Web-based supplementary materials for "The LZIP: A Bayesian latent factor model for correlated zero-inflated counts"

Brian Neelon* **and Dongjun Chung**

Medical University of South Carolina, Charleston, South Carolina, U.S.A

*email: neelon@musc.edu

Web Appendix A. Derivation of Cov(Z_{i1}, Z_{i2})

If we assume independent Ga(α,α) priors for $\xi_{il}\,(l=1,\ldots,L)$, then $\mathrm{V}(\xi_i)=\alpha^{-1}$ and $\mathrm{V}(\boldsymbol{\xi}_i)=\mathrm{diag}(\alpha^{-1})$, where $\boldsymbol{\xi}_i=(\xi_{i1},\ldots,\xi_{iL})'$. Therefore,

$$\begin{aligned}
\mathrm{Cov}(Z_{i1},Z_{i2}) &= \mathrm{Cov}_{\boldsymbol{\xi}_i}\left[\mathrm{E}(Z_{i1}|\boldsymbol{\xi}_i),\mathrm{E}(Z_{i2}|\boldsymbol{\xi}_i)\right]+\underbrace{\mathrm{E}_{\boldsymbol{\xi}_i}\left[\mathrm{Cov}(Z_{i1},Z_{i2}|\boldsymbol{\xi}_i)\right]}_{=0\text{ by conditional independence}}\\
&= \mathrm{Cov}_{\boldsymbol{\xi}_i}\left[\boldsymbol{\lambda}_1'\boldsymbol{\xi}_i\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_1),\boldsymbol{\lambda}_2'\boldsymbol{\xi}_i\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_2)\right]\\
&= \left[\boldsymbol{\lambda}_1'\mathrm{Var}(\boldsymbol{\xi}_i)\boldsymbol{\lambda}_2\right]\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_1)\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_2)\\
&= \left[\boldsymbol{\lambda}_1'\mathrm{diag}(\alpha^{-1})\boldsymbol{\lambda}_2\right]\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_1+\boldsymbol{x}_{ij}'\boldsymbol{\beta}_2)\\
&= \alpha^{-1}\left(\sum_{l=1}^{L}\lambda_{1l}\lambda_{2l}\right)\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_1+\boldsymbol{x}_{ij}'\boldsymbol{\beta}_2).
\end{aligned}$$

Setting $\alpha=1$, we obtain expression (8) in the manuscript.

Web Appendix B: Proof of Proposition 1

Let $\boldsymbol{\xi}_i=(\xi_{i1},\ldots,\xi_{iL})'$, where $\{\xi_{il}\}$ are independent Ga(α,α) random variables. Then,

$$\begin{aligned}
\mathrm{p}(z_{i11},\ldots,z_{iJ2}) &= \int_{\boldsymbol{\xi}_i}\prod_{j=1}^{J}\prod_{k=1}^{K}\mathrm{p}(z_{ijk}|\boldsymbol{\xi}_i)f(\boldsymbol{\xi}_i;\alpha)\,\mathrm{d}\boldsymbol{\xi}_i\\
&= \int_{\boldsymbol{\xi}_i}\prod_{j=1}^{J}\prod_{k=1}^{K}\mathrm{Poi}(z_{ijk}|\mu_{ijk})f(\boldsymbol{\xi}_i;\alpha)\,\mathrm{d}\boldsymbol{\xi}_i,\ \text{ where } \mu_{ijk}=\boldsymbol{\lambda}_{jk}'\boldsymbol{\xi}_i\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_{jk}).
\end{aligned}$$

From equation (13), we have $Z_{ijk}=\sum_{l=1}^{L}Z_{ijkl}$, where $Z_{ijkl}|\xi_{il}\overset{ind}{\sim}\mathrm{Poi}(\mu_{ijkl})$, $\mu_{ijkl}=\lambda_{jkl}\xi_{il}\exp(\boldsymbol{x}_{ij}'\boldsymbol{\beta}_{jk})$, and $\mu_{ijk}=\sum_{l=1}^{L}\mu_{ijkl}$. By moment generating function theory, $\mathrm{p}(z_{ijk}|\boldsymbol{\xi}_i)=\prod_{l=1}^{L}\ \mathrm{Poi}(z_{ijkl}|\mu_{ijkl})$, and hence, by independence of $\{\xi_{il}\}$, we have

$$\begin{aligned}
\mathrm{p}(z_{i11},\ldots,z_{iJ2}) &= \int_{\boldsymbol{\xi}_i} \prod_{j=1}^{J}\prod_{k=1}^{K}\prod_{l=1}^{L} \mathrm{Poi}(z_{ijkl}|\mu_{ijkl}) f(\xi_{il};\alpha)\,\mathrm{d}\xi_{il} \\
&= \prod_{l=1}^{L}\int_{\xi_{il}} \left\{ \prod_{j=1}^{J}\prod_{k=1}^{K} \frac{[\lambda_{jkl}\xi_{il}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{jk})]^{z_{ijkl}}\exp[-\lambda_{jkl}\xi_{il}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{jk})]}{z_{ijkl}!}\times \right.\\
&\qquad\qquad \left. \frac{\alpha^{\alpha}\xi_{il}^{\alpha-1}\exp(-\alpha\xi_{il})}{\Gamma(\alpha)} \right\}\mathrm{d}\xi_{il} \\
&= \prod_{l=1}^{L}\left\{ \frac{\prod_{j,k}\eta_{ijkl}^{z_{ijkl}}\alpha^{\alpha}}{\Gamma(\alpha)\prod_{j,k} z_{ijkl}!}\int_{\xi_{il}} \xi_{il}^{\sum_{j,k} z_{ijkl}+\alpha-1}\exp\left[\left(-\textstyle\sum_{j,k}\eta_{ijkl}+\alpha\right)\xi_{il}\right]\mathrm{d}\xi_{il}\right\},
\end{aligned}$$

where $\eta_{ijkl}=\lambda_{jkl}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{jk})$. The integrand can be recognized as the kernel of a $\mathrm{Ga}(z_{il}+\alpha,\eta_{il}+\alpha)$ distribution, where $z_{il}=\sum_{j,k} z_{ijkl}$ and $\eta_{il}=\sum_{j,k}\eta_{ijkl}$. Thus, we have

$$\begin{aligned}
\mathrm{p}(z_{i11},\ldots,z_{iJ2}) &= \prod_{l=1}^{L}\frac{\Gamma(z_{il}+\alpha)\prod_{j,k}\eta_{ijkl}^{z_{ijkl}}\alpha^{\alpha}}{\Gamma(\alpha)\prod_{j,k} z_{ijkl}!(\eta_{il}+\alpha)^{\sum_{j,k} z_{ijkl}+\alpha}}\underbrace{\int_{\xi_{il}}\mathrm{Ga}\left(z_{il}+\alpha,\eta_{il}+\alpha\right)\mathrm{d}\xi_{il}}_{=1} \\
&= \prod_{l=1}^{L}\frac{\Gamma(z_{il}+\alpha)}{\Gamma(\alpha)\prod_{j,k} z_{ijkl}!}\left(\frac{\alpha}{\eta_{il}+\alpha}\right)^{\alpha}\prod_{j,k}\left(\frac{\eta_{ijkl}}{\eta_{il}+\alpha}\right)^{z_{ijkl}},
\end{aligned}$$

which is the probability distribution function for the product of L independent $\mathrm{NegMult}(\alpha,\pi_{i11l},\ldots,\pi_{iJ2l})$ random variables, where $\pi_{ijkl}=\eta_{ijkl}/(\eta_{il}+\alpha)$. A similar approach can be used to show that any subset of $\mathrm{p}(z_{i11},\ldots,z_{iJ2})$ is also product negative multinomial. In particular, we can derive equation (11) in the manuscript by noting that $\mathrm{p}(z_{i11},z_{i21},\ldots,z_{iJ1})$ is the product of L independent $\mathrm{NegMult}(\alpha,\pi_{ij1l},\ldots,\pi_{iJ1l})$ random variables, and hence

$$\begin{aligned}
\psi_i &= 1-\prod_{l=1}^{L}\Pr(z_{i11l}=0,\ldots,z_{iJ1l}=0) = 1-\prod_{l=1}^{L}\left(\frac{\alpha}{\alpha+\sum_{j=1}^{J}\eta_{ij1l}}\right)^{\alpha} \\
&= 1-\prod_{l=1}^{L}\left[\frac{\alpha}{\alpha+\sum_{j=1}^{J}\lambda_{j1l}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{j1})}\right]^{\alpha},
\end{aligned}$$

as in equation (11). Setting α and J to 1 yields equation (6) as a special case. Using the above integration, we can also show that the univariate marginal distribution of Z_{ijkl} is $\mathrm{NegBin}[\alpha,\eta_{ijkl}/(\alpha+\eta_{ijkl})]$ with mean $\mathrm{E}(Z_{ijkl})=\eta_{ijkl}=\lambda_{jkl}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{jk})$. Thus, $\mathrm{E}(Z_{ij2})=\sum_{l=1}^{L}\mathrm{E}(Z_{ij2l})=\sum_{l=1}^{L}\eta_{ij2l}=\left(\sum_{l=1}^{L}\lambda_{j2l}\right)\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{j2})$ as in expression (12). Setting $J=1$ we obtain expression (5).

Web Appendix C: MCMC Algorithm

1. *Data Augmentation Step 1.* For all (i, j), we first introduce the latent Poisson random variables, Z_{ij1} and Z_{ij2}. The update for Z_{ij1}, the latent Poisson for the binary component of the LZIP, depends on the observed response y_{ij} and the current value of Z_{ij2}, the latent Poisson for the count component. In particular, the following sampling rules hold:

   (a) If $y_{ij} > 0$, then we know subject i is "at-risk" for outcome j and hence $Z_{ij1} > 0$. Therefore, update Z_{ij1} from $\text{TPoi}(\mu_{ij1})$, where $\text{TPoi}(\mu)$ denotes a Poisson distribution with mean μ truncated at 0, and μ_{ij1} is defined in equation (9) of the manuscript.

   (b) Next, consider the case where $y_{ij} = 0$. Note first that $y_{ij} = 0$ i.f.f. at least one (or both) of Z_{ij1} or Z_{ij2} equals zero. In the case where $y_{ij} = Z_{ij2} = 0$, then any non-negative integer value for Z_{ij1} is consistent with $y_{ij} = 0$. Therefore, update Z_{ij1} from $\text{Poi}(\mu_{ij1})$;

   (c) Otherwise, if $y_{ij} = 0$ and $Z_{ij2} > 0$, set $Z_{ij1} = 0$ to ensure $y_{ij} = 0$.

2. For all (i, j), the update for Z_{ij2} depends on the current value of Z_{ij1}:

   (a) Consider the case where $Z_{ij1} = 0$. By the contrapositive of 1(a) above, $Z_{ij1} = 0 \Rightarrow y_{ij} = 0$ (i.e., subject i is *not* at risk for outcome j and hence a zero must be observed). Because $Z_{ij1} = 0$, any count value for Z_{ij2} is consistent with $y_{ij} = 0$. Therefore, draw Z_{ij2} from a $\text{Poi}(\mu_{ij2})$ distribution, where μ_{ij2} is defined in equation (9);

   (b) If $Z_{ij1} > 0$, then subject i is at risk for outcome j. In this case, set $Z_{ij2} = y_{ij}$.

3. *Data Augmentation Step 2.* Assuming L latent factors, let

$$
\begin{aligned}
Z_{ijk} &= \sum_{l=1}^{L} Z_{ijkl}, \text{ where} \\
Z_{ijkl} &\overset{ind}{\sim} \text{Poi}(\mu_{ijkl})
\end{aligned}
$$

   and μ_{ijkl} is defined in equation (13). Next, for all $i = 1, \ldots, n$, $j = 1 \ldots, J$, and $k = 1, 2$, the joint full conditional for the L random variables $(Z_{ijk1}, \ldots, Z_{ijkL})$ given Z_{ijk} is

$$
\begin{aligned}
\Pr(Z_{ijk1} = z_{ij1}, \ldots, Z_{ijkL} = z_{ijkL} | Z_{ijk} = z_{ijk}, \text{rest}) &= \frac{\prod_{l=1}^{L} \left(\mu_{ijkl}^{z_{ijkl}} \mathrm{e}^{-\mu_{ijkl}} / z_{ijkl}! \right)}{\mu_{ijk}^{z_{ijk}} \mathrm{e}^{-\mu_{ijk}} / z_{ijk}!} \\
&= \frac{z_{ijk}! \prod_{l=1}^{L} \mu_{ijkl}^{z_{ijkl}}}{\left(\prod_{l=1}^{L} z_{ijkl}! \right) \mu_{ijk}^{z_{ijk}}} \\
&= \frac{z_{ijk}!}{\prod_{l=1}^{L} z_{ijkl}!} \prod_{l=1}^{L} \left(\frac{\mu_{ijkl}}{\mu_{ijk}} \right)^{z_{ijkl}} \\
&\sim \text{Multinom}(z_{ijk}, \pi_{ijk1}, \ldots, \pi_{ijkL}),
\end{aligned}
$$

   where $\pi_{ijkl} = \mu_{ijkl}/\mu_{ijk}$, μ_{ijkl} is defined in equation (13), and $\mu_{ijk} = \sum_{l=1}^{L} \mu_{ijkl}$ is defined in equation (9). In the case of a single latent factor, $Z_{ijk1} = Z_{ijk}$ and this step is omitted.

4. *Update* ξ_{il}. Conditional on the $2J \times 1$ vector $\boldsymbol{Z}_{il} = (Z_{i11l}, \ldots, Z_{iJ2l})'$, ξ_{il} $(i = 1, \ldots, n;\ l = 1, \ldots, L)$ has a gamma full conditional:

$$\begin{aligned}
\xi_{il} | \boldsymbol{Z}_{il} = \boldsymbol{z}_{il}, \text{rest} \quad &\propto \quad \xi_{il}^{\sum_{j,k} z_{ijkl}} \exp\Big(- \xi_{il} \underbrace{\sum_{j,k} \lambda_{jkl} \mathrm{e}^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{jk}}}_{\eta_{il} \text{ from eq. (10)}} \Big) \cdot \xi_{il}^{\alpha-1} \exp(-\alpha \xi_{il}) \\
&\sim \quad \mathrm{Ga}\left(\alpha + \sum_{j,k} z_{ijkl},\ \alpha + \eta_{il} \right),
\end{aligned}$$

where η_{il} is defined in equation (10), and the prior shape and rate parameter, α, is fixed at 1 to allow for unrestricted factor loadings.

5. *Update* λ_{jkl}. Assume a $\mathrm{Ga}(a, b)$ for λ_{jkl}. Conditional on the $n \times 1$ vector $\boldsymbol{Z}_{jkl} = (Z_{1jkl}, \ldots, Z_{njkl})'$, update λ_{jkl} $(j = 1, \ldots, J;\ k = 1, 2;\ l = 1, \ldots, L)$ from its gamma full conditional:

$$\begin{aligned}
\lambda_{jkl} | \boldsymbol{Z}_{jkl} = \boldsymbol{z}_{jkl}, \text{rest} \quad &\propto \quad \lambda_{jkl}^{\sum_{i=1}^{n} z_{ijkl}} \exp\left(-\lambda_{jkl} \sum_{i=1}^{n} \xi_{il} \mathrm{e}^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{jk}} \right) \cdot \lambda_{jkl}^{a-1} \exp(-b\lambda_{jkl}) \\
&\sim \quad \mathrm{Ga}\left(a + \sum_{i=1}^{n} z_{ijkl},\ b + \sum_{i=1}^{n} \xi_{il} \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_{jk}) \right).
\end{aligned}$$

6. *Update* β_{jkh}. Without loss of generality, assume identical covariates in the binary and count components of the LZIP; that is $\boldsymbol{x}_{ij1} = \boldsymbol{x}_{ij2} = \boldsymbol{x}_{ij}$. The update for the (jkh)-th regression parameter, β_{jkh} $(j = 1, \ldots, J;\ k = 1, 2;\ h = 1, \ldots, p)$, depends on whether the corresponding covariate, x_{ijh}, is discrete or continuous. For categorical predictors, a $\mathrm{Ga}(c, d)$ prior on $\exp(\beta_{jkh})$ is conditionally conjugate, allowing for straightforward Gibbs sampling. For example, if x_{ijh} is dichotomous taking values 0 and 1, the full conditional for $\exp(\beta_{jkh})$ is

$$\begin{aligned}
\exp(\beta_{jkh}) | \boldsymbol{Z}_{jk} = \boldsymbol{z}_{jk}, \text{rest} \quad &\propto \quad \prod_{i: x_{ijh}=1} \left\{ \prod_{l=1}^{L} \exp(\beta_{jkh})^{z_{ijkl}} \exp\left[- \left(\lambda_{jkl} \xi_{il} \mathrm{e}^{\tilde{\boldsymbol{x}}'_{ij}\tilde{\boldsymbol{\beta}}_{jk}} \right) \mathrm{e}^{\beta_{jkh}} \right] \right\} \times \\
&\qquad \exp(\beta_{jkh})^{c-1} \exp\left(-d \mathrm{e}^{\beta_{jkh}} \right) \\
&\sim \quad \mathrm{Ga}\left(c + \sum_{i: x_{ijh}=1} \underbrace{\sum_{l=1}^{L} z_{ijkl}}_{z_{ijk}},\ d + \sum_{i: x_{ijh}=1} \underbrace{\sum_{l=1}^{L} \lambda_{jkl} \xi_{il}}_{\boldsymbol{\lambda}'_{jk}\boldsymbol{\xi}_i} \exp(\tilde{\boldsymbol{x}}'_{ij}\tilde{\boldsymbol{\beta}}_{jk}) \right) \\
&\sim \quad \mathrm{Ga}\left(c + \sum_{i: x_{ijh}=1} z_{ijk},\ d + \sum_{i: x_{ijh}=1} \boldsymbol{\lambda}'_{jk}\boldsymbol{\xi}_i \exp(\tilde{\boldsymbol{x}}'_{ij}\tilde{\boldsymbol{\beta}}_{jk}) \right),
\end{aligned}$$

where $\boldsymbol{Z}_{jk} = (Z_{1jk}, \ldots, Z_{njk})'$, $\tilde{\boldsymbol{x}}_{ij}$ is $\boldsymbol{x}_{ij}$ with x_{ijh} removed, and $\tilde{\boldsymbol{\beta}}_{jk}$ is $\boldsymbol{\beta}_{jk}$ with β_{jkh} removed. When x_{ijh} has more than two categories, we introduce indicators for each category level; in this case, the update for the category-specific β's will have the same form as above. When x_{ijh} is ordinal or continuous, we update β_{jkh} using a random-walk Metropolis-Hastings step.

Web Appendix D: Web Tables

Web Table 1: Posterior means and 95% credible intervals (CrIs) for simulation study 3: bivariate LZIP model with a two latent factors. Results are for simulation with 40% zeros and $\text{Ga}(1,1)$ priors for both the factor loadings and the exponentiated regression coefficients for the binary predictor, x_{ij1}.

n	Outcome	Model Component	Parameter	Simulated Value	Posterior Mean (95% CrI)
500	Y_1	Binary	λ_{111}	2.50	$1.88\,(0.94, 3.31)$
			λ_{112}	0.00	$0.18\,(0.01, 0.49)$
			$\beta_{111}^{\dagger}$	1.00	$0.64\,(0.08, 1.17)$
			$\beta_{112}^{\ddagger}$	0.50	$0.39\,(0.21, 0.58)$
		Count	λ_{121}	0.00	$0.11\,(0.00, 0.31)$
			λ_{122}	2.50	$2.49\,(1.97, 3.13)$
			β_{121}	0.25	$0.16\,(-0.10, 0.42)$
			β_{122}	-0.25	$-0.18\,(-0.24, -0.12)$
	Y_2	Binary	λ_{211}	2.50	$1.12\,(0.50, 2.11)$
			λ_{212}	0.00	$0.37\,(0.08, 0.77)$
			β_{211}	0.75	$0.47\,(0.02, 0.92)$
			β_{212}	0.25	$0.00\,(-0.13, 0.14)$
		Count	λ_{221}	0.00	$0.13\,(0.01, 0.33)$
			λ_{222}	2.50	$2.60\,(2.05, 3.27)$
			β_{221}	0.50	$0.38\,(0.12, 0.63)$
			β_{222}	-0.50	$-0.45\,(-0.51, -0.38)$
5000	Y_1	Binary	λ_{111}	2.50	$2.31\,(1.82, 2.88)$
			λ_{112}	0.00	$0.00\,(0.00, 0.06)$
			β_{111}	1.00	$1.04\,(0.82, 1.24)$
			β_{112}	0.50	$0.54\,(0.47, 0.62)$
		Count	λ_{121}	0.00	$0.00\,(0.00, 0.11)$
			λ_{122}	2.50	$2.62\,(2.40, 2.84)$
			β_{121}	0.25	$0.17\,(0.08, 0.26)$
			β_{122}	-0.25	$-0.26\,(-0.28, -0.24)$
	Y_2	Binary	λ_{211}	2.50	$2.22\,(1.72, 2.59)$
			λ_{212}	0.00	$0.04\,(0.00, 0.10)$
			β_{211}	0.75	$0.70\,(0.50, 0.89)$
			β_{212}	0.25	$0.25\,(0.20, 0.31)$
		Count	λ_{221}	0.00	$0.04\,(0.00, 0.09)$
			λ_{222}	2.50	$2.51\,(2.30, 2.73)$
			β_{221}	0.50	$0.48\,(0.39, 0.57)$
			β_{222}	-0.50	$-0.51\,(-0.53, -0.49)$

* Estimates rounded to two decimal places.
† Regression coefficients for binary predictor, x_{ij1}, updated using conjugate Gibbs steps.
‡ Regression coefficients for continuous predictor, x_{ij2}, updated using random-walk Metropolis-Hastings steps.

Web Table 2: Posterior means and 95% credible intervals (CrIs) for simulation study 3: bivariate LZIP model with a two latent factors. Results are for simulation with 70% zeros and Ga(0.001, 0.001) priors for both the factor loadings and the exponentiated regression coefficients for the binary predictor, x_{ij1}.

n	Outcome	Model Component	Parameter	Simulated Value	Posterior Mean (95% CrI)
500	Y_1	Binary	λ_{111}	1.00	0.68 (0.44, 1.03)
			λ_{112}	0.00	0.00 (0.00, 0.00)*
			$\beta_{111}^{\dagger}$	0.75	1.15 (0.49, 1.89)
			$\beta_{112}^{\ddagger}$	−0.25	−0.33 (−0.53, −0.12)
		Count	λ_{121}	0.00	0.00 (0.00, 0.00)
			λ_{122}	1.50	1.13 (0.90, 1.42)
			β_{121}	−0.50	−0.65 (−0.99, −0.32)
			β_{122}	0.75	0.71 (0.61, 0.82)
	Y_2	Binary	λ_{211}	1.50	2.95 (1.60, 5.49)
			λ_{212}	0.00	0.00 (0.00, 0.00)
			β_{211}	−0.25	−0.60 (−1.38, 0.16)
			β_{212}	0.75	0.82 (0.59, 1.03)
		Count	λ_{221}	0.00	0.00 (0.00, 0.00)
			λ_{222}	1.00	0.82 (0.62, 1.06)
			β_{221}	0.75	0.71 (0.24, 1.08)
			β_{222}	−0.50	−0.55 (−0.65, −0.43)
5000	Y_1	Binary	λ_{111}	1.00	1.21 (1.00, 1.44)
			λ_{112}	0.00	0.00 (0.00, 0.00)
			$\beta_{111}^{\dagger}$	0.75	0.76 (0.56, 0.96)
			$\beta_{112}^{\ddagger}$	−0.25	−0.32 (−0.49, −0.24)
		Count	λ_{121}	0.00	0.00 (0.00, 0.00)
			λ_{122}	1.50	1.34 (1.22, 1.48)
			β_{121}	−0.50	−0.45 (−0.57, −0.35)
			β_{122}	0.75	0.78 (0.75, 0.87)
	Y_2	Binary	λ_{211}	1.50	1.59 (1.31, 1.93)
			λ_{212}	0.00	0.00 (0.00, 0.00)
			β_{211}	−0.25	−0.21 (−0.46, −0.01)
			β_{212}	0.75	0.80 (0.73, 0.85)
		Count	λ_{221}	0.00	0.00 (0.00, 0.00)
			λ_{222}	1.00	0.96 (0.87, 1.06)
			β_{221}	0.75	0.77 (0.65, 0.90)
			β_{222}	−0.50	−0.50 (−0.53, −0.46)

* Estimates rounded to two decimal places.

† Regression coefficients for binary predictor, x_{ij1}, updated using conjugate Gibbs steps.

‡ Regression coefficients for continuous predictor, x_{ij2}, updated using random-walk Metropolis-Hastings steps.

Web Table 3: Posterior means and 95% credible intervals (CrIs) for simulation study 3: bivariate LZIP model with a two latent factors. Results are for simulation with 70% zeros and Ga(1, 1) priors for both the factor loadings and the exponentiated regression coefficients for the binary predictor, x_{ij1}.

n	Outcome	Model Component	Parameter	Simulated Value	Posterior Mean (95% CrI)
500	Y_1	Binary	λ_{111}	1.00	0.49 (0.16, 0.90)
			λ_{112}	0.00	0.15 (0.00, 0.49)
			$\beta_{111}^{\dagger}$	0.75	0.77 (0.24, 1.29)
			$\beta_{112}^{\ddagger}$	−0.25	−0.26 (−0.47, −0.06)
		Count	λ_{121}	0.00	0.48 (0.07, 1.20)
			λ_{122}	1.50	0.99 (0.05, 1.62)
			β_{121}	−0.50	−0.70 (−1.05, −0.37)
			β_{122}	0.75	0.70 (0.56, 0.82)
	Y_2	Binary	λ_{211}	1.50	1.42 (0.24, 2.81)
			λ_{212}	0.00	0.41 (0.01, 1.61)
			β_{211}	−0.25	−0.41 (−1.04, 0.22)
			β_{212}	0.75	0.76 (0.50, 1.00)
		Count	λ_{221}	0.00	0.00 (0.00, 0.00)*
			λ_{222}	1.00	0.82 (0.62, 1.06)
			β_{221}	0.75	0.71 (0.24, 1.08)
			β_{222}	−0.50	−0.55 (−0.65, −0.43)
5000	Y_1	Binary	λ_{111}	1.00	0.89 (0.70, 1.10)
			λ_{112}	0.00	0.04 (0.00, 0.12)
			$\beta_{111}^{\dagger}$	0.75	0.82 (0.62, 1.02)
			$\beta_{112}^{\ddagger}$	−0.25	−0.28 (−0.35, −0.21)
		Count	λ_{121}	0.00	0.02 (0.00, 0.06)
			λ_{122}	1.50	1.44 (1.30, 1.59)
			β_{121}	−0.50	−0.57 (−0.68, −0.46)
			β_{122}	0.75	0.77 (0.74, 0.80)
	Y_2	Binary	λ_{211}	1.50	1.44 (1.07, 1.91)
			λ_{212}	0.00	0.05 (0.00, 0.18)
			β_{211}	−0.25	−0.36 (−0.58, −0.15)
			β_{212}	0.75	0.73 (0.62, 0.85)
		Count	λ_{221}	0.00	0.02 (0.00, 0.08)
			λ_{222}	1.00	0.96 (0.86, 1.07)
			β_{221}	0.75	0.69 (0.56, 0.81)
			β_{222}	−0.50	−0.53 (−0.57, −0.49)

* Estimate rounded to two decimal places.
† Regression coefficients for binary predictor, x_{ij1}, updated using conjugate Gibbs steps.
‡ Regression coefficients for continuous predictor, x_{ij2}, updated using random-walk Metropolis-Hastings steps.

Web Table 4: WAIC results for simulation studies.

True Model	Fitted Model: No Factors	One Factor	Two Factors
Simulation Study 1 (No-Factor Model)	**1598**	—*	—*
Simulation Study 2 (One-Factor Model)	2293	**1632**	1673
Simulation Study 3 (Two-Factor Model)	5527	3482	**3419**

* Fitted model did not converge.

† Preferred model in bold.

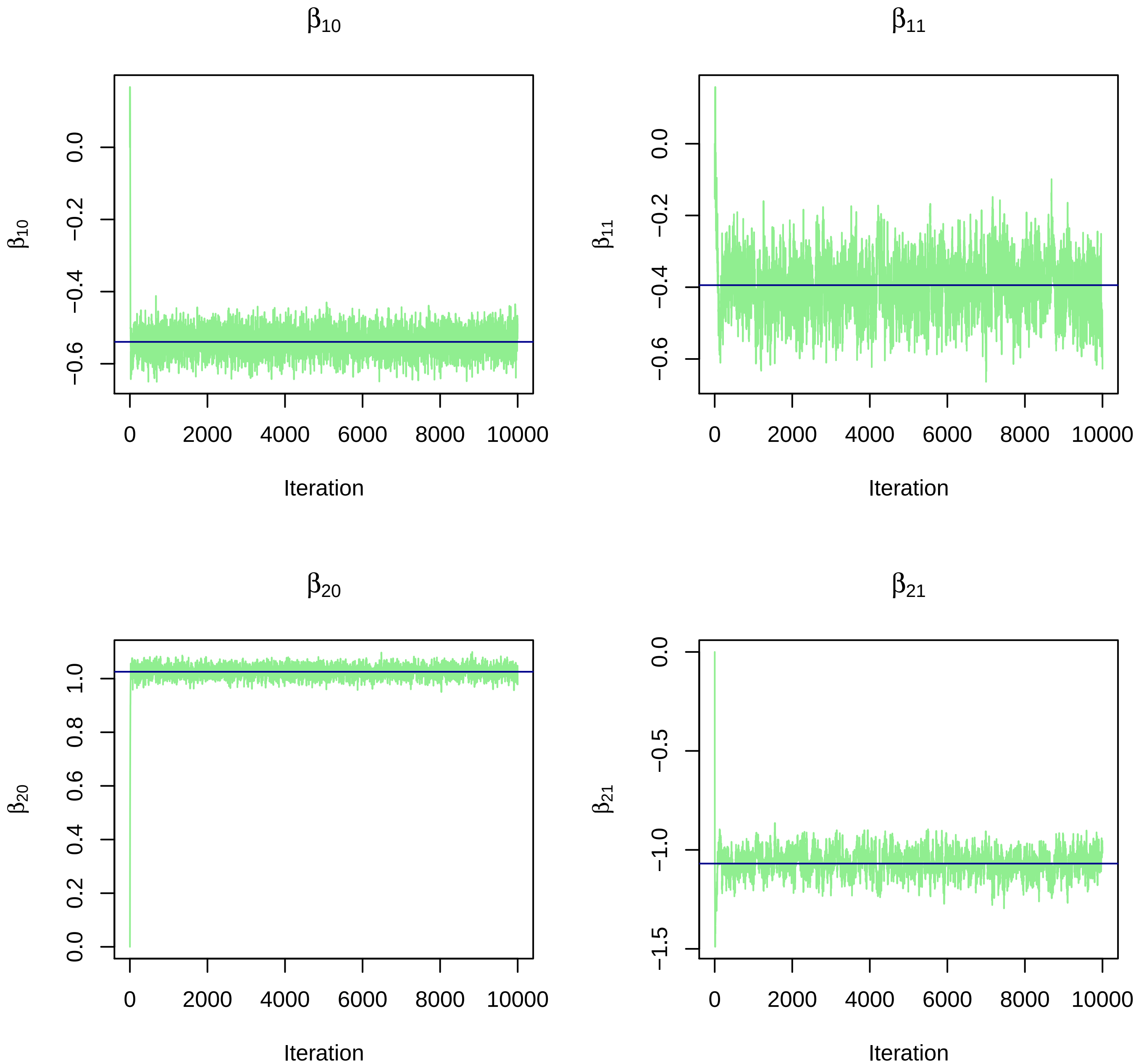
Web Table 5: Posterior means and 95% credible intervals (CrIs) for the one-factor model from the breast cancer genomics study.

Pathway	Model Component	Parameter	Parameter Name	Posterior Mean (95% CrI)
MAPK	Binary	λ_{11}	Factor Loading	$6.46\ (4.61, 9.13)$
		β_{111}	Stage 1 vs 2*	$-0.37\ (-0.99, 0.47)$
		β_{112}	Stage 2 vs 3†	$0.22\ (-0.40, 1.10)$
	Count	λ_{12}	Factor Loading	$4.90\ (4.47, 5.37)$
		β_{121}	Stage 1 vs 2	$-0.11\ (-0.30, 0.09)$
		β_{122}	Stage 2 vs 3	$0.10\ (-0.08, 0.27)$
CCR Interaction	Binary	λ_{21}	Factor Loading	$4.50\ (3.56, 5.81)$
		β_{211}	Stage 1 vs 2	$-0.29\ (-0.76, 0.21)$
		β_{212}	Stage 2 vs 3	$0.05\ (-0.39, 0.50)$
	Count	λ_{22}	Factor Loading	$6.56\ (6.00, 7.18)$
		β_{221}	Stage 1 vs 2	$-0.14\ (-0.33, 0.05)$
		β_{222}	Stage 2 vs 3	$0.10\ (-0.07, 0.28)$
Endocytosis	Binary	λ_{31}	Factor Loading	$6.92\ (4.96, 9.51)$
		β_{311}	Stage 1 vs 2	$-0.39\ (-0.93, 0.16)$
		β_{312}	Stage 2 vs 3	$0.03\ (-0.59, 0.57)$
	Count	λ_{32}	Factor Loading	$7.27\ (6.66, 57.96)$
		β_{321}	Stage 1 vs 2	$-0.12\ (-0.31, 0.07)$
		β_{322}	Stage 2 vs 3	$0.12\ (-0.04, 0.29)$

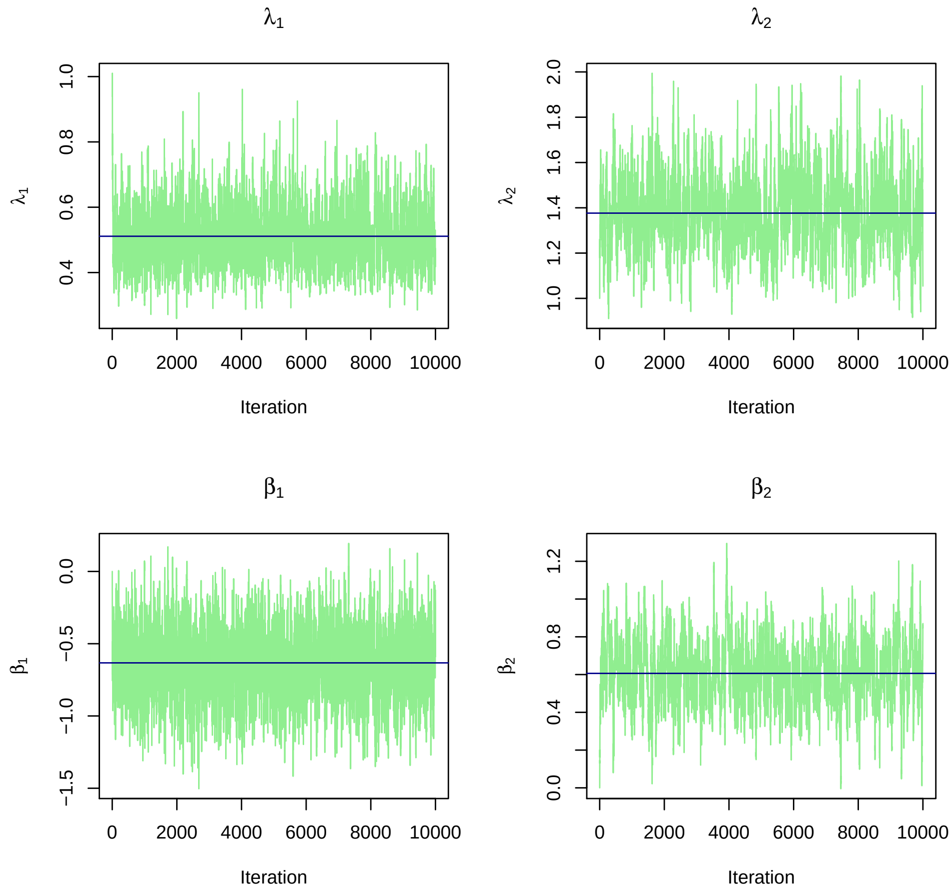
* Regression coefficient comparing stages 1 and 2, with stage 2 as reference group.

† Regression coefficient comparing stages 2 and 3, with stage 2 as reference group.

Web Appendix E: Web Figures



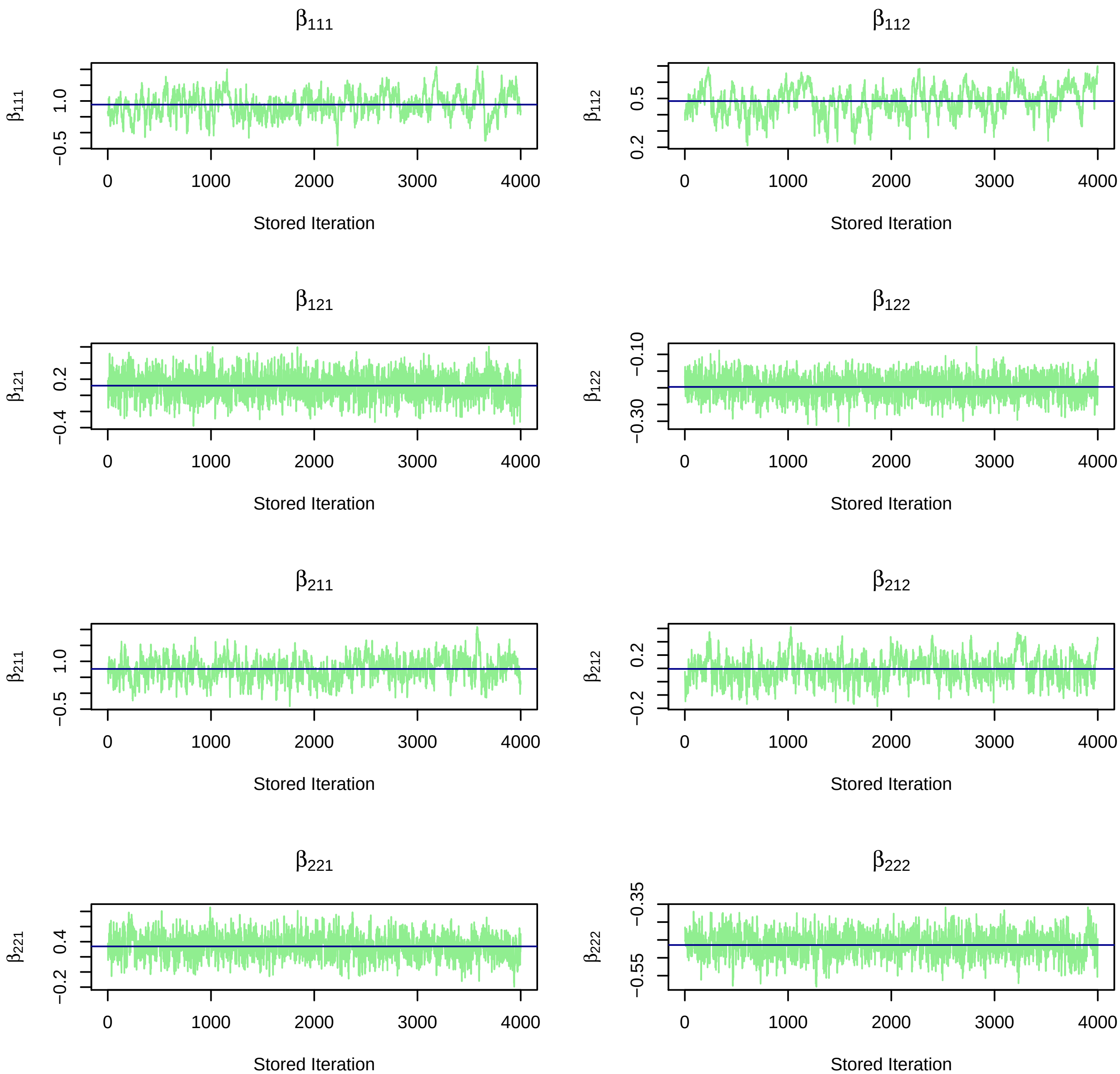
Web Figure 1: Trace plots for regression parameters for simulation study 1 with $n = 5000$, 70% zeros, and gamma hyperparameters $c = d = 0.001$. True coefficient values: $\beta_{10} = -0.50$, $\beta_{11} = -0.50$, $\beta_{20} = 1$, $\beta_{21} = -1$. Horizontal lines denote posterior means. All parameters initialized at 0.



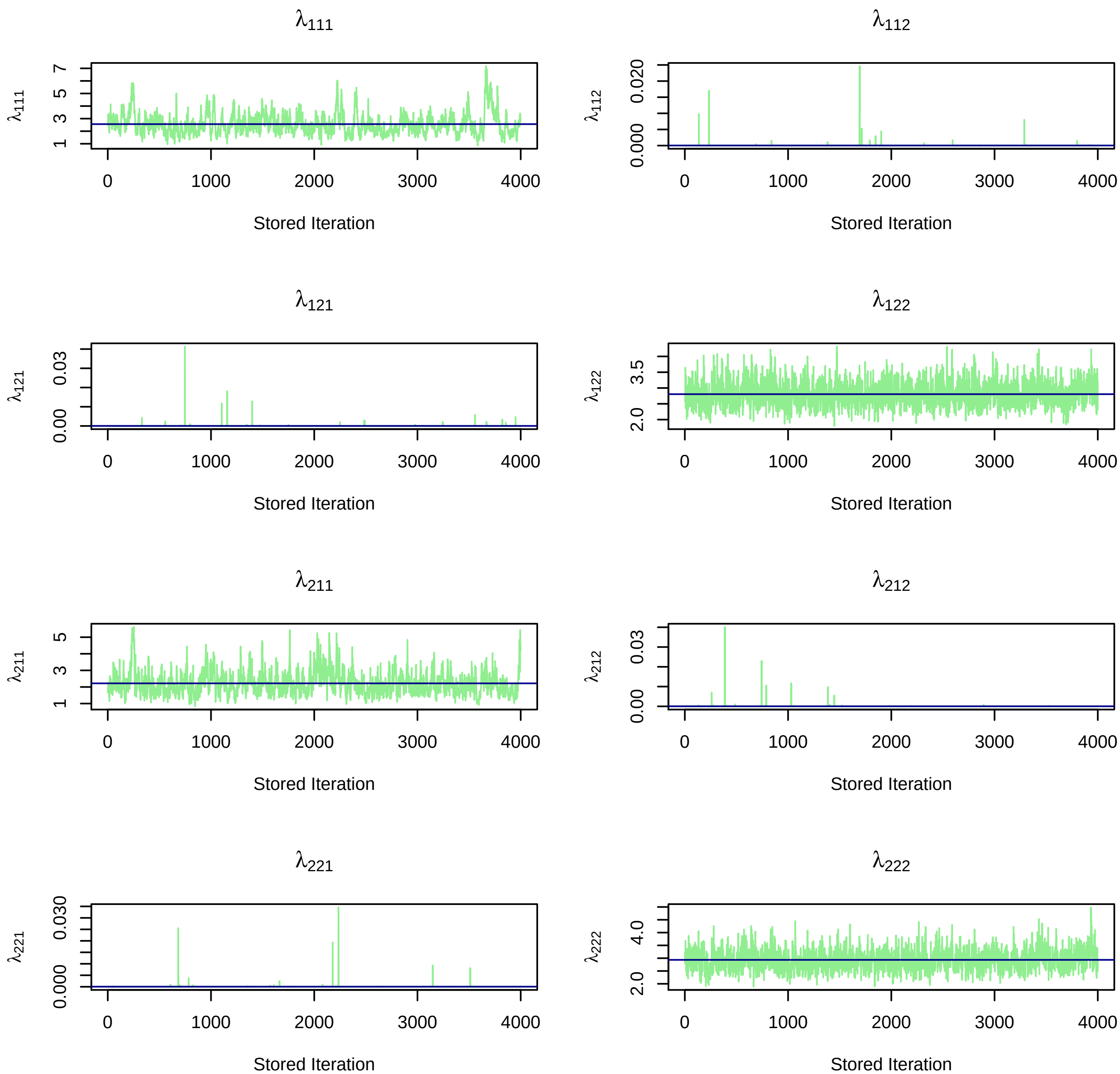
Web Figure 2: Trace plots for LZIP parameters in simulation study 2 with $n = 500$, 70% zeros, and Ga$(0.001, 0.001)$ hyperparameters. True parameter values: $\lambda_1 = 0.50$, $\lambda_2 = 1.50$, $\beta_1 = -0.50$, and $\beta_2 = 0.75$. Horizontal lines denote posterior means. λ_1 and λ_2 initialized at 1, and β_1 and β_2 initialized at 0.

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_{111} & \lambda_{112} \\ \lambda_{121} & \lambda_{122} \\ \lambda_{211} & \lambda_{212} \\ \lambda_{221} & \lambda_{222} \end{pmatrix} = \begin{pmatrix} \lambda_{111} & 0 \\ 0 & \lambda_{122} \\ \lambda_{211} & 0 \\ 0 & \lambda_{222} \end{pmatrix}$$

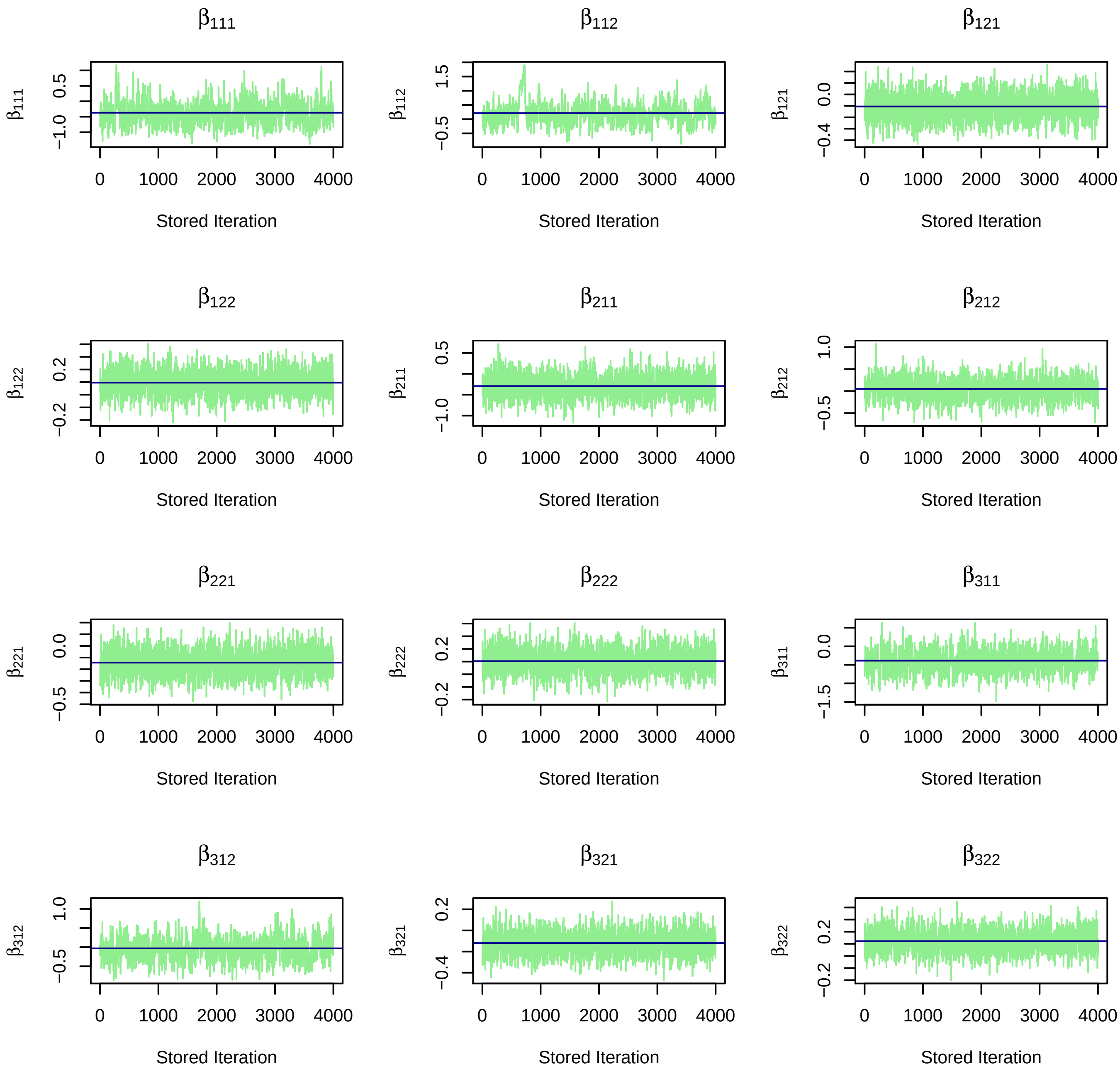
Web Figure 3: Factor loading matrix for simulation study 3: two-factor bivariate LZIP. Here, λ_{jkl} denotes the loading for the j-th outcome, k-th model component (binary versus count), and l-th factor. For this simulation, λ_{112}, λ_{121}, λ_{212}, and λ_{221} were set to the limiting value of 0. This represents no association between the binary and count components of the same outcome (i.e., no "within-outcome" association), but allows for dependence between 1) the binary components of the two outcomes and 2) the count components of the two outcomes (i.e., "between-outcome" association).



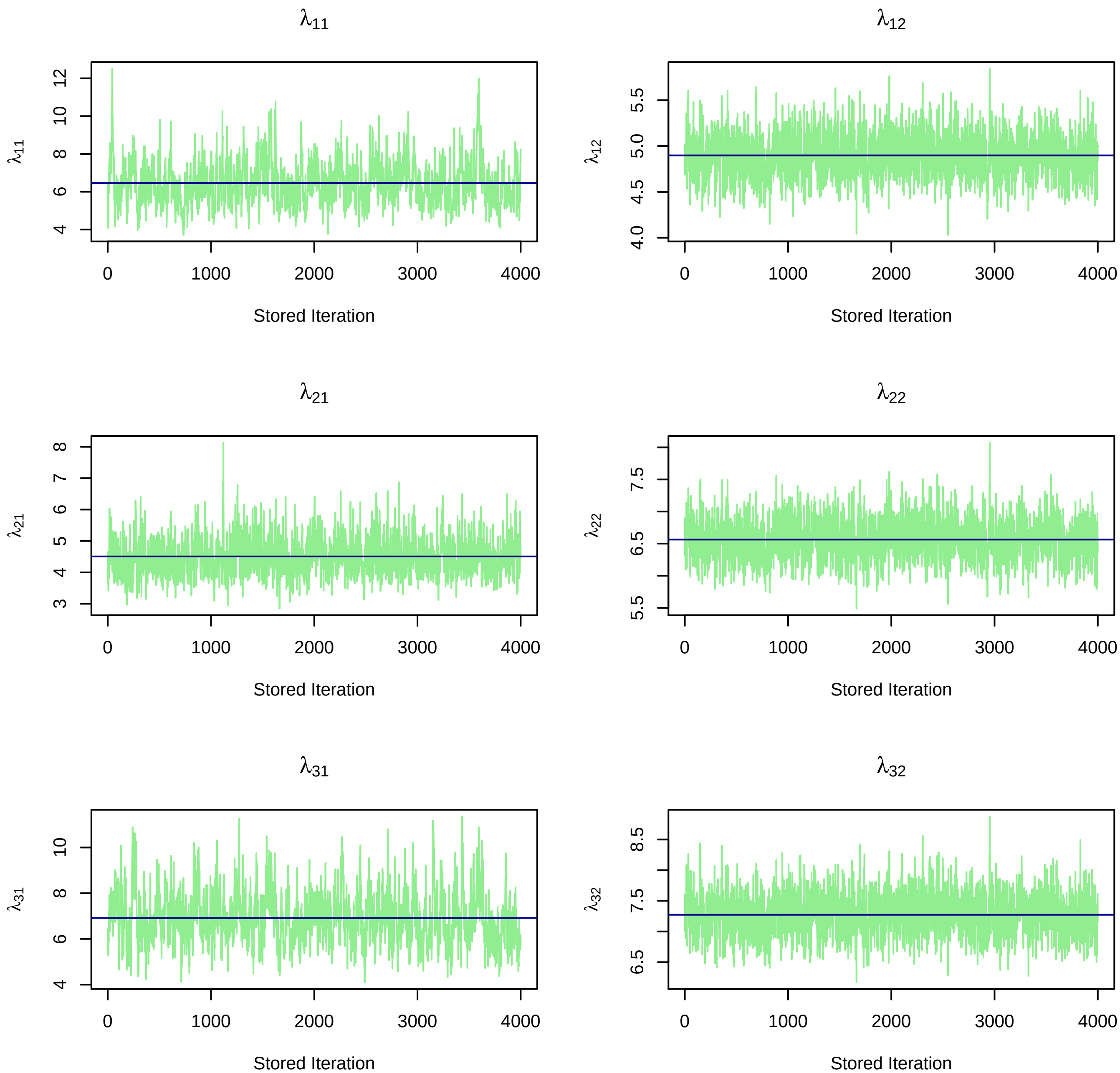
Web Figure 4: Trace plots for regression coefficients (β's) in simulation study 3 with $n = 500$, 40% zeros, and Ga$(0.001, 0.001)$ hyperparameters. True parameter values are $\beta_{111} = 1.00$, $\beta_{112} = 0.50$, $\beta_{121} = 0.25$ $\beta_{122} = -0.25$, $\beta_{211} = .75$, $\beta_{212} = 0.25$, $\beta_{221} = 0.50$, and $\beta_{222} = -0.50$.



Web Figure 5: Trace plots for the factor loadings (λ's) in simulation study 3 with $n = 500$, 40% zeros, and $\text{Ga}(0.001, 0.001)$ hyperparameters. True parameter values: 2.5 for λ_{111}, λ_{122}, λ_{211} and λ_{222}; 0 for remaining λ's.

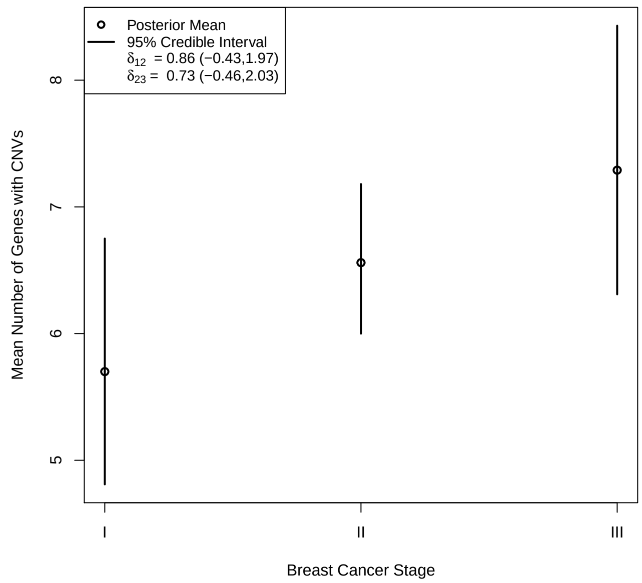


Web Figure 6: Trace plots for the regression coefficients (β's) in for the breast cancer genomics data analysis.

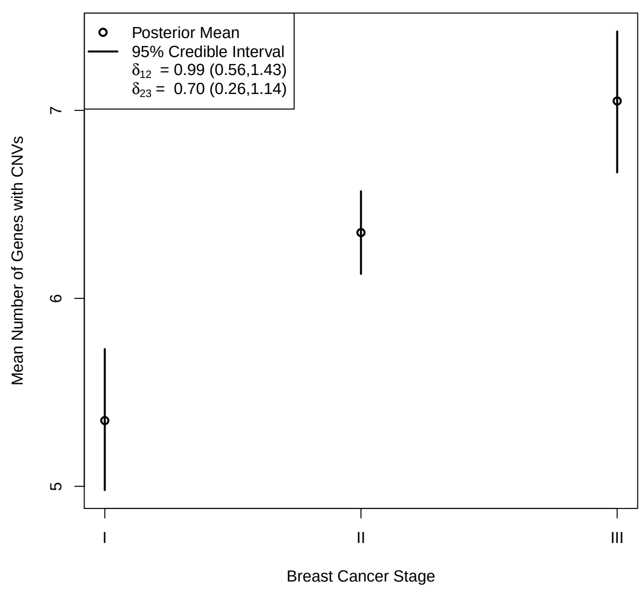


Web Figure 7: Trace plots for the factor loadings (λ's) for the breast cancer genomics data analysis.

(a)

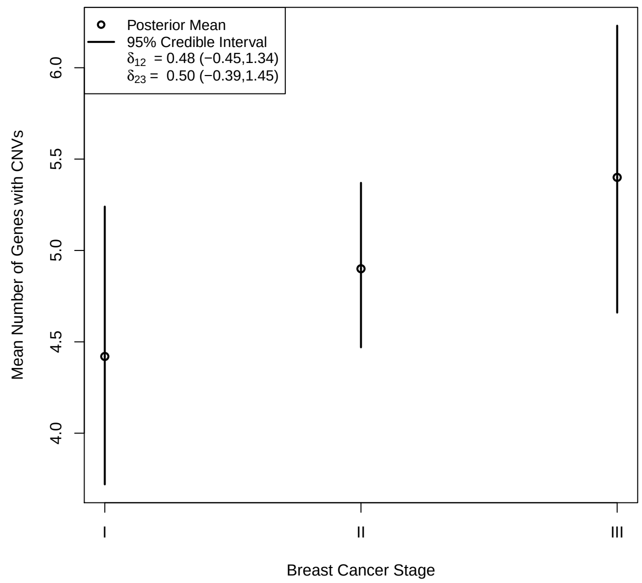


(b)

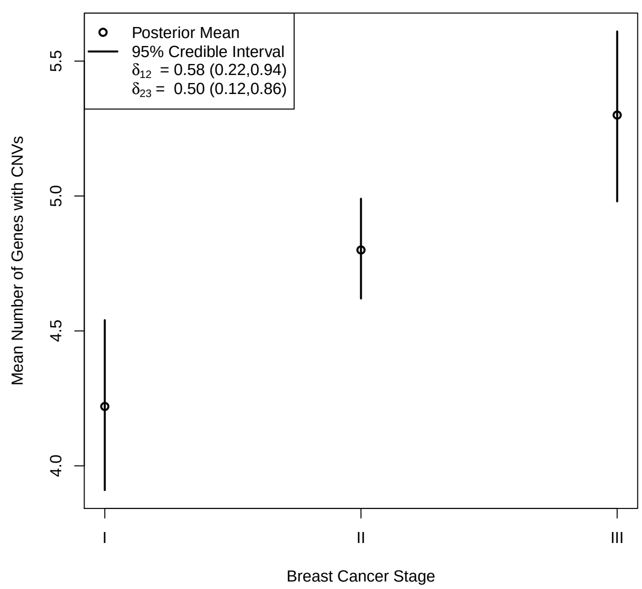


Web Figure 8: Gene activation results for CCR Interaction pathway. Panel (a): Population-average mean number of genes with CNVs, conditional on pathway activation. Panel (b): Population-average mean number of genes with CNVs among *all* patients (with and without pathway activation). Circles denote posterior mean estimates; solid lines are 95% credible intervals; and δ_{12} and δ_{23} are the differences between stages 1 and 2 and stages 2 and 3, respectively.

(a)

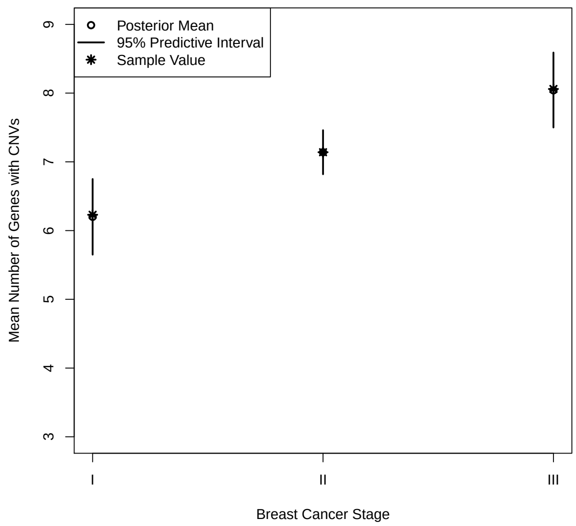


(b)

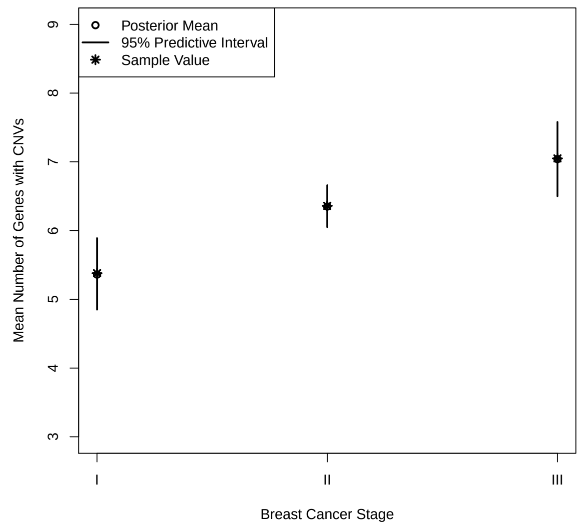


Web Figure 9: Gene activation results for Endocytosis pathway. Panel (a): Population-average mean number of genes with CNVs, conditional on pathway activation. Panel (b): Population-average mean number of genes with CNVs among *all* patients (with and without pathway activation). Circles denote posterior mean estimates; solid lines are 95% credible intervals; and δ_{12} and δ_{23} are the differences between stages 1 and 2 and stages 2 and 3, respectively.

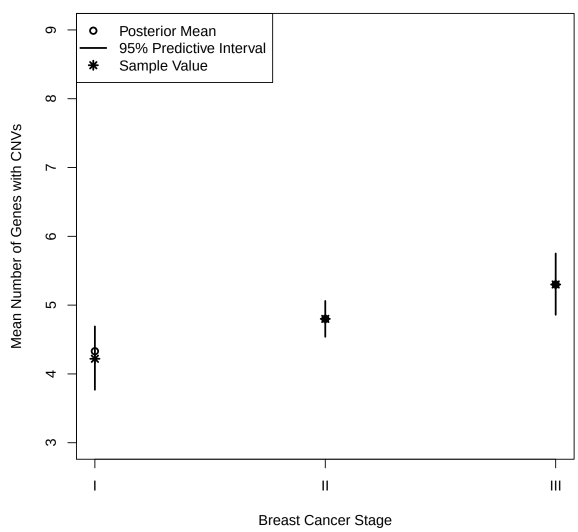
(a) MAPK



(b) CCR Interaction

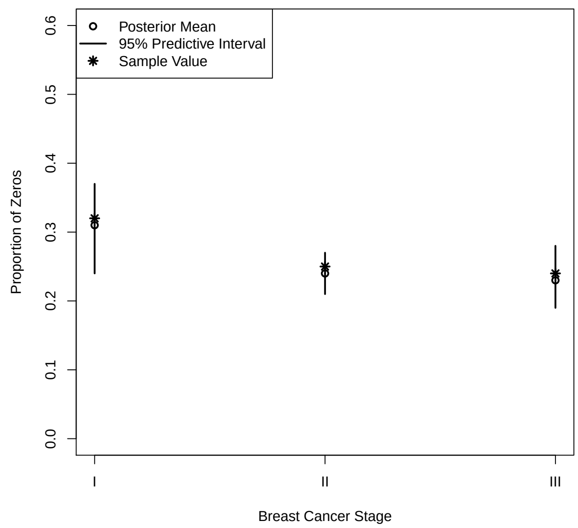


(c) Endocytosis

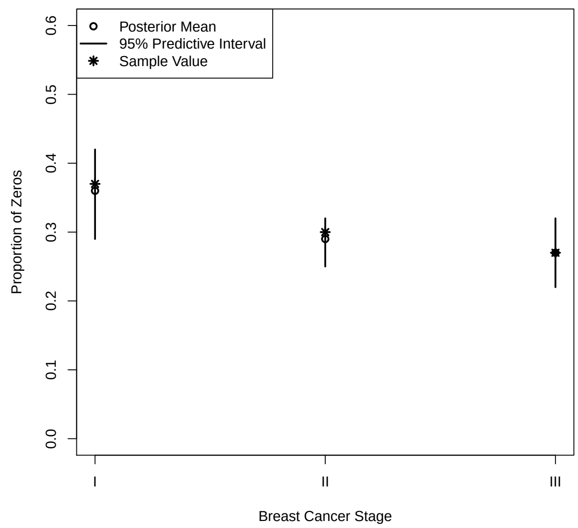


Web Figure 10: Posterior predictive checks based on the mean number of genes with CNVs among *all* patients (with and without pathway activation). Circles denote posterior predictive mean estimates; solid lines are 95% posterior predictive intervals; and asterisks denote observed sample values.

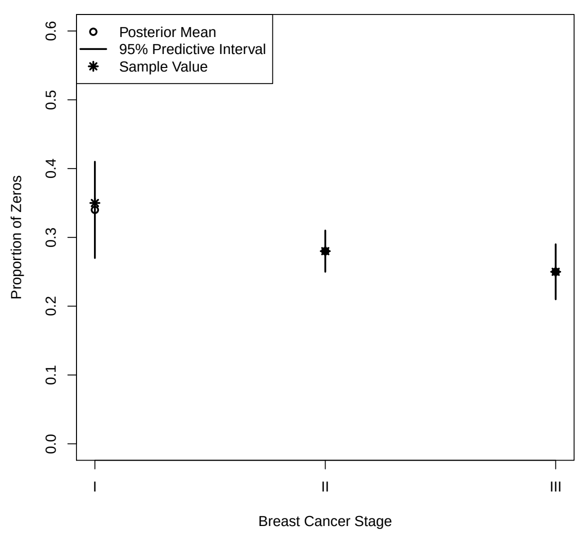
(a) MAPK



(b) CCR Interaction

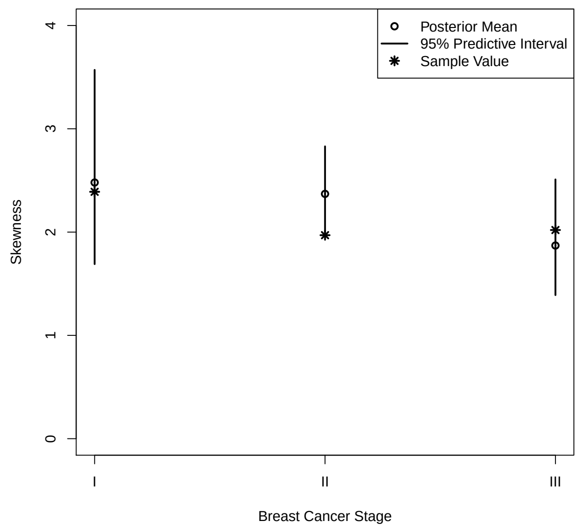


(c) Endocytosis

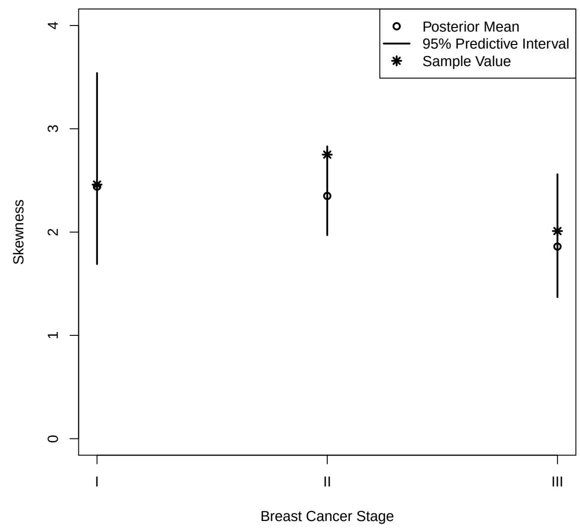


Web Figure 11: Posterior predictive checks based on the sample proportion of zeros. Circles denote posterior predictive mean estimates; solid lines are 95% posterior predictive intervals; and asterisks denote observed sample values.

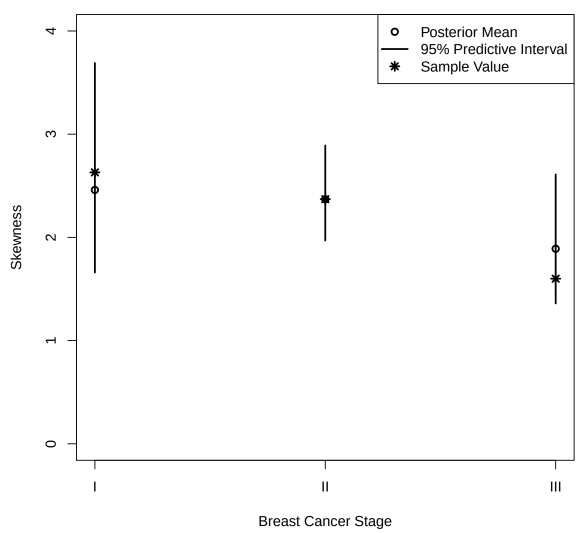
(a) MAPK



(b) CCR Interaction



(c) Endocytosis



Web Figure 12: Posterior predictive checks based on the sample skewness. Circles denote posterior predictive mean estimates; solid lines are 95% posterior predictive intervals; and asterisks denote observed sample values.