1 Online Supplement for *Detection of latent heteroscedasticity and group-based regression effects in linear models via Bayesian model selection*

1.1 Marginal Model Probability Calculations

To prevent numeric underflow we consider model probabilites on the logarithmic scale in all cases. Consider the set of all log-marginal probabilities $\mathcal{L} = \{\log P(Y|m_s^c)\}$. Let $\ell^\star = \max \mathcal{L}$ and $m^\star = \arg\max_m \mathcal{L}$. Transform $\mathcal{L}^\star = \mathcal{L} - \ell^\star$ to obtain the set $\exp[\mathcal{L}^\star] = \{\exp[\log P(Y|m_s^c) - \log P(Y|m^\star)]\} = \{\frac{P(Y|m_s^c)}{P(Y|m^\star)}\} = \{B^{m_s^c, m^\star}\}$, representing Bayes factors for each model relative to the most probable model. Note the untransformed model probabilities are given by

$$P = \{\frac{P(Y|m_s^c)P(m_s^c)}{\sum P(Y|m_s^c)P(m_s^c)}\} = \{\frac{P(Y|m_s^c)P(m_s^c)/P(Y|m^\star)}{\sum P(Y|m_s^c)P(m_s^c)/P(Y|m^\star)}\} = \{\frac{B^{m_s^c, m^\star}P(m_s^c)}{\sum B^{m_s^c, m^\star}P(m_s^c)}\};$$

thus we can easily use these transformed log-marginal probabilities to obtain posterior model probabilities. Our Laplace approximations generally provided values comparable to quadrature and cubature-based approximations at less computational expense.

1.2 Derivation of Model Probabilities: Noninformative Regression Effect Priors

Let $\boldsymbol{H_X} := \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$.

1.2.1 Homoscedastic Error Variance, No Group-Based Regression Effects

Let SSResid_g^c denote the residual sum of squares for the mth model for the cth class and gth group.

$$
\begin{aligned}
P(\boldsymbol{Y}|m) &= \int_{\Theta} \left[(2\pi)^{-\frac{N}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \cdot \exp\{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})^T \Sigma^{-1} (\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})\} \right]^b P(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \int_0^{\infty} \int_{-\infty}^{\infty} (2\pi)^{-\frac{Nb}{2}} \gamma^{\frac{Nb}{2}-1} \exp\{-b \cdot \frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} + \boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\boldsymbol{\beta}]\} d\boldsymbol{\beta} d\gamma \\
&= \int_0^{\infty} \int_{-\infty}^{\infty} (2\pi)^{-\frac{Nb}{2}} \gamma^{\frac{Nb}{2}-1} \exp\{-b \cdot \frac{\gamma}{2}[\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{Y} + \boldsymbol{Y}^T\boldsymbol{H}\boldsymbol{Y}]\} \times \\
&\qquad \exp\{-b \cdot \frac{\gamma}{2}[\boldsymbol{Y}^T\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{H}\boldsymbol{Y}]\} d\boldsymbol{\beta} d\gamma \\
&= \int_0^{\infty} (2\pi)^{-\frac{Nb}{2}} \gamma^{\frac{Nb}{2}-1} (2\pi)^{+\frac{p}{2}} |b^{-1}\gamma^{-1}(\boldsymbol{X}^T\boldsymbol{X})^{-1}|^{+\frac{1}{2}} \exp\{-b \cdot \frac{\gamma}{2}[\boldsymbol{Y}^T(I-\boldsymbol{H})\boldsymbol{Y}]\} d\gamma \\
&= (2\pi)^{-\frac{Nb-P}{2}} b^{-\frac{P}{2}} |\boldsymbol{X}^T\boldsymbol{X}|^{-\frac{1}{2}} \Gamma\left(\frac{Nb-P}{2}\right) \left(\frac{b \cdot \mathrm{SSResid}^{\mathrm{I}}}{2}\right)^{-\frac{Nb-P}{2}} \\
&= \pi^{-\frac{Nb-P}{2}} b^{-\frac{Nb}{2}} |\boldsymbol{X}^T\boldsymbol{X}|^{-\frac{1}{2}} \Gamma\left(\frac{Nb-P}{2}\right) (\mathrm{SSResid}^{\mathrm{I}})^{-\frac{Nb-P}{2}}. \ \square
\end{aligned}
$$

$$
\begin{aligned}
q^b(\boldsymbol{Y}|m) &= \frac{\int_{\Theta} P(Y|m^{\mathrm{I}}, \boldsymbol{\theta}) P(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta} P^b(Y|m^{\mathrm{I}}, \boldsymbol{\theta}) P(\boldsymbol{\theta}) d\boldsymbol{\theta}} \\
&= \frac{\pi^{-\frac{N-P}{2}} |\boldsymbol{X}^T\boldsymbol{X}|^{-\frac{1}{2}} \Gamma\left(\frac{N-P}{2}\right) (\mathrm{SSResid}^{\mathrm{I}})^{-\frac{N-P}{2}}.}{\pi^{-\frac{Nb-P}{2}} b^{-\frac{Nb}{2}} |\boldsymbol{X}^T\boldsymbol{X}|^{-\frac{1}{2}} \Gamma\left(\frac{Nb-P}{2}\right) (\mathrm{SSResid}^{\mathrm{I}})^{-\frac{Nb-P}{2}}} \\
&= \pi^{-\frac{N(1-b)}{2}} b^{\frac{Nb}{2}} (\mathrm{SSResid}^{\mathrm{I}})^{-\frac{N(1-b)}{2}} \frac{\Gamma\left(\frac{N-P}{2}\right)}{\Gamma\left(\frac{Nb-P}{2}\right)}. \ \square
\end{aligned}
$$

1.2.2 Homoscedastic Error Variance, Group-Based Regression Effects

$$
\begin{aligned}
P(\boldsymbol{Y}|m_s) &= \int_{\Theta} \left[(2\pi)^{-\frac{N}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \cdot \exp\{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})\} \right]^b P(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \int_0^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (2\pi)^{-\frac{Nb}{2}} \gamma^{\frac{Nb}{2}-1} \times \\
&\qquad \exp\{-b \cdot \frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} + \boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\boldsymbol{\beta}^T]\} d\boldsymbol{\beta}_1^T d\boldsymbol{\beta}_2^T d\gamma \\
&= \int_0^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (2\pi)^{-\frac{Nb}{2}} \gamma^{\frac{Nb}{2}-1} \exp\{-b \cdot \frac{\gamma}{2}[\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{Y} + \boldsymbol{Y}^T\boldsymbol{H}\boldsymbol{Y}]\} \times \\
&\qquad \exp\{-b \cdot \frac{\gamma}{2}[\boldsymbol{Y}^T\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{H}\boldsymbol{Y}]\} d\boldsymbol{\beta}_1 d\boldsymbol{\beta}_2 d\gamma \\
&= \int_0^{\infty} (2\pi)^{-\frac{Nb}{2}} \gamma^{\frac{Nb}{2}-1} (2\pi)^{+\frac{P}{2}} |b^{-1}\gamma^{-1}(\boldsymbol{X}_1\boldsymbol{X}_1)^{-1}|^{+\frac{1}{2}} |b^{-1}\gamma^{-1}(\boldsymbol{X}_2\boldsymbol{X}_2)^{-1}|^{+\frac{1}{2}} \times \\
&\qquad \exp\{-b \cdot \frac{\gamma}{2}[\boldsymbol{Y}_1^T(I-\boldsymbol{H}_1)\boldsymbol{Y}_1 + \boldsymbol{Y}_2^T(I-\boldsymbol{H}_2)\boldsymbol{Y}_2]\} d\gamma \\
&= (2\pi)^{-\frac{Nb-P}{2}} b^{-\frac{P}{2}} |\boldsymbol{X}_1^T\boldsymbol{X}_1|^{-\frac{1}{2}} |\boldsymbol{X}_2^T\boldsymbol{X}_2|^{-\frac{1}{2}} \Gamma\left(\frac{Nb-P}{2}\right) \times \\
&\qquad \left(\frac{b \cdot \left[\text{SSResid}_1^{\text{II}} + \text{SSResid}_2^{\text{II}}\right]}{2} \right)^{-\frac{Nb-P}{2}} \\
&= \pi^{-\frac{Nb-P}{2}} b^{-\frac{Nb}{2}} |\boldsymbol{X}_1^T\boldsymbol{X}_1|^{-\frac{1}{2}} |\boldsymbol{X}_2^T\boldsymbol{X}_2|^{-\frac{1}{2}} \Gamma\left(\frac{Nb-P}{2}\right) (\text{SSResid}_1^{\text{II}} + \text{SSResid}_2^{\text{II}})^{-\frac{Nb-P}{2}}. \square
\end{aligned}
$$

$$
\begin{aligned}
q^b(\boldsymbol{Y}|m_s) &= \frac{\int_{\Theta} P(Y|m_s^{\text{II}}, \boldsymbol{\theta}) P(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta} P^b(Y|m_s^{\text{II}}, \boldsymbol{\theta}) P(\boldsymbol{\theta}) d\boldsymbol{\theta}} \\
&= \frac{\pi^{-\frac{N-P}{2}} b^{-\frac{N}{2}} |\boldsymbol{X}_1^T\boldsymbol{X}_1|^{-\frac{1}{2}} |\boldsymbol{X}_2^T\boldsymbol{X}_2|^{-\frac{1}{2}} \Gamma\left(\frac{N-P}{2}\right) (\text{SSResid}_1^{\text{II}} + \text{SSResid}_2^{\text{II}})^{-\frac{N-P}{2}}}{\pi^{-\frac{Nb-P}{2}} b^{-\frac{Nb}{2}} |\boldsymbol{X}_1^T\boldsymbol{X}_1|^{-\frac{1}{2}} |\boldsymbol{X}_2^T\boldsymbol{X}_2|^{-\frac{1}{2}} \Gamma\left(\frac{Nb-P}{2}\right) (\text{SSResid}_1^{\text{II}} + \text{SSResid}_2^{\text{II}})^{-\frac{Nb-P}{2}}} \\
&= \pi^{-\frac{N(1-b)}{2}} b^{\frac{Nb}{2}} (\text{SSResid}_1^{\text{II}} + \text{SSResid}_2^{\text{II}})^{-\frac{N(1-b)}{2}} \frac{\Gamma\left(\frac{N-P}{2}\right)}{\Gamma\left(\frac{Nb-P}{2}\right)}. \square
\end{aligned}
$$

1.2.3 Heteroscedastic Error Variance

A Laplace approximation is used to evaluate $\int P(\boldsymbol{Y}|\boldsymbol{\varphi}, m)P(\boldsymbol{\varphi})d\boldsymbol{\varphi}$, parametrized with respect to the log-variances $\lambda_1 = \log\sigma_1^2$ and $\lambda_2 = \log\sigma_2^2$. Denote $\boldsymbol{\Lambda}$ as the log-variance matrix, $J_{\boldsymbol{\Lambda}}$ as the transformation Jacobian, $(\lambda_1^\star, \lambda_2^\star)$ as the mode of the log-marginal distribution, and $\nabla^\star$ as the Hessian evaluated at this mode. A subscript of b refers to the same quantities calculated with respect to the fractional exponentiated likelihood. The joint modes $(\lambda_1^\star, \lambda_2^\star)$ and $(\lambda_{b_1}^\star, \lambda_{b_2}^\star)$ are computed using the function `optim` in R; similarly, the Hessians $\nabla^\star$ and $\nabla_b^\star$ are evaluated at these values using the function `hessian` in the package `numderiv` (Gilbert and Varadhan, 2016; R Core Team, 2017).

We first integrate over the regression effects vector $\boldsymbol{\beta}$. Let $\boldsymbol{H}_\Phi := \Phi\boldsymbol{X}(\boldsymbol{X}^T\Phi\boldsymbol{X})^{-1}\boldsymbol{X}^T\Phi$.

$$
\begin{aligned}
P^b(\boldsymbol{Y}|m_s^{\mathrm{III}}, \varphi_1, \varphi_2) &= \int_\Theta \left[(2\pi)^{-\frac{N}{2}}|\Phi|^{\frac{1}{2}} \cdot \exp\{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})^T\Phi(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})\}\right]^b P(\boldsymbol{\theta})d\boldsymbol{\theta} \\
&= \int_{-\infty}^{\infty} (2\pi)^{-\frac{Nb}{2}} {\varphi_1}^{\frac{n_1 b}{2}-1}{\varphi_2}^{\frac{n_2 b}{2}-1}\times \\
&\qquad \exp\{-b\cdot\frac{1}{2}[\boldsymbol{Y}^T\Phi\boldsymbol{Y} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\Phi\boldsymbol{Y} + \boldsymbol{\beta}^T\boldsymbol{X}^T\Phi\boldsymbol{X}\boldsymbol{\beta}^T]\}d\boldsymbol{\beta} \\
&= \int_{-\infty}^{\infty} (2\pi)^{-\frac{Nb}{2}} {\varphi_1}^{\frac{n_1 b}{2}-1}{\varphi_2}^{\frac{n_2 b}{2}-1}\times \\
&\qquad \exp\{-b\cdot\frac{1}{2}[\boldsymbol{\beta}^T\boldsymbol{X}^T\Phi\boldsymbol{X}\boldsymbol{\beta} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{Y} + \boldsymbol{Y}^T\boldsymbol{H}_\Phi\boldsymbol{Y}]\}\times \\
&\qquad \exp\{-b\cdot\frac{1}{2}[\boldsymbol{Y}^T\Phi\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{H}_\Phi\boldsymbol{Y}]\}d\boldsymbol{\beta} \\
&= (2\pi)^{-\frac{Nb-P}{2}}{\varphi_1}^{\frac{n_1 b}{2}-1}{\varphi_2}^{\frac{n_2 b}{2}-1}b^{-\frac{P}{2}}|\boldsymbol{X}^T\Phi\boldsymbol{X}|^{-\frac{1}{2}}\exp\{-b\cdot\frac{1}{2}[\boldsymbol{Y}^T\Phi\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{H}_\Phi\boldsymbol{Y}]\}
\end{aligned}
$$

We reparametrize the precisions of $P^b(\boldsymbol{Y}|m_s^{\mathrm{III}}, \varphi_1, \varphi_2)$ and $P(\boldsymbol{Y}|m_s^{\mathrm{III}}, \varphi_1, \varphi_2)$ to log-variances λ_1 and λ_2 to elicit a shape more conducive to the Laplace approximation.

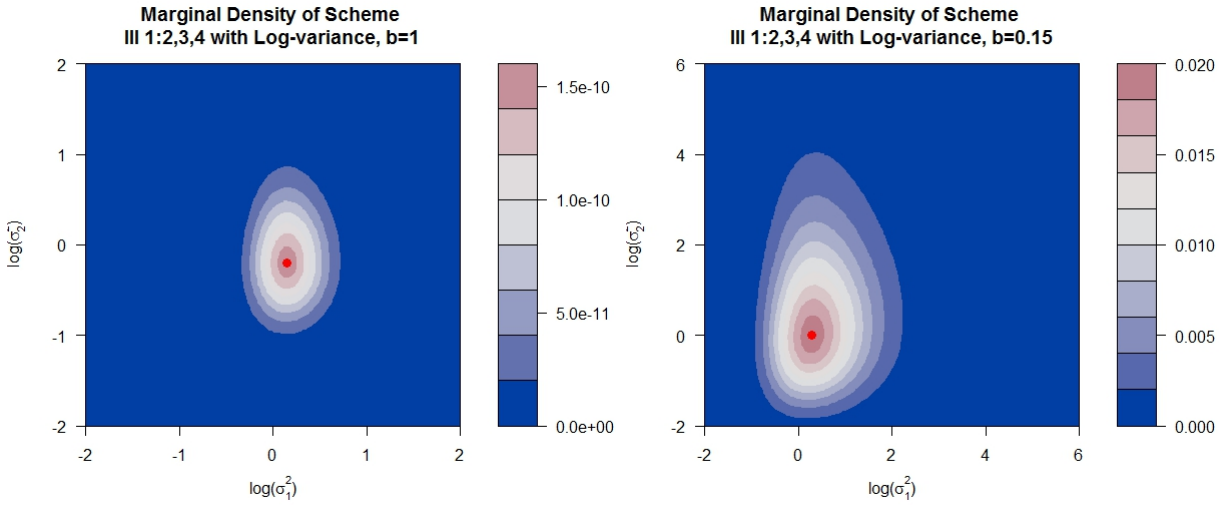

Figure 1: The marginal densities $P(\boldsymbol{Y}|m^{\mathrm{II}}_{1:2,3,4}, \lambda_1, \lambda_2)$ and $P^b(\boldsymbol{Y}|m^{\mathrm{III}}_{1:2,3,4}, \lambda_1, \lambda_2)$ for one particular grouping of an ANOVA layout with $K = 4$ levels of $n_k = 10$ observations each show an overall shape conducive to the Laplace approximation in both the raw and exponentiated likelihood cases.

For $\lambda_1 = \ln \varphi_1^{-1}$ and $\lambda_2 = \ln \varphi_2^{-1}$, $P(\boldsymbol{Y}|m^{\mathrm{III}}_s, \lambda_1, \lambda_2) = P(\boldsymbol{Y}|m^{\mathrm{III}}_s, \varphi_1, \varphi_2) \cdot |J_{\boldsymbol{\Lambda}}|$ where $\frac{\partial \varphi_1}{\partial \lambda_1} = -\exp\{-\lambda_1\}$ and $\frac{\partial \varphi_2}{\partial \lambda_2} = -\exp\{-\lambda_2\}$, so $|J_{\boldsymbol{\Lambda}}| = \exp\{-(\lambda_1 + \lambda_2)\}$.

Thus the Laplace approximation for the fractional marginal model probability is given by

$$q^b(\boldsymbol{Y}|m^{\mathrm{III}}_s) \approx \frac{(2\pi)|-\nabla^\star|^{-\frac{1}{2}} \cdot P(\boldsymbol{Y}|m^{\mathrm{III}}_s, \varphi^\star_{b_1}, \varphi^\star_{b_2})|J^\star_{\boldsymbol{\Lambda}}|}{(2\pi)|-\nabla^\star_b|^{-\frac{1}{2}} \cdot P^b(\boldsymbol{Y}|m^{\mathrm{III}}_s, \varphi^\star_{b_1}, \varphi^\star_{b_2})|J^\star_{b_{\boldsymbol{\Lambda}}}|}. \; \square$$

Empirical comparisons with cubature-based approximations, in addition to the shapes of the marginal densities $P(\boldsymbol{Y}|m^{\mathrm{III}}_s, \lambda_1, \lambda_2)$ and $P^b(\boldsymbol{Y}|m^{\mathrm{III}}_s, \lambda_1, \lambda_2)$, indicate that this Laplace approximation is satisfactory.

1.2.4 Special Case: Heteroscedastic Error Variance with no Regression Effects Spanning Two Variances

$$
\begin{aligned}
P(\boldsymbol{Y}|m_s) &= \int_{\Theta} \left[(2\pi)^{-\frac{N}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \cdot \exp\{-\frac{1}{2}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta})\} \right]^b P(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \int_0^{\infty}\int_0^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (2\pi)^{-\frac{Nb}{2}} |\boldsymbol{\Sigma}|^{-\frac{b}{2}} \exp\{-b \cdot \frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} + \boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\boldsymbol{\beta}]\} \\
&\qquad d\boldsymbol{\beta}_1 d\boldsymbol{\beta}_2 d\gamma_1 d\gamma_2 \\
&= \int_0^{\infty}\int_0^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (2\pi)^{-\frac{Nb}{2}} \gamma^{\frac{n_1 b}{2}-1}\gamma^{\frac{n_2 b}{2}-1} \times \\
&\qquad \exp\{-b \cdot \frac{1}{2}[\boldsymbol{Y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} + \boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\boldsymbol{\beta}]\} d\boldsymbol{\beta}_1 d\boldsymbol{\beta}_2 d\gamma_1 d\gamma_2 \\
&= \int_0^{\infty}\int_0^{\infty} (2\pi)^{-\frac{Nb}{2}} \gamma^{\frac{n_1 b}{2}-1}\gamma^{\frac{n_2 b}{2}-1} (2\pi)^{+\frac{P}{2}} |b^{-1}\gamma^{-1}(\boldsymbol{X}_1^T\boldsymbol{X}_1)^{-1}|^{+\frac{1}{2}} |b^{-1}\gamma^{-1}(\boldsymbol{X}_2^T\boldsymbol{X}_2)^{-1}|^{+\frac{1}{2}} \times \\
&\qquad \exp\{-b \cdot \frac{\gamma}{2}[\boldsymbol{Y}_1^T(I-\boldsymbol{H}_1)\boldsymbol{Y}_1 + \boldsymbol{Y}_2^T(I-\boldsymbol{H}_2)\boldsymbol{Y}_2]\} d\gamma_1 d\gamma_2 \\
&= (2\pi)^{-\frac{Nb-P}{2}} b^{-\frac{P}{2}} |\boldsymbol{X}_1^T\boldsymbol{X}_1|^{-\frac{1}{2}} |\boldsymbol{X}_2^T\boldsymbol{X}_2|^{-\frac{1}{2}} \Gamma\left(\frac{Nb-P}{2}\right) \times \\
&\qquad \left(\frac{b \cdot \left[\mathrm{SSResid}_1^{\mathrm{IV}} + \mathrm{SSResid}_2^{\mathrm{IV}}\right]}{2}\right)^{-\frac{Nb-P}{2}} \\
&= (2\pi)^{-\frac{Nb-P}{2}} b^{-\frac{P}{2}} |\boldsymbol{X}_1^T\boldsymbol{X}_1|^{-\frac{1}{2}} |\boldsymbol{X}_2^T\boldsymbol{X}_2|^{-\frac{1}{2}} \Gamma\left(\frac{n_1 b - p_1}{2}\right) \Gamma\left(\frac{n_2 b - p_2}{2}\right) \times \\
&\qquad \left(\frac{b \cdot \mathrm{SSResid}_1^{\mathrm{IV}}}{2}\right)^{-\frac{n_1 b - p_1}{2}} \left(\frac{b \cdot \mathrm{SSResid}_2^{\mathrm{IV}}}{2}\right)^{-\frac{n_2 b - p_2}{2}}.
\end{aligned}
$$

$$
\begin{aligned}
q^b(\boldsymbol{Y}|m_s) &= \frac{\int\limits_{\Theta} P(Y|m_s,\boldsymbol{\theta})P(\boldsymbol{\theta})d\boldsymbol{\theta}}{\int\limits_{\Theta} P^b(Y|m_s^{\mathrm{IV}},\boldsymbol{\theta})P(\boldsymbol{\theta})d\boldsymbol{\theta}} \\
&= \frac{(2\pi)^{-\frac{N-P}{2}}|\boldsymbol{X}_1\boldsymbol{X}_1|^{-\frac{1}{2}}|\boldsymbol{X}_2\boldsymbol{X}_2|^{-\frac{1}{2}}\Gamma\left(\frac{n_1-p_1}{2}\right)\Gamma\left(\frac{n_2-p_2}{2}\right)}{(2\pi)^{-\frac{Nb-P}{2}}b^{-\frac{P}{2}}|\boldsymbol{X}_1\boldsymbol{X}_1|^{-\frac{1}{2}}|\boldsymbol{X}_2\boldsymbol{X}_2|^{-\frac{1}{2}}\Gamma\left(\frac{n_1b-p_1}{2}\right)\Gamma\left(\frac{n_2b-p_2}{2}\right)}\times \\
&\qquad \frac{\left(\frac{\mathrm{SSResid}_1^{\mathrm{IV}}}{2}\right)^{-\frac{n_1-p_1}{2}}\left(\frac{\mathrm{SSResid}_2^{\mathrm{IV}}}{2}\right)^{-\frac{n_2-p_2}{2}}}{\left(\frac{b\cdot\mathrm{SSResid}_1^{\mathrm{IV}}}{2}\right)^{-\frac{n_1b-p_1}{2}}\left(\frac{b\cdot\mathrm{SSResid}_2^{\mathrm{IV}}}{2}\right)^{-\frac{n_2b-p_2}{2}}} \\
&= \pi^{-\frac{N(1-b)}{2}}b^{\frac{Nb}{2}}(\mathrm{SSResid}_1^{\mathrm{IV}})^{-\frac{n_1(1-b)}{2}}(\mathrm{SSResid}_2^{\mathrm{IV}})^{-\frac{n_2(1-b)}{2}}\frac{\Gamma\left(\frac{n_1-p_1}{2}\right)}{\Gamma\left(\frac{n_1b-p_1}{2}\right)}\frac{\Gamma\left(\frac{n_2-p_2}{2}\right)}{\Gamma\left(\frac{n_2b-p_2}{2}\right)}.\ \square
\end{aligned}
$$

1.3 Derivation of Model Probabilities: Mixture g Regression Effect Priors

$$
\begin{aligned}
P(\boldsymbol{Y}, g|m_s) &= \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty P^b(\boldsymbol{Y}|\alpha, \beta, \varphi, m_s) P(\alpha, \varphi) P(\beta|\varphi, g) P(g) d\alpha d\beta d\varphi dg \\
&= \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty (2\pi)^{-\frac{Nb}{2}} \varphi^{\frac{Nb}{2}} \exp\{-\frac{\varphi b}{2}(\boldsymbol{Y} - \mathbf{1}^T\alpha - \boldsymbol{X}\boldsymbol{\beta})^T(\boldsymbol{Y} - \mathbf{1}^T\alpha - \boldsymbol{X}\boldsymbol{\beta})\} \times \\
&\quad (2\pi)^{-\frac{P}{2}} |g\varphi^{-1}(\boldsymbol{X}^T\boldsymbol{X})^{-1}|^{-\frac{1}{2}} \exp\{-\frac{\varphi}{2}\frac{\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta}}{g}\} \varphi^{-1} P(g) d\alpha d\boldsymbol{\beta} d\varphi dg \\
&= \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty (2\pi)^{-\frac{Nb+P}{2}} \varphi^{\frac{Nb}{2}-1} \times \\
&\quad \exp\{-\frac{\varphi}{2}(\alpha^2 b \mathbf{1}^T\mathbf{1} - 2\alpha b(\mathbf{1}^T\boldsymbol{Y} - \mathbf{1}^T\boldsymbol{X}\boldsymbol{\beta}) + b(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})\boldsymbol{H_1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}))\} \times \\
&\quad |g\varphi^{-1}(\boldsymbol{X}^T\boldsymbol{X})^{-1}|^{-\frac{1}{2}} \exp\{-\frac{\varphi}{2}(\boldsymbol{\beta}^T\left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{g} + b\boldsymbol{X}^T\boldsymbol{X}\right)\boldsymbol{\beta} - 2\boldsymbol{\beta}^T b\boldsymbol{X}^T\boldsymbol{Y})\} \times \\
&\quad \exp\{-\frac{\varphi}{2}(b\boldsymbol{Y}^T\boldsymbol{Y} - b(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})\boldsymbol{H_1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}))\} P(g) d\alpha d\boldsymbol{\beta} d\varphi dg \\
&= \int_0^\infty \int_0^\infty \int_{-\infty}^\infty (2\pi)^{-\frac{Nb+P-1}{2}} \varphi^{\frac{Nb+P-1}{2}-1} b^{-\frac{1}{2}} |\boldsymbol{X}^T\boldsymbol{X}|^{-\frac{1}{2}} N^{-\frac{1}{2}} g^{-\frac{P}{2}} \times \\
&\quad \exp\{-\frac{\varphi}{2}(\boldsymbol{\beta}^T\left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{g} + b\boldsymbol{X}^T\boldsymbol{X}\right)\boldsymbol{\beta} - 2\boldsymbol{\beta}^T b\boldsymbol{X}^T\boldsymbol{Y} + \frac{b^2 g}{1+bg}\boldsymbol{Y}^T\boldsymbol{H_X}\boldsymbol{Y})\} \times \\
&\quad \exp\{-\frac{\varphi}{2}(b\boldsymbol{Y}^T\boldsymbol{Y} - b\boldsymbol{Y}^T\boldsymbol{H_1}\boldsymbol{Y} - \frac{b^2 g}{1+bg}\boldsymbol{Y}^T\boldsymbol{H_X}\boldsymbol{Y})\} P(g) d\boldsymbol{\beta} d\varphi dg \\
&= \int_0^\infty \int_0^\infty (2\pi)^{-\frac{Nb-1}{2}} \varphi^{\frac{Nb-1}{2}-1} b^{-\frac{1}{2}} N^{-\frac{1}{2}} (1+bg)^{-\frac{P}{2}} \times \\
&\quad \exp\{-\frac{\varphi}{2}(b\boldsymbol{Y}^T\boldsymbol{Y} - b\boldsymbol{Y}^T\boldsymbol{H_1}\boldsymbol{Y} - \frac{b^2 g}{1+bg}\boldsymbol{Y}^T\boldsymbol{H_X}\boldsymbol{Y})\} d\varphi dg \\
&= \int_0^\infty \frac{\Gamma\left(\frac{Nb-1}{2}\right)}{\sqrt{\pi}^{Nb-1}\sqrt{N}} b^{-\frac{Nb}{2}} (1+bg)^{-\frac{Nb-P-1}{2}} \mathrm{SST}^{-\frac{Nb-1}{2}} [1 + bg(1-R^2)]^{-\frac{Nb-1}{2}} P(g) dg. \ \square
\end{aligned}
$$

After integrating over α, $\boldsymbol{\beta}$, and $\boldsymbol{\varphi}$, we obtain

$$P(\boldsymbol{Y}, g|m_s^c) = \int_0^\infty \frac{\Gamma(\frac{Nb-1}{2}) b^{-\frac{Nb}{2}}}{\sqrt{\pi}^{Nb-1}\sqrt{N}} \frac{\mathrm{SST}^{-\frac{Nb-1}{2}}}{(1+bg)^{\frac{Nb-P-1}{2}}} [1+bg(1-R^2)]^{-\frac{Nb-1}{2}} g^{-1.5} \exp\{-\frac{N}{2g}\} dg \quad (1)$$

To execute the Laplace approximation over g, we must obtain the mode

$$g^\star = \underset{g}{\operatorname{argmax}}\ P(\boldsymbol{Y}, g|m_s^c),$$

the root of the equation

$$-Qb^2(P+3)g^3 + (b(Nb-P-4)-2Q)g^2 + (Nb(2-R^2)-3)g + N := 0 \quad (2)$$

where $Q = 1-R^2$. We also require the Hessian evaluated at the mode,

$$\begin{aligned} \boldsymbol{H}^\star &= \frac{\partial^2}{\partial g^2}[\log((1+bg)^{\frac{Nb-P-1}{2}}(1+Qbg)^{-\frac{Nb-1}{2}} g^{-\frac{3}{2}} \exp\{-\frac{N}{2g}\})]\Big|^{g=g^\star} \\ &= \frac{1}{2}\left[\frac{(Nb-1)b^2Q^2}{(1+Qbg^\star)^2} - \frac{(Nb-P-1)b^2}{(1+bg^\star)^2} + \frac{3}{(g^\star)^2} - \frac{2N}{(g^\star)^3}\right]. \end{aligned}$$

These expressions are appropriate to use in homoscedastic classes with either global or distinct regression effects; we simply compute the corresponding R^2 and use the appropriate P based on the model under consideration.

In classes with heteroscedasticity, the integral is intractable over both $\boldsymbol{\varphi}$ and g; thus we must employ a three-dimensional Laplace approximation to evaluate this integral. For computational ease and to improve the accuracy of the approximation, we again parametrize with respect to the log-variance; let $\boldsymbol{\Lambda}$ represent the log-variance matrix. Denote $\boldsymbol{H}_{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda}\mathbf{1}(\mathbf{1}^T\boldsymbol{\Lambda}\mathbf{1})^{-1}\mathbf{1}^T\boldsymbol{\Lambda}$; then integrating out the global intercept and regression effects yields an expression for $P^b(\boldsymbol{Y}, \lambda_1, \lambda_2, g|m_s^c)$

$$= (2\pi)^{-\frac{Nb+P-1}{2}} \lambda_1^{\frac{n_1 b}{2}-1} \lambda_2^{\frac{n_2 b}{2}-1} g^{-\frac{P}{2}} b^{-\frac{P+1}{2}} |\boldsymbol{X}^T\boldsymbol{\Lambda}\boldsymbol{X}|^{\frac{1}{2}} |\mathbf{1}^T\boldsymbol{\Lambda}\mathbf{1}|^{-\frac{1}{2}} | \left(\frac{1+bg}{bg}\right) \boldsymbol{X}^T\boldsymbol{\Lambda}\boldsymbol{X} - \boldsymbol{X}^T\boldsymbol{H}_{\boldsymbol{\Lambda}}\boldsymbol{X}|^{-\frac{1}{2}} \times$$

$$J_{\boldsymbol{\Lambda}} \cdot \exp\{-\frac{b}{2}[\boldsymbol{Y}\boldsymbol{\Lambda}\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{H}_{\boldsymbol{\Lambda}}\boldsymbol{Y} - \boldsymbol{Y}^T(\boldsymbol{\Lambda} - \boldsymbol{H}_{\boldsymbol{\Lambda}})^T\boldsymbol{X}(\frac{1+bg}{bg}\boldsymbol{X}^T\boldsymbol{\Lambda}\boldsymbol{X} - \boldsymbol{X}^T\boldsymbol{H}_{\boldsymbol{\Lambda}}\boldsymbol{X})\boldsymbol{X}^T(\boldsymbol{\Lambda} - \boldsymbol{H}_{\boldsymbol{\Lambda}})\boldsymbol{Y}^T]\}$$

The joint mode $(\lambda_1^\star, \lambda_2^\star, g^\star)$ is computed using the function `optim` in R; similarly, the Hessian is computed at this value using the function `hessian` in the package `numderiv`.

1.3.1 Heteroscedastic Zellner-Siow Cauchy Result

Let $\boldsymbol{\beta}|\Phi, g \sim N(\mathbf{0}, g(\boldsymbol{X}^T\Phi\boldsymbol{X})^{-1})$ and $g \sim \text{IG}(\frac{1}{2}, \frac{N}{2})$. Then

$P(\boldsymbol{\beta}|g, \Phi) = (2\pi)^{-\frac{P}{2}} |g(\boldsymbol{X}^T\Phi\boldsymbol{X})^{-1}|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\boldsymbol{\beta}\frac{\boldsymbol{X}^T\Phi\boldsymbol{X}}{g}\boldsymbol{\beta}\}$ and $P(g) = \frac{\left(\frac{N}{2}\right)^{\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} g^{-\frac{3}{2}} \exp\{-\frac{N}{2g}\}$.

$$\begin{aligned}
P(\boldsymbol{\beta}|\boldsymbol{\varphi}) &= \int P(\boldsymbol{\beta}|\Phi, g)P(g)dg \\
&= \int (2\pi)^{-\frac{P}{2}} |g(\boldsymbol{X}^T\Phi\boldsymbol{X})^{-1}|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\boldsymbol{\beta}\frac{\boldsymbol{X}^T\Phi\boldsymbol{X}}{g}\boldsymbol{\beta}\} \times \frac{\left(\frac{N}{2}\right)^{\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} g^{-\frac{3}{2}} \exp\{-\frac{N}{2g}\}dg \\
&= \int (2\pi)^{-\frac{P}{2}} |\boldsymbol{X}^T\Phi\boldsymbol{X}|^{\frac{1}{2}} (\frac{N}{2})^{\frac{1}{2}} \pi^{-\frac{1}{2}} g^{-\frac{P+3}{2}} \exp\{-\frac{1}{2}\boldsymbol{\beta}^T\frac{\boldsymbol{X}^T\Phi\boldsymbol{X}}{g}\boldsymbol{\beta} - \frac{1}{2}\cdot\frac{N}{g}\}dg \\
&= \int (2\pi)^{-\frac{P}{2}} |\boldsymbol{X}^T\Phi\boldsymbol{X}|^{\frac{1}{2}} N^{\frac{1}{2}} 2^{-\frac{1}{2}} \pi^{-\frac{1}{2}} \times g^{-\frac{P+1}{2}-1} \exp\{-\frac{\boldsymbol{\beta}^T\boldsymbol{X}^T\Phi\boldsymbol{X}\boldsymbol{\beta}+N/2}{g}\}dg \\
&= 2^{-\frac{P+1}{2}} \pi^{-\frac{P+1}{2}} |\boldsymbol{X}^T\Phi\boldsymbol{X}|^{\frac{1}{2}} N^{\frac{1}{2}} [(\boldsymbol{\beta}^T\boldsymbol{X}^T\Phi\boldsymbol{X}\boldsymbol{\beta} + N)^{\frac{P+1}{2}} 2^{-\frac{P+1}{2}} \Gamma\left(\frac{P+1}{2}\right)]^{-1} \\
&= \pi^{-\frac{P+1}{2}} |\boldsymbol{X}^T\Phi\boldsymbol{X}|^{\frac{1}{2}} N^{\frac{1}{2}} (\boldsymbol{\beta}^T\boldsymbol{X}^T\Phi\boldsymbol{X}\boldsymbol{\beta} + N)^{-\frac{P+1}{2}} \Gamma\left(\frac{P+1}{2}\right)^{-1} \\
&= \pi^{-\frac{P+1}{2}} |\boldsymbol{X}^T\Phi\boldsymbol{X}|^{\frac{1}{2}} N^{\frac{1}{2}} (\frac{\boldsymbol{\beta}^T\boldsymbol{X}^T\Phi\boldsymbol{X}\boldsymbol{\beta} + N}{N})^{-\frac{P+1}{2}} N^{-\frac{P+1}{2}} \Gamma\left(\frac{P+1}{2}\right)^{-1} \\
&= \pi^{-\frac{P+1}{2}} |\boldsymbol{X}^T\Phi\boldsymbol{X}|^{\frac{1}{2}} N^{-\frac{P}{2}} (1 + \boldsymbol{\beta}^T\frac{\boldsymbol{X}^T\Phi\boldsymbol{X}}{N}\boldsymbol{\beta})^{-\frac{P+1}{2}} \Gamma\left(\frac{P+1}{2}\right)^{-1} \\
&= \pi^{-\frac{p+1}{2}} \Gamma\left(\frac{P+1}{2}\right)^{-1} \times |\frac{\boldsymbol{X}^T\Phi\boldsymbol{X}}{N}|^{\frac{1}{2}} (1 + \boldsymbol{\beta}^T\frac{\boldsymbol{X}^T\Phi\boldsymbol{X}}{N}\boldsymbol{\beta})^{-\frac{P+1}{2}} \\
&\sim \text{MVCauchy}_P(\text{location} = 0, \text{scale} = \left(\frac{\boldsymbol{X}^T\Phi\boldsymbol{X}}{N}\right)^{-1}). \ \square
\end{aligned}$$

1.4 Supplemental ANCOVA Simulation Study

We let $n_k := 10$ for each level of the SLGF. The parameter settings are provided in Table 1.

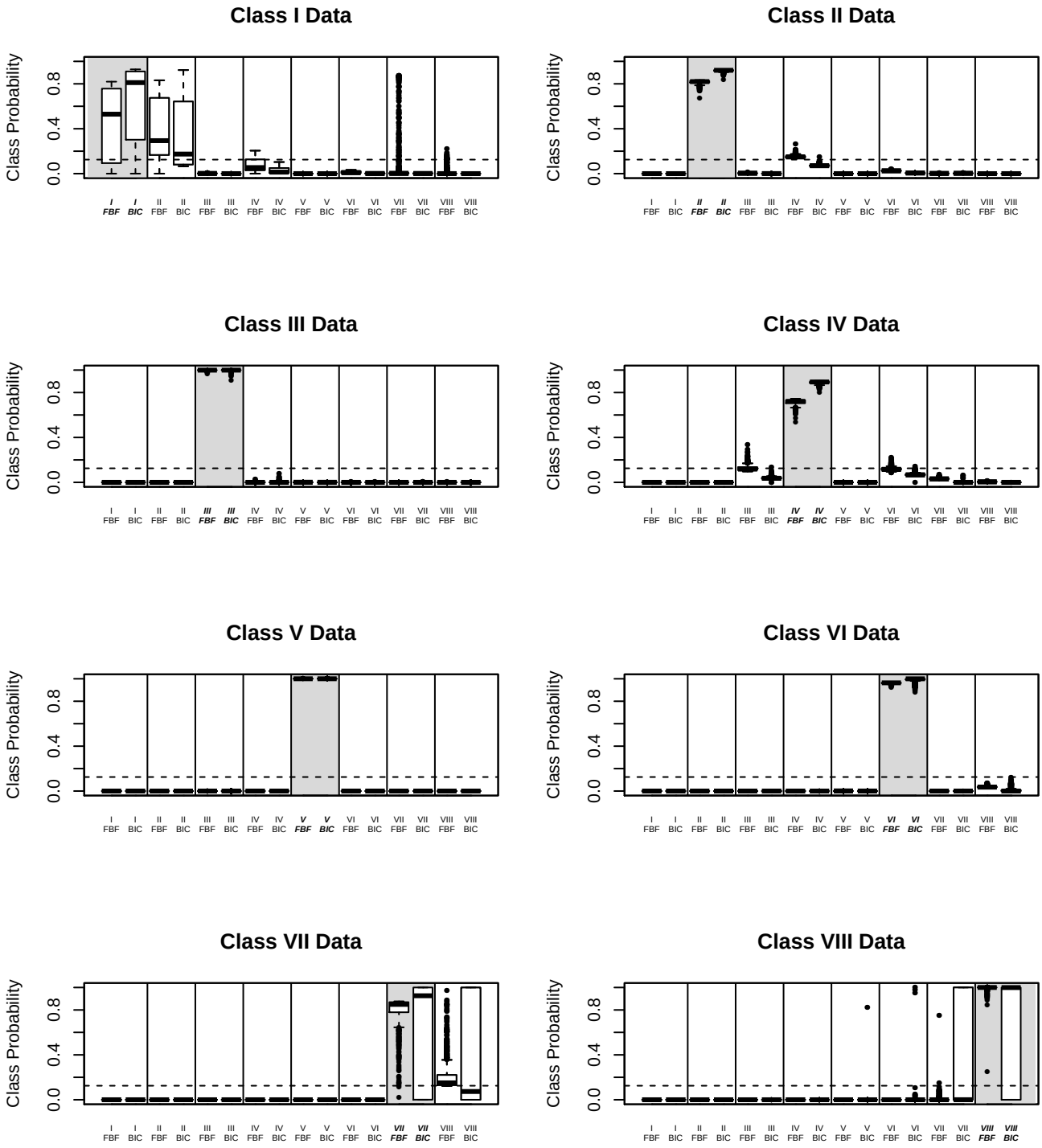

Figure 2: Posterior probabilities (y-axis) by class based on 1000 Monte Carlo data sets with $K = 4$ levels of the SLGF, each with $n_k = 50$ observations, for $N = 200$ total observations. The left boxplot in each class shows the posterior model probability based on FBF; the right boxplot in each class shows a BIC-approximated posterior model probability. The true model class is emphasized by the gray shaded region. The dashed line indicates the prior by model class.

Class	Parameters
I (Null)	$\alpha = 0$, $\sigma^2 = 1$
II (SLR)	$\alpha = 0$, $\tau = 10$, $\sigma^2 = 1$
III (ANCOVA)	$\alpha = 0.75$, $\nu = (1.50, 2.25, 3.00)$, $\tau = 0.25$, $\sigma^2 = 1$
IV (Group-Contracted ANCOVA)	$\alpha = 0$, $\tilde{\nu} = 5$, $\tau = 4$, $\sigma^2 = 1$
V (Interaction ANCOVA)	$\alpha = 3$, $\nu = (3, 6, 9)$, $\rho = (-2, -4, -6)$, $\tau = 2$, $\sigma^2 = 1$
VI (Group-Interaction ANCOVA)	$\alpha = 0$, $\tilde{\nu} = 0.8$, $\tilde{\rho} = -0.5$, $\tau = 1$, $\sigma^2 = 1$
VII (Heteroscedastic Group-Contracted ANCOVA)	$\alpha = 0$, $\tilde{\nu} = 3$, $\tau = -0.5$, $\sigma_1^2 = 1$, $\sigma_2^2 = 0.25$
VIII (Heteroscedastic Group-Interaction ANCOVA)	$\alpha = 0$, $\tilde{\nu} = 3$, $\tau = 2$, $\tilde{\rho} = 2$, $\sigma_1^2 = 1$, $\sigma_2^2 = 0.25$

Table 1: Settings for the eight model classes in the ANCOVA simulation studies where $K = 4$.

1.5 Supplemental Two-way Layout Simulation Study

We provide three additional simulation studies in the twoway layout scenario: 10×5 layouts with a smaller effect size than the study provided in Section **??**, as well as 5×5 studies with larger and smaller effect sizes.

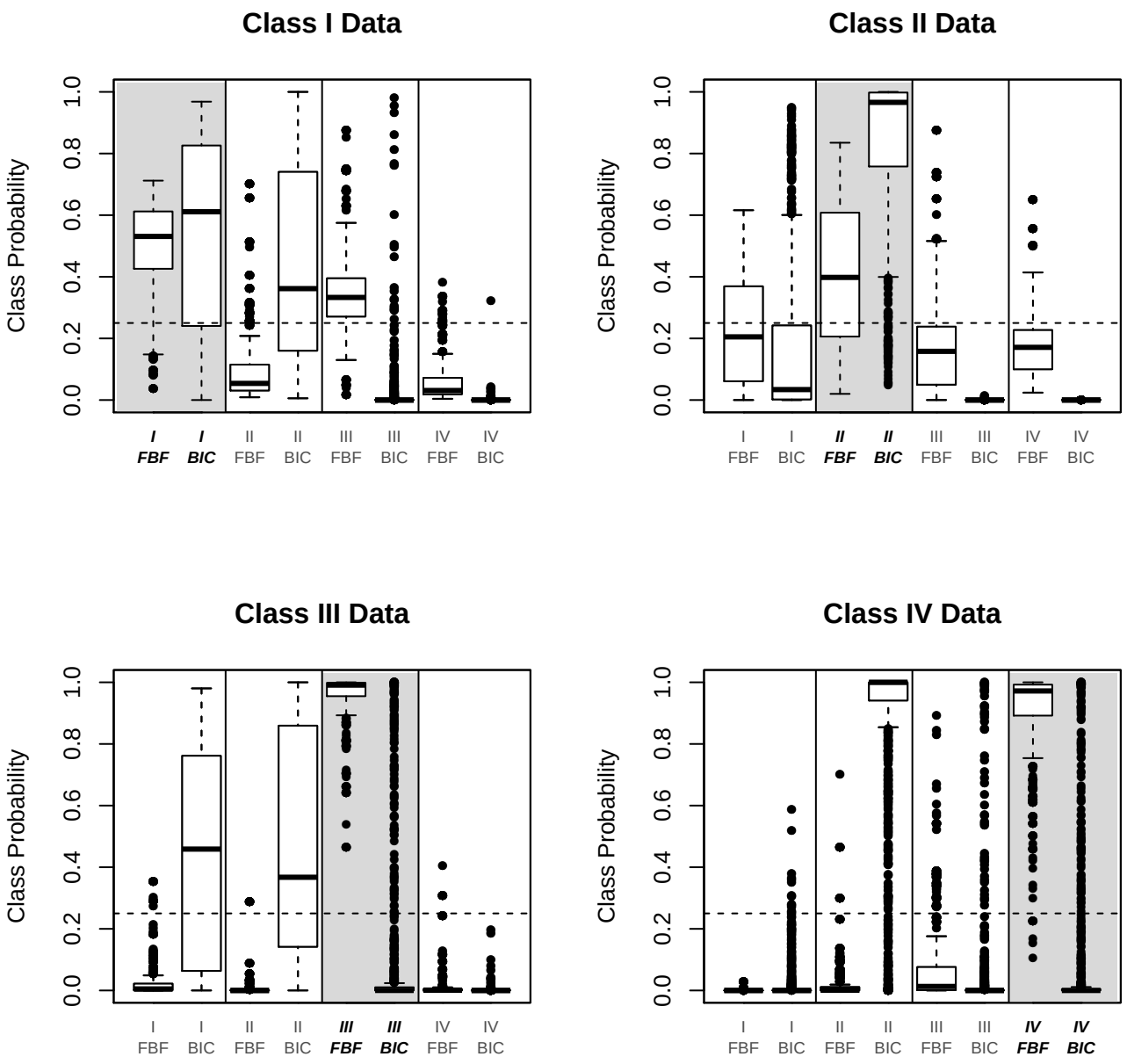

Figure 3: Posterior probabilities (y-axis) by class based on 1000 Monte Carlo 10×5 layouts with smaller effect size. The true model class is emphasized by the gray shaded region. The dashed line indicates the prior by model class.

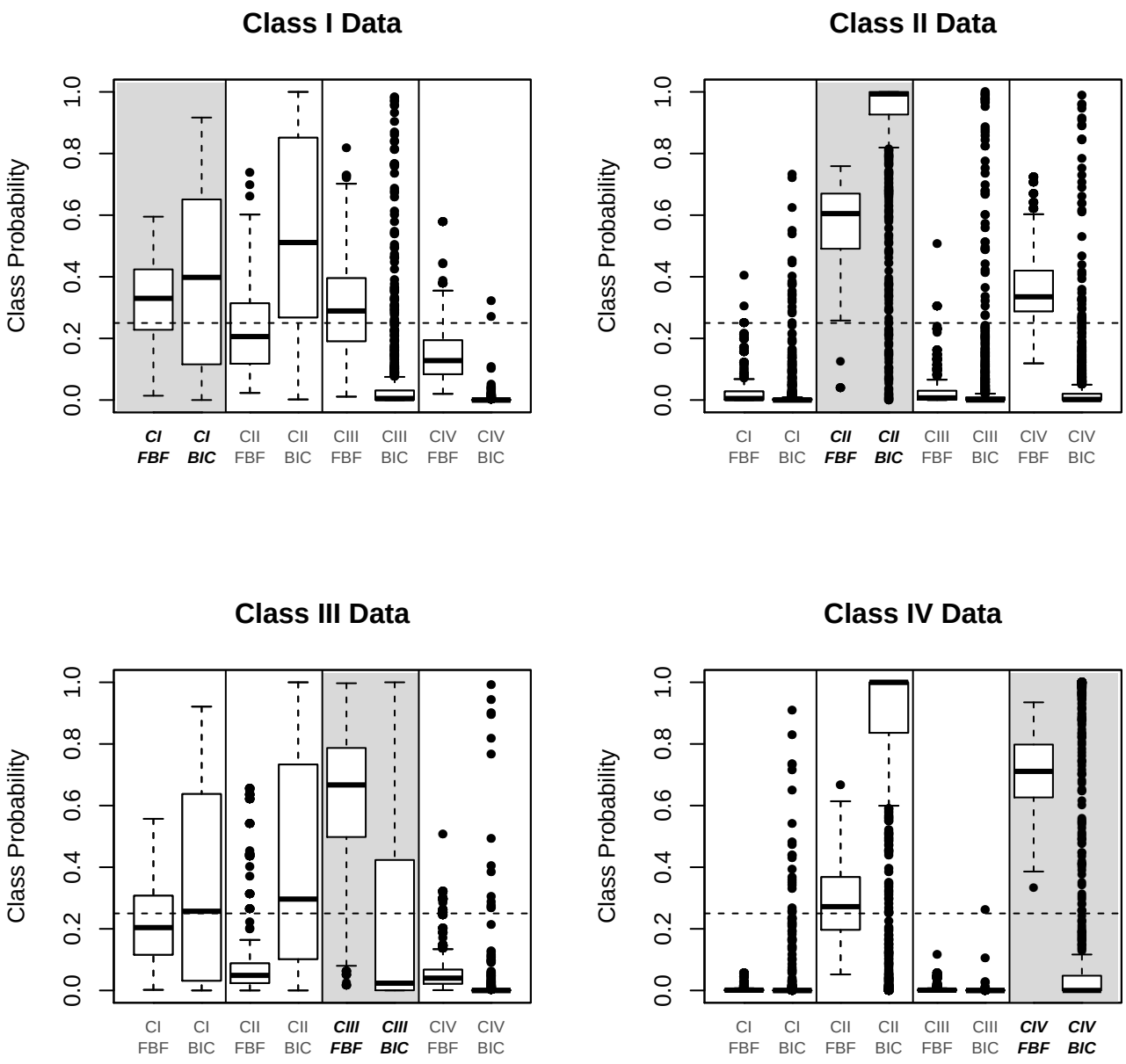

Figure 4: Posterior probabilities (y-axis) by class based on 1000 Monte Carlo 5×5 layouts with larger effect size. The true model class is emphasized by the gray shaded region. The dashed line indicates the prior by model class.

Class	Parameters
I (Additive Model)	$\alpha = 1$, $\boldsymbol{\nu} \in \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $\boldsymbol{\tau} \in \{1, 2, 3, 4, 5\}$, $\sigma^2 = 1$
II (Group-by-Column Interaction)	$\alpha = 1$, $\boldsymbol{\nu} \in \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $\boldsymbol{\tau}_1 \in \{1.0, 1.8, 2.6, 3.4, 4.2\}$, $\boldsymbol{\tau}_2 \in \{4.2, 3.4, 2.6, 1.8, 1.0\}$, $\sigma^2 = 1$
III (Heteroscedastic Additive)	$\alpha = 1$, $\boldsymbol{\nu} \in \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $\boldsymbol{\tau} \in \{1, 2, 3, 4, 5\}$, $\sigma_1^2 = 1.0$, $\sigma_2^2 = 0.10$
IV (Heteroscedastic Group-by-Column Interaction)	$\alpha = 1$, $\boldsymbol{\nu} \in \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $\boldsymbol{\tau}_1 \in \{1.0, 1.8, 2.6, 3.4, 4.2\}$, $\boldsymbol{\tau}_2 \in \{4.2, 3.4, 2.6, 1.8, 1.0\}$, $\sigma_1^2 = 1.0$, $\sigma_2^2 = 0.10$

Table 2: Settings for the four model classes in the two-way unreplicated layout simulation study with 10×5 layouts.

1.6 Analysis of Outliers

Detection of and response to outliers involves subjective decision-making. On the basis of Cook's distance based on the ANCOVA model, an analyst might flag observations 39, 40, and 41, all of which are potato, as influential. In the absence of specific additional knowledge regarding these data points, the analyst might consider removing the observations, which may potentially obfuscate a key component of the potato effect. Our method provides an alternative interpretation. We illustrate our method's sensitivity to outliers in this context with the Flurry data set. First, we review the five most probable models according to our method, with all data points in consideration:

Alternatively, we analyze the dataset with observations 39, 40, and 41 removed, obtaining the following five most probable models:

Class	Parameters
I (Additive Model)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5, 7, 8, 9, 10\}$, $\boldsymbol{\tau} \in \{1, 2, 3, 4, 5\}$, $\sigma^2 = 1$
II (Group-by-Column Interaction)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5, 7, 8, 9, 10\}$, $\boldsymbol{\tau}_1 \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$, $\boldsymbol{\tau}_2 \in \{3.0, 2.5, 2.0, 1.5, 1.0\}$, $\sigma^2 = 1$
III (Heteroscedastic Additive)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5, 7, 8, 9, 10\}$, $\boldsymbol{\tau} \in \{1, 2, 3, 4, 5\}$, $\sigma_1^2 = 1.0$, $\sigma_2^2 = 0.25$
IV (Heteroscedastic Group-by-Column Interaction)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5, 7, 8, 9, 10\}$, $\boldsymbol{\tau}_1 \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$, $\boldsymbol{\tau}_2 \in \{3.0, 2.5, 2.0, 1.5, 1.0\}$, $\sigma_1^2 = 1.0$, $\sigma_2^2 = 0.25$

Table 3: Settings for the four model classes in the 10×5 two-way layout simulation study with smaller effect size.

Thus we do of course see a major change in the results. Although this may reflect our method‘s sensitivity to outliers to some extent, we have nevertheless provided the researcher with meaningful information regarding the potato observations. We have indeed identified that they vary differently that the other levels. When analyzed without the outliers, we can still provide information that there may be a latent interaction with the scheme corn:canna,potato.

The preceding analysis offers the analyst a new choice. Rather than removing outliers and potentially artificially diminishing the evidence for heteroscedasticity, our method allows for a model-based interpretation for observed variability in instances where outliers appear.

Class	Parameters
I (Additive Model)	$\boldsymbol{\nu} \in \{1, 1.5, 2, 2.5, 3\}$, $\boldsymbol{\tau} \in \{1, 2, 3, 4, 5\}$, $\sigma^2 = 1$
II (Group-by-Column Interaction)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5\}$, $\boldsymbol{\tau}_1 \in \{1.0, 1.8, 2.6, 3.4, 4.2\}$, $\boldsymbol{\tau}_2 \in \{4.2, 3.4, 2.6, 1.8, 1.0\}$, $\sigma^2 = 1$
III (Heteroscedastic Additive)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5\}$, $\boldsymbol{\tau} \in \{1, 2, 3, 4, 5\}$, $\sigma_1^2 = 1.0$, $\sigma_2^2 = 0.25$
IV (Heteroscedastic Group-by-Column Interaction)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5\}$, $\boldsymbol{\tau}_1 \in \{1.0, 1.8, 2.6, 3.4, 4.2\}$, $\boldsymbol{\tau}_2 \in \{4.2, 3.4, 2.6, 1.8, 1.0\}$, $\sigma_1^2 = 1.0$, $\sigma_2^2 = 0.25$

Table 4: Settings for the four model classes in the two-way layout simulation study with 5×5 layouts with larger effect size.

Class	Parameters
I (Additive Model)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5\}$, $\boldsymbol{\tau} \in \{1, 2, 3, 4, 5\}$, $\sigma^2 = 1$
II (Group-by-Column Interaction)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5\}$, $\boldsymbol{\tau}_1 \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$, $\boldsymbol{\tau}_2 \in \{3.0, 2.5, 2.0, 1.5, 1.0\}$, $\sigma^2 = 1$
III (Heteroscedastic Additive)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5\}$, $\boldsymbol{\tau} \in \{1, 2, 3, 4, 5\}$, $\sigma_1^2 = 1.0$, $\sigma_2^2 = 0.25$
IV (Heteroscedastic Group-by-Column Interaction)	$\boldsymbol{\nu} \in \{1, 2, 3, 4, 5\}$, $\boldsymbol{\tau}_1 \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$, $\boldsymbol{\tau}_2 \in \{3.0, 2.5, 2.0, 1.5, 1.0\}$, $\sigma_1^2 = 1.0$, $\sigma_2^2 = 0.25$

Table 5: Settings for the four model classes in the two-way layout simulation study with 5×5 layouts with smaller effect size.

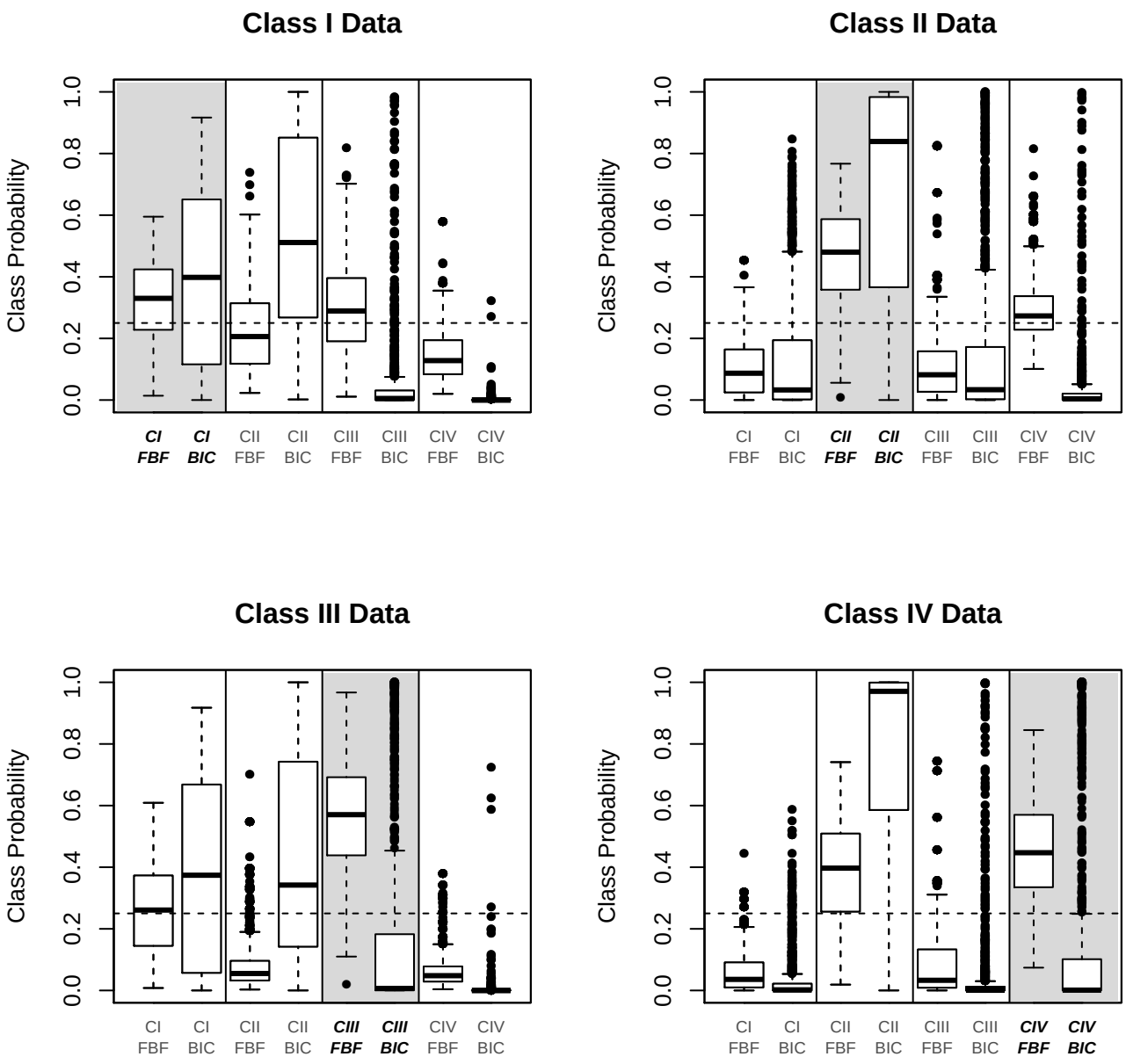

Figure 5: Posterior probabilities (y-axis) by class based on 1000 Monte Carlo 5×5 layouts with smaller effect size. The true model class is emphasized by the gray shaded region. The dashed line indicates the prior by model class.

Model	Scheme	Variance	$P(m_s^c\|\boldsymbol{Y})$
strength∼film+group	potato:canna,corn	Heteroscedastic	0.68755253
strength∼film+group+film*group	potato:canna,corn	Heteroscedastic	0.27411868
strength∼film	None	Homoscedastic	0.01314097
strength∼film+group+film*group	corn:canna,potato	Homoscedastic	0.00864268
strength∼film+starch+film*starch	None	Homoscedastic	0.00455560

Table 6: The five most probable models for the Flurry (1939) dataset, where outliers are included in the analysis. Our method strongly favors heteroscedasticity with the scheme potato:canna,corn.

Model	Scheme	Variance	$P(m_s^c\|\boldsymbol{Y})$
strength∼film+group+film*group	corn:canna,potato	Homoscedastic	0.84545662
strength∼film+starch+film*starch	None	Homoscedastic	0.15245441
strength∼film	None	Homoscedastic	0.00092031
strength∼film+group	canna:corn,potato	Homoscedastic	0.00026691
strength∼film+starch	None	Homoscedastic	0.00023384

Table 7: The five most probable models for the Flurry (1939) dataset, where outliers have been excluded. Our method now favors a homoscedastic model with a film, group, and film by group interaction with scheme corn:canna,potato.

References

Flurry, M. S. (1939). Breaking strength, elongation and folding endurance of films of starches and gelatin used in sizing. *Technical Bulletin (United States Department of Agriculture)*, 674:1–36.

Gilbert, P. and Varadhan, R. (2016). *numDeriv: Accurate Numerical Derivatives*. R package version 2016.8-1.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.