{p_n}\left(\frac{1}{g_n+1}\left(\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n\right), \frac{g_n}{g_n+1}\sigma^2\left(\boldsymbol{X}_n^T\boldsymbol{X}_n\right)^{-1}\right).
$$

*Then by Lemma 6, there exists an* $\epsilon > 0$ *and a subsequence* $k_n$ *of* $n$ *such that, a.s.*$(P_0)$, $P_M(||\boldsymbol{\beta}_{k_n} - \hat{\boldsymbol{\beta}}_{k_n}||_\infty > \epsilon \mid \sigma^2, \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}) > 1/2$ *for every* $n$ *and every* $\sigma^2 > 0$. *Then*

$$
\begin{aligned}
&P_M\left(||\boldsymbol{\beta}_{k_n} - \hat{\boldsymbol{\beta}}_{k_n}||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right) \\
&= E_M\left[P_M\left(||\boldsymbol{\beta}_{k_n} - \hat{\boldsymbol{\beta}}_{k_n}||_\infty > \epsilon \mid \sigma^2, \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right) \;\middle|\; \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right] \\
&\geq 1/2 \quad \text{for every } n \ \text{ a.s.}(P_0).
\end{aligned}
$$

*Therefore* $P_M(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n) \nrightarrow 0$, *so posterior consistency does not occur.*

*For the remaining cases, suppose* $(g_n+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty \to 0$. *Then since* $||\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n||_\infty \leq ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty + ||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty$ *and* $||\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_{0n}||_\infty \to 0$ *a.s.*$(P_0)$ *by Lemma 1, it follows that* $(g_n+1)^{-1}||\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n||_\infty \to 0$ *a.s.*$(P_0)$. *Then*

$$
\begin{aligned}
&P_M\left(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B||_\infty > 2\epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) - P_M\left(||\hat{\boldsymbol{\beta}}_n^B - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\quad \leq P_M\left(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\quad\quad \leq P_M\left(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B||_\infty > \epsilon/2 \mid \hat{\boldsymbol{\beta}}_n, S_n\right) + P_M\left(||\hat{\boldsymbol{\beta}}_n^B - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon/2 \mid \hat{\boldsymbol{\beta}}_n, S_n\right)
\end{aligned}
$$

*by the triangle inequality. Note that*

$$
P_M\left(||\hat{\boldsymbol{\beta}}_n^B - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) = I(||\hat{\boldsymbol{\beta}}_n^B - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon),
$$

*where* $I(\cdot)$ *denotes the indicator function. But* $\hat{\boldsymbol{\beta}}_n^B - \hat{\boldsymbol{\beta}}_n = (g_n+1)^{-1}(\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n)$, *so this indicator is zero for all sufficiently large* $n$ *a.s.*$(P_0)$. *Therefore, posterior consistency occurs in Cases 2–3 below if and only if* $P_M(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 0$ *a.s.*$(P_0)$ *for every* $\epsilon > 0$. *We now consider the individual cases.*

*Case 2: Suppose that* $(g_n+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty \to 0$, *and also suppose that* $g_n(g_n+1)^{-2}(\log p_n)n^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to 0$. *Observe that*

$$\begin{aligned}
&P_M\left(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\quad = E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B\right|\right|_\infty > \epsilon \mid \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \,\Big|\, \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\quad \le E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B\right|\right|_\infty > \epsilon \mid \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) I\left(\sigma^2 \le \frac{2\widetilde{\theta}_{0n}}{n}\right) \,\Bigg|\, \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\qquad + P_M\left(\sigma^2 > \frac{2\widetilde{\theta}_{0n}}{n} \,\Bigg|\, \hat{\boldsymbol{\beta}}_n, S_n\right).
\end{aligned}$$

*We immediately have that* $P_M(\sigma^2 > 2\widetilde{\theta}_{0n}/n \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 0$ *a.s.*$(P_0)$ *by Lemma 9, so it suffices to work with the first term to establish posterior consistency. Let* $v_{n,ij}$ *denote the* $ij$*th element of* $n(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1}$, *and note specifically that the diagonal elements may be bounded by* $\lambda_{\min} \le v_{n,ii} \le \lambda_{\max}$ *for all* $n$ *and* $i$. *Also recall that* $\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B \mid \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n \sim N_{p_n}(\boldsymbol{0}_{p_n}, g_n(g_n+1)^{-1}\sigma^2(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1})$ *under* $P_M$. *Now let* $\epsilon > 0$, *and bound the aforementioned first term by*

$$\begin{aligned}
&E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B\right|\right|_\infty > \epsilon \mid \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) I\left(\sigma^2 \le \frac{2\widetilde{\theta}_{0n}}{n}\right) \,\Bigg|\, \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\quad \le E_M\left[\sum_{i=1}^{p_n} P_M\left(\left|\beta_{n,i} - \hat{\beta}_{n,i}^B\right| > \epsilon \mid \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) I\left(\sigma^2 \le \frac{2\widetilde{\theta}_{0n}}{n}\right) \,\Bigg|\, \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\quad \le E_M\left[\sum_{i=1}^{p_n} 2\Phi\left(-\sqrt{\frac{\epsilon^2(g_n+1)n}{g_n v_{n,ii}\sigma^2}}\right) I\left(\sigma^2 \le \frac{2\widetilde{\theta}_{0n}}{n}\right) \,\Bigg|\, \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\quad \le 2p_n E_M\left[\Phi\left(-\sqrt{\frac{\epsilon^2(g_n+1)n^2}{2g_n\lambda_{\max}\widetilde{\theta}_{0n}}}\right) \,\Bigg|\, \hat{\boldsymbol{\beta}}_n, S_n\right] = 2p_n\Phi\left(-\sqrt{\frac{\epsilon^2(g_n+1)n^2}{2g_n\lambda_{\max}\widetilde{\theta}_{0n}}}\right)
\end{aligned}$$

*where* $\Phi(\cdot)$ *denotes the standard normal cdf. Then by the Mills ratio,*

$$2p_n\Phi\left(-\sqrt{\frac{\epsilon^2(g_n+1)n^2}{2g_n\lambda_{\max}\widetilde{\theta}_{0n}}}\right) \le 2p_n\sqrt{\frac{g_n\lambda_{\max}\widetilde{\theta}_{0n}}{\pi\epsilon^2(g_n+1)n^2}}\,\exp\left(-\frac{\epsilon^2(g_n+1)n^2}{4g_n\lambda_{\max}\widetilde{\theta}_{0n}}\right).$$

*This expression clearly tends to zero if* $\widetilde{\theta}_{0n}/n$ *is bounded above, so we may instead assume that* $\widetilde{\theta}_{0n}/n \to \infty$, *which by inspection occurs if and only if* $(g_n+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to \infty$. *Then* $\widetilde{\theta}_{0n} \le 2n\breve{\lambda}_{0n}^{-1}(g_n+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2$ *for all sufficiently large* $n$, *and hence*

$$\begin{aligned}
&2p_n\Phi\left(-\sqrt{\frac{\epsilon^2(g_n+1)n^2}{2g_n\lambda_{\max}\widetilde{\theta}_{0n}}}\right) \\
&\quad \le 2p_n\sqrt{\frac{2\lambda_{\max}\, g_n\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2}{\pi\breve{\lambda}_{0n}\epsilon^2(g_n+1)^2\,n}}\,\exp\left(-\frac{\breve{\lambda}_{0n}\epsilon^2(g_n+1)^2\,n}{8\lambda_{\max}\,g_n\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2}\right)
\end{aligned}$$

$$\leq \sqrt{\frac{8\lambda_{\max}\, g_n\, ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\, \log p_n}{\pi \breve{\lambda}_{0n}\epsilon^2 (g_n+1)^2\, n}} \times \exp\left[\left(1 - \frac{\breve{\lambda}_{0n}\epsilon^2 (g_n+1)^2\, n}{8\lambda_{\max}\, g_n\, ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\, \log p_n}\right)\log p_n\right] \to 0 \ \text{ for every } \epsilon > 0$$

*by the assumption that* $g_n(g_n+1)^{-2}(\log p_n)n^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to 0$. *Therefore, posterior consistency occurs.*

*Case 3: Suppose* $(g_n+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty \to 0$, *but now suppose that* $g_n(g_n+1)^{-2}(\log p_n)n^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \nrightarrow 0$. *Then there exist a subsequence* $k_n$ *of* $n$ *and a constant* $\delta > 0$ *such that* $g_{k_n}(g_{k_n}+1)^{-2}(\log p_{k_n})k_n^{-1}||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 > \delta$ *for all* $n$. *Note that posterior inconsistency of the subsequence* $P_M(\boldsymbol{\beta}_{k_n} \mid \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n})$ *implies posterior inconsistency of the overall sequence* $P_M(\boldsymbol{\beta}_n \mid \hat{\boldsymbol{\beta}}_n, S_n)$, *so we may assume without loss of generality that* $k_n = n$ *for notational convenience. Also, define* $\boldsymbol{\Sigma}_n$ *to be the* $p_n \times p_n$ *matrix with elements* $\Sigma_{n,ij} := v_{n,ij}/\sqrt{v_{n,ii}v_{n,jj}}$, *where* $v_{n,ij}$ *denotes the* $ij$*th element of* $n(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1}$ *as before. Then*

$$\begin{aligned} &P_M\left(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) \\ &\geq E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B\right|\right|_\infty > \epsilon \mid \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) I\left(\sigma^2 \geq \frac{\widetilde{\theta}_{0n}}{2n}\right) \,\middle|\, \hat{\boldsymbol{\beta}}_n, S_n\right] \\ &\geq E_M\left[P_M\left(\max_{1\leq i\leq p_n}\left|\beta_{n,i} - \hat{\beta}_{n,i}^B\right| > \sqrt{\epsilon^2\frac{v_{n,i}}{\lambda_{\min}}} \,\middle|\, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) I\left(\sigma^2 \geq \frac{\widetilde{\theta}_{0n}}{2n}\right) \,\middle|\, \hat{\boldsymbol{\beta}}_n, S_n\right]. \end{aligned}$$

*Then we may write*

$$\begin{aligned} &P_M\left(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) \\ &\qquad \geq E_M\left[P_M\left(\max_{1\leq i\leq p_n}|Z_i| > \sqrt{\frac{(g_n+1)\epsilon^2 n}{g_n\lambda_{\min}\sigma^2}} \,\middle|\, \sigma^2\right) I\left(\sigma^2 \geq \frac{\widetilde{\theta}_{0n}}{2n}\right) \,\middle|\, \hat{\boldsymbol{\beta}}_n, S_n\right], \end{aligned}$$

*where* $\boldsymbol{Z}_n \sim N_{p_n}(\boldsymbol{0}_{p_n}, \boldsymbol{\Sigma}_n)$ *and is independent of* $\sigma^2$ *under* $P_M$. *Now note that the innermost conditional probability is a nondecreasing function of* $\sigma^2$, *which implies that*

$$\begin{aligned} &P_M\left(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n^B||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n\right) \\ &\qquad \geq E_M\left[P_M\left(\max_{1\leq i\leq p_n}|Z_i| > \sqrt{\frac{2(g_n+1)\epsilon^2 n^2}{g_n\lambda_{\min}\widetilde{\theta}_{0n}}} \,\middle|\, \sigma^2\right) I\left(\sigma^2 \geq \frac{\widetilde{\theta}_{0n}}{2n}\right) \,\middle|\, \hat{\boldsymbol{\beta}}_n, S_n\right] \\ &\qquad = P_M\left(\max_{1\leq i\leq p_n}|Z_i| > \sqrt{\frac{2(g_n+1)\epsilon^2 n^2}{g_n\lambda_{\min}\widetilde{\theta}_{0n}}}\right) P_M\left(\sigma^2 \geq \frac{\widetilde{\theta}_{0n}}{2n} \,\middle|\, \hat{\boldsymbol{\beta}}_n, S_n\right), \end{aligned}$$

*since the entries of* $\boldsymbol{\Sigma}_n$ *depend only on* $\boldsymbol{X}_n^T\boldsymbol{X}_n$. *Then Lemma 9 immediately implies that* $P_M(\sigma^2 \geq \widetilde{\theta}_{0n}/2n \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 1$ *a.s.*$(P_0)$, *so it suffices to show that the first* *term is bounded away from zero for all sufficiently large n. Now define* $\eta_{0n} := [2(g_n + 1)\epsilon^2 n^2 / g_n \lambda_{\min} \widetilde{\theta}_{0n}]^{1/2}$ *and* $\widetilde{\Sigma}_{n,i} := Var(Z_i \mid Z_{i+1}, \ldots, Z_{p_n})$*. Then*

$$P_M\left(\max_{1\le i\le p_n}|Z_i| \le \eta_{0n}\right) = E_M\left[P_M\left(|Z_1| \le \eta_{0n} \mid Z_2, Z_3, \ldots, Z_{p_n}\right)\prod_{i=2}^{p_n} I_{\{|Z_i|\le \eta_{0n}\}}\right]$$
$$\le \left[1 - 2\Phi\left(-\eta_{0n}/\sqrt{\widetilde{\Sigma}_{n,1}}\right)\right] E_M\left[\prod_{i=2}^{p_n} I_{\{|Z_i|\le\eta_{0n}\}}\right]$$

*by Lemma 7 and the fact that* $\widetilde{\Sigma}_{n,1}$ *does not depend on* $Z_2, Z_3, \ldots, Z_{p_n}$*. By repeated conditioning on* $Z_{i+1}, Z_{i+2}, \ldots, Z_{p_n}$ *for* $i = 2, 3, \ldots, p_n - 1$ *and application of Lemma 7 as above, we find that*

$$P_M\left(\max_{1\le i\le p_n}|Z_i| \le \eta_{0n}\right) \le \prod_{i=1}^{p_n}\left[1 - 2\Phi\left(-\eta_{0n}/\sqrt{\widetilde{\Sigma}_{n,i}}\right)\right].$$

*Note that*

$$\widetilde{\theta}_{0n} \ge \frac{n\breve{\lambda}_{0n}^{-1}}{g_n + 1}\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \ge \frac{\delta(g_n+1)n^2}{\lambda_{\max}\, g_n \log p_n},$$

*which implies that*

$$\eta_{0n} \le \sqrt{\frac{2\lambda_{\max}\epsilon^2 \log p_n}{\delta\lambda_{\min}}}.$$

*The eigenvalues of* $\boldsymbol{\Sigma}_n$ *are bounded below by* $\lambda_{\min}/\lambda_{\max}$*, so* $\inf_{n,i}\widetilde{\Sigma}_{n,i} \ge \lambda_{\min}/\lambda_{\max}$ *by Lemma 8. Then it follows that*

$$P_M\left(\max_{1\le i\le p_n}|Z_i| \le \eta_{0n}\right) \le \left[1 - 2\Phi\left(-\sqrt{\frac{2\lambda_{\max}^2\epsilon^2\log p_n}{\delta\lambda_{\min}^2}}\right)\right]^{p_n}$$
$$\le \exp\left[-2p_n\Phi\left(-\sqrt{\frac{2\lambda_{\max}^2\epsilon^2\log p_n}{\delta\lambda_{\min}^2}}\right)\right].$$

*Notice that if any subsequence of* $p_n$ *is bounded above, then the quantity*

$$\Phi[-(2\lambda_{\max}^2\epsilon^2\log p_n/\delta\lambda_{\min}^2)^{1/2}]$$

*is bounded away from zero along that subsequence, and thus posterior inconsistency follows immediately. So we may instead assume that* $p_n \to \infty$*. Then*

$$2\lambda_{\max}^2\epsilon^2\log p_n/\delta\lambda_{\min}^2 \to \infty,$$

*in which case the inequality*

$$1 - \Phi(t) \ge (t^{-1} - t^{-3})(2\pi)^{-1/2}\exp(-t^2/2) \ge (2t)^{-1}(2\pi)^{-1/2}\exp(-t^2/2)$$

*for large $t$ may be applied for all sufficiently large $n$, yielding*

$$P_M\left(\max_{1\le i\le p_n}|Z_i|>\eta_{0n}\right)\ge 1-\exp\left[-p_n\sqrt{\frac{\delta\lambda_{\min}^2}{4\pi\lambda_{\max}^2\epsilon^2\log p_n}}\;\exp\left(-\frac{\lambda_{\max}^2\epsilon^2\log p_n}{\delta\lambda_{\min}^2}\right)\right]$$
$$=1-\exp\left\{-\sqrt{\frac{\delta\lambda_{\min}^2}{4\pi\lambda_{\max}^2\epsilon^2\log p_n}}\;\exp\left[\left(1-\frac{\lambda_{\max}^2\epsilon^2}{\delta\lambda_{\min}^2}\right)\log p_n\right]\right\}$$
$$\to 1 \;\; \textit{for } \epsilon<\sqrt{\frac{\delta\lambda_{\min}^2}{\lambda_{\max}^2}}.$$

*Therefore posterior consistency does not occur.*

**Proof** (Proof of Lemma 10)**.** *Define $\delta:=\liminf_{n\to\infty}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2$, and assume $\delta>0$. Then*

$$\frac{T_n}{p_n}=\frac{T_n}{\theta_{0n}}\left(\sigma_0^2+\frac{n}{p_n\check{\lambda}_{0n}}\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\right)>\frac{T_n}{\theta_{0n}}\left(\sigma_0^2+\frac{\delta}{2\lambda_{\max}}\right)>\sigma_0^2+\frac{\delta}{4\lambda_{\max}}$$

*for all sufficiently large $n$ a.s.$(P_0)$, since $T_n/\theta_{0n}\to 1$ a.s.$(P_0)$ by Lemma 4. Then*

$$\liminf_{n\to\infty}\hat{g}_n^{EB}>\liminf_{n\to\infty}\left[\left(\frac{n-p_n+a-2}{S_n+b}\right)\left(\sigma_0^2+\frac{\delta}{4\lambda_{\max}}\right)-1\right]=\frac{\delta}{4\lambda_{\max}\sigma_0^2}>0 \quad a.s.(P_0)$$

*since $(n-p_n+a-2)/(S_n+b)\to 1/\sigma_0^2$ a.s.$(P_0)$ by Lemma 3.*

**Proof** (Proof of Theorem 2)**.** *By Theorem 1, we immediately have that posterior consistency occurs if and only if both*

$$\frac{||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty}{\hat{g}_n^{EB}+1}\to 0 \;\; \textit{and} \;\; \frac{\hat{g}_n^{EB}\log p_n}{(\hat{g}_n^{EB}+1)^2\,n}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\to 0 \;\; a.s.(P_0). \tag{9}$$

*We now consider three cases.*

*Case 1: Suppose there do not exist a subsequence $k_n$ of $n$ and a constant $A>0$ such that $||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_2^2\to A$ and $||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_\infty\nrightarrow 0$. Now let $k_n$ be a subsequence of $n$, and consider two sub-cases.*

*Case 1.1: Suppose $||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0n}||_\infty\to 0$. Then clearly the first condition in (9) is satisfied trivially. Note that for any further subsequence $m_n$ of $k_n$ for which $||\boldsymbol{\gamma}_{m_n}-\boldsymbol{\beta}_{0m_n}||_2^2\to 0$, the second condition in (9) is satisfied trivially as well, so we may instead assume $\liminf_{n\to\infty}||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_2^2>0$. Then for all sufficiently large $n$ a.s.$(P_0)$,*

$$\begin{aligned}&\frac{\hat{g}_{k_n}^{EB}\log p_{k_n}}{(\hat{g}_{k_n}^{EB}+1)^2\,k_n}||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_2^2\\ &\le\frac{\log k_n}{k_n}\left(\frac{S_{k_n}+b}{k_n-p_{k_n}+a-2}\right)\left(\frac{\theta_{0k_n}}{T_{k_n}}\right)\left(\frac{p_{k_n}||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_2^2}{\theta_{0k_n}}\right)\\ &\le\frac{\log k_n}{k_n}\left(\frac{S_{k_n}+b}{k_n-p_{k_n}+a-2}\right)\left(\frac{\theta_{0k_n}}{T_{k_n}}\right)\frac{p_{k_n}\lambda_{\max}}{k_n}\to 0 \quad a.s.(P_0)\end{aligned} \tag{10}$$

*by Lemmas 3, 4, and 10. Thus, both conditions in (9) hold along the subsequence $k_n$.*

*Case 1.2: Note that Case 1.1 can be applied to any further subsequence $m_n$ of $k_n$ for which $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_\infty \to 0$, so we may suppose for Case 1.2 that $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0$. Note also that in this case, there cannot exist any further subsequence $m_n$ of $k_n$ for which $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_2^2$ converges to a nonzero constant, since this would contradict the original supposition of Case 1. Then since $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \geq \liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty^2 > 0$, it follows that $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to \infty$. Then for all sufficiently large $n$ a.s.$(P_0)$,*

$$\begin{aligned}
&\frac{1}{\hat{g}_{k_n}^{EB}+1}||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \\
&= \left(\frac{S_{k_n}+b}{k_n - p_{k_n} + a - 2}\right)\left(\frac{\theta_{0k_n}}{T_{k_n}}\right)\left(\frac{p_{k_n}||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty}{\theta_{0k_n}}\right) \\
&\leq \left(\frac{S_{k_n}+b}{k_n - p_{k_n} + a - 2}\right)\left(\frac{\theta_{0k_n}}{T_{k_n}}\right)\left(\frac{p_{k_n}\lambda_{\max}}{k_n||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2}\right) \to 0 \quad a.s.(P_0)
\end{aligned} \tag{11}$$

*by Lemmas 3, 4, and 10, while (10) also holds by the same lemmas. Thus, both conditions hold along the subsequence $k_n$. Since Cases 1.1 and 1.2 together establish that both conditions hold along any subsequence $k_n$, they hold for the whole sequence, and therefore posterior consistency occurs.*

*Case 2: Now suppose there exist a subsequence $k_n$ of $n$ and a constant $A > 0$ such that $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to A > 0$ and $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \nrightarrow 0$, and suppose $\alpha = 0$. Note that Case 1.1 can be applied to any further subsequence $m_n$ of $k_n$ for which $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_\infty \to 0$, so we may suppose for Case 2 that $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0$. Then (10) and (11) still hold by Lemmas 3, 4, and 10 since $p_{k_n}/k_n \to 0$ and $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2 \geq \liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0$. Hence, the two conditions hold for every subsequence, and consequently for the overall sequence. Therefore posterior consistency occurs.*

*Case 3: Now suppose there exist a subsequence $k_n$ of $n$ and a constant $A > 0$ such that $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to A > 0$ and $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \nrightarrow 0$, but suppose $\alpha > 0$. As in Case 2, we may suppose for Case 3 that $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0$. Then for all sufficiently large $n$ a.s.$(P_0)$,*

$$\begin{aligned}
\frac{||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty}{\hat{g}_{k_n}^{EB}+1} &= \left(\frac{S_{k_n}+b}{k_n - p_{k_n} + a - 2}\right)\left(\frac{\theta_{0k_n}}{T_{k_n}}\right)\left(\frac{p_{k_n}||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty}{\theta_{0k_n}}\right) \\
&\geq \left(\frac{S_{k_n}+b}{k_n - p_{k_n} + a - 2}\right)\left(\frac{\theta_{0k_n}}{T_{k_n}}\right)\left(\frac{p_{k_n}\lambda_{\min}}{k_n||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2}\right) \\
&\qquad \times \liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \\
&\to \frac{\sigma_0^2\alpha\lambda_{\min}}{A}\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0 \qquad a.s.(P_0)
\end{aligned}$$

*by Lemmas 3, 4, and 10. The first condition fails for the subsequence $k_n$ and hence for the overall sequence. Therefore posterior consistency does not occur.*

**Proof** (Proof of Lemma 11)**.** *Assume that $n^{-3}\,T_n^2\,E_M[g^2(g+1)^{-4} \mid \hat{\boldsymbol{\beta}}_n, S_n] \to 0$ a.s.$(P_0)$. By Lemma 2, to determine whether posterior consistency occurs, it suffices to consider* *whether* $P_M(||\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 0$ *a.s.*$(P_0)$ *for every* $\epsilon > 0$*. By iterated expectation and the triangle inequality,*

$$\begin{aligned}
&E_M\left[P_M\left(\left|\left|\tilde{\boldsymbol{\beta}}_n^B(g) - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > 2\epsilon \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&- E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^B(g)\right|\right|_\infty > \epsilon \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\quad \leq E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\quad \leq E_M\left[P_M\left(\left|\left|\tilde{\boldsymbol{\beta}}_n^B(g) - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon/2 \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\qquad + E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^B(g)\right|\right|_\infty > \epsilon/2 \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right].
\end{aligned} \tag{12}$$

*Consider* $P_M(||\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^B(g)||_\infty > \epsilon \mid g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n)$ *for some arbitrary* $\epsilon > 0$ *and* $g \geq 0$*. Under* $P_M$*,*

$$\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^B(g) \;\Big|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n \sim N_{p_n}\left(\mathbf{0}, \frac{g}{g+1}\sigma^2\left(\boldsymbol{X}_n^T\boldsymbol{X}_n\right)^{-1}\right).$$

*Let* $v_{n,11}, \ldots, v_{n,p_np_n}$ *denote the diagonal elements of* $n(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1}$*, and write*

$$\begin{aligned}
&P_M\left(\left|\left|\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^B(g)\right|\right|_\infty > \epsilon \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\leq \sum_{i=1}^{p_n} P_M\left(\left|\beta_{n,i} - \tilde{\beta}_{n,i}^B(g)\right| > \epsilon \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\leq \sum_{i=1}^{p_n} P_M\left(\left[\beta_{n,i} - \tilde{\beta}_{n,i}^B(g)\right]^4 > \epsilon^4 \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\leq \sum_{i=1}^{p_n} \frac{3g^2\sigma^4 v_{n,ii}}{(g+1)^2 n^2\epsilon^4} \leq \frac{3\lambda_{\max} g^2\sigma^4}{(g+1)^2 n\epsilon^4}.
\end{aligned}$$

*Then*

$$\begin{aligned}
&E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^B(g)\right|\right|_\infty > \epsilon \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\leq \frac{3\lambda_{\max}}{n\epsilon^4} E_M\left(\frac{g^2\sigma^4}{(g+1)^2} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&= \frac{3\lambda_{\max}}{n\epsilon^4} E_M\left[\frac{g^2}{(g+1)^2} E_M\left(\sigma^4 \mid g, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right].
\end{aligned}$$

*Observe from the form of the posterior in (1) that under* $P_M$*,*

$$\sigma^2 \mid g, \hat{\boldsymbol{\beta}}_n, S_n \sim \mathit{InverseGamma}\left(\frac{n+a-2}{2}, \frac{S_n + b + (g_n+1)^{-1}T_n}{2}\right).$$

*Therefore,*

$$
\begin{aligned}
&E_M\left[P_M\left(\left|\left|\boldsymbol{\beta}_n - \tilde{\boldsymbol{\beta}}_n^B(g)\right|\right|_\infty > \epsilon \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&= \frac{3\lambda_{\max}}{n\epsilon^4} E_M\left[\frac{g^2\left[S_n + b + (g_n+1)^{-1}T_n\right]^2}{(g+1)^2(n+a-4)(n+a-6)} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\le \frac{6\lambda_{\max}}{n\epsilon^4}\left(\frac{S_n+b}{n+a-6}\right)^2 + \frac{6\lambda_{\max}}{n\epsilon^4}\frac{T_n^2}{(n+a-6)^2} E_M\left[\frac{g^2}{(g+1)^4} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \to 0 \quad a.s.(P_0)
\end{aligned}
$$

*by Lemma 3 and the initial assumption. Then this result and the inequalities in (12) imply that posterior consistency occurs if and only if*

$$
E_M\left[P_M\left(\left|\left|\tilde{\boldsymbol{\beta}}_n^B(g) - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon \;\middle|\; g, \sigma^2, \hat{\boldsymbol{\beta}}_n, S_n\right) \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \to 0 \;\; a.s.(P_0)
$$

*for every* $\epsilon > 0$*. Since* $\tilde{\boldsymbol{\beta}}_n^B(g) - \hat{\boldsymbol{\beta}}_n = (g+1)^{-1}(\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n)$*, we may equivalently state that posterior consistency occurs if and only if* $P_M[(g+1)^{-1}||\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n) \to 0 \;\; a.s.(P_0)$ *for every* $\epsilon > 0$*. But again by the triangle inequality,*

$$
\begin{aligned}
&P_M\left(\frac{1}{g+1}\left|\left|\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}\right|\right|_\infty > 2\epsilon \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) - P_M\left(\frac{1}{g+1}\left|\left|\boldsymbol{\beta}_{0n} - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\le P_M\left(\frac{1}{g+1}\left|\left|\boldsymbol{\gamma}_n - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) \qquad (13) \\
&\le P_M\left(\frac{1}{g+1}\left|\left|\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}\right|\right|_\infty > \epsilon/2 \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\qquad + P_M\left(\frac{1}{g+1}\left|\left|\boldsymbol{\beta}_{0n} - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon/2 \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right).
\end{aligned}
$$

*For any arbitrary* $\epsilon > 0$*,*

$$
\begin{aligned}
P_M\left(\frac{1}{g+1}\left|\left|\boldsymbol{\beta}_{0n} - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) &\le P_M\left(\left|\left|\boldsymbol{\beta}_{0n} - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&= I\left(\left|\left|\boldsymbol{\beta}_{0n} - \hat{\boldsymbol{\beta}}_n\right|\right|_\infty > \epsilon\right) \to 0 \;\; a.s.(P_0)
\end{aligned}
$$

*by Lemma 1, where* $I(\cdot)$ *denotes the usual indicator function. Then this result and (13) together imply that posterior consistency occurs if and only if* $P_M[(g+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n] \to 0 \;\; a.s.(P_0)$ *for every* $\epsilon > 0$*.*

**Proof** (Proof of Lemma 12)**.** *From the form of the posterior in (3) and the transformation in (4),*

$$
\begin{aligned}
&\frac{T_n^2}{n^3} E_M\left[\frac{g^2}{(g+1)^4} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\le \frac{T_n^2}{n^3} E_M\left[\frac{1}{(g+1)^2} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right]
\end{aligned}
$$

$$= \frac{T_n^2 \displaystyle\int_0^\infty (g+1)^{(n-p_n+a-c-6)/2} \left[(g+1)(S_n+b)+T_n\right]^{-(n+a-2)/2} \, dg}{n^3 \displaystyle\int_0^\infty (g+1)^{(n-p_n+a-c-2)/2} \left[(g+1)(S_n+b)+T_n\right]^{-(n+a-2)/2} \, dg}$$

$$= \frac{(S_n+b)^2 \displaystyle\int_{W_n}^1 u^{(n-p_n+a-c-6)/2}(1-u)^{(p_n+c)/2} \, du}{n^3 \displaystyle\int_{W_n}^1 u^{(n-p_n+a-c-2)/2}(1-u)^{(p_n+c-4)/2} \, du}.$$

*Now let* $H_n \sim Beta((n-p_n+a-c-4)/2,\ (p_n+c-2)/2)$ *and* $\widetilde{H}_n \sim Beta((n-p_n+a-c)/2,\ (p_n+c-2)/2)$ *with both independent of* $\hat{\boldsymbol{\beta}}_n$ *and* $S_n$ *under* $P_M$*, and observe that* $H_n$ *is stochastically smaller than* $\widetilde{H}_n$ *under* $P_M$*. Also let* $\Gamma(\cdot)$ *denote the usual gamma function. Continuing, we have that*

$$\frac{T_n^2}{n^3}\, E_M\left[\frac{g^2}{(g+1)^4} \,\middle|\, \hat{\boldsymbol{\beta}}_n, S_n\right]$$

$$= \frac{(S_n+b)^2\, \Gamma\left(\frac{n-p_n+a-c-4}{2}\right)\Gamma\left(\frac{p_n+c+2}{2}\right) P_M\left(H_n > W_n \mid \hat{\boldsymbol{\beta}}_n, S_n\right)}{n^3\, \Gamma\left(\frac{n-p_n+a-c}{2}\right)\Gamma\left(\frac{p_n+c-2}{2}\right) P_M\left(\widetilde{H}_n > W_n \mid \hat{\boldsymbol{\beta}}_n, S_n\right)}$$

$$\leq \frac{1}{n}\left(\frac{S_n+b}{n}\right)^2 \frac{(p_n+c)(p_n+c-2)}{(n-p_n+a-c-2)(n-p_n+a-c-4)} \to 0 \quad a.s.(P_0)$$

*by Lemma 3.*

**Proof** (Proof of Lemma 13)**.** *Assume* $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \geq \delta$ *for some* $\delta > 0$*. Then*

$$\limsup_{n\to\infty} W_n$$

$$= \limsup_{n\to\infty}\left(1+\frac{T_n}{S_n+b}\right)^{-1}$$

$$\leq \left(1 + \liminf_{n\to\infty}\left[\frac{p_n\sigma_0^2 + n\breve{\lambda}_{0n}^{-1}\, ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2}{n-p_n}\right] \liminf_{n\to\infty}\left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-1}$$

$$\leq \left(1 + \frac{\alpha}{1-\alpha} + \frac{\delta}{\lambda_{\max}(1-\alpha)\sigma_0^2}\right)^{-1} = \frac{(1-\alpha)\lambda_{\max}\sigma_0^2}{\delta+\lambda_{\max}\sigma_0^2} < 1-\alpha \quad a.s.(P_0)$$

*by Lemmas 3 and 4.*

**Proof** (Proof of Lemma 14)**.** *Assume* $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to \infty$*, and let* $\epsilon > 0$*. Then*

$$W_n = \left(1+\frac{T_n}{S_n+b}\right)^{-1}$$

$$= \left(1 + \frac{p_n\sigma_0^2 + n\breve{\lambda}_{0n}^{-1}\, ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2}{n-p_n}\left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-1} \to 0 \quad a.s.(P_0)$$

*since the term in square brackets converges to* $1/\sigma_0^2$ *a.s.*$(P_0)$ *by Lemmas 3 and 4. This establishes (i). Now write*

$$\frac{\epsilon^{-1}\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty\,(S_n+b)}{\epsilon^{-1}\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty\,(S_n+b)+T_n}=\left(1+\frac{\epsilon\,T_n}{||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty\,(S_n+b)}\right)^{-1}$$

*and observe that*

$$\begin{aligned}&\left(1+\frac{\epsilon\,T_n}{||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty\,(S_n+b)}\right)^{-1}\\ &=\left(1+\frac{\epsilon\left(p_n\sigma_0^2+n\breve{\lambda}_{0n}^{-1}\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\right)}{(n-p_n)\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty}\left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-1}\\ &\le\left(1+\frac{\epsilon\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2}{\lambda_{\max}}\left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-1}\to 0\quad a.s.(P_0)\end{aligned}$$

*since, once again, the term in square brackets converges to* $1/\sigma_0^2$ *a.s.*$(P_0)$ *by Lemmas 3 and 4. It then follows immediately that* $L_n(\epsilon)\to 0$ *a.s.*$(P_0)$*, establishing (ii).*

**Proof** (Proof of Lemma 15)**.** *Assume that* $||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\to A>0$ *and* $\liminf_{n\to\infty}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty>0$*. Let* $\epsilon>0$*. Then* $\limsup_{n\to\infty}W_n\le(1-\alpha)\lambda_{\max}\sigma_0^2/(A+\lambda_{\max}\sigma_0^2)<1$ *a.s.*$(P_0)$ *by Lemma 13, and*

$$\begin{aligned}\limsup_{n\to\infty}\widetilde{L}_n(\epsilon)&=\limsup_{n\to\infty}\left(1+\frac{\epsilon\,T_n}{||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty\,(S_n+b)}\right)^{-1}\\ &=\limsup_{n\to\infty}\left(1+\frac{\epsilon\left(p_n\sigma_0^2+n\breve{\lambda}_{0n}^{-1}\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\right)}{(n-p_n)\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty}\left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-1}\\ &\le\limsup_{n\to\infty}\left(1+\frac{\epsilon\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2}{\lambda_{\max}}\left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-1}\\ &=\left(1+\frac{\epsilon\sqrt{A}}{\lambda_{\max}\sigma_0^2}\right)^{-1}=\frac{\lambda_{\max}\sigma_0^2}{A^{1/2}\epsilon+\lambda_{\max}\sigma_0^2}<1\quad a.s.(P_0)\end{aligned}$$

*since the term in square brackets converges to* $1/\sigma_0^2$ *a.s.*$(P_0)$ *by Lemmas 3 and 4. Define*

$$L^\star(\epsilon)=\max\left\{\frac{(1-\alpha)\lambda_{\max}\sigma_0^2}{A+\lambda_{\max}\sigma_0^2},\ \frac{\lambda_{\max}\sigma_0^2}{A^{1/2}\epsilon+\lambda_{\max}\sigma_0^2}\right\}<1,$$

*and observe that* $\limsup_{n\to\infty}L_n(\epsilon)\le L^\star(\epsilon)$ *a.s.*$(P_0)$*. This establishes (i).*

*Now define* $\widetilde{A}:=\liminf_{n\to\infty}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty>0$*, and note that*

$$\begin{aligned}&\liminf_{n\to\infty}L_n(\epsilon)\\ &\ge\liminf_{n\to\infty}\left(1+\frac{\epsilon\,T_n}{||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty\,(S_n+b)}\right)^{-1}\end{aligned}$$

$$\geq \left(1+\limsup_{n\to\infty}\left[\frac{\epsilon\left(p_n\sigma_0^2+n\breve{\lambda}_{0n}^{-1}\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_2^2\right)}{(n-p_n)\,||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty}\right]\limsup_{n\to\infty}\left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-1},$$

*which implies that*

$$\liminf_{n\to\infty} L_n(\epsilon)\geq\left(1+\frac{\epsilon\left(\alpha\sigma_0^2+A/\lambda_{\min}\right)}{(1-\alpha)\widetilde{A}\sigma_0^2}\right)^{-1}\quad a.s.(P_0)$$

*by Lemmas 3 and 4. Then it can be seen that for any $\zeta<1$, there exists $\epsilon_\zeta>0$ such that $\liminf_{n\to\infty} L_n(\epsilon_\zeta)>\zeta$ $a.s.(P_0)$, establishing (ii).*

**Proof** (Proof of Lemma 16)**.** *Let $\epsilon>0$. Note that $E(Z_n)=a_n/(a_n+b_n)\to 1-\alpha$, and thus $|a_n/(a_n+b_n)-(1-\alpha)|\leq\epsilon/2$ for all sufficiently large $n$. Also note that $Var(Z_n)=a_nb_n/[(a_n+b_n)^2(a_n+b_n+1)]\leq 1/a_n<2/[n(1-\alpha)]$ for all sufficiently large $n$. Then for all sufficiently large $n$,*

$$\begin{aligned}&P\left(1-\alpha-\epsilon\leq Z_n\leq 1-\alpha+\epsilon\right)\\&\quad=P\left(1-\alpha-\frac{a_n}{a_n+b_n}-\epsilon\leq Z_n-\frac{a_n}{a_n+b_n}\leq 1-\alpha-\frac{a_n}{a_n+b_n}+\epsilon\right)\\&\quad\geq P\left(-\frac{\epsilon}{2}\leq Z_n-\frac{a_n}{a_n+b_n}\leq\frac{\epsilon}{2}\right)\geq 1-\frac{4}{\epsilon^2}\,Var(Z_n)\geq 1-\frac{8}{n(1-\alpha)\epsilon^2}\to 1,\end{aligned}$$

*where the second of the three inequalities is Chebyshev's inequality.*

**Proof** (Proof of Theorem 3)**.** *By Lemmas 11 and 12, posterior consistency occurs if and only if $P_M[(g+1)^{-1}||\boldsymbol{\gamma}_n-\boldsymbol{\beta}_{0n}||_\infty>\epsilon\mid\hat{\boldsymbol{\beta}}_n,S_n]\to 0$ $a.s.(P_0)$ for every $\epsilon>0$, which by (5) occurs if and only if $P_M[W_n<U_n<L_n(\epsilon)\mid\hat{\boldsymbol{\beta}}_n,S_n]/P_M(U_n>W_n\mid\hat{\boldsymbol{\beta}}_n,S_n)\to 0$ $a.s.(P_0)$ for every $\epsilon>0$. We now consider the same three cases as in the proof of Theorem 2.*

*Case 1: Suppose there do not exist a subsequence $k_n$ of $n$ and a constant $A>0$ such that $||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_2^2\to A$ and $||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_\infty\nrightarrow 0$. Let $k_n$ be a subsequence of $n$, and let $\epsilon>0$. Now consider two sub-cases.*

*Case 1.1: Suppose $||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_\infty\to 0$. Then $\epsilon^{-1}||\boldsymbol{\gamma}_{k_n}-\boldsymbol{\beta}_{0k_n}||_\infty<1$ for all sufficiently large $n$. This implies that $L_{k_n}(\epsilon)=W_{k_n}$ for all sufficiently large $n$ $a.s.(P_0)$, and therefore$P_M[W_{k_n}<U_{k_n}<L_{k_n}(\epsilon)\mid\hat{\boldsymbol{\beta}}_n,S_n]=0$ for all sufficiently large $n$ $a.s.(P_0)$. Also, $P_M(U_{k_n}>W_{k_n}\mid\hat{\boldsymbol{\beta}}_{k_n},S_{k_n})>0$ for all $n$ $a.s.(P_0)$ since $W_{k_n}<1$ for all $n$ $a.s.(P_0)$. Thus,*

$$\frac{P_M\left[W_{k_n}<U_{k_n}<L_{k_n}(\epsilon)\mid\hat{\boldsymbol{\beta}}_{k_n},S_{k_n}\right]}{P_M(U_{k_n}>W_{k_n}\mid\hat{\boldsymbol{\beta}}_n,S_n)}\to 0\quad a.s.(P_0)$$

*by the combination of our results for its numerator and denominator.*

*Case 1.2: Note that Case 1.1 can be applied to any further subsequence $m_n$ of $k_n$ for which $||\boldsymbol{\gamma}_{m_n}-\boldsymbol{\beta}_{0m_n}||_\infty\to 0$, so we may suppose for Case 1.2 that*

$\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0$. *Note also that in this case, there cannot exist any further subsequence* $m_n$ *of* $k_n$ *for which* $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_2^2$ *converges to a nonzero constant, since this would contradict the original supposition of Case 1. Then since* $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \geq \liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty^2 > 0$, *it follows that* $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to \infty$. *Then Lemma 14 implies that both* $W_{k_n} \to 0$ *a.s.*$(P_0)$ *and* $L_n(\epsilon) \to 0$ *a.s.*$(P_0)$, *which in turn implies that both* $W_{k_n} < (1-\alpha)/2$ *and* $L_{k_n}(\epsilon) < (1-\alpha)/2$ *for all sufficiently large* $n$ *a.s.*$(P_0)$. *Then for all sufficiently large* $n$ *a.s.*$(P_0)$,

$$\frac{P_M\left[W_{k_n} < U_{k_n} < L_{k_n}(\epsilon) \mid \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right]}{P_M\left(U_{k_n} > W_{k_n} \mid \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right)} \leq \frac{P_M\left[U_{k_n} < \frac{1-\alpha}{2} \,\middle|\, \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right]}{P_M\left[U_{k_n} > \frac{1-\alpha}{2} \,\middle|\, \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right]} \to 0 \quad a.s.(P_0)$$

*by Lemma 16. Finally, since Cases 1.1 and 1.2 together establish that the relevant condition holds along any subsequence* $k_n$, *it holds for the whole sequence, and therefore posterior consistency occurs.*

*Case 2: Now suppose there exist a subsequence* $k_n$ *of* $n$ *and a constant* $A > 0$ *such that* $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to A > 0$ *and* $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \nrightarrow 0$, *and suppose* $\alpha = 0$. *Note that Case 1 can be applied to any subsequence* $m_n$ *of* $n$ *for which either* $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_2^2$ *does not converge to any nonzero constant or* $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_\infty \to 0$, *so it suffices to show that the relevant condition holds along the subsequence* $k_n$. *Note also that this means we may suppose for Case 2 that* $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0$. *Now let* $\epsilon > 0$. *By Lemma 13,* $\limsup_{n\to\infty} W_{k_n} \leq \lambda_{\max}\sigma_0^2/(A + \lambda_{\max}\sigma_0^2)$ *a.s.*$(P_0)$, *which implies that* $W_{k_n} < 2\lambda_{\max}\sigma_0^2/(A + 2\lambda_{\max}\sigma_0^2)$ *for all sufficiently large* $n$ *a.s.*$(P_0)$. *Moreover, by Lemma 15, there exists* $L^\star(\epsilon) < 1$ *such that* $\limsup_{n\to\infty} L_{k_n}(\epsilon) \leq L^\star(\epsilon)$ *a.s.*$(P_0)$, *which implies that* $L_{k_n}(\epsilon) < [1 + L^\star(\epsilon)]/2$ *for all sufficiently large* $n$ *a.s.*$(P_0)$. *Then for all sufficiently large* $n$ *a.s.*$(P_0)$,

$$\begin{aligned} &\frac{P_M\left[W_{k_n} < U_{k_n} < L_{k_n}(\epsilon) \mid \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right]}{P_M\left(U_{k_n} > W_{k_n} \mid \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right)} \\ &\leq \frac{P_M\left[U_{k_n} < \frac{1 + L^\star(\epsilon)}{2} \,\middle|\, \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right]}{P_M\left(U_{k_n} > \frac{2\lambda_{\max}\sigma_0^2}{A + 2\lambda_{\max}\sigma_0^2} \,\middle|\, \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right)} \to 0 \quad a.s.(P_0) \end{aligned}$$

*by Lemma 16. Therefore posterior consistency occurs.*

*Case 3: Now suppose there exist a subsequence* $k_n$ *of* $n$ *and a constant* $A > 0$ *such that* $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to A > 0$ *and* $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \nrightarrow 0$, *but suppose* $\alpha > 0$. *By Lemma 13,* $\limsup_{n\to\infty} W_{k_n} \leq (1-\alpha)\lambda_{\max}\sigma_0^2/(A + \lambda_{\max}\sigma_0^2)$ *a.s.*$(P_0)$, *which implies that* $W_{k_n} < 2(1-\alpha)\lambda_{\max}\sigma_0^2/(A + 2\lambda_{\max}\sigma_0^2)$ *for all sufficiently large* $n$ *a.s.*$(P_0)$. *By Lemma 15, there exists* $\epsilon_{1-\alpha/4} > 0$ *such that* $\liminf_{n\to\infty} L_{k_n}(\epsilon_{1-\alpha/4}) \geq 1 - \alpha/4$ *a.s.*$(P_0)$, *which implies that* $L_{k_n}(\epsilon_{1-\alpha/4}) > 1 - \alpha/2$ *for all sufficiently large* $n$ *a.s.*$(P_0)$. *Then for all sufficiently* *large $n$ a.s.$(P_0)$,*

$$\begin{aligned}
&\frac{P_M\left[W_{k_n} < U_{k_n} < L_{k_n}(\epsilon_{1-\alpha/4}) \;\middle|\; \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right]}{P_M\left(U_{k_n} > W_{k_n} \mid \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right)} \\
&\geq P_M\left[W_{k_n} < U_{k_n} < L_{k_n}(\epsilon_{1-\alpha/4}) \;\middle|\; \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right] \\
&\geq P_M\left[\frac{2(1-\alpha)\lambda_{\max}\sigma_0^2}{A+2\lambda_{\max}\sigma_0^2} < U_{k_n} < 1-\frac{\alpha}{2} \;\middle|\; \hat{\boldsymbol{\beta}}_{k_n}, S_{k_n}\right] \\
&\to 1 \quad a.s.(P_0)
\end{aligned}$$

*by Lemma 16. Since the relevant condition fails to hold for the subsequence $k_n$, it fails to hold for the overall sequence. Therefore posterior consistency does not occur.*

**Proof** (Proof of Lemma 17)**.** *Consider two cases.*

*Case 1: Suppose $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to 0$. Then $\theta_{0n} = p_n\sigma_0^2 + n\check{\lambda}_{0n}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \leq n\sigma_0^2$ for all sufficiently large $n$. This result and (8) imply that*

$$(\mu_4)_0\,(T_n) := E_0\left[(T_n-\theta_{0n})^4\right] \leq 48\sigma_0^4\theta_{0n}^2 + 192\sigma_0^6\theta_{0n} \leq 96n^2\sigma_0^8$$

*for all sufficiently large $n$. Then there exists $N$ such that*

$$\sum_{n=N}^{\infty} P_0\left(T_n > 2n\sigma_0^2\right) \leq \sum_{n=N}^{\infty} P_0\left(|T_n-\theta_{0n}| > n\sigma_0^2\right) \leq \sum_{n=N}^{\infty}\frac{96n^2\sigma_0^8}{n^4\sigma_0^8} = 96\sum_{n=N}^{\infty}\frac{1}{n^2} < \infty$$

*by Markov's inequality applied to $(T_n-\theta_{0n})^4$, which in turn implies by the Borel-Cantelli lemma that $\limsup_{n\to\infty}(T_n/n) \leq 2\sigma_0^2$ a.s.$(P_0)$. Therefore, $n^{-3}\,T_n^2\,E_M[g^2(g+1)^{-4} \mid \hat{\boldsymbol{\beta}}_n, S_n] \leq n^{-3}\,T_n^2 \to 0$ a.s.$(P_0)$.*

*Case 2: Note immediately that Case 1 can be applied to any subsequence $k_n$ of $n$ for which $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to 0$, so we may suppose for Case 2 that $\liminf_{n\to\infty}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 > 0$. Then $\limsup_{n\to\infty} W_n < 1-\alpha$ by Lemma 13. Define $\psi_n(u) := I_{(W_n,1)}(u)\exp[-nW_n(1-u)/2(u-W_n)]$, where $I$ denotes the usual indicator function, and note that this is a nondecreasing function of $u$ on the interval $(0,1)$. Using the form of the posterior in (6) and the transformation in (4), we may write*

$$\begin{aligned}
&\frac{T_n^2}{n^3}E_M\left[\frac{g^2}{(g+1)^4} \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\quad= \frac{T_n^2\displaystyle\int_0^\infty \frac{(g+1)^{(n-p_n+a-10)/2}}{\left[(g+1)(S_n+b)+T_n\right]^{-(n+a-2)/2}}\,g^{1/2}\exp\left(-\frac{n}{2g}\right)\,dg}{n^3\displaystyle\int_0^\infty \frac{(g+1)^{(n-p_n+a-2)/2}}{\left[(g+1)(S_n+b)+T_n\right]^{-(n+a-2)/2}}\,g^{-3/2}\exp\left(-\frac{n}{2g}\right)\,dg} \\
&\quad= \frac{(S_n+b)^4\displaystyle\int_0^1 u^{(n-p_n+a-c-10)/2}(1-u)^{(p_n+4)/2}\left[\frac{u-W_n}{W_n(1-u)}\right]^{1/2}\psi_n(u)\,du}{n^3\,T_n^2\displaystyle\int_0^1 u^{(n-p_n+a-c-2)/2}(1-u)^{(p_n-4)/2}\left[\frac{u-W_n}{W_n(1-u)}\right]^{-3/2}\psi_n(u)\,du}
\end{aligned}$$

$$\leq \frac{(S_n + b)^2 \int_0^1 u^{(n-p_n+a-c-9)/2}(1-u)^{(p_n+3)/2}\ \psi_n(u)\ du}{n^3(1-W_n)^2 \int_0^1 u^{(n-p_n+a-c-5)/2}(1-u)^{(p_n-1)/2}\ \psi_n(u)\ du}.$$

*Now let $h_n$ and $\widetilde{h}_n$ denote the densities with respect to Lebesgue measure of Beta($(n - p_n + a - 7)/2,\ (p_n + 5)/2$) and Beta($(n - p_n + a - 3)/2,\ (p_n + 1)/2$) random variables, respectively. Then we may continue by writing*

$$\begin{aligned}
&\frac{T_n^2}{n^3} E_M\left[\frac{g^2}{(g+1)^4}\ \middle|\ \hat{\boldsymbol{\beta}}_n, S_n\right] \\
&\leq \frac{(S_n+b)^2\ \Gamma\left(\dfrac{n-p_n+a-7}{2}\right)\Gamma\left(\dfrac{p_n+5}{2}\right)\ \int_0^1 h_n(u)\ \psi_n(u)\ du}{n^3(1-W_n)^2\ \Gamma\left(\dfrac{n-p_n+a-3}{2}\right)\Gamma\left(\dfrac{p_n+1}{2}\right)\ \int_0^1 \widetilde{h}_n(u)\ \psi_n(u)\ du} \\
&\leq \frac{(S_n+b)^2\ (p_n+3)(p_n+1)}{n^3(1-W_n)^2\ (n-p_n+a-5)(n-p_n+a-7)} \to 0 \quad a.s.(P_0).
\end{aligned}$$

*Note that the last inequality holds because a random variable with density $h_n$ is stochastically smaller than a random variable with density $\widetilde{h}_n$ and because $\psi_n$ is nondecreasing on $(0,1)$, while the almost sure convergence to zero is by Lemma 3 and the fact that $\limsup_{n\to\infty} W_n < 1 - \alpha \leq 1$ a.s.($P_0$) by Lemma 13.*

**Proof** (Proof of Lemma 18)**.** *Note immediately that both (i) and (ii) are trivial if $\xi = 0$ or $\xi \geq 1$, so assume $0 < \xi < 1$. Next, by Stirling's approximation, we may bound the normalizing constant by*

$$\log \frac{\Gamma(a_n+b_n)}{\Gamma(a_n)\Gamma(b_n)} \leq \log \frac{(a_n+b_n)^{a_n+b_n-1/2}}{a_n^{a_n-1/2}\ b_n^{b_n-1/2}}$$

*for all sufficiently large $n$. We may rewrite this as*

$$\log \frac{\Gamma(a_n+b_n)}{\Gamma(a_n)\Gamma(b_n)} \leq a_n \log\left(\frac{a_n+b_n}{a_n}\right) + b_n \log\left(\frac{a_n+b_n}{b_n}\right) + \frac{1}{2}\log\left(\frac{a_n b_n}{a_n+b_n}\right)$$

*for all sufficiently large $n$. Then*

$$\begin{aligned}
&P(Z_n \leq \xi) \\
&= \frac{\Gamma(a_n+b_n)}{\Gamma(a_n)\Gamma(b_n)} \int_0^\xi z^{a_n-1}(1-z)^{b_n-1}\ dz \\
&\leq \frac{\Gamma(a_n+b_n)}{\Gamma(a_n)\Gamma(b_n)} \int_0^\xi z^{a_n-1}\ dz = \frac{\Gamma(a_n+b_n)\ \xi^{a_n}}{\Gamma(a_n)\Gamma(b_n)\ a_n} \\
&\leq \exp\left[a_n \log\xi - \log a_n + a_n \log\left(\frac{a_n+b_n}{a_n}\right) + b_n \log\left(\frac{a_n+b_n}{b_n}\right) + \frac{1}{2}\log\left(\frac{a_n b_n}{a_n+b_n}\right)\right]
\end{aligned}$$

*for all sufficiently large $n$. Now observe that*

$$\begin{aligned}
&\frac{1}{n}\log P(Z_n \le \xi) \\
&\le \frac{a_n}{n}\log\xi - \frac{1}{n}\log a_n + \frac{a_n}{n}\log\left(\frac{a_n+b_n}{a_n}\right) + \frac{b_n}{n}\log\left(\frac{a_n+b_n}{b_n}\right) + \frac{1}{2n}\log\left(\frac{a_n b_n}{a_n+b_n}\right) \\
&\to \begin{cases} (1-\alpha)\log\xi - (1-\alpha)\log(1-\alpha) - \alpha\log\alpha & \textit{if } \alpha > 0, \\ \log\xi & \textit{if } \alpha = 0. \end{cases}
\end{aligned}$$

*If $\alpha > 0$, then $(1-\alpha)\log(1-\alpha)+\alpha\log\alpha \ge -\log 2$, and thus $\limsup_{n\to\infty} n^{-1}\log P(Z_n \le \xi) \le (1-\alpha)\log\xi + \log 2$. Then $n^{-1}\log P(Z_n \le \xi) \le (1-\alpha)\log\xi + \log 4$ for all sufficiently large $n$, which implies (i). If instead $\alpha = 0$, then $\limsup_{n\to\infty} n^{-1}\log P(Z_n \le \xi) \le \log\xi$, so $n^{-1}\log P(Z_n \le \xi) \le \frac{1}{2}\log\xi$ for all sufficiently large $n$ (noting that $\log\xi < 0$). This implies (ii).*

**Proof** (Proof of Lemma 19)**.** *Let $\delta = \liminf_{n\to\infty} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 > 0$. Then by Lemma 13, $\limsup_{n\to\infty} W_n \le (1-\alpha)\lambda_{\max}\sigma_0^2/(\delta + \lambda_{\max}\sigma_0^2)$ a.s.$(P_0)$, which implies that $W_n < 2(1-\alpha)\lambda_{\max}\sigma_0^2/(\delta + 2\lambda_{\max}\sigma_0^2) < 1-\alpha$ for all sufficiently large $n$ a.s.$(P_0)$. Then for all sufficiently large $n$ a.s.$(P_0)$,*

$$\begin{aligned}
R_n &\ge \int_{(1-\alpha+W_n)/2}^{1} f_n(u)\left[\frac{u-W_n}{(1-u)}\right]^{-3/2} \exp\left[-\frac{nW_n(1-u)}{2(u-W_n)}\right]\, du \\
&\ge \exp\left(-\frac{nW_n}{1-\alpha-W_n}\right)\int_{(1-\alpha+W_n)/2}^{1} f_n(u)\left[\frac{u}{(1-u)}\right]^{-3/2} du \\
&= \frac{\Gamma\left(\frac{n-p_n+a-3}{2}\right)\Gamma\left(\frac{p_n+1}{2}\right)}{\Gamma\left(\frac{n-p_n+a}{2}\right)\Gamma\left(\frac{p_n-2}{2}\right)} \exp\left(-\frac{nW_n}{1-\alpha-W_n}\right) \\
&\quad \times P_M\left(\frac{1-\alpha+W_n}{2} < \widetilde{U}_n < 1 \,\middle|\, \hat{\boldsymbol{\beta}}_n, S_n\right) \\
&\ge \frac{\Gamma\left(\frac{n-p_n+a-3}{2}\right)\Gamma\left(\frac{p_n+1}{2}\right)}{\Gamma\left(\frac{n-p_n+a}{2}\right)\Gamma\left(\frac{p_n-2}{2}\right)} \exp\left(-\frac{nW_n}{1-\alpha-W_n}\right) \\
&\quad \times P_M\left[\left(\frac{\delta+4\lambda_{\max}\sigma_0^2}{2\delta+4\lambda_{\max}\sigma_0^2}\right)(1-\alpha) < \widetilde{U}_n < 1\right] \qquad (14)
\end{aligned}$$

*where $\widetilde{U}_n \sim$ Beta$((n-p_n+a-3)/2, (p_n+1)/2)$, independent of $\hat{\boldsymbol{\beta}}_n$ and $S_n$, under $P_M$. For all sufficiently large $n$, Stirling's approximation yields that*

$$
\begin{aligned}
&\frac{\Gamma\left(\frac{n-p_n+a-3}{2}\right)\Gamma\left(\frac{p_n+1}{2}\right)}{\Gamma\left(\frac{n-p_n+a}{2}\right)\Gamma\left(\frac{p_n-2}{2}\right)} \\
&\geq \frac{\left(\frac{n-p_n+a-3}{2}\right)^{(n-p_n+a-4)/2}\exp\left(-\frac{n-p_n+a-3}{2}\right)}{2\left(\frac{n-p_n+a}{2}\right)^{(n-p_n+a-1)/2}\exp\left(-\frac{n-p_n+a}{2}\right)} \\
&\quad\times\frac{\left(\frac{p_n+1}{2}\right)^{p_n/2}\exp\left(-\frac{p_n+1}{2}\right)}{\left(\frac{p_n-2}{2}\right)^{(p_n-3)/2}\exp\left(-\frac{p_n-2}{2}\right)} \\
&= \frac{1}{2}\left(\frac{n-p_n+a-3}{n-p_n+a}\right)^{(n-p_n+a-4)/2}\left(\frac{p_n+1}{p_n-2}\right)^{(p_n-3)/2}\left(\frac{p_n+1}{n-p_n+a}\right)^{3/2}.
\end{aligned}
$$

*Then for all sufficiently large $n$,*

$$
\frac{\Gamma\left(\frac{n-p_n+a-3}{2}\right)\Gamma\left(\frac{p_n+1}{2}\right)}{\Gamma\left(\frac{n-p_n+a}{2}\right)\Gamma\left(\frac{p_n-2}{2}\right)} \geq \frac{1}{4}\left(\frac{p_n+1}{n-p_n+a}\right)^{3/2} \geq 2\,(4n)^{-3/2}. \tag{15}
$$

*Now observe that*

$$
P_M\left[\left(\frac{\delta+4\lambda_{\max}\sigma_0^2}{2\delta+4\lambda_{\max}\sigma_0^2}\right)(1-\alpha) < \widetilde{U}_n < 1\right] \to 1
$$

*by Lemma 16, which implies that*

$$
P_M\left[\left(\frac{\delta+4\lambda_{\max}\sigma_0^2}{2\delta+4\lambda_{\max}\sigma_0^2}\right)(1-\alpha) < \widetilde{U}_n < 1\right] > \frac{1}{2} \tag{16}
$$

*for all sufficiently large $n$. Then by combining Inequalities 14, 15, and 16, we have that for all sufficiently large $n$ a.s.$(P_0)$,*

$$
R_n \geq (4n)^{-3/2}\exp\left(-\frac{nW_n}{1-\alpha-W_n}\right) = \exp\left\{-n\left[\frac{W_n}{1-\alpha-W_n}+\frac{3}{2n}\log(4n)\right]\right\}.
$$

*Finally, take $K = 2\limsup_{n\to\infty}[W_n/(1-\alpha-W_n)]$. Observe that $K<\infty$ a.s.$(P_0)$ due to the fact that $\limsup_{n\to\infty} W_n \leq (1-\alpha)\lambda_{\max}\sigma_0^2/(\delta+\lambda_{\max}\sigma_0^2) < 1-\alpha$ a.s.$(P_0)$. Then $R_n \geq \exp(-nK)$ for all sufficiently large $n$ a.s.$(P_0)$.*

**Proof** (Proof of Lemma 20)**.** *Let $\epsilon > 0$, and assume $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to \infty$. Then by Lemma 14, $W_n \to 0$ a.s.$(P_0)$ and $L_n(\epsilon) \to 0$ a.s.$(P_0)$. Next, observe that the last two terms of the integrand in $Q_n(\epsilon)$ comprise an unnormalized InverseGamma$(1/2,\ nW_n/2)$*

*density in $(u-W_n)/(1-u)$, the mode of which occurs at $nW_n/3$. Then for all sufficiently large $n$,*

$$Q_n(\epsilon) \leq \int_{W_n}^{L_n(\epsilon)} f_n(u) \left(\frac{nW_n}{3}\right)^{-3/2} \exp\left(-\frac{3}{2}\right)\, du \leq 2\,(nW_n)^{-3/2} \int_0^{L_n(\epsilon)} f_n(u)\, du$$
$$\leq 2^{2n+1}\,(nW_n)^{-3/2}\,[L_n(\epsilon)]^{n(1-\alpha)}$$

*by Lemma 18. Now note that if $\epsilon^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty \leq 1$, then $L_n(\epsilon) = W_n$, in which case $Q_n(\epsilon) = 0$ and the result is trivial. So instead assume that $\epsilon^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty > 1$, which in turn implies that $L_n(\epsilon) \leq \epsilon^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty W_n$. Then*

$$\begin{aligned}
Q_n(\epsilon) &\leq 2^{2n+1} \left(\frac{n(S_n+b)}{S_n+b+T_n}\right)^{-3/2} \left(\frac{\epsilon^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty (S_n+b)}{\epsilon^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty (S_n+b) + T_n}\right)^{n(1-\alpha)} \\
&\leq 2^{2n+1}\, n^{-3/2} \left(1 + \frac{\left(p_n\sigma_0^2 + n\breve{\lambda}_{0n}^{-1}\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right)}{n-p_n} \left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{3/2} \\
&\quad\times \left(1 + \frac{\epsilon\left(p_n\sigma_0^2 + n\breve{\lambda}_{0n}^{-1}\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2\right)}{(n-p_n)\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty} \left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-n(1-\alpha)} \\
&\leq 2^{2n+1}\, n^{-3/2} \left(\frac{4\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2}{(1-\alpha)\breve{\lambda}_{0n}\sigma_0^2}\right)^{3/2} \left(\frac{\epsilon\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2}{2(1-\alpha)\breve{\lambda}_{0n}\sigma_0^2}\right)^{-n(1-\alpha)} \\
&= 2^{n(3-\alpha)+4}(n\epsilon)^{-3/2}\left(\epsilon^{-1}(1-\alpha)\lambda_{\max}\sigma_0^2\right)^{n(1-\alpha)-3/2} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^{-n(1-\alpha)+3}
\end{aligned}$$

*for all sufficiently large $n$ a.s.$(P_0)$ by Lemmas 3 and 4 since the quantity in square brackets converges to $1/\sigma_0^2$ a.s.$(P_0)$. Now continue by writing that for all sufficiently large $n$ a.s.$(P_0)$,*

$$\begin{aligned}
&Q_n(\epsilon) \\
&\leq \exp\left\{-n\left[\left(1-\alpha-\frac{3}{n}\right)\log\left(||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2\right) - \left(1-\alpha-\frac{3}{2n}\right)\log\left(\epsilon^{-1}(1-\alpha)\lambda_{\max}\sigma_0^2\right)\right.\right. \\
&\qquad\qquad \left.\left. + \frac{3}{2n}\log(n\epsilon) - \left(3-\alpha+\frac{4}{n}\right)\log 2\right]\right\} \\
&= \exp\left[-n\kappa_n(\epsilon)\right],
\end{aligned}$$

*where $\kappa_n(\epsilon) \to \infty$ is defined to be the quantity in square brackets.*

**Proof** (Proof of Lemma 21)**.** *Assume $||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2 \to A > 0$, $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty > 0$, and $\alpha = 0$. Let $\epsilon > 0$. Note that $R_n > 0$ for all $n$ a.s.$(P_0)$ since $W_n < 1$ for all $n$ a.s.$(P_0)$. Then whenever $L_n(\epsilon) \leq W_n$, we immediately have that $Q_n(\epsilon)/R_n = 0$ exactly, so we may instead assume that $L_n(\epsilon) > W_n$ for all $n$. By Lemma 15, there exists $L^\star(\epsilon) < 1$ such that $\limsup_{n\to\infty} L_n(\epsilon) \leq L^\star(\epsilon)$ a.s.$(P_0)$, which implies that $L_n(\epsilon) < [1+L^\star(\epsilon)]/2$ for all sufficiently large $n$ a.s.$(P_0)$. Then we may write that for all sufficiently* *large* $n$ *a.s.*$(P_0)$,

$$R_n \geq \int_{[1+L_n(\epsilon)]/2}^{1} f_n(u) \left[\frac{u - W_n}{(1-u)}\right]^{-3/2} \exp\left[-\frac{nW_n(1-u)}{2(u-W_n)}\right] du$$
$$\geq \exp\left\{-\frac{nW_n\,[1-L_n(\epsilon)]}{2\,[1+L_n(\epsilon)-2W_n]}\right\} \int_{[3+L^\star(\epsilon)]/4}^{1} f_n(u) \left[\frac{u}{(1-u)}\right]^{-3/2} du$$
$$\geq \frac{\Gamma\left(\frac{n-p_n+a-3}{2}\right)\Gamma\left(\frac{p_n+1}{2}\right)}{\Gamma\left(\frac{n-p_n+a}{2}\right)\Gamma\left(\frac{p_n-2}{2}\right)} \exp\left\{-\frac{nW_n\,[1-L_n(\epsilon)]}{4\,[L_n(\epsilon)-W_n]}\right\}$$
$$\times P_M\left(\frac{3+L^\star(\epsilon)}{4} < \widetilde{U}_n < 1 \;\middle|\; \hat{\boldsymbol{\beta}}_n, S_n\right)$$
$$\geq (4n)^{-3/2} \exp\left\{-\frac{nW_n\,[1-L_n(\epsilon)]}{4\,[L_n(\epsilon)-W_n]}\right\} \tag{17}$$

*by Inequalities 15 and 16. Next, write* $Q_n(\epsilon)$ *as*

$$Q_n(\epsilon) = \int_{W_n}^{L_n(\epsilon)} f_n(u) \left[\frac{u-W_n}{(1-u)}\right]^{-3/2} \exp\left[-\frac{W_n(1-u)}{2(u-W_n)}\right] \exp\left[-\frac{(n-1)W_n(1-u)}{2(u-W_n)}\right] du.$$

*The second and third terms of the integrand comprise an unnormalized*

$$InverseGamma(1/2, W_n/2)$$

*density in* $(u-W_n)/(1-u)$*, which has mode* $W_n/3$*. Then*

$$Q_n(\epsilon) \leq \left(\frac{W_n}{3}\right)^{-3/2} \exp\left(-\frac{3}{2}\right) \exp\left\{-\frac{(n-1)W_n\,[1-L_n(\epsilon)]}{2\,[L_n(\epsilon)-W_n]}\right\} \int_0^{L_n(\epsilon)} f_n(u)\, du,$$
$$\leq (2W_n)^{-3/2} \exp\left\{-\frac{(n-1)W_n\,[1-L_n(\epsilon)]}{2\,[L_n(\epsilon)-W_n]}\right\} [L_n(\epsilon)]^{n/2}$$

*by Lemma 18. Then this result and Inequality 17 together yield that for all sufficiently large* $n$ *a.s.*$(P_0)$,

$$\frac{Q_n(\epsilon)}{R_n} \leq \left(\frac{W_n}{2n}\right)^{-3/2} [L_n(\epsilon)]^{n/2} \exp\left\{-\frac{nW_n\,[1-L_n(\epsilon)]}{2\,[L_n(\epsilon)-W_n]}\left(\frac{n-1}{n}-\frac{1}{2}\right)\right\}$$
$$\leq \left(\frac{2n}{W_n}\right)^{3/2} \exp\left\{-\frac{nW_n\,[1-L^\star(\epsilon)]}{16}\right\}. \tag{18}$$

*Now observe that*

$$\liminf_{n\to\infty} W_n$$
$$= \liminf_{n\to\infty} \left(1 + \frac{T_n}{S_n+b}\right)^{-1}$$

$$= \liminf_{n\to\infty} \left(1 + \frac{p_n\sigma_0^2 + n\check{\lambda}_{0n}^{-1}\,||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_2^2}{n - p_n}\left[\frac{(n-p_n)\,T_n}{(S_n+b)\,\theta_{0n}}\right]\right)^{-1}$$
$$\geq \left(1 + A/\lambda_{\min}\sigma_0^2\right)^{-1} = \frac{\lambda_{\min}\sigma_0^2}{A+\lambda_{\min}\sigma_0^2} \quad a.s.(P_0),$$

*which implies that $W_n > \lambda_{\min}\sigma_0^2/(2A+\lambda_{\min}\sigma_0^2)$ for all sufficiently large $n$ a.s.$(P_0)$. We may combine this with Inequality 18 to yield that for all sufficiently large $n$ a.s.$(P_0)$,*

$$\frac{Q_n(\epsilon)}{R_n} \leq \left[\frac{2n(2A+\lambda_{\min}\sigma_0^2)}{\lambda_{\min}\sigma_0^2}\right]^{3/2} \exp\left\{-\frac{n\lambda_{\min}\sigma_0^2\,[1-L^\star(\epsilon)]}{16(2A+\lambda_{\min}\sigma_0^2)}\right\} \to 0 \quad a.s.(P_0)$$

*since $L^\star(\epsilon) < 1$.*

**Proof** (Proof of Theorem 4)**.** *By Lemmas 11 and 17, posterior consistency occurs if $P_M[(g+1)^{-1}||\boldsymbol{\gamma}_n - \boldsymbol{\beta}_{0n}||_\infty > \epsilon \mid \hat{\boldsymbol{\beta}}_n, S_n] \to 0$ a.s.$(P_0)$ for every $\epsilon > 0$. Then by (7), this occurs if $Q_n(\epsilon)/R_n \to 0$ a.s.$(P_0)$ for every $\epsilon > 0$. We now proceed according to cases similar to those in the proofs of the previous theorems.*

*Case 1: Suppose there do not exist a subsequence $k_n$ of $n$ and a constant $A > 0$ such that $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to A$ and $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \nrightarrow 0$. Let $k_n$ be a subsequence of $n$, and let $\epsilon > 0$. Now consider two sub-cases.*

*Case 1.1: Suppose $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0n}||_\infty \to 0$. Then $\epsilon^{-1}||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty < 1$ for all sufficiently large $n$ a.s.$(P_0)$. This implies that $L_{k_n}(\epsilon) = W_{k_n}$ and $Q_{k_n}(\epsilon) = 0$ for all sufficiently large $n$ a.s.$(P_0)$. Also, $R_{k_n} > 0$ for all $n$ a.s.$(P_0)$ since $W_{k_n} < 1$ a.s.$(P_0)$. Therefore, $Q_{k_n}(\epsilon)/R_{k_n} \to 0$ a.s.$(P_0)$.*

*Case 1.2: Note that Case 1.1 can be applied to any further subsequence $m_n$ of $k_n$ for which $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_\infty \to 0$, so we may suppose for Case 1.2 that $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0$. Note also that in this case, there cannot exist any further subsequence $m_n$ of $k_n$ for which $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_2^2$ converges to a nonzero constant, since this would contradict the original supposition of Case 1. Then since $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \geq \liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty^2 > 0$, it follows that $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to \infty$. Observe that by Lemmas 19 and 20, there exist a constant $K$ and a sequence of constants $\kappa_n(\epsilon) \to \infty$ such that $Q_{k_n}(\epsilon)/R_{k_n} \leq \exp\{-n\,[\kappa_n(\epsilon) - K]\} \to 0$ a.s.$(P_0)$. Finally, since Cases 1.1 and 1.2 together establish that $Q_{k_n}(\epsilon)/R_{k_n} \to 0$ a.s.$(P_0)$ for every subsequence $k_n$, it follows that $Q_n(\epsilon)/R_n \to 0$ a.s.$(P_0)$, and therefore posterior consistency occurs.*

*Case 2: Now suppose there exist a subsequence $k_n$ of $n$ and a constant $A > 0$ such that $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_2^2 \to A > 0$ and $||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty \nrightarrow 0$, and suppose $\alpha = 0$. Note that Case 1 can be applied to any subsequence $m_n$ of $n$ for which either $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_2^2$ does not converge to any nonzero constant or $||\boldsymbol{\gamma}_{m_n} - \boldsymbol{\beta}_{0m_n}||_\infty \to 0$, so it suffices to show that $Q_{k_n}(\epsilon)/R_{k_n} \to 0$ a.s.$(P_0)$. Note also that this means we may suppose for Case 2 that $\liminf_{n\to\infty} ||\boldsymbol{\gamma}_{k_n} - \boldsymbol{\beta}_{0k_n}||_\infty > 0$. Now let $\epsilon > 0$. Then we immediately have that $Q_{k_n}(\epsilon)/R_n \to 0$ a.s.$(P_0)$ by Lemma 21. Therefore posterior consistency occurs.*

## References

Armagan, A., Dunson, D., Lee, J., Bajwa, W., and Strawn, N. (2013). "Posterior consistency in linear models under shrinkage priors." *Biometrika*, 100(4): 1011–1018. 628, 630

Barron, A., Schervish, M., and Wasserman, L. (1999). "The consistency of posterior distributions in nonparametric problems." *Annals of Statistics*, 27: 536–561. 627

Berger, J., Ghosh, J., and Mukhopadhyay, N. (2003). "Approximations and consistency of Bayes factors as model dimension grows." *Journal of Statistical Planning and Inference*, 112: 241–258. 628

Bernstein, S. (1934). *Theory of Probability*. GTTI, Moscow. 631

Bontemps, D. (2011). "Bernstein-von Mises theorems for Gaussian regression with increasing number of regressors." *Annals of Statistics*, 39: 2557–2584. 628

Diaconis, P. and Freedman, D. (1986). "On the consistency of Bayes estimates." *Annals of Statistics*, 14: 1–26. 631

Doob, J. (1948). "Application of the theory of martingales." *Colloques Internationaux du Centre National de la Recherche Scientifique*, 13: 23–27. 631

Fernandez, C., Ley, E., and Steel, M. (2001). "Benchmark priors for Bayesian model averaging." *Journal of Econometrics*, 100: 381–427. 629, 634

George, E. and Foster, D. (2000). "Calibration and empirical Bayes variable selection." *Biometrika*, 87: 731–747. 627, 641

Ghosal, S. (1999). "Asymptotic normality of posterior distributions in high-dimensional linear models." *Bernoulli*, 5: 315–331. 628

Ghosal, S., Ghosh, J., and van der Vaart, A. (2000). "Convergence rates of posterior distributions." *Annals of Statistics*, 28: 500–531. 627

Ghosh, J., B. Sinha, B., and Joshi, S. (1982). "Expansions for posterior probability and integrated Bayes risk." In Gupta, S. and Berger, J. (eds.), *Statistical Decision Theory and Related Topics*, volume 3, 403–456. Academic Press, New York. 627, 631

Jiang, W. (2007). "Bayesian variable selection for high-dimensional generalized linear models: Convergence rates of the fitted densities." *Annals of Statistics*, 35: 1487–1511. 628

Kass, R. and Wasserman, L. (1995). "A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion." *Journal of the American Statistical Association*, 90: 928–934. 634

Laplace, P. (1774). "Mémoire sur la probabilité des causes par les événements." *Mémoire Academy Recherche Scientifique*, 6: 621–656. 631

LeCam, L. (1982). "On the risk of Bayes estimates." In Gupta, S. and Berger, J. (eds.), *Statistical Decision Theory and Related Topics*, volume 3, 121–138. Academic Press, New York. 631

Lee, J. and Oh, H.-S. (2013). "Bayesian regression based on principal components for high-dimensional data." *Journal of Multivariate Analysis*, 117: 175–192. 628

Liang, F., Paulo, R., Molina, G., Clyde, M., and Berger, J. (2008). "Mixtures of $g$ priors for Bayesian variable selection." *Journal of the American Statistical Association*, 103: 410–423. 627, 629, 637, 641

Mallick, B., Ghosh, D., and Ghosh, M. (2005). "Bayesian classification of tumors using gene expression data." *Journal of the Royal Statistical Society B*, 67: 219–234. 627

Moreno, E., Giron, F., and Casella, G. (2010). "Consistency of objective Bayes factors as the model dimension grows." *Annals of Statistics*, 38: 1937–1952. 628

Shang, Z. and Clayton, K. (2011). "Consistency of Bayesian linear model selection with a growing number of parameters." *Journal of Statistical Planning and Inference*, 141: 3463–3474. 629

von Mises, R. (1964). *Mathematical Theory of Probability and Statistics*. Academic Press, New York. 631

Zellner, A. (1986). "On assessing prior distributions and Bayesian regression analysis with $g$-prior distributions." In Goel, P. and Zellner, A. (eds.), *Bayesian Inference and Decision Techniques*, 233–243. North-Holland, Amsterdam. 627

Zellner, A. and Siow, A. (1980). "Posterior odds ratios for selected regression hypotheses." In J. Bernardo, D. L., M. DeGroot and Smith, A. (eds.), *Bayesian Statistics*, 585–604. University Press, Valencia. 627, 629, 639, 641

Zhang, Z., Jordan, M., and Yeung, D. (2009). "Posterior consistency of the Silverman g-prior in Bayesian model choice." *Advances in Neural Information Processing Systems*, 22. 629

**Acknowledgments**

This work was supported in part by NSF Grants DMS–1106084 and DMS–1007494.