SUPPLEMENTAL APPENDIX for
A Bayesian Split Population Survival Model for Duration Data With Misclassified Failure Events

February 5, 2019

Benjamin E. Bagozzi,* Minnie M. Joo,† Bomin Kim,‡ Bumba Mukherjee§

*Assistant Professor, Dept. of Political Science & International Relations, University of Delaware. Email: bagozzib@udel.edu

†Ph.D. Candidate, Dept. of Political Science, Penn State. Email: mxj222@psu.edu

‡Ph.D. Candidate, Dept. of Statistics, Penn State. Email: bzk147@psu.edu

§Professor, Dept. of Political Science, Penn State. Email: sxm73@psu.edu

Contents

SUPPLEMENTAL APPENDIX

This Supplemental Appendix is divided into five main sections. In the first section, we derive the closed form of the full conditional distributions of $\pi(\Sigma_\beta|\boldsymbol{\beta}_i)$ and $\pi(\Sigma_\gamma|\boldsymbol{\gamma}_i)$ that are required for Gibbs sampling in Step 1 of the MCMC method (with the slice sampling scheme) used for estimating the Bayesian Misclassified Failure (MF) Weibull model. This section also presents the steps for slice sampling for $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and ρ. The second section provides a detailed derivation of the log-likelihood function of our parametric MF model with time-varying covariates that is stated in equation (9) of the main paper. This is followed by a discussion of the main properties of this parametric MF model, and the description of the MF exponential model and its estimation via MCMC methods. The third section presents and discusses the additional results (including tables and figures) from the Monte Carlo simulation analyses that were mentioned—but not presented in detail to save space—in the main paper. The fourth section presents all the additional tables, figures and convergence diagnostic checks generated from the application of our Bayesian MF Weibull model to the Buhaug *et al* (2009) data. The fifth section discusses in detail the Bayesian MF Weibull model results (including all tables and figures) for the Reenock, Bernhard and Sobek (2007) democratic survival application.

I Full Conditional Distributions and Slice Sampling

We first derive the closed form of the full conditional distributions of $\pi(\Sigma_\beta|\boldsymbol{\beta_i})$ and $\pi(\Sigma_\gamma|\boldsymbol{\gamma_i})$ in Step 1 of the MCMC estimation of our MF Wiebull model:

1. Σ_β:

$$
\begin{aligned}
\pi(\Sigma_\beta|\boldsymbol{\beta}) &\propto \pi(\boldsymbol{\beta}|\boldsymbol{\mu}_\beta = \mathbf{0}, \Sigma_\beta) \times \pi(\Sigma_\beta) \\
&\propto |\Sigma_\beta|^{-\frac{1}{2}} \exp\Big\{ -\frac{1}{2}\Big(\boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\beta}\Big)\Big\} \times |\Sigma_\beta|^{-\frac{\nu_1+p_1+1}{2}} \exp\Big\{ -\frac{1}{2}\mathrm{tr}\Big(\mathbf{I}_{p_1}\Sigma_\beta^{-1}\Big)\Big\} \\
&= |\Sigma_\beta|^{-\frac{1+\nu_1+p_1+1}{2}} \exp\Big\{ -\frac{1}{2}\mathrm{tr}\Big(\boldsymbol{\beta}\boldsymbol{\beta}'\Sigma_\beta^{-1}\Big)\Big\} \times \exp\Big\{ -\frac{1}{2}\mathrm{tr}\Big(\mathbf{I}_{p_1}\Sigma_\beta^{-1}\Big)\Big\} \qquad \text{(A.1)} \\
&= |\Sigma_\beta|^{-\frac{1+\nu_1+p_1+1}{2}} \exp\Big\{ -\frac{1}{2}\mathrm{tr}\Big((\boldsymbol{\beta}\boldsymbol{\beta}' + \mathbf{I}_{p_1})\Sigma_\beta^{-1}\Big)\Big\} \\
&\sim \mathrm{IW}(\boldsymbol{\beta}\boldsymbol{\beta}' + \mathbf{I}_{p_1}, \quad 1+\nu_1)
\end{aligned}
$$

2. Σ_γ:

$$
\begin{aligned}
\pi(\Sigma_\gamma|\boldsymbol{\gamma}) &\propto \pi(\boldsymbol{\gamma}|\boldsymbol{\mu}_\gamma = \mathbf{0}, \Sigma_\gamma) \times \pi(\Sigma_\gamma) \\
&\propto |\Sigma_\gamma|^{-\frac{1}{2}} \exp\Big\{ -\frac{1}{2}\Big(\boldsymbol{\gamma}'\Sigma_\gamma^{-1}\boldsymbol{\gamma}\Big)\Big\} \times |\Sigma_\gamma|^{-\frac{\nu_1+p_2+1}{2}} \exp\Big\{ -\frac{1}{2}\mathrm{tr}\Big(\mathbf{I}_{p_2}\Sigma_\gamma^{-1}\Big)\Big\} \\
&= |\Sigma_\gamma|^{-\frac{1+\nu_2+p_2+1}{2}} \exp\Big\{ -\frac{1}{2}\mathrm{tr}\Big(\boldsymbol{\gamma}\boldsymbol{\gamma}'\Sigma_\gamma^{-1}\Big)\Big\} \times \exp\Big\{ -\frac{1}{2}\mathrm{tr}\Big(\mathbf{I}_{p_2}\Sigma_\gamma^{-1}\Big)\Big\} \qquad \text{(A.2)} \\
&= |\Sigma_\gamma|^{-\frac{1+\nu_2+p_2+1}{2}} \exp\Big\{ -\frac{1}{2}\mathrm{tr}\Big((\boldsymbol{\gamma}\boldsymbol{\gamma}' + \mathbf{I}_{p_2})\Sigma_\gamma^{-1}\Big)\Big\} \\
&\sim \mathrm{IW}(\boldsymbol{\gamma}\boldsymbol{\gamma}' + \mathbf{I}_{p_2}, \quad 1+\nu_2)
\end{aligned}
$$

We next turn to describe the slice sampling algorithm for estimating $\boldsymbol{\beta}$,$\boldsymbol{\gamma}$ and ρ. Following the current practice in Bayesian mixture survival models, we use univariate slice sampler with stepout and shrinkage (Neal 2003) in Step 2, where the closed form of the full conditional distribution is intractable. We also follow the modifications made in 'BayesMixSurv' R package (Mahani, Mansour, and Mahani 2016). Below are the steps to perform slice sampling for $\boldsymbol{\beta}$ (note that slice sampling for $\boldsymbol{\gamma}$ and ρ is done in the exact same manner and hence not described here to avoid repetition):

For $\boldsymbol{\beta}_p$, $p = \{1, ..., P\}$,

- **Step 0.** Choose an arbitrary starting point β_{p_0} and size of the slice w, and set $i = 0$.
- **Step 1.** Draw y from Uniform$(0, f(\beta_{p_0}))$ defining slice $S = \{\beta_p : y < f(\beta_p)\}$, where

$$
\begin{aligned}
f(\beta_p) &\propto \pi(\beta_p|\boldsymbol{\beta}_{-p}, \mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \mathbf{t0}, \boldsymbol{\gamma}) && \text{if exponential} \\
&\propto \pi(\beta_p|\boldsymbol{\beta}_{-p}, \mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \mathbf{t0}, \boldsymbol{\gamma}, \rho) && \text{if Weibull.}
\end{aligned} \qquad \text{(A.3)}
$$

- **Step 2.** Find an interval, $I = (L, R)$, around β_{p_0} that contains all, or much, of the slice, where the initial interval is determined as:

$$\begin{aligned} & u \sim \text{Uniform}(0, w) \\ & L = \beta_{p_0} - u \qquad , \\ & R = \beta_{p_0} + (w - u) \end{aligned} \tag{A.4}$$

and expand the interval until its ends are outside the slice or until the limit on steps (limit on steps = m) is reached ("stepping-out" procedure), by comparing y and $(f(L), f(R))$:

$$\begin{aligned} & J = \text{Floor}(\text{Uniform}(0, m)) \\ & K = (m - 1) - J \\ & \text{Repeat while } J > 0 \text{ and } y < f(L): \\ & L = L - w, J = J - 1 \\ & \text{Repeat while } K > 0 \text{ and } y < f(R): \\ & R = R + w, K = K - 1 \end{aligned} \tag{A.5}$$

- **Step 3.** Draw a new point β_{p_1} from the part of the slice within this interval I, and shrink the interval on each rejection ("shrinkage" procedure):

$$\begin{aligned} \text{Repeat:} \qquad & \beta_{p1} \sim \text{Uniform}(L, R) \\ & \text{if } y < f(\beta_{p_1}), \text{ accept } \beta_{p_1} \text{ and exit loop} \\ & \text{if } \beta_{p_1} < \beta_{p_0}, \text{ then } L = \beta_{p_1} \\ & \qquad\qquad\qquad \text{else } R = \beta_{p_1}. \end{aligned} \tag{A.6}$$

- **Step 4.** Set $i = i + 1$, $\beta_{p_0} = \beta_{p_1}$, and go to Step 1.
- **Step 5.** After N iterations, summarize the parameter estimates using all sampled values (via, e.g., credible intervals or posterior means).

II Time-Varying Misclassified Failure Model

Log-Likelihood and Properties

The likelihood function of the general parametric MF model that is developed and defined from equations (1)-(5) in the paper is given by:

$$L = \prod_{i=1}^{N} [\alpha_i + (1 - \alpha_i) f(t_i | \mathbf{X}_i, \boldsymbol{\beta})]^{C_i} [(1 - \alpha_i) S(t_i | \mathbf{X}_i \boldsymbol{\beta})]^{1 - C_i}, \tag{A.7}$$

and the model's log-likelihood is

$$\ln L = \sum_{i=1}^{N} \{\widetilde{C}_i \ln[\alpha_i + (1-\alpha_i) f(t_i|\mathbf{X}_i, \boldsymbol{\beta})] + (1-\widetilde{C}_i) \ln[(1-\alpha_i) S(t_i|\mathbf{X}_i, \boldsymbol{\beta})]\}. \tag{A.8}$$

Further, recall from the text that the log-likelihood function of the standard general parametric survival model with time-varying covariates is

$$\ln L = \sum_{i=1}^{N} \left\{ \tilde{C}_{ij} \ln \left[\frac{f(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})}{S(t0_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})} \right] + (1-\tilde{C}_{ij}) \ln \left[\frac{S(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})}{S(t0_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})} \right] \right\}, \tag{A.9}$$

where $\widetilde{C}_{ij} = 0$ denotes all censored observations that are correctly recored, while $\widetilde{C}_{ij} = 1$ are non-censored (i.e., "failed") observations, which may be contaminated with cases that are actually censored cases (i.e., $C_{ij} = 0$). Hence, given $\widetilde{C}_{ij} = 0$ and $\widetilde{C}_{ij} = 1$, we can define the probability of misclassification as:

$$\alpha_{ij} = Pr(\tilde{C}_{ij} = 1 | C_{ij} = 0). \tag{A.10}$$

Incorporating α_{ij}, the unconditional density of an event happening is

$$Pr(\tilde{C}_{ij} = 1 | C_{ij} = 0) + Pr(\tilde{C}_{ij} = 0 | C_{ij} = 0) \Pr(t_{ij} \leq T_{ij}) = \alpha_{ij} + (1-\alpha_{ij}) \frac{f(t_{ij})}{S(t0_{ij})}, \tag{A.11}$$

with a corresponding unconditional survival function of

$$Pr(\tilde{C}_{ij} = 0 | C_{ij} = 0) \Pr(t_{ij} > T_{ij}) = (1-\alpha_{ij}) \frac{S(t_{ij})}{S(t0_{ij})}. \tag{A.12}$$

Combining these two parts and using equation A.9, the log-likelihood function of the parametric MF model with time-varying covariates is defined as:

$$\ln L = \sum_{i=1}^{N} \left\{ \widetilde{C}_{ij} \ln \left[\alpha_{ij} + (1-\alpha_{ij}) \frac{f(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})}{S(t0_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})} \right] + (1-\widetilde{C}_{ij}) \ln \left[(1-\alpha_{ij}) \frac{S(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})}{S(t0_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})} \right] \right\}, \tag{A.13}$$

where $\alpha_{ij} = \dfrac{\exp(\mathbf{Z}_{ij}\boldsymbol{\gamma})}{1+\exp(\mathbf{Z}_{ij}\boldsymbol{\gamma})}$.

Building on the preceding discussion, note that if one were to define a probability of non-misclassification as $\delta_{ij} = 1 - \alpha_{ij}$ and substitute this quantity into Equation A.11, the log-likelihood would be defined as:

$$\ln L = \sum_{i=1}^{N} \left\{ C_{ij} \ln \left[(1-\delta_{ij}) + \delta_{ij} \frac{f(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})}{S(t0_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})} \right] + (1-C_{ij}) \ln \left[\delta_{ij} \frac{S(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})}{S(t0_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})} \right] \right\}, \tag{A.14}$$

which is symmetric to the log likelihood of the split-population survival model:

$$\ln L = \sum_{i=1}^{N} \left\{ \tilde{C}_{ij} \ln \left[\delta_{ij} \frac{f(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})}{S(t0_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})} \right] + (1 - \tilde{C}_{ij}) \ln \left[(1 - \delta_i) + \delta_{ij} \frac{S(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})}{S(t0_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta})} \right] \right\}. \quad \text{(A.15)}$$

This implies that some properties of the cure model also hold for the MF model, including (i) the reduction of the latter to a normal parametric model whenever $\delta = 1$ or $\alpha = 0$ and (ii) parameter identification even in the case where identical covariates are included in $\mathbf{Z}$ and $\mathbf{X}$.[1]

Misclassified Failure Exponential Model

To develop the MF exponential model, we require to first define the density function and survival function in this case which are respectively:

$$\begin{aligned} f(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta}) &= \exp(\mathbf{X}_{ij}\boldsymbol{\beta})\exp(-\exp(\mathbf{X}_{ij}\boldsymbol{\beta})t_{ij}) \\ S(t_{ij}|\mathbf{X}_{ij}, \boldsymbol{\beta}) &= \exp(-\exp(\mathbf{X}_{ij}\boldsymbol{\beta})t_{ij}). \end{aligned} \quad \text{(A.16)}$$

Following the steps taken to define the log-likelihood function in equation (9) in the main paper, the log-likelihood function of the MF exponential model with time varying covariates is:

$$\begin{aligned} &\ln L(\beta, \gamma) \\ &= \sum_{i=1}^{N} \left\{ \widetilde{C}_{ij} \ln \left[\alpha_{ij} + (1 - \alpha_{ij}) \frac{\exp(\mathbf{X}_{ij}\boldsymbol{\beta})\exp(-\exp(\mathbf{X}_{ij}\boldsymbol{\beta})t_{ij})}{\exp(-\exp(\mathbf{X}_{ij}\boldsymbol{\beta})t0_{ij})} \right] \right. \\ &\qquad \left. + (1 - \widetilde{C}_{ij}) \ln \left[(1 - \alpha_{ij}) \frac{\exp(-\exp(\mathbf{X}_{ij}\boldsymbol{\beta})t_{ij})}{\exp(-\exp(\mathbf{X}_{ij}\boldsymbol{\beta})t0_{ij})} \right] \right\}, \end{aligned} \quad \text{(A.17)}$$

where $\mathbf{X}_{ij}$ is the i^{th} row of the covariate matrix $\mathbf{X}$ at time j and $\alpha_{ij} = \dfrac{\exp(\mathbf{Z}_{ij}\boldsymbol{\gamma})}{1 + \exp(\mathbf{Z}_{ij}\boldsymbol{\gamma})}$. As shown in (A.17), the time-varying MF exponential model accounts for the probability of misclassified failure while estimating the effect of the covariates that influence the survival of the event of interest (assumed to be exponentially distributed).

While the MF exponential model with time-varying covariates can be estimated by maximum likelihood using, for example, BFGS,[2] we estimate this model via the MCMC algorithm employed for Bayesian inference. We thus label our model as Bayesian MF exponential model

[1] See Box-Steffensmeier and Zorn (1999, 5) for a discussion of these properties in the context of the split-population model.

[2] The Broyden, Fletcher, Goldfarb, Shannon (BFGS) method in the R *optim* function. In our Monte Carlo analysis, we briefly assess the properties of the MF exponential model estimated by BFGS.

given the use of MCMC estimation. To conduct Bayesian inference, we need to assign a prior for each of the MF exponential model's two parameters – $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ – and then define the conditional posterior distribution of these parameters. Following standard practice, we assign the multivariate normal prior to $\boldsymbol{\beta} = \{\beta_1, ..., \beta_{p_1}\}$ and $\boldsymbol{\gamma} = \{\gamma_1, ..., \gamma_{p_2}\}$:

$$\begin{aligned} \boldsymbol{\beta} &\sim \mathrm{MVN}_{p_1}(\mu_\beta, \Sigma_\beta), \qquad \boldsymbol{\gamma} \sim \mathrm{MVN}_{p_2}(\mu_\gamma, \Sigma_\gamma) \\ \Sigma_\beta &\sim \mathrm{IW}(S_\beta, \nu_\beta), \qquad \Sigma_\gamma \sim \mathrm{IW}(S_\gamma, \nu_\gamma). \end{aligned} \tag{A.18}$$

where we fix $\mu_\beta = \mathbf{0}$ and $\mu_\gamma = \mathbf{0}$ and S_β, ν_β, S_γ, ν_γ are the hyper parameters. Note that we use hierarchical Bayesian modeling to estimate Σ_β and Σ_γ using the Inverse-Wishart (IW) distribution. Given these prior specifications and the hyperparameters, the conditional posterior distributions for $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ parameters in the Bayesian MF exponential model (with time-varying covariates) are:

$$\begin{aligned} P(\boldsymbol{\beta}|\mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \mathbf{t0}, \boldsymbol{\gamma}) &\propto P(\boldsymbol{\beta}|\mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \mathbf{t0}, \boldsymbol{\gamma}) \times P(\boldsymbol{\beta}|\ \boldsymbol{\mu}_\beta, \Sigma_\beta) \\ P(\boldsymbol{\gamma}|\mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \mathbf{t0}, \boldsymbol{\beta}) &\propto P(\boldsymbol{\gamma}|\mathbf{C}, \mathbf{X}, \mathbf{Z}, \mathbf{t}, \mathbf{t0}, \boldsymbol{\beta}) \times P(\boldsymbol{\gamma}|\boldsymbol{\mu}_\gamma, \Sigma_\gamma). \end{aligned} \tag{A.19}$$

For the sampling scheme, since closed forms for the posterior distributions of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are not available, we use the same MCMC methods with the slice sampling algorithm described in the main paper for the Bayesian MF Weibull Model. The only difference is that, unlike the Weibull model, we ignore the slice sampling for ρ in the Bayesian MF exponential model. The closed form of the full conditional distributions of $P(\Sigma_\beta|\boldsymbol{\beta}_i)$ and $P(\Sigma_\gamma|\boldsymbol{\gamma}_i)$ as well as the slice sampling scheme for $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ is derived and described above (and hence not repeated here).

III Monte Carlo Simulation Results

This section provides a complete presentation of the Monte Carlo (MC) results that are referenced in the main paper. Recall that we conduct 15 MC experiments in total. Experiments 1 and 2 are primarily presented and discussed within the main paper, and evaluate the relative performance of a Bayesian Weibull model and a Bayesian MF Weibull model when one's true data generating process (d.g.p.) is (i) standard Weibull (Experiment 1) or (ii) MF Weibull with a 5% misclassified failure rate (Experiment 2). Experiments 3-4 instead assess the performance of maximum likelihood (specifically, BFGS) estimated versions of the Weibull and MF-Weibull models for these same non-MF Weibull (Experiment 3) and MF-Weibull (Experiment 4) outcome variables. Experiments 5-8 consider an exponentially distributed outcome variable (Experiments 5 and 7), or a MF exponential outcome variable (Experiments 6 and 8), and evaluate the performance of either (i) *Bayesian* Weibull, MF-exponential, and MF Weibull models (Experiments 5-6) or (ii) *BFGS* exponential, Weibull, MF exponential and MF Weibull models (Experiments 7-8).

Experiments 9, 10, and 11 revisit the Bayesian Weibull and Bayesian MF Weibull comparisons that we conduct in Experiment 2 under conditions where one's misclassified failure rate is increased from 5% to 8%, 12%, and 15%, respectively. Experiments 12 and 13 reevaluate our primary experiments (i.e., Experiments 1 and 2) when the estimated Bayesian Weibull and Bayesian MF Weibull models are specified using a (very) weakly (in other words, least-) informative multivariate Cauchy prior, which is distinct from the weakly-informative multivariate Normal prior that is employed in all other Bayesian MC experiments. Experiments 14 and 15 simulate a log-logistic survival outcome variable (Experiment 14) or MF log-logistic survival outcome variable, and compare the performance of our Bayesian (MF) Weibull models to that of a Bayesian Cox Proportional Hazard (PH) survival model for the corresponding non-Weibull distributed outcome variables.

For all fifteen MC experiments, we compare each relevant model under conditions of $N = 1,000$, $N = 1,500$, and $N = 2,000$. Below, we first present the plotted $\hat{\beta}$ and $\hat{\gamma}$ values for Experiments 1-2 (Figures A.1-A.3) which are referenced in the main text. We then provide a more in-depth interpretation of Experiments 3-15, which includes our reporting of each experiment's corresponding (averaged) parameter estimates, (MC)SE's, RMSEs, and 95% (confidence/credible) coverage probabilities (CPs); and parameter estimate distribution plots.

Experiment 3 compares the performance of (i) a BFGS Weibull model and (ii) a BFGS MF-Weibull model when the true outcome variable's d.g.p. is Weibull and the resultant survival

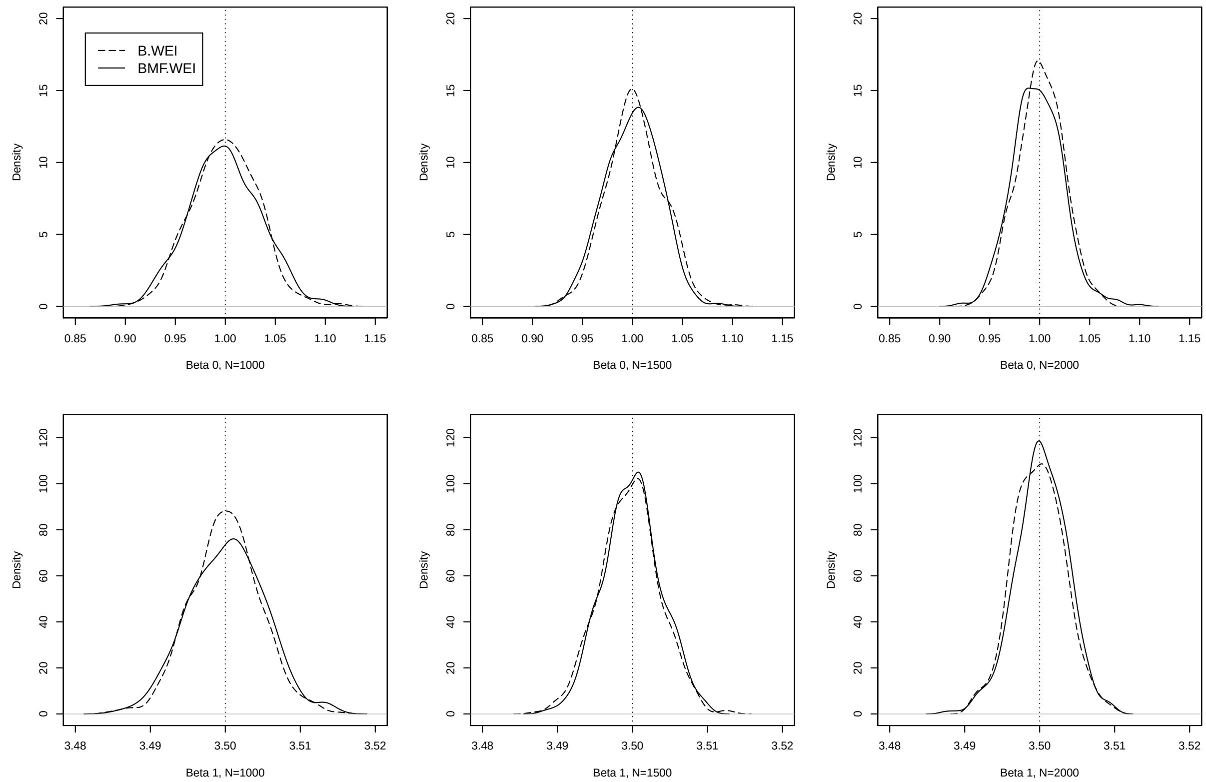


Figure A.1: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 1

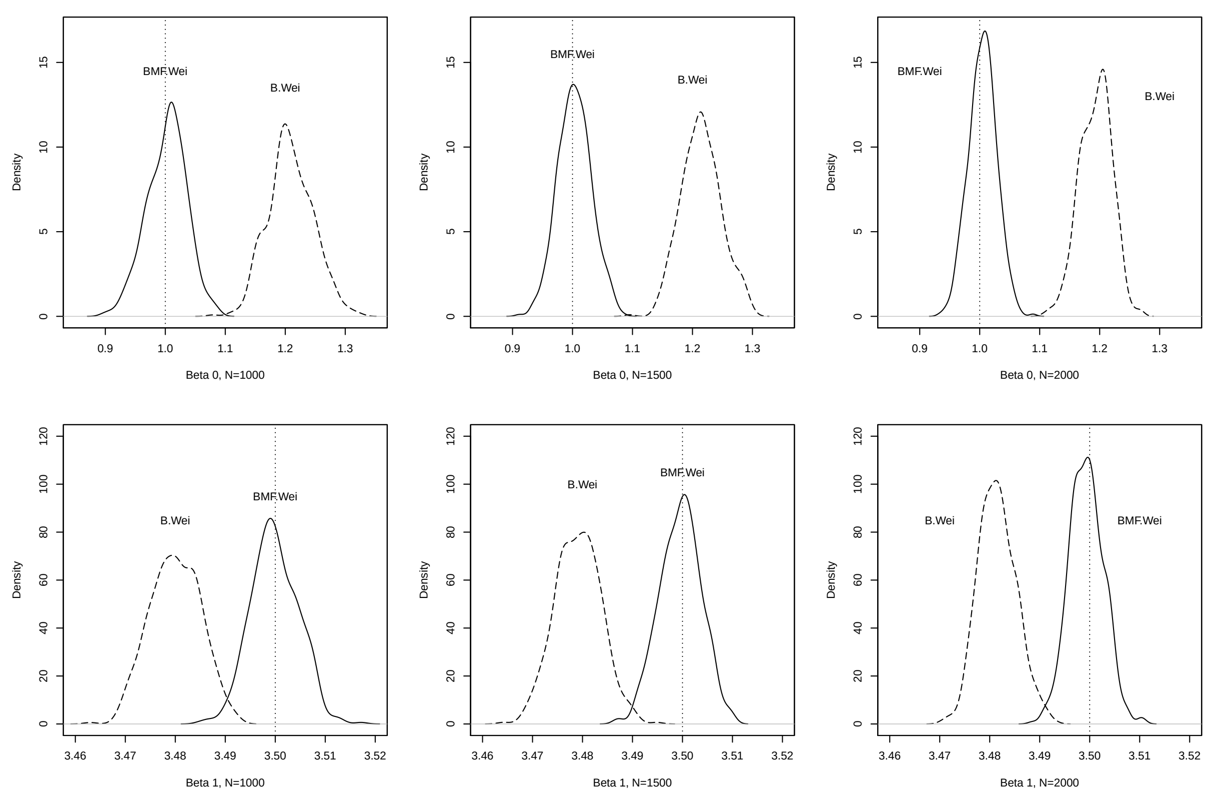


Figure A.2: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 2

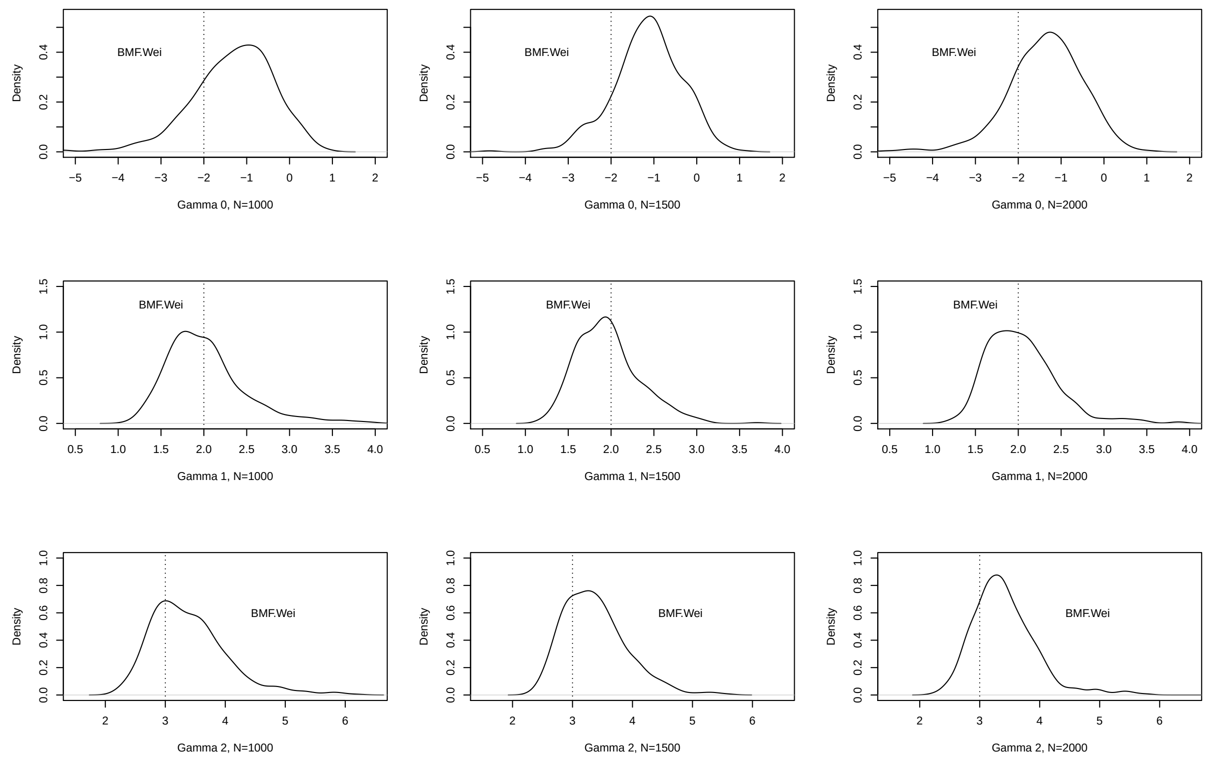


Figure A.3: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 2

outcome variable contains no instances of misclassified failures. We report this experiment's survival stage MC results in the top portion of Table A.1, and in Figure A.4. The results obtained in Experiment 3 are largely comparable to those obtained in Experiment 1. In circumstances where a researcher encounters a non-MF Weibull-distributed outcome variable, the BFGS MF Weibull and BFGS non-MF Weibull estimators each perform comparably with respect to bias and efficiency. For instance, one can note in Figure A.4 that each of the MF and non-MF $\hat{\beta}$ distributions virtually identical for each N evaluated. Likewise, the BFGS Weibull and BFGS MF Weibull models' SEs and RMSEs (reported in the top half of Table A.1) are comparable to the third decimal place for each parameter, and N, of interest.

At the same time, the 95% empirical CP values reported in Experiment 3 consistently favor the BFGS Weibull model over the BFGS MF Weibull to a larger degree than was the case for our comparable models in Experiment 1. Notably, we find in Experiment 3 that our BFGS Weibull CPs are generally 2 to 4 percentage points higher than those of the BFGS MF Weibull CP, for each parameter and each N considered. Nevertheless both sets of CPs appear commensurate, in that they range from 93.2%-96.2% in the case of the BFGS MF Weibull and from 93.4-96.2% in the case of the BFGS Weibull. Hence, there does not appear to be a substantial risk in (mis)applying a BFGS MF Weibull model to a non-MF Weibull-distributed outcome variable.

Table A.1: Maximum Likelihood β-Estimates for Experiments 3 and 4

	Experiment 3: Non-MF Weibull D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	$SE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$SE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
1000	BFGS Weibull	1.000	0.033	0.028	0.934	3.500	0.005	0.004	0.942
	BFGS MF Weibull	1.000	0.033	0.028	0.932	3.500	0.005	0.004	0.938
1500	BFGS Weibull	1.001	0.027	0.022	0.962	3.500	0.004	0.003	0.962
	BFGS MF Weibull	1.001	0.027	0.022	0.960	3.500	0.004	0.003	0.962
2000	BFGS Weibull	0.998	0.024	0.020	0.946	3.500	0.003	0.003	0.946
	BFGS MF Weibull	0.998	0.024	0.020	0.946	3.500	0.003	0.003	0.944
	Experiment 4: MF Weibull D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	$SE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$SE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
1000	BFGS Weibull	1.210	0.042	0.210	0.000	3.480	0.006	0.020	0.080
	BFGS MF Weibull	1.002	0.034	0.027	0.950	3.500	0.005	0.004	0.966
1500	BFGS Weibull	1.217	0.035	0.217	0.000	3.479	0.005	0.021	0.002
	BFGS MF Weibull	1.003	0.027	0.022	0.936	3.500	0.004	0.003	0.944
2000	BFGS Weibull	1.196	0.029	0.196	0.000	3.481	0.004	0.019	0.008
	BFGS MF Weibull	1.005	0.024	0.019	0.946	3.499	0.003	0.003	0.950

Note: True parameter values are $\beta_0 = 1$ and $\beta_1 = 3.5$.

Table A.2: Maximum Likelihood γ-Estimates for Experiment 4

	Experiment 4: MF Weibull D.G.P.												
#Obs.	Model	$\hat{\gamma}_0$	$SE(\hat{\gamma}_0)$	$RMSE(\hat{\gamma}_0)$	$CP(\hat{\gamma}_0)$	$\hat{\gamma}_1$	$SE(\hat{\gamma}_1)$	$RMSE(\hat{\gamma}_1)$	$CP(\hat{\gamma}_1)$	$\hat{\gamma}_2$	$SE(\hat{\gamma}_2)$	$RMSE(\hat{\gamma}_2)$	$CP(\hat{\gamma}_2)$
1000	BFGS MF Weibull	-1.648	1.153	0.926	0.960	2.106	0.439	0.377	0.940	3.401	0.624	0.595	0.942
1500	BFGS MF Weibull	-1.377	0.937	0.875	0.896	2.048	0.352	0.295	0.946	3.433	0.520	0.520	0.936
2000	BFGS MF Weibull	-1.511	0.877	0.781	0.898	2.067	0.333	0.291	0.956	3.399	0.464	0.478	0.928

Note: True parameter values are $\gamma_0 = -2$, $\gamma_1 = 2$, and $\gamma_2 = 3$.

On the other hand, MC Experiment 4 suggests that there *is* a non-negligible risk in (mis)applying a standard BGFS Weibull model to a MF Weibull-distributed outcome variable. Specifically, we observe in Figure A.5, and in the bottom half of Table A.1, that the BFGS Weibull model's $\hat{\beta}$'s generally overestimate β_0 and underestimate β_1, relative to the BFGS MF Weibull estimator. As a result, the corresponding RMSEs reported for Experiment 4 in Table A.1 consistently favor the BFGS MF Weibull model over the BFGS Weibull model by a factor of five to ten. The BFGS Weibull's CPs likewise largely fail to encompass the true parameter values (with a range of 0%-8% across all N's and parameters), which contrasts notably with

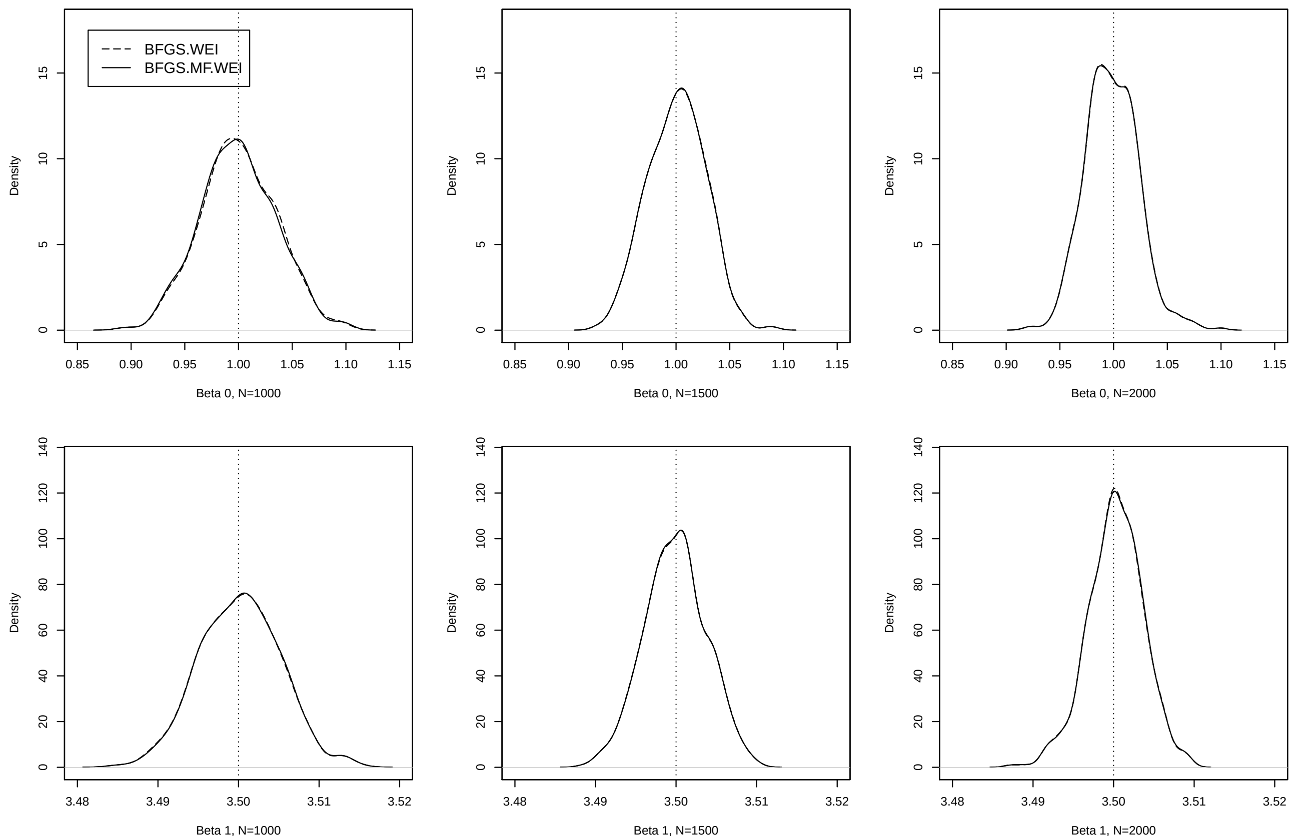


Figure A.4: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 3

those of the BFGS MF Weibull (with a range of 94%-96%) in this Experiment. The BFGS Weibull model's averaged SEs in Table A.1 are each also noticeably larger than those of the BFGS MF Weibull, no matter the β parameter evaluated, or the number of observations considered. Together this suggests that the BFGS MF Weibull model—as was the case for the Bayesian MF Weibull model — is preferable to the standard Weibull estimator when misclassified failure cases exist within one's Weibull distributed outcome variable. Turning to Table A.2 and Figure A.6, we can also briefly note here that the BFGS MF Weibull model's $\hat{\gamma}$'s generally exhibit higher bias, and lower efficiency, than either the BFGS MF Weibull $\hat{\beta}$ results in Table A.1, or the Bayesian MF Weibull $\hat{\gamma}$ results obtained in Experiment 2. The latter finding lends support to our main paper's contention that the Bayesian MF Weibull model is likely preferable to the BFGS MF Weibull model for applied research.

We next turn to MC Experiments 5–6, which compare the performance of the Bayesian

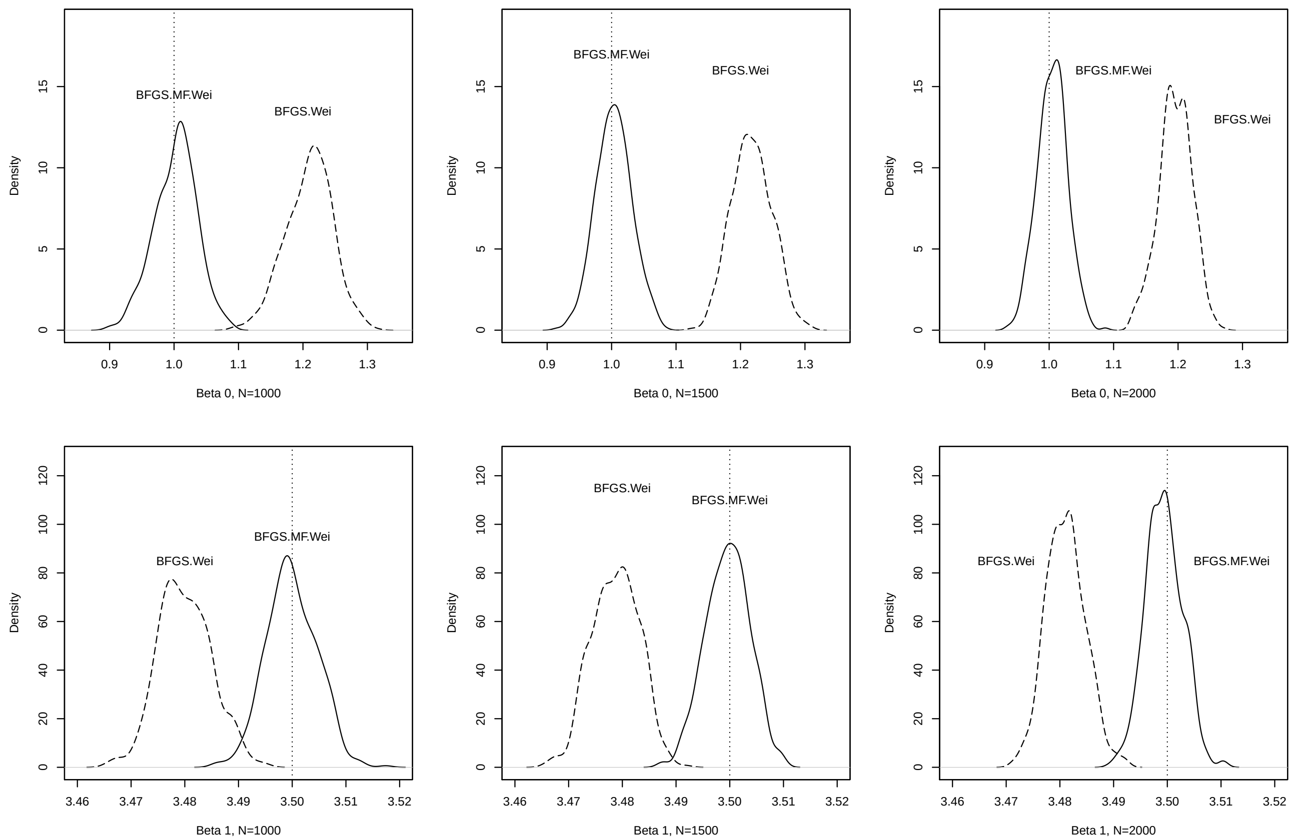


Figure A.5: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 4

(MF) Weibull models to that of the Bayesian MF exponential models when the true d.g.p. is either (i) exponential (Experiment 5) or (ii) MF exponential (Experiment 6). We report the results from these two additional MC experiments in Tables A.3-A.4 and in Figures A.7-A.9.

Beginning first with Experiment 5, we find here that the Bayesian MF estimators again perform comparably to our non-MF Bayesian estimators when the true d.g.p. contains no misclassified failure cases. For example, we find in Table A.3 that there are many case where we obtain slightly lower RMSEs (and hence less bias), and/or slightly superior CPs, within our Bayesian MF exponential and Bayesian MF Weibull model parameter estimates than in the case of the Bayesian Weibull model. The corresponding averaged $\hat{\beta}$ values reported in Table A.3, and the plots of each $\hat{\beta}$ in Figure A.7, strongly support these conclusions. Nevertheless, and as was the case in Experiments 1 and 2, we do find that the non-MF Bayesian Weibull model exhibits consistently lower MCSEs than either the Bayesian MF exponential or the Bayesian

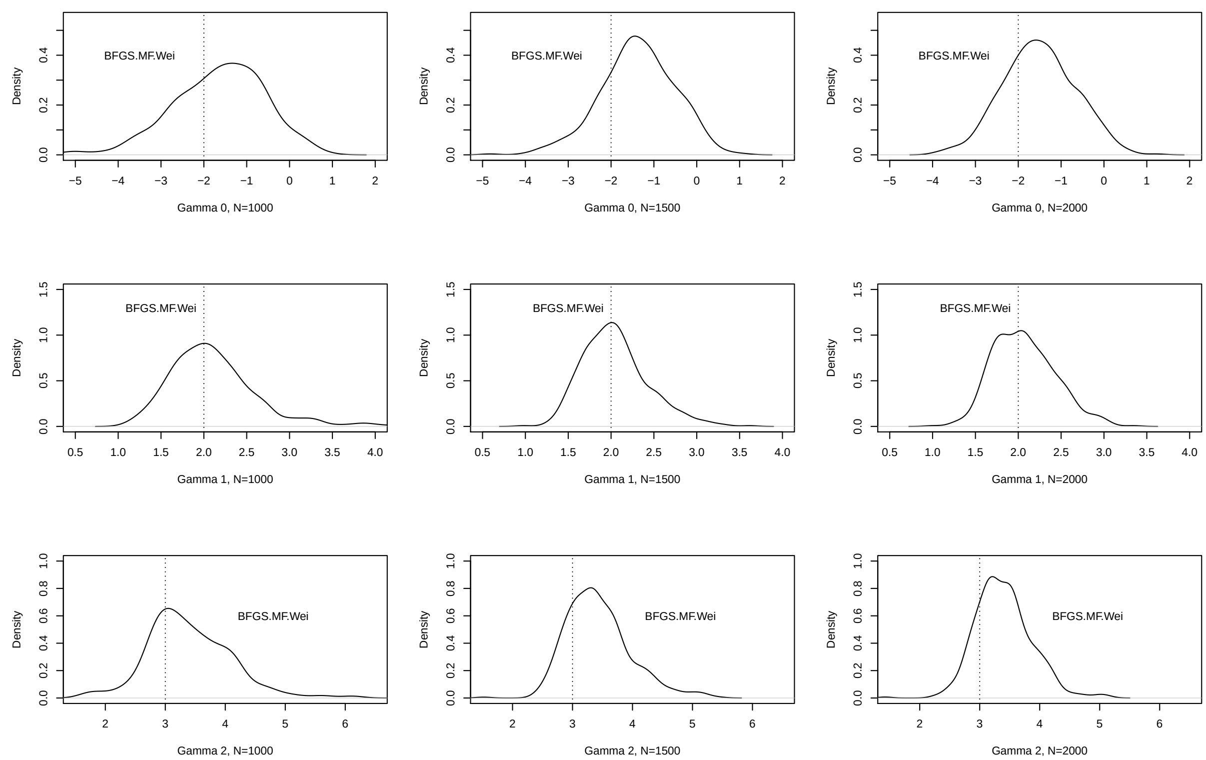


Figure A.6: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 4

MF Weibull within the top half of Table A.3. This suggests that when no misclassified failure cases exist, non-MF Bayesian survival models are preferable to MF Bayesian survival models for reasons of efficiency and parsimony.

With regards to MC Experiment 6, Tables A.3-A.4 and Figures A.8-A.9 reaffirm the conclusions obtained in Experiment 2. For example, the $\hat{\beta}$'s and RMSEs for the Bayesian MF exponential and Bayesian MF Weibull models in Table A.3 consistently indicate that these Bayesian MF survival models exhibit little to no bias in recovered survival-stage parameter estimates when misclassified failure rates exist, whereas the non-MF Bayesian Weibull model consistently overestimates $\hat{\beta}_0$ and underestimates $\hat{\beta}_1$. We find in this case that the corresponding Bayesian MF survival model RMSEs are generally 4-7 times smaller than those of the Bayesian Weibull estimator in Experiment 6, whereas each model's MCSEs are fairly similar across each N evaluated. Similarly, the 95% empirical CPs for each model evaluated consistently favor the Bayesian MF models in every instance, with our Bayesian MF Weibull and Exponential

Table A.3: Markov Chain Monte Carlo (MCMC) β-Estimates for Experiments 5 and 6

	Experiment 5: Non-MF Exponential D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	*MCSE*($\hat{\beta}_0$)	*RMSE*($\hat{\beta}_0$)	*CP*($\hat{\beta}_0$)	$\hat{\beta}_1$	*MCSE*($\hat{\beta}_1$)	*RMSE*($\hat{\beta}_1$)	*CP*($\hat{\beta}_1$)
1000	Bayes Weibul	1.002	0.009	0.052	0.946	3.499	0.001	0.007	0.940
	Bayes MF Exponential	0.999	0.065	0.053	0.950	3.502	0.009	0.008	1.000
	Bayes MF Weibull	0.995	0.067	0.054	1.000	3.502	0.009	0.008	1.000
1500	Bayes Weibull	1.002	0.007	0.041	0.940	3.499	0.001	0.006	0.942
	Bayes MF Exponential	1.001	0.053	0.044	0.922	3.500	0.008	0.006	0.926
	Bayes MF Weibull	1.001	0.053	0.044	0.926	3.500	0.008	0.006	0.948
2000	Bayes Weibull	1.001	0.006	0.037	0.944	3.500	0.001	0.006	0.926
	Bayes MF Exponential	0.999	0.046	0.037	0.950	3.500	0.007	0.005	0.942
	Bayes MF Weibull	0.999	0.046	0.038	0.944	3.500	0.007	0.005	0.948
	Experiment 6: MF Exponential D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	*MCSE*($\hat{\beta}_0$)	*RMSE*($\hat{\beta}_0$)	*CP*($\hat{\beta}_0$)	$\hat{\beta}_1$	*MCSE*($\hat{\beta}_1$)	*RMSE*($\hat{\beta}_1$)	*CP*($\hat{\beta}_1$)
1000	Bayes Weibul	1.309	0.009	0.309	0.002	3.467	0.001	0.033	0.114
	Bayes MF Exponential	1.012	0.009	0.052	0.932	3.499	0.001	0.008	0.930
	Bayes MF Weibull	1.012	0.009	0.053	0.930	3.499	0.001	0.008	0.932
1500	Bayes Weibull	1.317	0.007	0.317	0.000	3.465	0.001	0.035	0.014
	Bayes MF Exponential	1.004	0.008	0.045	0.940	3.499	0.001	0.007	0.936
	Bayes MF Weibull	1.003	0.007	0.046	0.936	3.499	0.001	0.007	0.926
2000	Bayes Weibull	1.290	0.006	0.290	0.000	3.469	0.001	0.031	0.006
	Bayes MF Exponential	1.010	0.007	0.040	0.932	3.499	0.001	0.006	0.918
	Bayes MF Weibull	1.010	0.007	0.041	0.942	3.499	0.001	0.006	0.930

Note: True parameter values are $\beta_0 = 1$ and $\beta_1 = 3.5$.

Table A.4: Markov Chain Monte Carlo (MCMC) γ-Estimates for Experiment 6

	Experiment 6: MF Exponential D.G.P.												
#Obs.	Model	$\hat{\gamma}_0$	*MCSE*($\hat{\gamma}_0$)	*RMSE*($\hat{\gamma}_0$)		*CP*($\hat{\gamma}_0$)$\hat{\gamma}_1$	*MCSE*($\hat{\gamma}_1$)	*RMSE*($\hat{\gamma}_1$)	*CP*($\hat{\gamma}_1$)	$\hat{\gamma}_2$	*MCSE*($\hat{\gamma}_2$)	*RMSE*($\hat{\gamma}_2$)	*CP*($\hat{\gamma}_2$)
1000	Bayes MF Exponential	-1.372	0.312	1.012	0.864	2.033	0.165	0.396	0.846	3.396	0.208	0.611	0.884
	Bayes MF Weibull	-1.350	0.306	1.038	0.872	2.016	0.165	0.389	0.854	3.373	0.205	0.592	0.888
1500	Bayes MF Exponential	-1.169	0.252	1.007	0.778	1.954	0.129	0.329	0.850	3.343	0.163	0.512	0.830
	Bayes MF Weibull	-1.178	0.259	0.994	0.786	1.963	0.129	0.330	0.836	3.357	0.160	0.525	0.854
2000	Bayes MF Exponential	-1.329	0.249	0.879	0.824	1.976	0.128	0.289	0.832	3.300	0.148	0.444	0.862
	Bayes MF Weibull	-1.322	0.246	0.877	0.788	1.976	0.129	0.297	0.836	3.305	0.149	0.456	0.886

Note: True parameter values are $\gamma_0 = -2$, $\gamma_1 = 2$, and $\gamma_2 = 3$.

models exhibiting empirical CPs that consistently lie in the 92%-94% range. By comparison, and across all N's and parameters of interest, our Bayesian Weibull models' CPs fall in the 0%-11% range for this experiment.

Taken together, these findings reaffirm Experiment 2's conclusion that the Bayesian MF survival models are preferable to the Bayesian Weibull when one's d.g.p. contains misclassified failure cases. In these regards, we can further note in Table A.4 and Figures A.8-A.9 that we cannot draw similar conclusions with regards to the preferability of the Bayesian MF exponential over the Bayesian MF Weibull (or vice-versa), as neither model consistently exhibits superior RMSEs (or MCSEs) over the other in these cases. In light of this, and given that the MF

Weibull nests the MF exponential, but more flexibly handles circumstances where one's hazard rate is non-constant, we can conclude that the Bayesian MF Weibull should typically be favored over the Bayesian MF exponential in applied research.

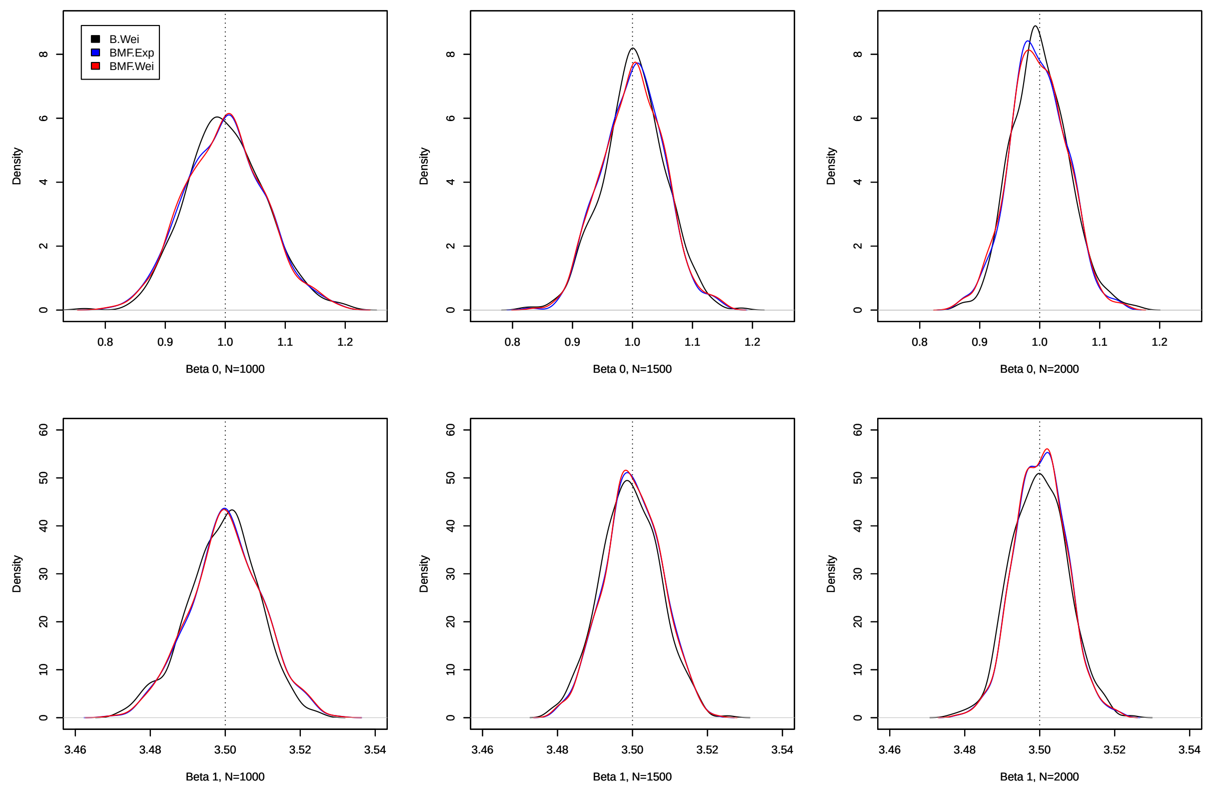


Figure A.7: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 5

We next discuss Experiments 7-8, which evaluate the performance of the BFGS (MF) Weibull and exponential models when the true d.g.p. corresponds to either (i) a simple exponential survival process with no instances of misclassified failures (Experiment 7) or an exponential process with 5% misclassified failure cases (Experiment 8). We report our $\hat{\beta}$ results for these two MC experiments within a Table A.5, and also plot the full distributions of these β parameter estimates within Figures A.10-A.11. In cases where a researcher encounters a non-MF exponential-distributed outcome variable, we find in the top-half of Table A.5 (and in Figure A.10) that our BFGS MF survival models perform commensurately with respect to efficiency, coverage, and accuracy. For both β parameters, and across each N evaluated, the BFGS MF exponential and BFGS MF Weibull models' averaged parameter estimates and 95%

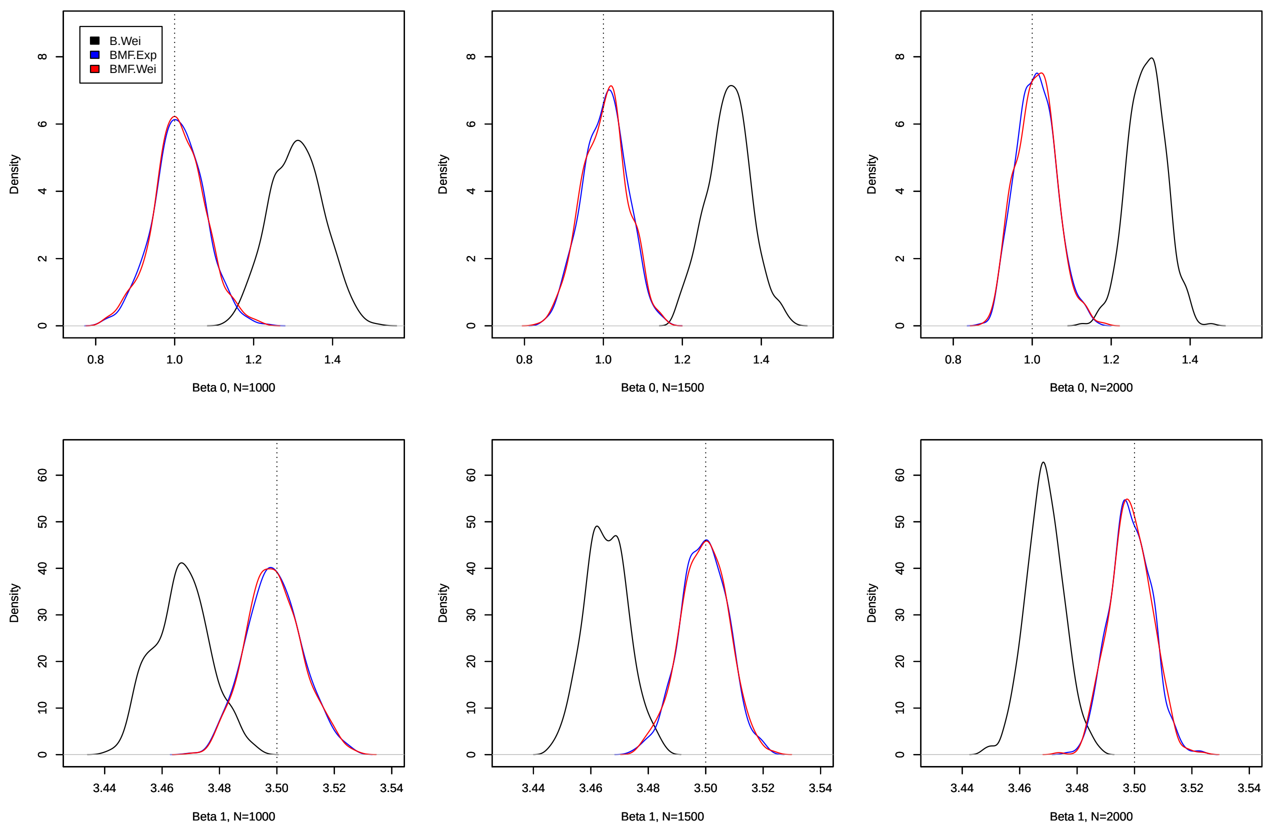


Figure A.8: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 6

CPs are virtually identical to those of the standard BFGS exponential and Weibull models. These similarities between the MF and non-MF survival models evaluated within Experiment 7 are also reflected in the SEs and RMSEs reported in the top half of Table A.5, which are generally identical (to the third decimal place) for each BFGS model pair considered.

We can also note in Table A.5 that in cases where one's true d.g.p. is exponential, the BFGS MF Weibull and BFGS MF exponential models recover very similar averaged $\hat{\beta}$ values, and exhibit near-identical 95% CPs, RMSEs and SEs. These latter findings generally hold true under Experiment 8 as well, when the true d.g.p. is MF exponential. For instance, each of the $\hat{\beta}$ values reported for the BFGS MF Weibull and BFGS MF exponential models in the lower half of A.5 (and in Figure A.11) are comparable, and the corresponding SEs and RMSEs for these two BFGS MF models are effectively equivalent. The same can be said for the $\hat{\gamma}$'s recovered by the BFGS MF exponential and BFGS MF Weibull models in Experiment 8, which are reported

Table A.5: Maximum Likelihood β-Estimates for Experiments 7 and 8

#Obs.	Model	$\hat{\beta}_0$	$SE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$SE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
	Experiment 7: Non-MF Exponential D.G.P.								
1000	BFGS Exponential	1.002	0.065	0.054	0.950	3.502	0.009	0.007	1.000
	BFGS Weibull	1.000	0.066	0.053	1.000	3.501	0.009	0.007	1.000
	BFGS MF Exponential	1.001	0.065	0.054	1.000	3.502	0.009	0.007	1.000
	BFGS MF Weibull	0.999	0.066	0.053	1.000	3.502	0.009	0.007	1.000
1500	BFGS Exponential	1.002	0.053	0.044	0.938	3.500	0.008	0.006	0.946
	BFGS Weibull	1.001	0.054	0.044	0.932	3.500	0.008	0.006	0.946
	BFGS MF Exponential	1.002	0.053	0.044	0.926	3.500	0.008	0.006	0.943
	BFGS MF Weibull	1.001	0.054	0.044	0.933	3.500	0.008	0.006	0.942
2000	BFGS Exponential	1.000	0.047	0.037	0.958	3.500	0.007	0.005	0.956
	BFGS Weibull	1.000	0.047	0.037	0.956	3.500	0.007	0.005	0.956
	BFGS MF Exponential	1.000	0.047	0.037	0.955	3.500	0.007	0.005	0.953
	BFGS MF Weibull	0.999	0.047	0.037	0.953	3.500	0.007	0.005	0.953
	Experiment 8: MF Exponential D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	$SE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$SE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
1000	BFGS Exponential	1.245	0.061	0.245	0.028	3.468	0.009	0.032	0.078
	BFGS Weibull	1.318	0.071	0.318	0.008	3.466	0.011	0.034	0.104
	BFGS MF Exponential	1.013	0.067	0.051	0.942	3.499	0.010	0.008	0.946
	BFGS MF Weibull	1.012	0.067	0.052	0.938	3.499	0.010	0.008	0.950
1500	BFGS Exponential	1.243	0.049	0.243	0.004	3.466	0.008	0.034	0.006
	BFGS Weibull	1.321	0.059	0.321	0.000	3.465	0.009	0.035	0.012
	BFGS MF Exponential	1.005	0.054	0.045	0.948	3.499	0.008	0.007	0.942
	BFGS MF Weibull	1.004	0.054	0.045	0.950	3.499	0.008	0.007	0.940
2000	BFGS Exponential	1.228	0.043	0.228	0.000	3.469	0.007	0.031	0.008
	BFGS Weibull	1.294	0.050	0.294	0.000	3.468	0.008	0.032	0.012
	BFGS MF Exponential	1.010	0.047	0.040	0.946	3.499	0.007	0.006	0.940
	BFGS MF Weibull	1.010	0.047	0.040	0.946	3.499	0.007	0.006	0.940

Note: True parameter values are $\beta_0 = 1$ and $\beta_1 = 3.5$.

Table A.6: Maximum Likelihood γ-Estimates for Experiment 8

#Obs.	Model	$\hat{\gamma}_0$	$SE(\hat{\gamma}_0)$	$RMSE(\hat{\gamma}_0)$	$CP(\hat{\gamma}_0)$	$\hat{\gamma}_1$	$SE(\hat{\gamma}_1)$	$RMSE(\hat{\gamma}_1)$	$CP(\hat{\gamma}_1)$	$\hat{\gamma}_2$	$SE(\hat{\gamma}_2)$	$RMSE(\hat{\gamma}_2)$	$CP(\hat{\gamma}_2)$
	Experiment 8: MF Weibull D.G.P.												
1000	BFGS MF Exponential	-1.756	1.232	0.966	0.968	2.153	0.498	0.395	0.970	3.447	0.723	0.622	0.982
	BFGS MF Weibull	-1.748	1.226	0.960	0.968	2.136	0.494	0.394	0.964	3.417	0.715	0.627	0.974
1500	BFGS MF Exponential	-1.421	0.978	0.878	0.930	2.041	0.382	0.304	0.950	3.391	0.575	0.524	0.966
	BFGS MF Weibull	-1.426	0.974	0.872	0.930	2.032	0.381	0.308	0.946	3.372	0.571	0.529	0.960
2000	BFGS MF Exponential	-1.587	0.912	0.736	0.942	2.056	0.361	0.278	0.950	3.336	0.512	0.449	0.952
	BFGS MF Weibull	-1.584	0.913	0.737	0.940	2.056	0.361	0.278	0.952	3.338	0.513	0.449	0.958

Note: True parameter values are $\gamma_0 = -2$, $\gamma_1 = 2$, and $\gamma_2 = 3$.

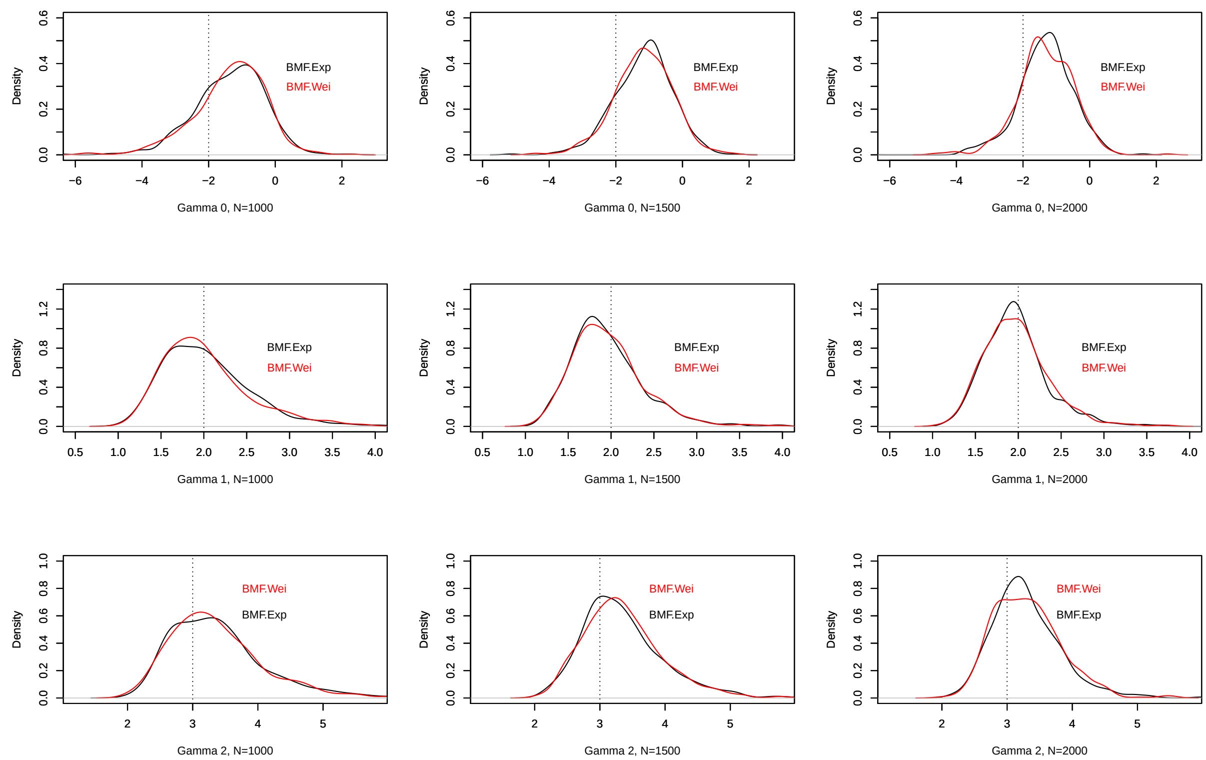


Figure A.9: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 6

in Table A.6 and Figure A.12. In this case, the BFGS MF exponential generally exhibits slightly higher bias than the BFGS MF Weibull when $N = 1,000$, but slightly lower or equivalent level of bias compared to the BFGS MF Weibull when $N = 1,500$ or $N = 2,000$. Returning to Table A.5 and Figure A.11, we can also note that our (BFGS) MF survival models substantially outperform the (BFGS) non-MF survival models when the true d.g.p. includes misclassified failure cases. However, we further find in this regard that our BFGS MF models' estimates of uncertainty and (to a lesser extent) RMSEs in Tables A.5-A.6 are generally inferior to those obtained for the Bayesian MF models in Tables A.3-A.4. As was the case for Experiments 1-4, these latter patterns suggest that Bayesian MF models are preferable to BFGS MF models for applied research.

Experiments 9-11 reevaluate the performance of the Bayesian Weibull and Bayesian MF Weibull models when applied to a MF Weibull-distributed outcome variable that exhibits a noticeably higher MF rate (i.e., α) than was the case for Experiments 2, 4, 6, and 8. More

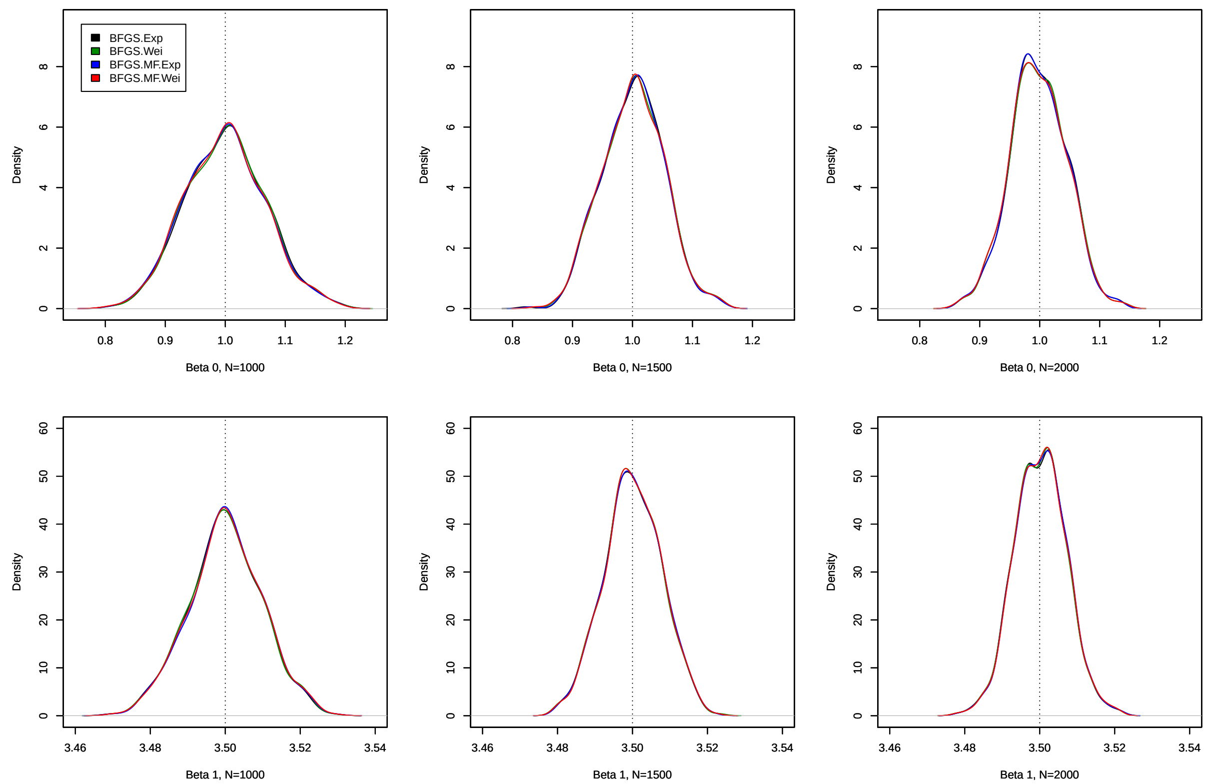


Figure A.10: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 7

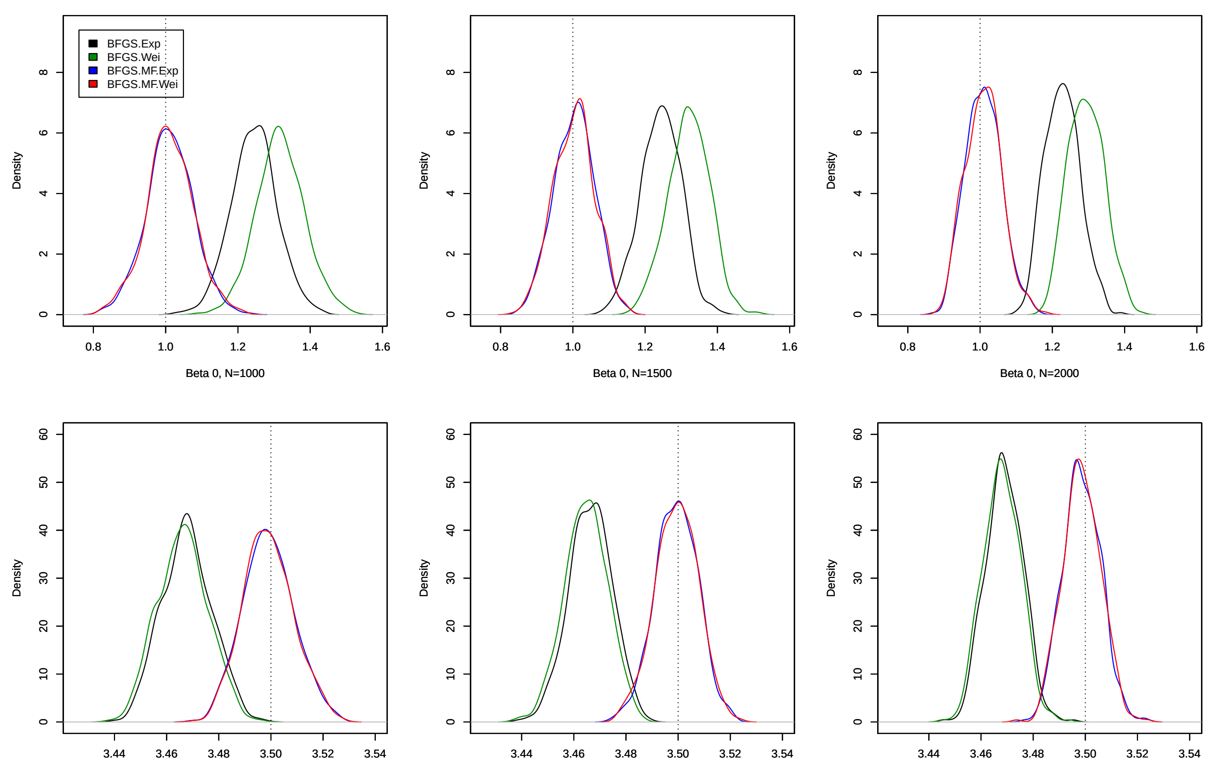


Figure A.11: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 8

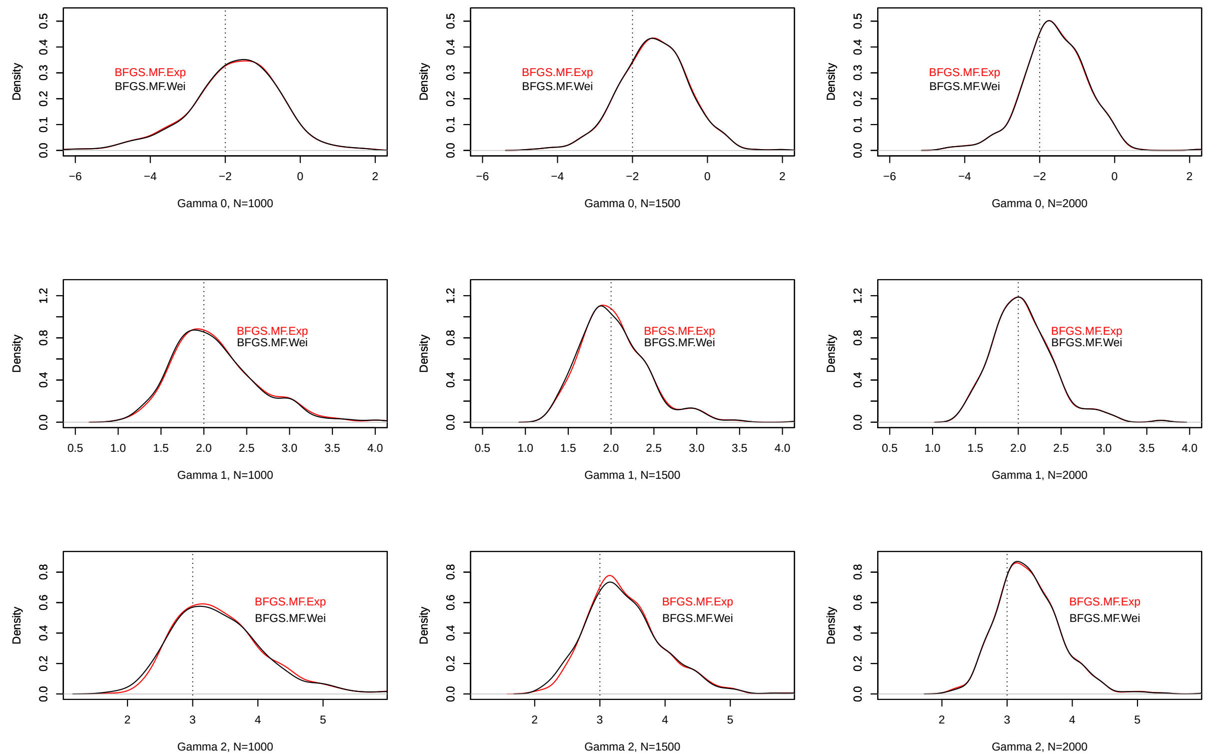


Figure A.12: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 8

specifically, Experiments 9, 10, and 11 compare the relative performance of the Bayesian Weibull and Bayesian MF Weibull models when one's MF rate increases (from 5%) to 8%, 12%, and 15%, respectively. We report these MC results in Table A.7 (β parameters) and in Table A.8 (γ parameters). We also provide the full distributions of each estimated β and γ parameter (i.e., across all 500 simulations) within Figures A.13-A.15 and Figures A.16-A.18, respectively. These tables and figures together demonstrate that the previously identified advantages of the Bayesian MF Weibull model (i.e., over the Bayesian Weibull model) become even more notable as one increases a Weibull-distributed outcome variable's MF rate from 5% to 8-15%. Beginning with Table A.7, we can note for example that the standard Bayesian Weibull model's mean $\hat{\beta}_0$'s substantially (and increasingly) overestimate β_0 as one's MF rate increases from 8%-to-15%; whereas the standard Bayesian Weibull model's $\hat{\beta}_1$'s increasingly underestimate β_1 under these same conditions.

By contrast, the Bayesian MF Weibull model's averaged β estimates do not exhibit any

Table A.7: Markov Chain Monte Carlo (MCMC) β-Estimates for Experiments 9-11

	Experiment 9: MF Weibull D.G.P. with $\alpha = 8\%$								
#Obs.	Model	$\hat{\beta}_0$	$MCSE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$MCSE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
1000	Bayes Weibull	1.371	0.005	0.371	0.000	3.462	0.001	0.038	0.000
	Bayes MF Weibull	1.000	0.003	0.026	0.938	3.500	0.000	0.004	0.898
1500	Bayes Weibull	1.361	0.004	0.361	0.000	3.463	0.001	0.037	0.000
	Bayes MF Weibull	1.000	0.003	0.023	0.936	3.500	0.000	0.003	0.922
2000	Bayes Weibull	1.376	0.003	0.376	0.000	3.461	0.000	0.039	0.000
	Bayes MF Weibull	1.000	0.002	0.018	0.914	3.500	0.000	0.003	0.928
	Experiment 10: MF Weibull D.G.P. with $\alpha = 12\%$								
#Obs.	Model	$\hat{\beta}_0$	$MCSE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$MCSE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
1000	Bayes Weibull	1.489	0.004	0.489	0.000	3.446	0.001	0.054	0.000
	Bayes MF Weibull	1.001	0.003	0.027	0.828	3.500	0.000	0.004	0.900
1500	Bayes Weibull	1.473	0.003	0.473	0.000	3.447	0.001	0.053	0.000
	Bayes MF Weibull	1.005	0.003	0.022	0.672	3.499	0.000	0.003	0.874
2000	Bayes Weibull	1.462	0.003	0.462	0.000	3.449	0.000	0.051	0.000
	Bayes MF Weibull	1.005	0.002	0.020	0.592	3.499	0.000	0.003	0.870
	Experiment 11: MF Weibull D.G.P. with $\alpha = 15\%$								
#Obs.	Model	$\hat{\beta}_0$	$MCSE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$MCSE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
1000	Bayes Weibull	1.559	0.004	0.559	0.000	3.435	0.001	0.065	0.000
	Bayes MF Weibull	1.011	0.004	0.030	0.880	3.499	0.001	0.004	0.868
1500	Bayes Weibull	1.552	0.004	0.552	0.000	3.437	0.001	0.063	0.000
	Bayes MF Weibull	1.010	0.003	0.023	0.722	3.499	0.000	0.003	0.892
2000	Bayes Weibull	1.557	0.003	0.557	0.000	3.436	0.000	0.064	0.000
	Bayes MF Weibull	1.009	0.002	0.020	0.782	3.499	0.000	0.003	0.924

Note: True parameter values are $\beta_0 = 1$ and $\beta_1 = 3.5$.

Table A.8: Markov Chain Monte Carlo (MCMC) γ-Estimates for Experiments 9-11

	Experiment 9: MF Weibull D.G.P. with $\alpha = 8\%$												
#Obs.	Model	$\hat{\gamma}_0$	$MCSE(\hat{\gamma}_0)$	$RMSE(\hat{\gamma}_0)$	$CP(\hat{\gamma}_0)$	$\hat{\gamma}_1$	$MCSE(\hat{\gamma}_1)$	$RMSE(\hat{\gamma}_1)$	$CP(\hat{\gamma}_1)$	$\hat{\gamma}_2$	$MCSE(\hat{\gamma}_2)$	$RMSE(\hat{\gamma}_2)$	$CP(\hat{\gamma}_2)$
1000	Bayes MF Weibull	2.003	0.278	0.913	0.948	1.084	0.086	0.272	0.950	4.208	0.159	0.578	0.942
1500	Bayes MF Weibull	1.962	0.192	0.707	0.948	1.077	0.049	0.192	0.938	4.167	0.116	0.442	0.936
2000	Bayes MF Weibull	2.139	0.206	0.672	0.938	1.030	0.057	0.182	0.952	4.169	0.097	0.387	0.938
	Experiment 10: MF Weibull D.G.P. with $\alpha = 12\%$												
#Obs.	Model	$\hat{\gamma}_0$	$MCSE(\hat{\gamma}_0)$	$RMSE(\hat{\gamma}_0)$	$CP(\hat{\gamma}_0)$	$\hat{\gamma}_1$	$MCSE(\hat{\gamma}_1)$	$RMSE(\hat{\gamma}_1)$	$CP(\hat{\gamma}_1)$	$\hat{\gamma}_2$	$MCSE(\hat{\gamma}_2)$	$RMSE(\hat{\gamma}_2)$	$CP(\hat{\gamma}_2)$
1000	Bayes MF Weibull	-1.886	0.418	1.571	0.930	1.920	0.176	0.447	0.956	5.706	0.218	0.940	0.952
1500	Bayes MF Weibull	-1.363	0.270	1.689	0.872	1.800	0.121	0.375	0.954	5.730	0.175	0.886	0.956
2000	Bayes MF Weibull	-1.441	0.261	1.636	0.840	1.842	0.118	0.331	0.930	5.776	0.161	0.867	0.932
	Experiment 11: MF Weibull D.G.P. with $\alpha = 15\%$												
#Obs.	Model	$\hat{\gamma}_0$	$MCSE(\hat{\gamma}_0)$	$RMSE(\hat{\gamma}_0)$	$CP(\hat{\gamma}_0)$	$\hat{\gamma}_1$	$MCSE(\hat{\gamma}_1)$	$RMSE(\hat{\gamma}_1)$	$CP(\hat{\gamma}_1)$	$\hat{\gamma}_2$	$MCSE(\hat{\gamma}_2)$	$RMSE(\hat{\gamma}_2)$	$CP(\hat{\gamma}_2)$
1000	Bayes MF Weibull	7.148	1.334	3.439	0.516	-1.028	0.232	0.675	0.924	9.424	0.720	4.424	0.932
1500	Bayes MF Weibull	7.568	0.802	3.275	0.272	-1.179	0.135	0.478	0.938	8.977	0.502	3.977	0.942
2000	Bayes MF Weibull	6.759	0.518	2.406	0.168	-1.013	0.094	0.343	0.936	8.385	0.278	3.385	0.926

Note: True parameter values are $\gamma_0 = 2$, $\gamma_1 = 1$, & $\gamma_2 = 4$ (Experiment 9); $\gamma_0 = -3$, $\gamma_1 = 2$, & $\gamma_2 = 5$ (Experiment 10); and $\gamma_0 = 4.5$, $\gamma_1 = -1$, & $\gamma_2 = 5$ (Experiment 11).

notable trends in increasing (or decreasing) size as one increases the MF rate beyond 5%, suggesting that whereas Bayesian Weibull model's estimates become more biased as the MF rate increases, the Bayesian MF Weibull model remains comparatively unbiased. This contention is reinforced by the reported RMSEs in Table A.7, and Figures A.13-A.15. With regards to the former quantities, for example, we find in Table A.7 that our Bayesian MF Weibull model's $\hat{\beta}_0$'s exhibit RMSEs that are generally 16 times smaller than those of the Bayesian Weibull model when α=8%, and RMSEs that are generally 24 times smaller than those of the Bayesian Weibull model when α=15%. The findings for $\hat{\beta}_1$ are similar, and demonstrate that one's Bayesian MF Weibull exhibits RMSEs that are 12 times smaller than those of the Bayesian Weibull model when α=8%, and RMSEs that are generally 21 times smaller than those of the Bayesian Weibull model when α=15%. At the same time, the standard Bayesian Weibull model's empirical CPs consistently fail to encompass our true β's, whereas the Bayesian Weibull exhibits 95% credible intervals that encompass our relevant β parameters in 93%-95% of all simulations evaluated under Experiments 9-11.

Turning next to the MF Weibull γ estimates for Experiments 9-11 (Table A.8 and Figures A.16-A.18), we can note that our averaged $\hat{\gamma}$ values for these Bayesian MF Weibull models generally recover one's true γ values at comparable rates to those of the Bayesian MF Weibull models in Experiments 2, 4, 6, and 8. However, we can again note that the Bayesian MF Weibull model's $\hat{\gamma}$'s in Table A.8 exhibit higher bias, and lower efficiency, than was the case for these same Bayesian MF Weibulls' $\hat{\beta}$'s in Table A.7.

Experiments 12-13 return to our primary experiments (i.e., Experiment 1 and Experiment 2), and reevaluate the performance of the Bayesian Weibull and Bayesian MF Weibull models for these two Experiments when each Bayesian model uses a (very-) weakly informative multivariate Cauchy prior, which is distinct from our preferred weakly-informative multivariate normal prior. The results from these additional exercises are presented in Tables A.9-A.10 and Figures A.19-A.21.

Turning first to Experiment 12—which evaluates the performance of our alternate-prior Bayesian Weibull and Bayesian MF Weibull models when the true d.g.p. is Weibull with no misclassified failures—we find in Table A.9 and Figure A.19 that the Bayesian MF Weibull again exhibits comparable performance to our standard Bayesian Weibull estimator. To this end, we can note that the averaged parameter estimates reported for all models on the top half of Table A.9 are effectively identical, as our the empirical 95% CPs and our reported parameter

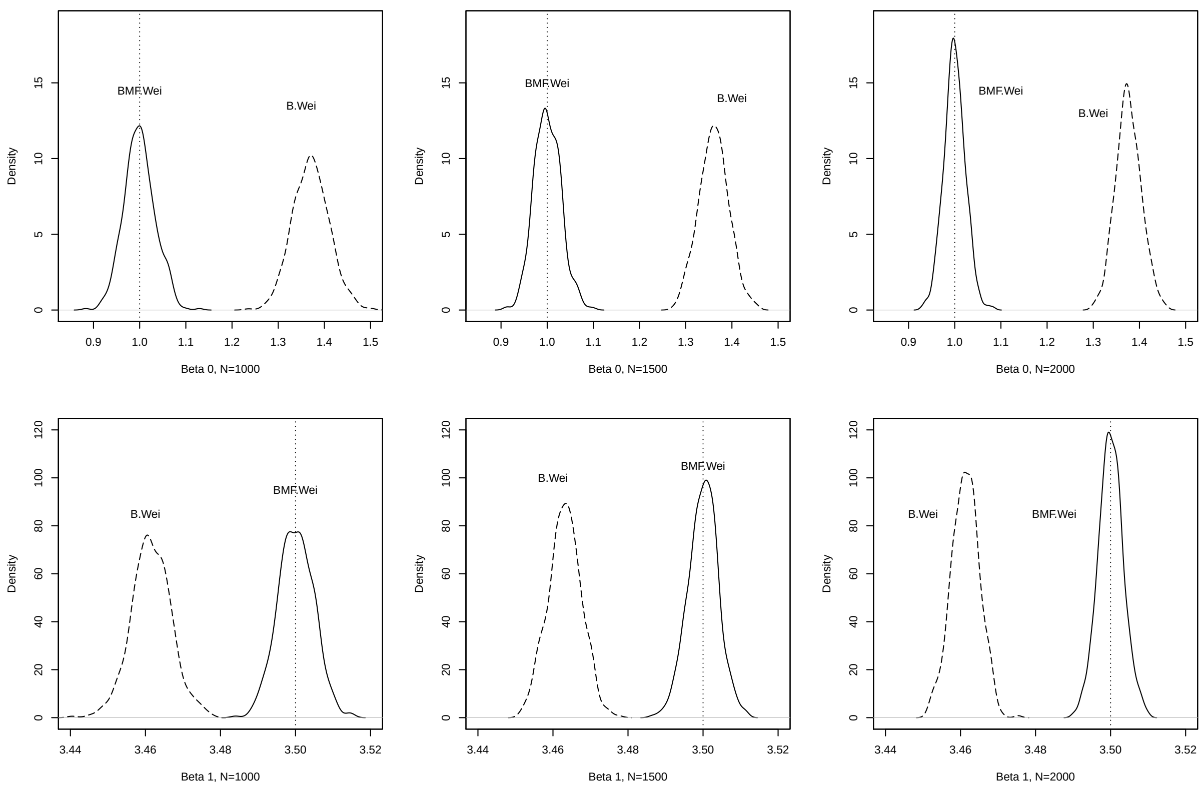


Figure A.13: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 9

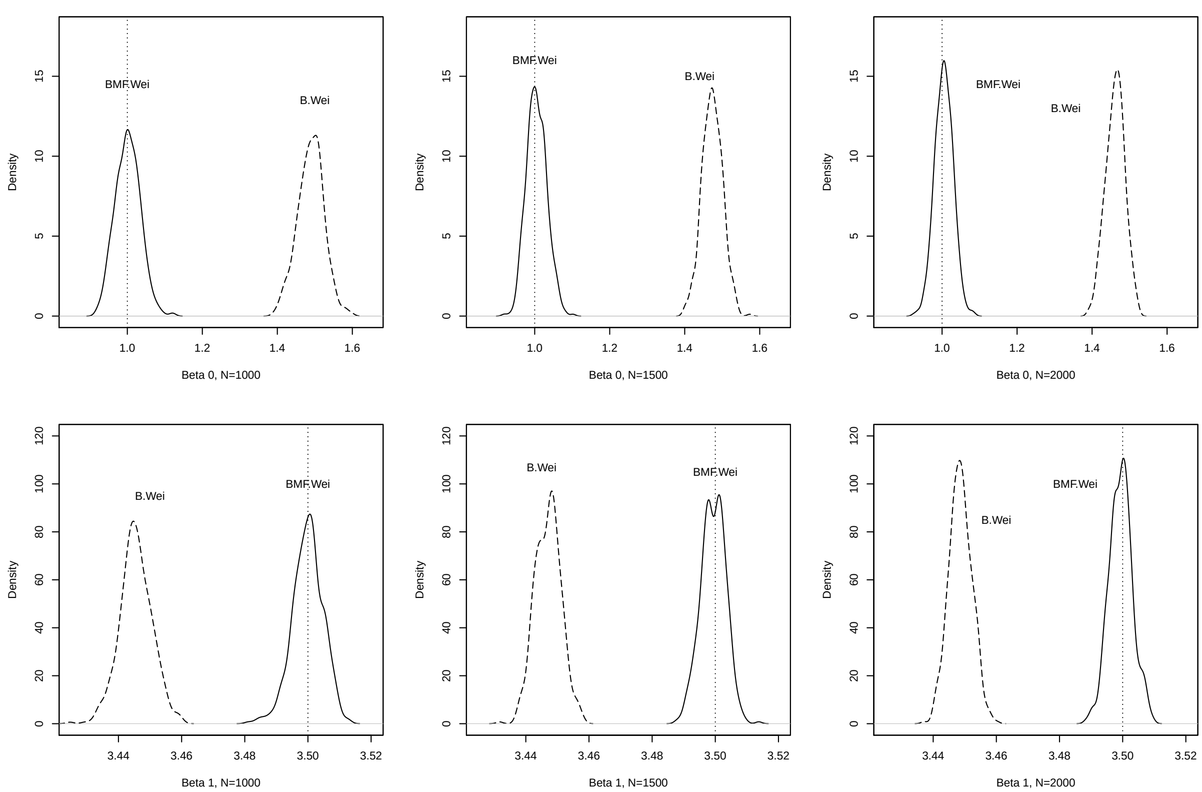


Figure A.14: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 10

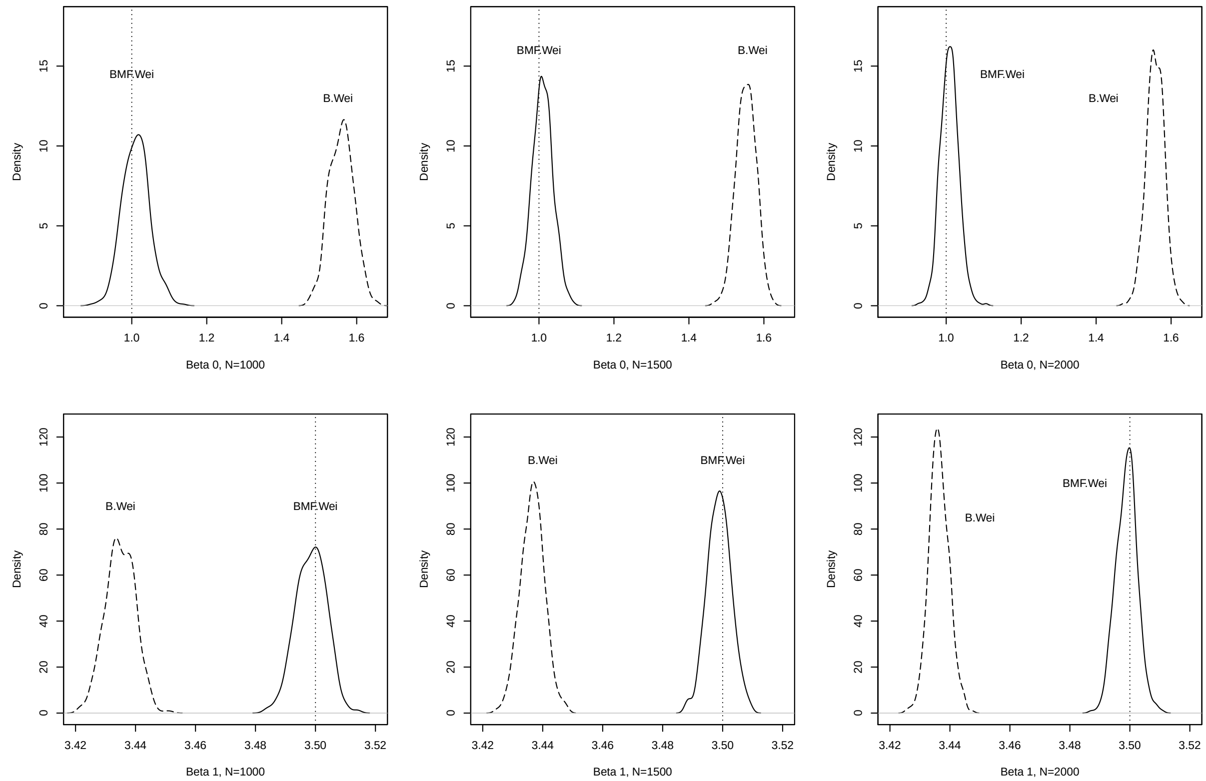


Figure A.15: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 11

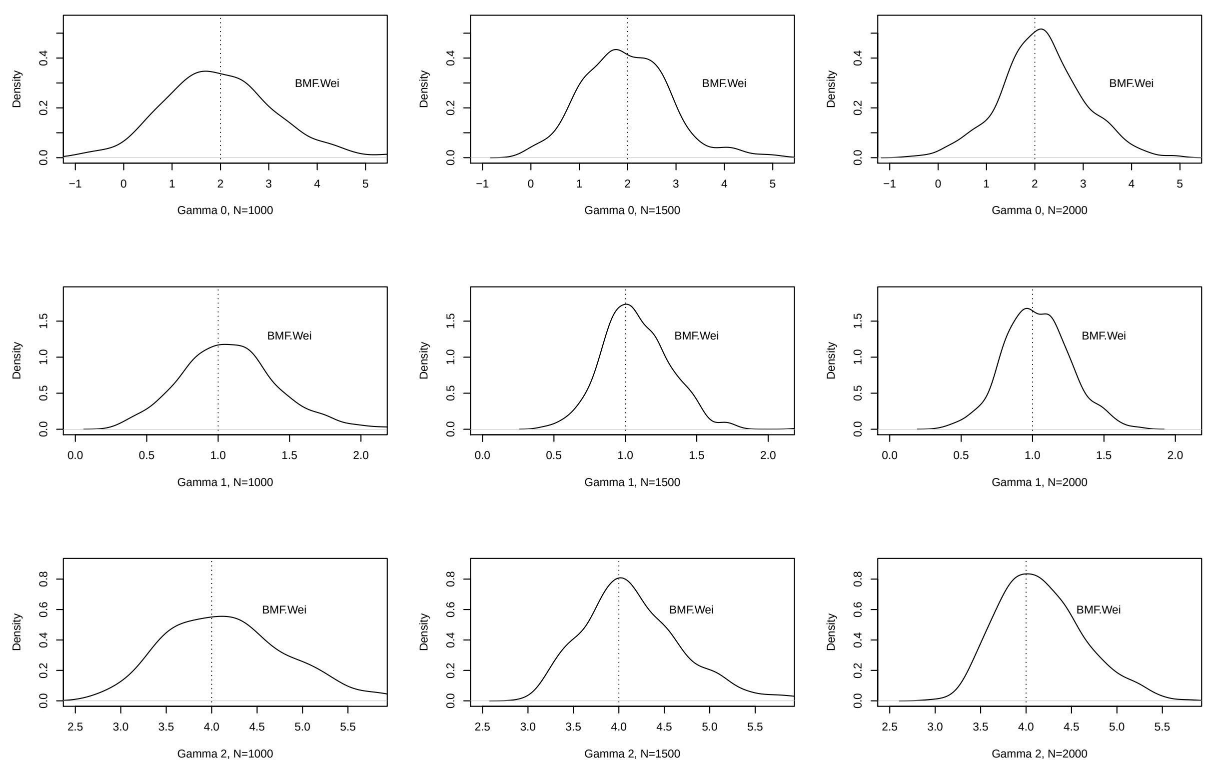


Figure A.16: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 9

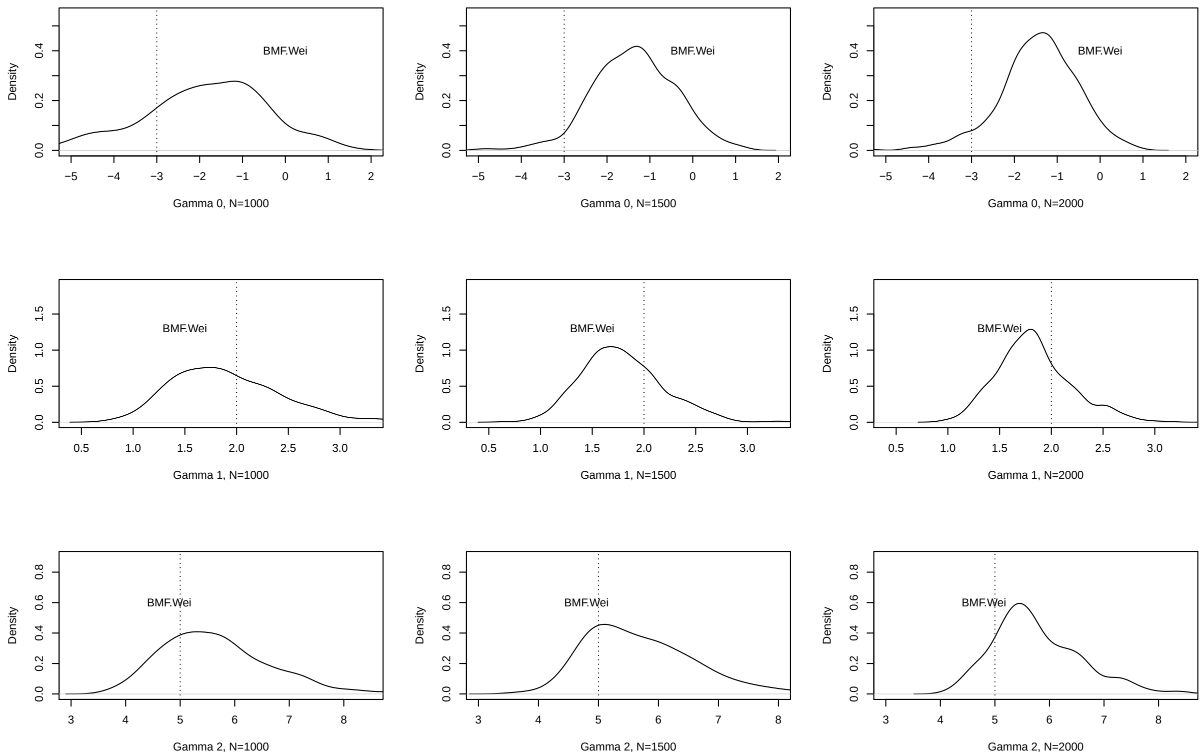


Figure A.17: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 10

estimates corresponding RMSEs. At the same time, although both models consistently exhibit low MCSEs, the Bayesian Weibull model's MCSEs are generally slightly smaller than those of the Bayesian MF Weibull. If we then compare these results to those obtained under Experiment 1—which examined a comparable d.g.p.–we find in the case of Experiment 12 that the use of a multivariate Cauchy prior (relative to the use of a multivariate normal prior) does not appear to have a consistent effect on either bias or coverage; as our RMSEs and CPs in Table A.9 are at times identical to those reported in Table 1 of the main paper; and at other times are slightly smaller or slightly larger than the corresponding Table 1 values. This is the case no matter whether one examines the Bayesian Weibull or the Bayesian MF Weibull models. Altogether these results suggest that our parameter estimates and empirical coverage probabilities are not overly sensitive to the choices of prior specification considered here, at least for the current experimental conditions evaluated.

With regards to Experiment 13—which evaluates the performance of our alternate-prior

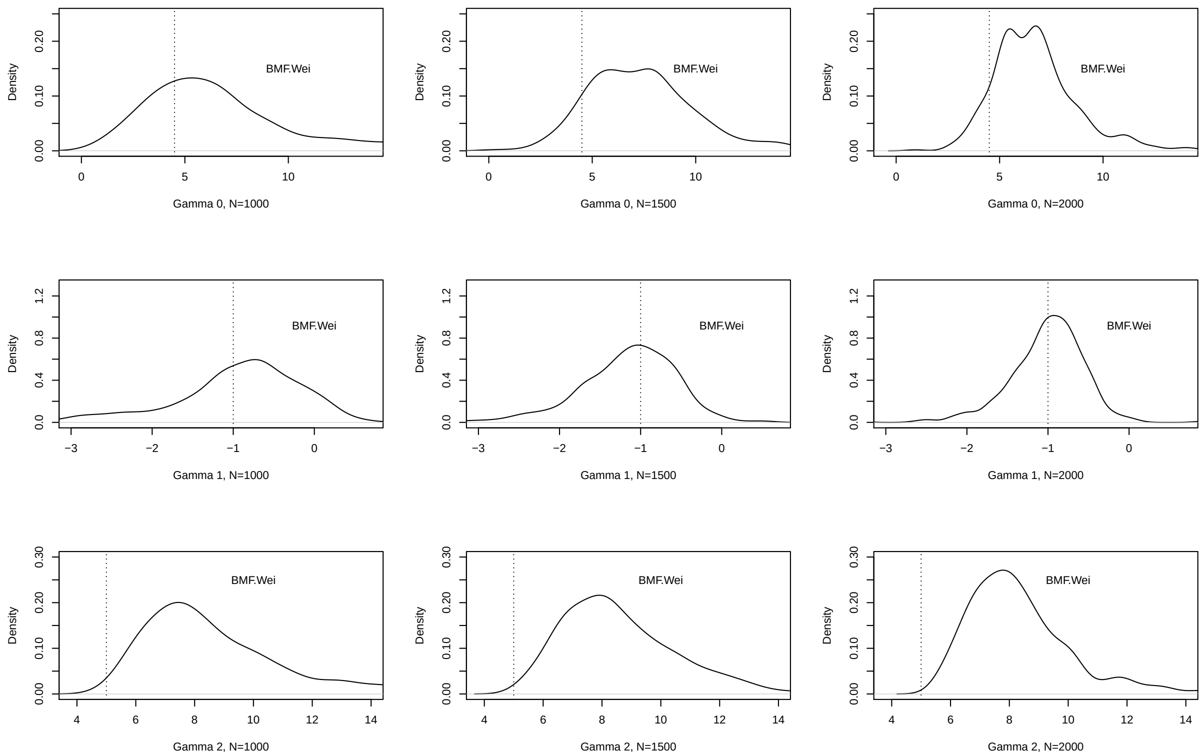


Figure A.18: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 11

Table A.9: Markov Chain Monte Carlo (MCMC) β-Estimates for Experiments 12-13

	Experiment 12: Non-MF Weibull D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	$MCSE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$MCSE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
1000	Bayes Weibull	1.001	3.27E-03	0.027	0.946	3.500	4.67E-04	0.004	0.918
	Bayes MF Weibull	0.999	3.24E-03	0.028	0.094	3.500	4.63E-04	0.004	1.000
1500	Bayes Weibull	0.999	2.63E-03	0.021	0.954	3.500	3.84E-04	0.003	0.938
	Bayes MF Weibull	1.000	2.64E-03	0.022	0.930	3.500	3.85E-04	0.003	0.998
2000	Bayes Weibull	0.998	2.31E-03	0.020	0.926	3.500	3.35E-04	0.003	0.926
	Bayes MF Weibull	1.000	2.32E-03	0.019	0.292	3.500	3.36E-04	0.003	0.998
	Experiment 13: MF Weibull D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	$MCSE(\hat{\beta}_0)$	$RMSE(\hat{\beta}_0)$	$CP(\hat{\beta}_0)$	$\hat{\beta}_1$	$MCSE(\hat{\beta}_1)$	$RMSE(\hat{\beta}_1)$	$CP(\hat{\beta}_1)$
1000	Bayes Weibull	1.212	3.88E-03	0.212	0.000	3.480	5.81E-04	0.020	0.042
	Bayes MF Weibull	1.003	3.30E-03	0.029	0.934	3.500	4.73E-04	0.004	0.940
1500	Bayes Weibull	1.216	3.19E-03	0.216	0.000	3.479	4.86E-04	0.021	0.006
	Bayes MF Weibull	1.003	2.67E-03	0.023	0.930	3.499	3.87E-04	0.003	0.944
2000	Bayes Weibull	1.193	2.76E-03	0.193	0.000	3.482	4.16E-04	0.018	0.004
	Bayes MF Weibull	1.003	2.35E-03	0.019	0.934	3.500	3.38E-04	0.003	0.946

Note: True parameter values are $\beta_0 = 1$ and $\beta_1 = 3.5$.

Bayesian Weibull and Bayesian MF Weibull models when the true d.g.p. is MF Weibull with 5% misclassified failures—we find in A.9-A.10 and Figures A.20-A.21 that the Bayesian MF Weibull model now clearly outperforms the Bayesian Weibull estimator. This finding is consistent with Experiment 2, and can be most readily observed in the bottom half of Table A.9 via the averaged RMSEs and 95% CPs reported therein. To this end, we can observe that our Bayesian Weibull model RMSEs are generally 5-7 times larger than those of the Bayesian MF Weibull model. Concurrently, the Bayesian Weibull model in the Experiment 13 portion of Table A.9 exhibits CPs that range from 0%-4%. Given that our Bayesian MF Weibull model's comparable CPs lie within the 93%-94% range, this suggest that the Bayesian Weibull model substantially under-performs relative to the Bayesian MF Weibull in empirical coverage within Experiment 13. Next, comparing our models across the choice of prior specification, we find that our Experiment 13 results are highly similar (in RMSEs and CPs) to those reported in Experiment 2; suggesting again that the choice of prior specification in this case has little effect on bias or coverage. Experiment 13's MF Weibull γ estimates (Table A.10 and Figure A.21) largely reinforce these conclusions. However, we do find that the use of a multivariate Cauchy prior tends to lead to slightly higher bias in one's γ estimates relative to our primary (multivariate normal) Bayesian MF Weibull survival models.

Table A.10: Markov Chain Monte Carlo (MCMC) γ-Estimates for Experiment 13

	Experiment 13: MF Weibull D.G.P.												
#Obs.	Model	$\hat{\gamma}_0$	*MCSE*($\hat{\gamma}_0$)	*RMSE*($\hat{\gamma}_0$)	*CP*($\hat{\gamma}_0$)	$\hat{\gamma}_1$	*MCSE*($\hat{\gamma}_1$)	*RMSE*($\hat{\gamma}_1$)	*CP*($\hat{\gamma}_1$)	$\hat{\gamma}_2$	*MCSE*($\hat{\gamma}_2$)	*RMSE*($\hat{\gamma}_2$)	*CP*($\hat{\gamma}_2$)
1000	Bayes MF Weibull	-1.437	0.287	1.078	0.870	2.109	0.207	0.465	0.888	3.504	0.250	0.674	0.918
1500	Bayes MF Weibull	-1.177	0.188	0.965	0.812	1.974	0.105	0.302	0.890	3.388	0.123	0.504	0.902
2000	Bayes MF Weibull	-1.349	0.249	0.914	0.844	2.094	0.163	0.359	0.876	3.516	0.197	0.628	0.874

Note: True parameter values are $\gamma_0 = -2$, $\gamma_1 = 2$, and $\gamma_2 = 3$.

How do our proposed Bayesian (MF) Weibull models perform when applied to a (MF) d.g.p. that is explicitly non-Weibull,[3] and how do our Bayesian (MF) Weibull models compare under such circumstances to an estimator that makes no assumptions about the shape of one's baseline hazard function? To begin to answer these questions, Experiments 14-15 generate a log-logistic (Experiment 14) or MF log-logistic (Experiment 15) survival dependent variable in a

[3]Note that the exponential d.g.p.'s evaluated above correspond to a Weibull d.g.p. with $\rho = 1$.

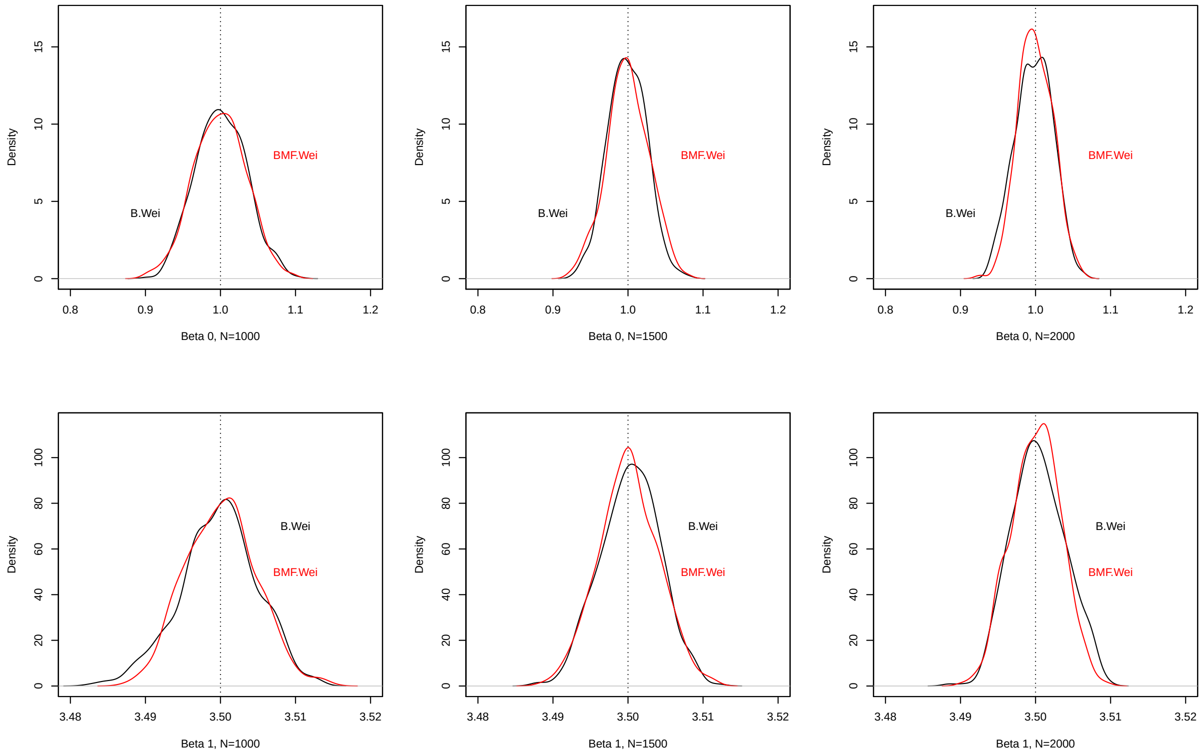


Figure A.19: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 12

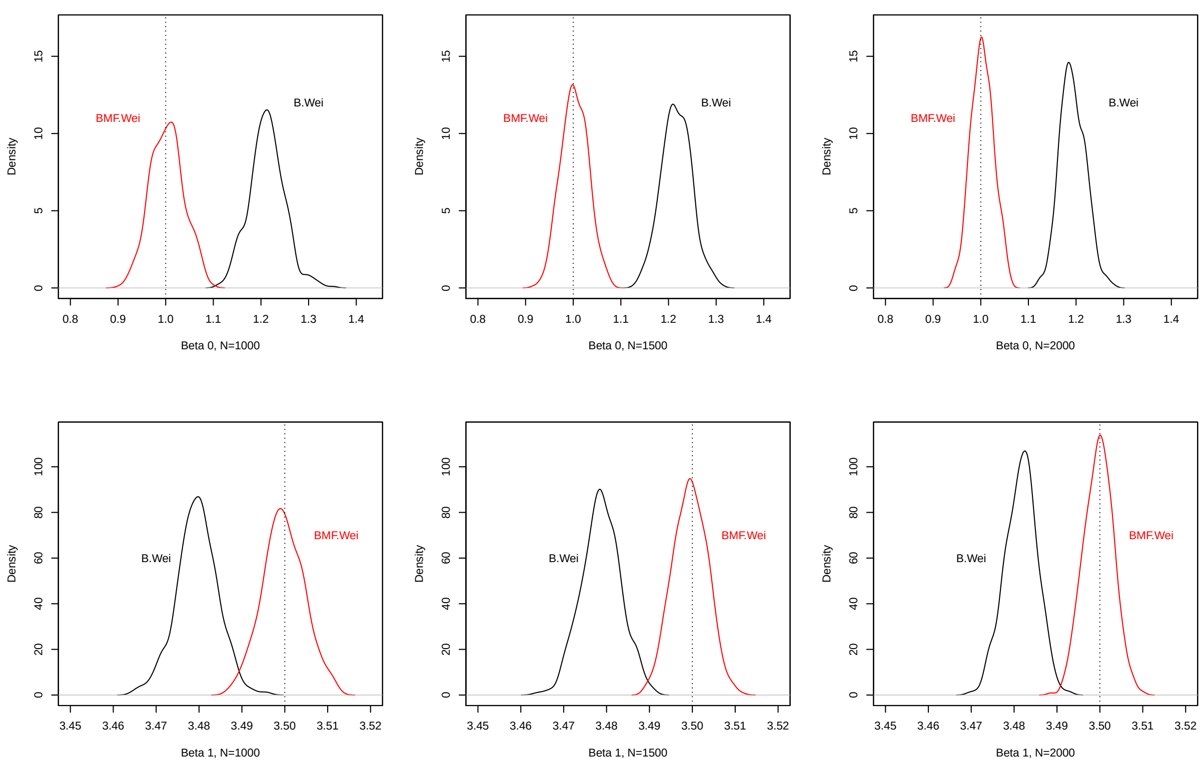


Figure A.20: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 13

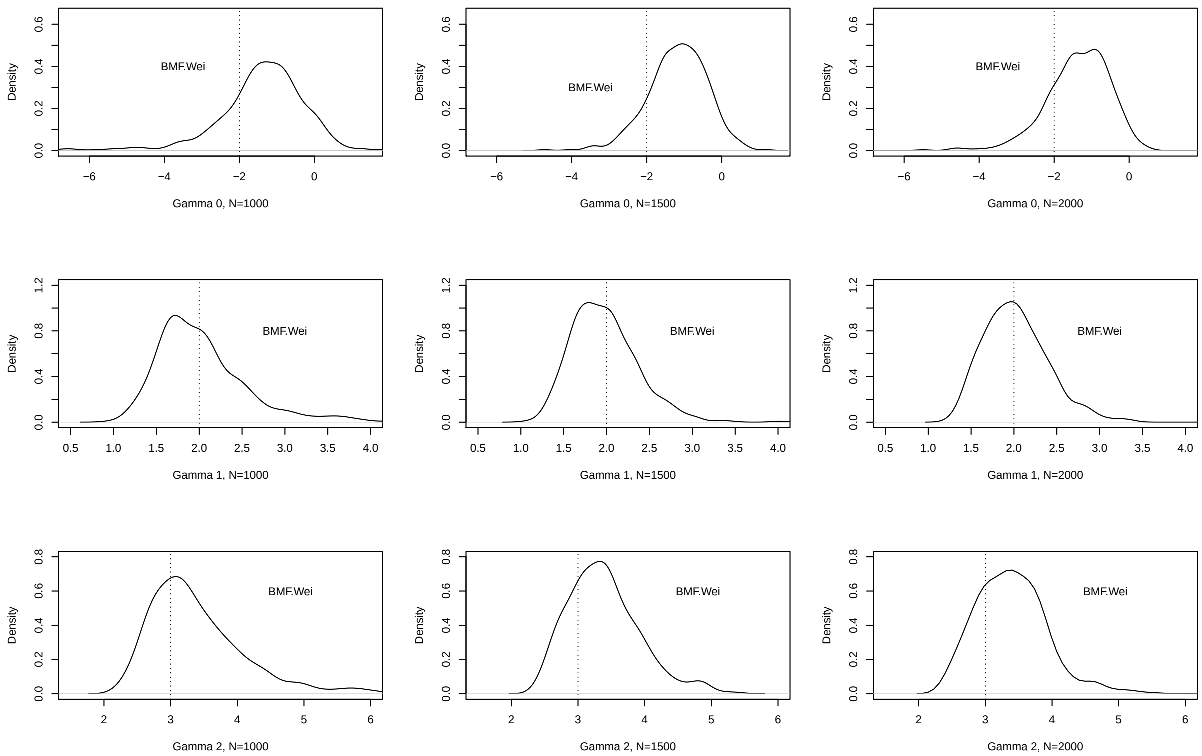


Figure A.21: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 13

comparable fashion to that described for Experiments 1-2 in the main paper.[4] We then compare the performance of the Bayesian Weibull estimator, the Bayesian MF Weibull estimator, and a Bayesian Cox proportional hazards (PH) estimator in each experiment. The results from Experiments 14-15 appear in Tables A.11-A.12 and Figures A.22-A.24.

With regards to Experiment 14—which evaluates the performance of the aforementioned models when the true d.g.p. is log-logistic with no misclassified failures—one can observe in Table A.11 and Figure A.22 that the Bayesian (MF) Weibull models exhibit RMSEs that are roughly double in size to the comparable RMSE values estimated in Experiment 1. Similarly, the Bayesian Weibull and Bayesian MF Weibull 95% CPs reported for Experiment 14 in Table A.11 have declined relative to the comparable values reported in Experiment 1; from a range of 92%-95% in the case of Experiment 1 to a range of 69%-84% in the case of Experiment 14. By comparison, the Bayesian Cox PH model generally exhibits lower bias (as measured

[4]E.g., for Experiment 15, we continue to employ a MF rate of approximately 5%.

via RMSE) than either the Bayesian Weibull model or the Bayesian MF Weibull model for this experiment—albeit with consistently *lower* empirical CPs than either of the Weibull-based models considered here. Altogether, these Experiment 14 results hence suggest that applications of the Bayesian (MF) Weibull models to non-Weibull distributed survival outcomes that exhibit no levels of misclassified failures will generally yield higher bias than either (i) a Bayesian Cox PH model applied to this same non-Weibull distributed outcome or (ii) comparable Bayesian (MF) Weibull models applied to a similar, but Weibull-distributed, outcome variable.

Table A.11: Markov Chain Monte Carlo (MCMC) β-Estimates for Experiments 14 and 15

	Experiment 14: Non-MF Log-logistic D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	*MCSE*($\hat{\beta}_0$)	*RMSE*($\hat{\beta}_0$)	*CP*($\hat{\beta}_0$)	$\hat{\beta}_1$	*MCSE*($\hat{\beta}_1$)	*RMSE*($\hat{\beta}_1$)	*CP*($\hat{\beta}_1$)
	Bayes Weibull	1.040	0.007	0.077	0.758	3.500	0.001	0.011	0.802
1000	Bayes Cox PH	–	–	–	–	3.506	0.000	2.506	0.754
	Bayes MF Weibull	1.081	0.008	0.092	0.638	3.495	0.001	0.010	0.816
	Bayes Weibull	1.040	0.006	0.070	0.726	3.500	0.001	0.010	0.746
1500	Bayes Cox PH	–	–	–	–	3.506	0.000	2.506	0.740
	Bayes MF Weibull	1.077	0.007	0.086	0.148	3.495	0.001	0.009	0.540
	Bayes Weibull	1.039	0.005	0.064	0.692	3.500	0.001	0.009	0.754
2000	Bayes Cox PH	–	–	–	–	3.509	0.000	2.509	0.704
	Bayes MF Weibull	1.083	0.006	0.086	0.480	3.495	0.001	0.008	0.768
	Experiment 15: MF Log-logistic D.G.P.								
#Obs.	Model	$\hat{\beta}_0$	*MCSE*($\hat{\beta}_0$)	*RMSE*($\hat{\beta}_0$)	*CP*($\hat{\beta}_0$)	$\hat{\beta}_1$	*MCSE*($\hat{\beta}_1$)	*RMSE*($\hat{\beta}_1$)	*CP*($\hat{\beta}_1$)
	Bayes Weibull	1.294	0.007	0.294	0.020	3.472	0.001	0.028	0.224
1000	Bayes Cox PH	–	–	–	–	3.478	0.000	2.478	0.218
	Bayes MF Weibull	1.092	0.009	0.104	0.864	3.493	0.001	0.012	0.508
	Bayes Weibull	1.309	0.006	0.309	0.000	3.470	0.001	0.030	0.084
1500	Bayes Cox PH	–	–	–	–	3.483	0.001	2.483	0.120
	Bayes MF Weibull	1.087	0.007	0.092	0.556	3.494	0.001	0.009	0.792
	Bayes Weibull	1.280	0.005	0.280	0.002	3.473	0.001	0.027	0.068
2000	Bayes Cox PH	–	–	–	–	3.488	0.000	2.488	0.142
	Bayes MF Weibull	1.088	0.006	0.090	0.494	3.494	0.001	0.009	0.738

Note: True parameter values are $\beta_0 = 1$ and $\beta_1 = 3.5$.

Table A.12: Maximum Likelihood γ-Estimates for Experiment 15

	Experiment 15: MF Log-Logistic D.G.P.												
#Obs.	Model	$\hat{\gamma}_0$	*SE*($\hat{\gamma}_0$)	*RMSE*($\hat{\gamma}_0$)	$\hat{\gamma}_1$	*SE*($\hat{\gamma}_1$)	*RMSE*($\hat{\gamma}_1$)	$\hat{\gamma}_2$	*SE*($\hat{\gamma}_2$)	*RMSE*($\hat{\gamma}_2$)			
1000	Bayes MF Weibull	-1.657	0.293	0.906	0.480	1.581	0.164	0.669	0.618	2.348	0.166	0.983	0.768
1500	Bayes MF Weibull	-1.367	0.209	0.813	0.794	1.429	0.117	0.646	0.414	2.190	0.128	0.950	0.416
2000	Bayes MF Weibull	-1.543	0.282	0.846	0.746	1.480	0.189	0.709	0.384	2.171	0.212	1.060	0.358

Note: True parameter values are $\gamma_0 = -2$, $\gamma_1 = 2$, and $\gamma_2 = 3$.

With regards to Experiment 15—which evaluates the performance of the aforementioned models when the true d.g.p. is MF log-logistic with roughly 5% misclassified failures—we find

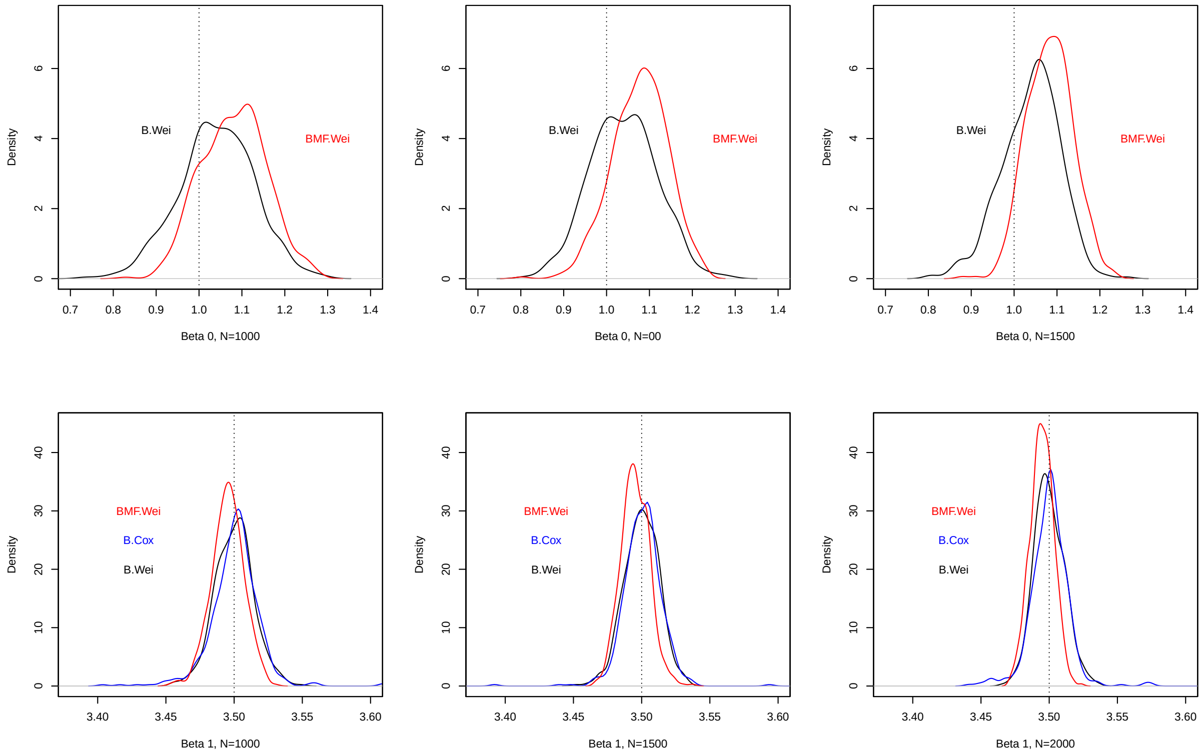


Figure A.22: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 14

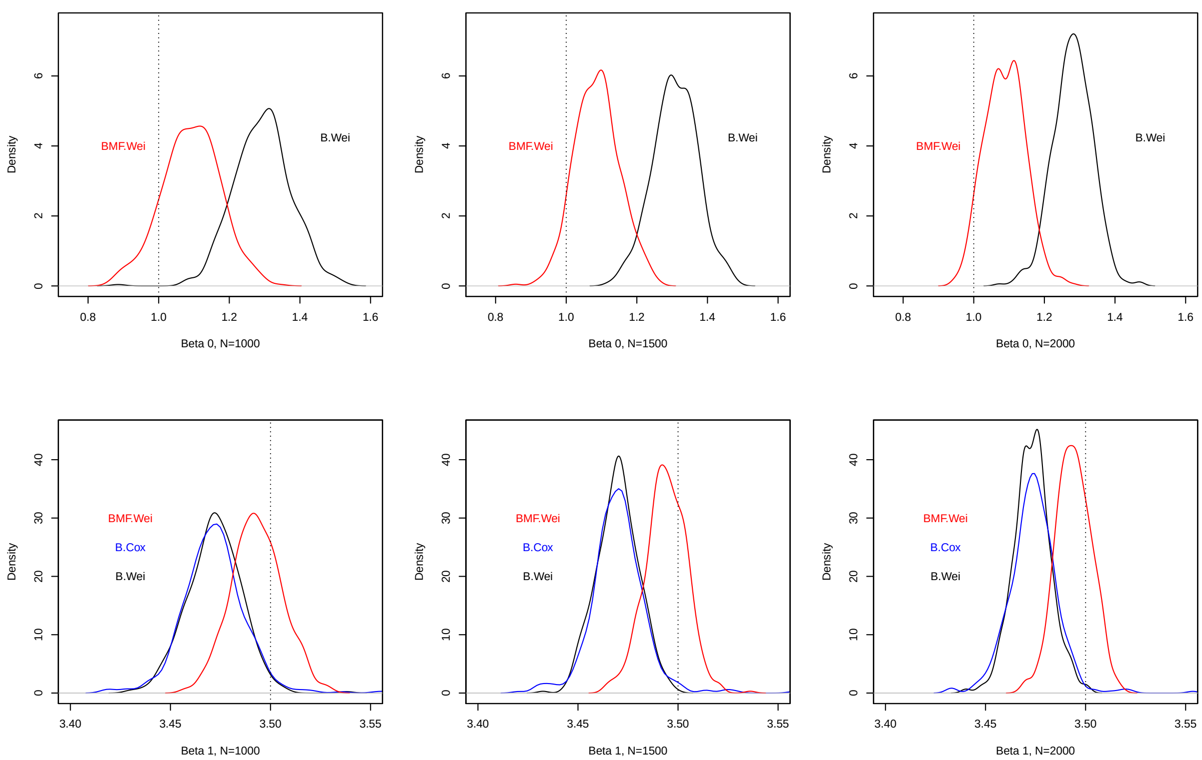


Figure A.23: Distributions of $\hat{\beta}$'s Across 500 Simulations for Experiment 15

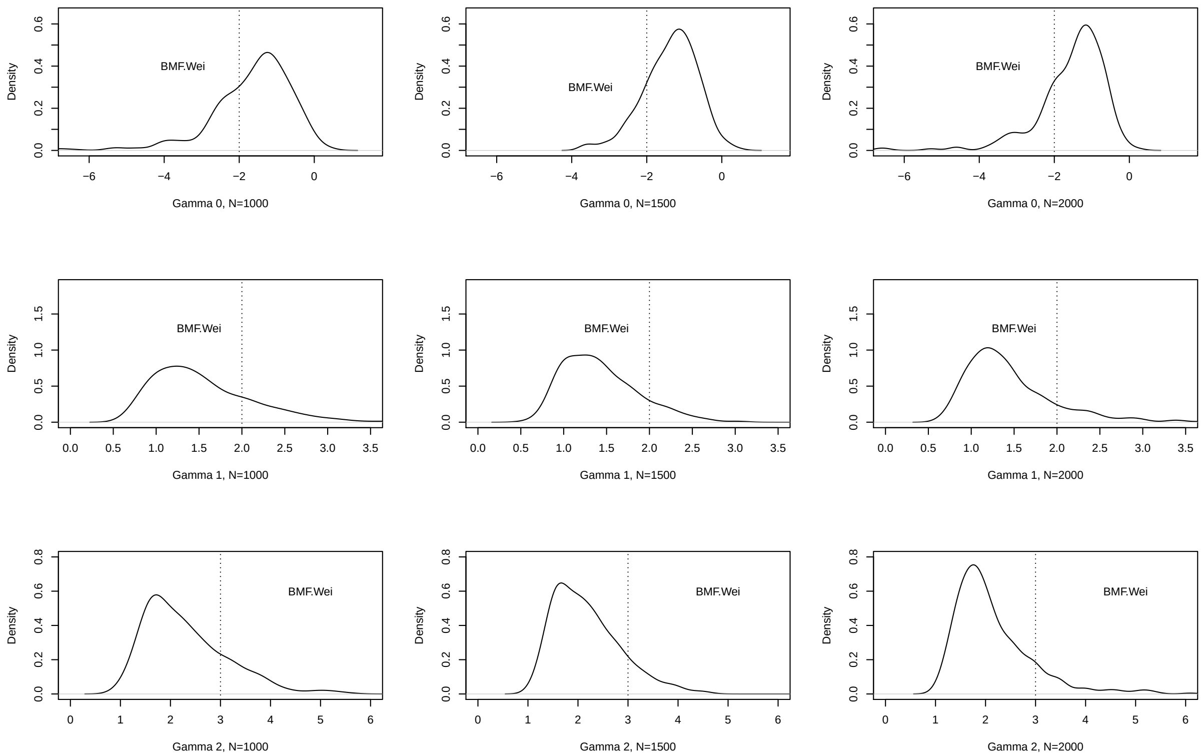


Figure A.24: Distributions of $\hat{\gamma}$'s Across 500 Simulations for Experiment 15

in Tables A.11-A.12 and Figures A.23-A.24 that the Bayesian (MF) Weibull models continue to exhibit higher RMSEs and worse CPs, relative to the equivalent Bayesian (MF) Weibull models discussed under Experiment 2 (i.e., the MF Weibull d.g.p. with roughly 5% misclassified failures). At the same time, the Bayesian MF Weibull model outperforms the Bayesian Weibull model for both parameters of interest, wherein the former model exhibits RMSEs that are roughly three times smaller than the latter model; and where the Bayesian MF Weibull model's 95% credible intervals continue to uniquely recover our true parameter values at commensurate rates. Most notably, however, are the comparisons between the Bayesian MF Weibull model and the Bayesian Cox PH model in the lower half of Table A.11. Here, we find that the Bayesian MF Weibull model's estimates of β_1 consistently exhibit lower levels of bias than those of the Bayesian Cox PH model, as measured via RMSEs. Likewise, the Bayesian MF Weibull model's 95% CPs for β_1 in Experiment 15 range from 61%-79%—far higher than those of the Bayesian Cox PH model for this same parameter (12%-22%). These comparisons suggest that in instances where one encounters a non-Weibull distributed outcome variable that exhibits a low-to-modest level of misclassified failures, the Bayesian MF Weibull model will often remain a superior choice over non-MF semi-parametric alternatives.

IV Civil War Application: Buhaug *et al* (2009)

This section is divided into two parts. The first part presents an additional table discussed in the text that focuses on our analysis of the Buhaug *et al* (2009) data. This part also presents the additional figures that are derived from the misclassification and survival stage of the Bayesian (MF) Weibull models that are estimated on the Buhaug *et al* (2009) data (these figures are listed and discussed in the paper). The second part first presents results from convergence diagnostic checks from the main Bayesian MF Weibull specification estimated on the Buhaug *et al* (2009) data. This part then presents and discusses the substantive results (illustrated as figures) and convergence diagnostics obtained from several additional Bayesian MF Weibull models applied to the Buhaug *et al* (2009) data as robustness tests.

Tables and Figures: Buhaug *et al* (2009) Application

To begin with, recall that we mentioned in the text that there are numerous misclassified civil conflict "failure" cases in the Buhaug *et al* (2009) data. Examples of these misclassified failure cases in which the civil conflict is coded as "terminated" in the Buhaug *et al* (2009) data but which persisted beyond their terminated date are listed below in Table A.13. Next, we mentioned in the paper that the dot-whisker plots of the misclassification stage covariates from the third and fourth Bayesian MF specification's misclassification stage are illustrated in the Supplemental Appendix. We therefore present these plots below (Figures A.25a–A.25b).

Further, we noted in footnote 8 in the text that the dot-whisker plots of the constant (i.e., intercept) and the *Border* $\times$ *Distance (ln)* interaction term from all the Bayesian MF Weibull survival stage specifications will be presented in the Supplemental Appendix. These plots are illustrated below.

We next present below additional figures that were discussed in the main paper for the

Table A.13: Civil Conflict Examples that Persisted Beyond "Failed" Date in Sample

Country	Civil War Case and Rebel Group(s)	Country	Civil War Case and Rebel Group(s)
Chad	Mouvement pour Démocratie et Développement-Forces Armées Occidentales (MDD-FAO)	Liberia	National Patriotic Front of Liberia (NPFL)
Philippines	Abu Sayyaf	Niger	Union des Fronts de la Résistance Armée (UFRA)
Philippines	Moro Islamic Liberation Front (MILF)	Peru	Sendero Luminoso
India	Tripura National Volunteers (TNV)	Ethiopia	Ogaden National Liberation Front (ONLF)
India	United Liberation Front of Assam (ULFA)	Egypt	El Gama'a El Islamiyya
India	Kuki National Front (KNF)	Pakistan	Muttahida Qaumi Movement (MQM)
Myanmar	Kachin Independence Organization (KIO)	Papua New Guinea	Bougainville Revolutionary Army (BRA)
Myanmar	Communist National Party (BCP)	DR Congo	Ninjas
Mozambique	RENAMO	Mali	Arab Islamic Front of Azawad (FIAA)
Sri Lanka	Liberation Tigers of Tamil Eelam (LTTE)	Indonesia	Freitlin: Revolutionary Front for an Independent East Timor

Figure A.25: Dot-Whisker Plots for Gamma Covariates in Bayesian MF Weibull Models

(a) Specification 3 (b) Specification 4

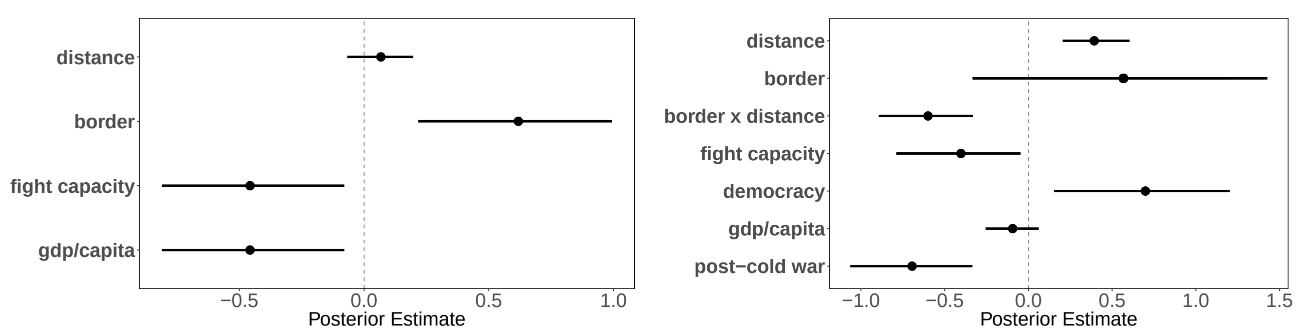
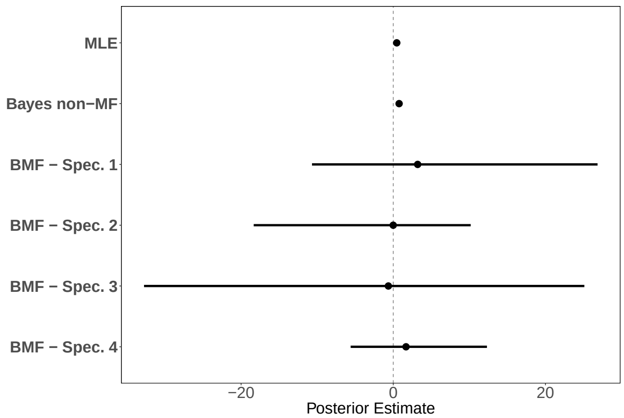
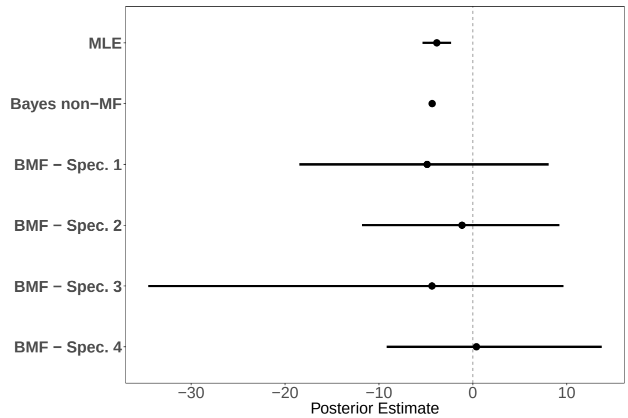


Figure A.26: Dot-Whisker Plots for Beta Covariates in MLE and Bayesian (MF) Weibull Models

(a) Border × Distance



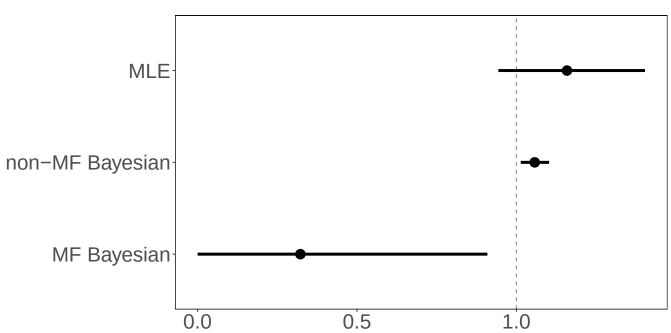
(b) Constant (intercept)



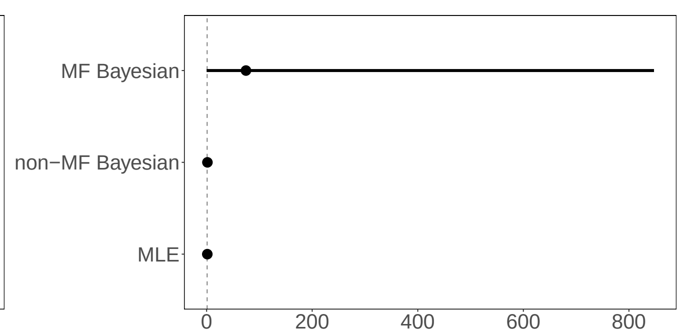
Buhaug *et al* (2009) application but not presented due to space constraints. Figures A.27a–A.27b below illustrate the hazard ratio plots derived from the estimate of the following key survival stage covariates in the MLE Weibull model, Bayesian non-MF Weibull model and the third Bayesian MF Weibull specification: *GDP capita at onset* (*ln*) and *post Cold War*. Recall that these figures illustrate the marginal effect of each of these two covariates on the hazard of civil conflict termination across the MLE and Bayesian (MF) Weibull models. We do not discuss the details of the hazard ratio results illustrated in these figures as they were discussed in the text.

Figure A.27: Hazard Ratios for GDP/capita at onset and Post-Cold War

(a) Hazard Ratio for GDP/capita at onset



(b) Hazard Ratio for Post-Cold War



Convergence Diagnostics & Robustness Tests

We present our analysis of convergence diagnostic checks of the Bayesian MF Weibull model estimated on the Buhaug *et al* (2009) data. As noted in the text, we specifically conduct the following three main exercises (described in more detail below) for this analysis in order to assess convergence of the MCMC parameter estimates in both the misclassification and survival stage in the Bayesian MF Weibull specifications that are applied to the Buhaug *et al* (2009) civil war duration data: autocorrelation plots, the Geweke (1992) diagnostic test, and the Heidelberger and Welch (1983) test of stationarity. To save space, we focus here on presenting the results from each of these three exercises conducted for the main Bayesian MF Weibull specification of interest in which the survival stage covariates repeats the survival stage used by Buhaug *et al* (2009; Table 1, Column 5), while the specification's misclassification stage includes the three covariates: *GDP capita at onset* (*ln*), *distance to capital* (*ln*) and *conflict at border*.

More importantly, for the first exercise, we extract and assess the autocorrelation plots of all the misclassification and survival stage parameters in the main Bayesian MF Weibull specification that incorporates the three aforementioned covariates in the misclassification stage. For the second exercise, we employ the Geweke (1992) convergence diagnostic test to assess convergence of the misclassification and survival stage parameters in the aforementioned Bayesian MF Weibull specification. The Geweke convergence diagnostic (Geweke, 1992) test that we employ in essence compares the location of the sampled parameter on two different time intervals of the Markov chain in order to assess convergence. If the mean values of the parameter in the two time intervals are approximately close to each other, then it is safe to assume that the two different parts of the Markov chain have similar locations in the state space, and hence that the two samples come from the same distribution. Usually one compares the last half of the chain, which is assumed to have converged, against some smaller interval in the beginning of the chain.[5] Because many of the estimated survival stage parameters in the Bayesian MF Weibull specification exhibits "slow mixing", we compare the first 20% of the Markov chain and the last 50% of the chain. The Geweke (1992) convergence diagnostic method summarized

[5]The Geweke diagnostic in fact uses spectral density estimation for the analysis described above.

here computes a *z-statistic* where the difference in the two sample means is divided by the asymptotic standard error of their difference.

For the third exercise, we use the Heidelberger and Welch (1983) test of stationarity which is a convergence diagnostic that determines whether or not the last part of a Markov chain of each parameter has stabilized (stated more technically, this test of stationarity determines whether the trace of simulated values arises from a stationary stochastic process).[6] This test uses the Cramer-von-Mises statistic to assess evidence of non-stationarity for each parameter in the model. We turn to present the autocorrelation plot (illustrated for lags 1-50) of each parameter in first the misclassification stage (see Figure A.28 below) in the main Bayesian MF Weibull specification that includes the three theoretically-identified covariates (*GDP capita at onset* (*ln*), *distance to capital* (*ln*), *conflict at border*). The autocorrelation plots illustrated in the top row of this figure include—from left to right—the following covariates: *constant*, *distance to capital* (*ln*), and *conflict at border*; the autocorrelation plot in the bottom row of this figure is *GDP capita at onset* (*ln*). We next present the autocorrelation plots (again illustrated for lags 1-50) of each parameter in the Bayesian MF Weibulls specification's survival stage (see Figure A.29 below). The autocorrelation plots illustrated in the top row of figure A.29 includes—from left to right—the following survival stage covariates: *constant*, *distance to capital* (*ln*), and *conflict at border.* The autocorrelation plots illustrated in the middle row of figure A.29 includes, again from left to right, the following survival stage covariates: *distance* x *border*, *fight capacity*, and *democracy at onset.* The autocorrelation plots in the bottom row of figure A.29 includes *GDP capita at onset* (*ln*) and *post cold war* respectively. Autocorrelation plots of each parameter in the Bayesian MF Weibull's misclassification and survival stage in the figures listed above generally reveals convergence and indicates that there is no high degree of autocorrelation for the respective posterior samples.

Next, consider the results from the Geweke (1992) convergence diagnostic tests for each misclassification and survival stage parameter in the same Bayesian MF Weibull specification

[6] We use the Heidelberger-Welch (1983) test of stationarity since it requires only one realization of the MCMC to use.

Figure A.28: Autocorrelation Plots for γ Covariates

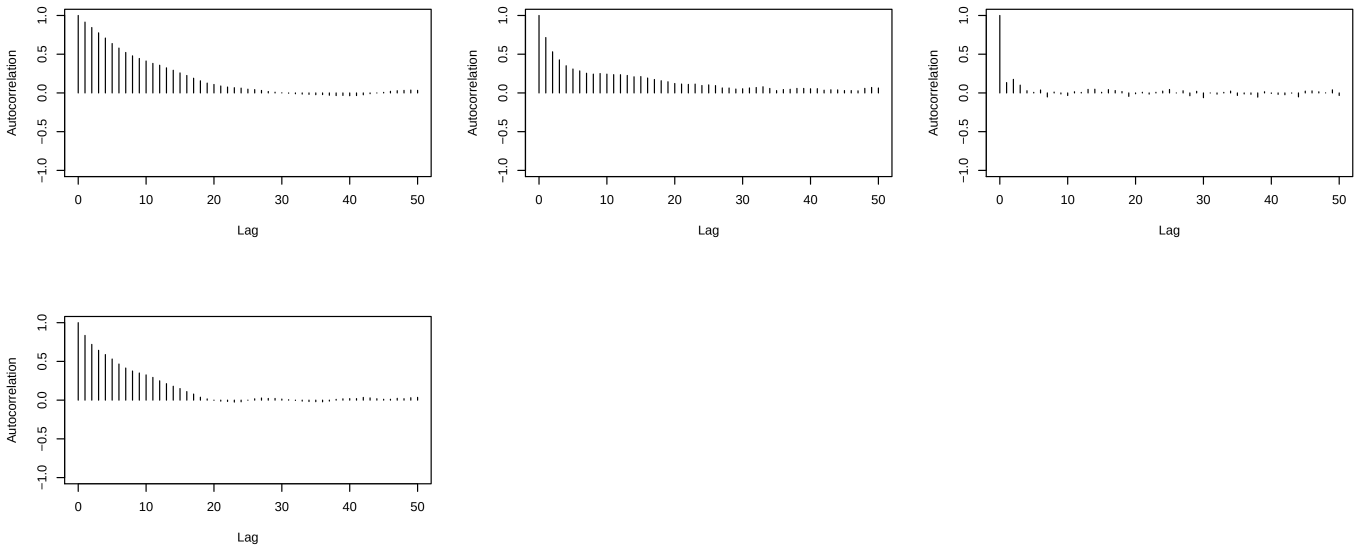
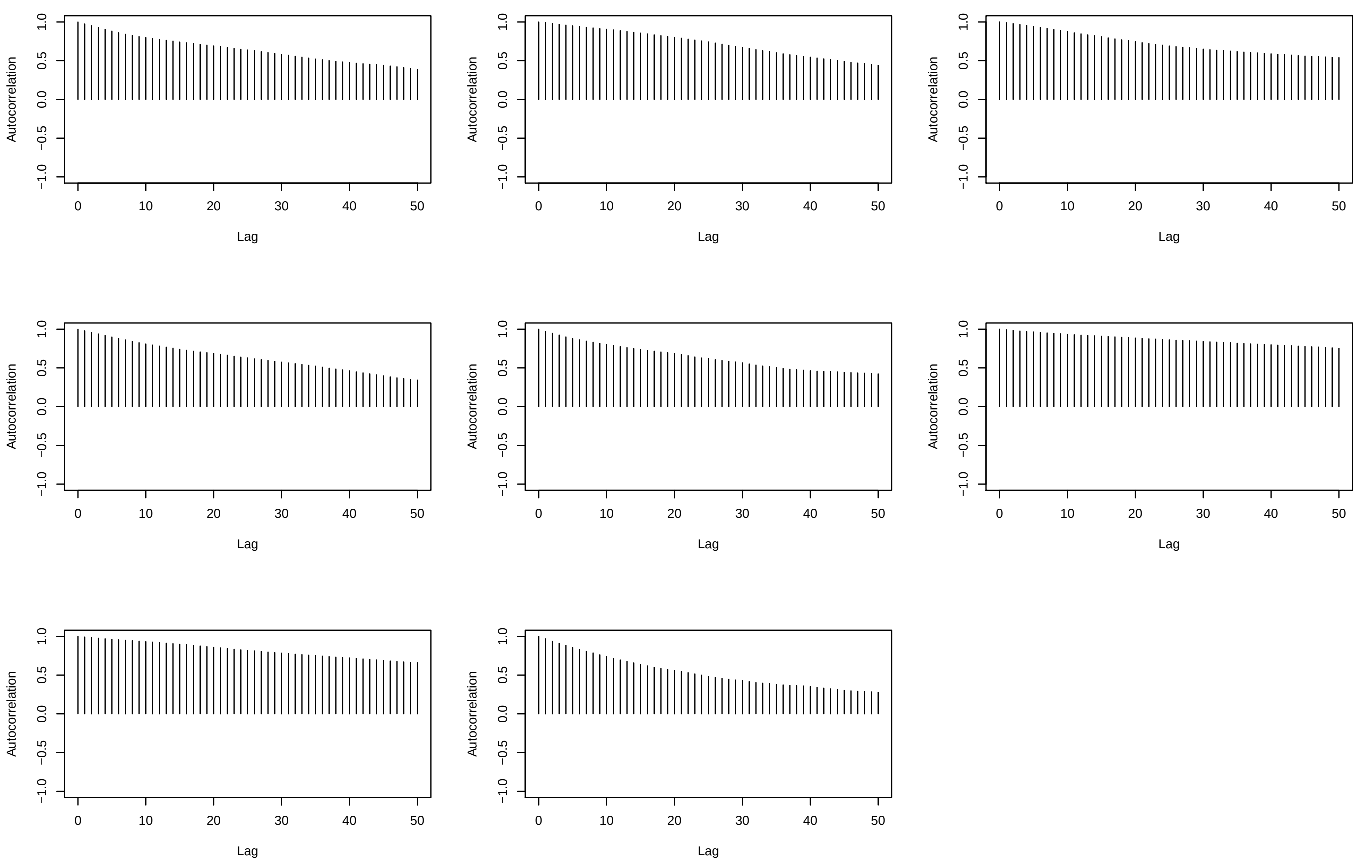


Figure A.29: Autocorrelation Plots for β Covariates



mentioned above. We first report below z-scores from the Geweke convergence diagnostics for this specification's (i) misclassification stage covariates in Table A.14 and (ii) survival stage covariates in Table A.15. We can clearly see from the Geweke diagnostics in Table A.14 that *none* of the misclassification stage parameters in the Bayeisan MF Weibull model produced significant z-scores, which indicates that there is no evidence against convergence for each of these parameters. The Geweke diagnostics in Table A.15 also shows that barring *distance* x *border*, none of the other survival stage parameters (including the main parameters of interest such as *GDP capita at onset* (*ln*) and *post cold war*) produced significant z-scores; this also indicates that there is no evidence against convergence for these survival stage parameters. Additionally, as indicated below in the last three columns of Table A.14, the Heidelberger and Welch (1983) test of stationarity *fails* to reject the null hypothesis that the Markov chain of each misclassification stage parameter in the main Bayesian MF Weibull specification of interest is from a stationary distribution. Further, the last three columns of Table A.15 shows that the Heidelberger and Welch (1983) stationarity test also *fails* to reject the null hypothesis that the Markov chain of each survival stage parameter in the Bayesian MF Weibull specification is from a stationary distribution. Hence, the Heidelberger and Welch (1983) test shows that all of the misclassification and survival stage parameters have passed the test of stationarity in Bayesian MF Weibull specification that includes the three theoretically-identified misclassification covariates. Altogether, however, the Geweke (1992) and Heidelberger and Welch (1983) tests indicates that all the parameters in the key Bayesian MF Weibull specification (estimated on the Buhaug *et al* (2009) data) has converged.

Table A.14: Convergence Diagnostics for γ Covariates

	Geweke	Heidelberger & Welch		
	Z-score	stationary test	start iteration	p-value
distance	-0.539	1	181	0.147
border	-0.432	1	1	0.831
gdp/capita	0.134	1	1	0.929

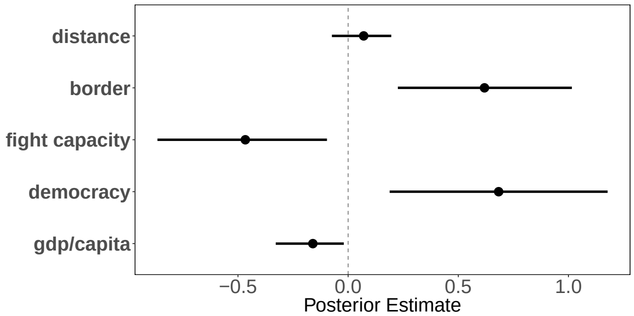
Table A.15: Convergence Diagnostics for β Covariates

	Geweke	Heidelberger & Welch		
	Z-score	stationary test	start iteration	p-value
border	0.578	1	1	0.974
distance	1.195	1	1	0.13
border×distance	-4.704	1	181	0.218
fight capacity	1.474	1	1	0.143
democracy	-0.233	1	1	0.913
gdp/capita	-0.092	1	1	0.444
post-cold war	-0.701	1	1	0.577

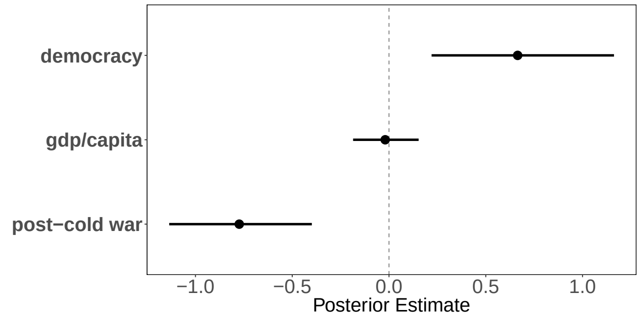
To complete the presentation of all the results from the Buhaug *et al* (2009) empirical application, we next illustrate the results from robustness test specifications for the Bayesian MF Weibull models estimated on the Buhaug *et al* (2009) data. For these robustness tests, the survival stage specification covariates in each additional Bayesian MF Weibull model applied to the Buhaug *et al* (2009) data are exactly the same as those reported in the Bayesian MF Weibull's survival stage in Figures 3a-3f presented in the paper and Figures A.26a–A.26b in the Supplemental Appendix. However, we vary the number of covariates in the misclassification stage of each "robustness test" Bayesian MF Weibull specification presented here. For example, we include *democracy score at onset* and *rebel fighting capacity* in addition to the three theoretically-identified misclassification stage covariates in the second Bayesian MF Weibull specification presented in the main paper (that includes all the survival stage covariates from the Buhaug *et al* (2009) model we focus on). The misclassification stage and survival stage estimates from this Bayesian MF Weibull robustness check specification are illustrated (as dot-whisker plots) respectively in Figures A.30a and A.32 below. Next, we include only *GDP capita (ln) at onset*, *democracy score at onset*, *post-Cold War years* to the misclassification stage of (again) the full Bayesian MF specification whose misclassification and survival stage estimates are illustrated in Figures A.30b and A.32 respectively. Finally, for the third specification robustness test, we include *GDP capita (ln) at onset*, *democracy score at onset*, *post-Cold War years*, *rebel fighting capacity, distance to capital* (*ln*), and *conflict at border.*

Figure A.30: Dot-Whisker Plots for Robustness Test γ Covariates in Bayesian MF Weibull Models

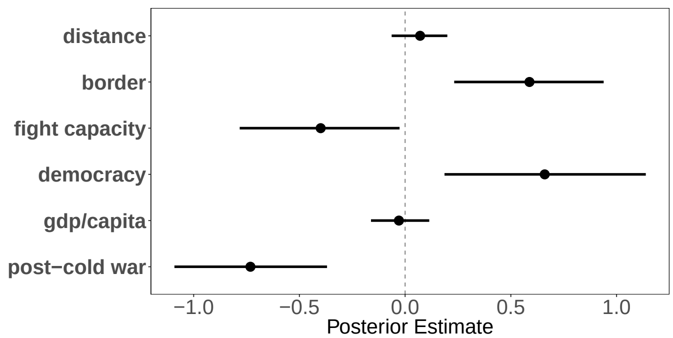
(a) Robustness Specification 1



(b) Robustness Specification 2



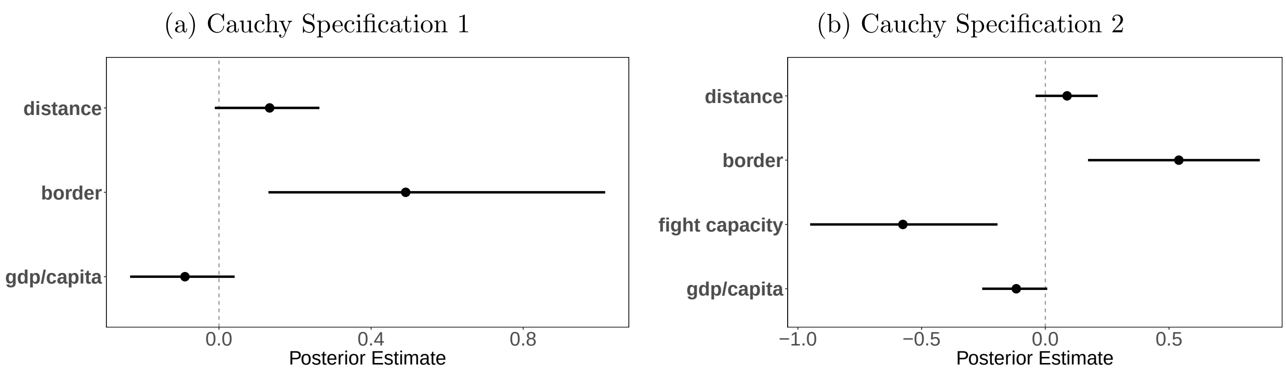
(c) Robustness Specification 3



The results for the misclassification stage covariates illustrated above in these additional specifications are almost exactly the same as those reported for these covariates in the paper. Moreover, the survival stage results for the covariates in each of the additional Bayesian MF Weibull specifications estimated for robustness tests (see Figures A.32a–A.32g) are largely the same as the results for these Bayesian MF Weibull covariates presented in the paper. For instance, the posterior mean estimate of *GDP capita at onset* (*ln*) is consistently negative and the 95% BCI of this covariate frequently excludes zero in the survival stage of all the Bayesian MF Weibull models listed above for the robustness tests. The mean estimate of *Post-Cold War years* years in the survival stage is almost always negative, although the 95%BCI of this estimate always includes zero. All other survival stage covariate results in these additional specifications are, as mentioned earlier, largely similar to the results obtained for these covariates as discussed in the text. Finally, analysis of the results from the autocorrelation plots, the Geweke (1992) convergence diagnostic test, and the Heidelberg and Welch (1983) stationarity test in the Supplemental Appendix show that in the Bayesian MF Weibull model "robustness test" specifications presented here, (i) all the misclassification stage parameter estimates has converged properly and (ii) almost all the parameter estimates in the survival stage has also converged properly (not reported to save space, but available on request).

We next present results from an additional exercise in which we estimate the Bayes MF Weibull model on the Buhaug *et al* (20009) data using the multivariate Cauchy (instead of the multivariate normal) prior. As shown below in Figures A.31 and A.33, the Bayes MF Weibull model's misclassification and survival stage results that use the multivariate Cauchy prior are similar to those obtained from the Bayes MF Weibull model that uses the multivariate normal prior.

Figure A.31: Dot-Whisker Plots for Cauchy-Prior γ Covariates in Bayesian MF Weibull Models



V Democratic Survival Application: RBS (2007)

As noted in the text, our second empirical application focuses on Reenock, Bernhard and Sobek's (2007) (hereafter RBS) study of democratic regime survival for the years 1961-1995. We present in full this empirical application of our Bayesian MF Weibull models to RBS's study of democratic regime survival below. To start with, the main aim of the RBS (2007) study is to explain how the deprivation of basic needs of civilians (specifically, food insecurity) impacts the survival of democratic regimes at various levels of economic development. RBS (2007) argue that the level of deprivation—conceptualized and operationalized "as the reciprocal

Figure A.32: Dot-Whisker Plots for Robustness Test β Covariates in MLE and Bayesian (MF) Weibull Models

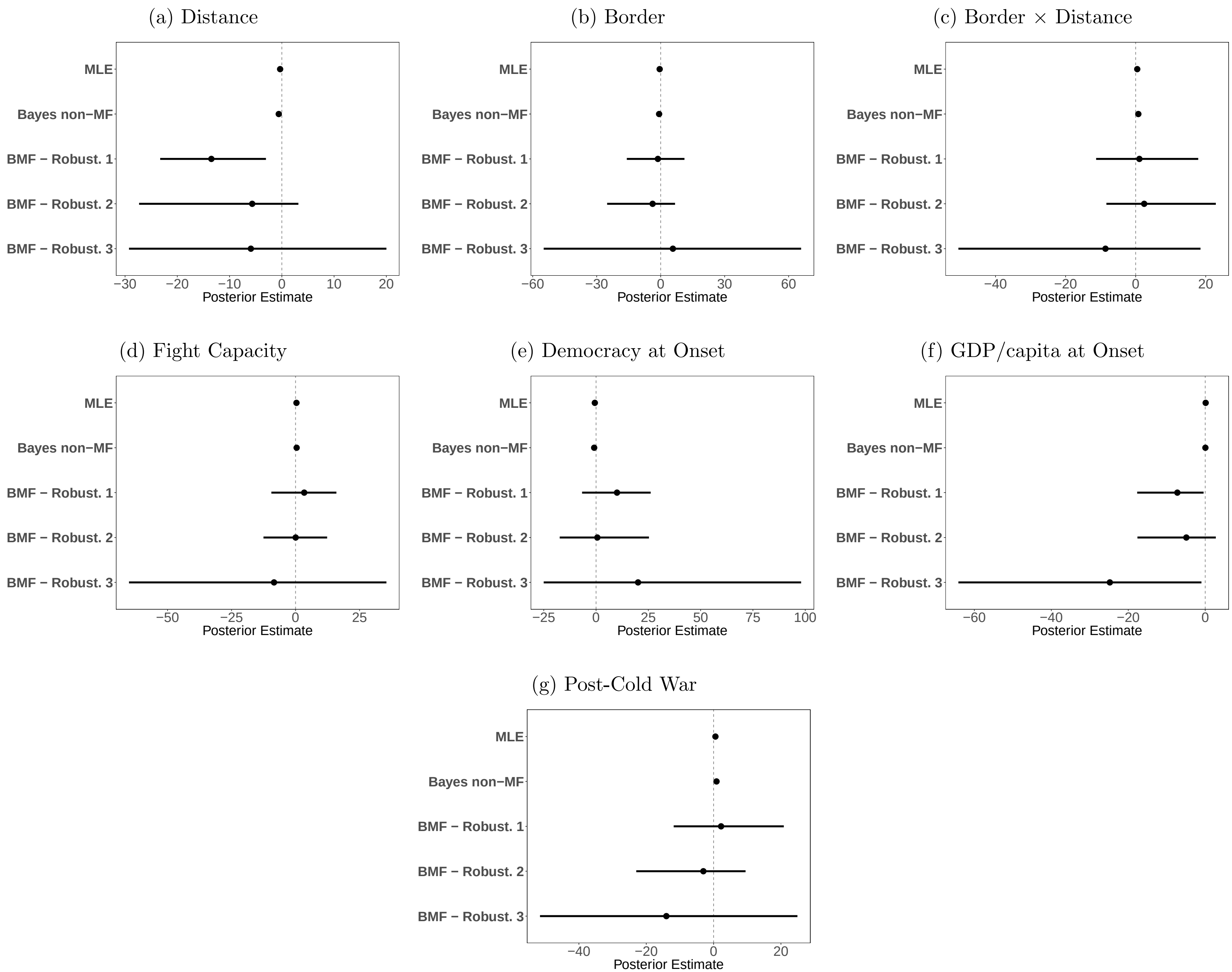
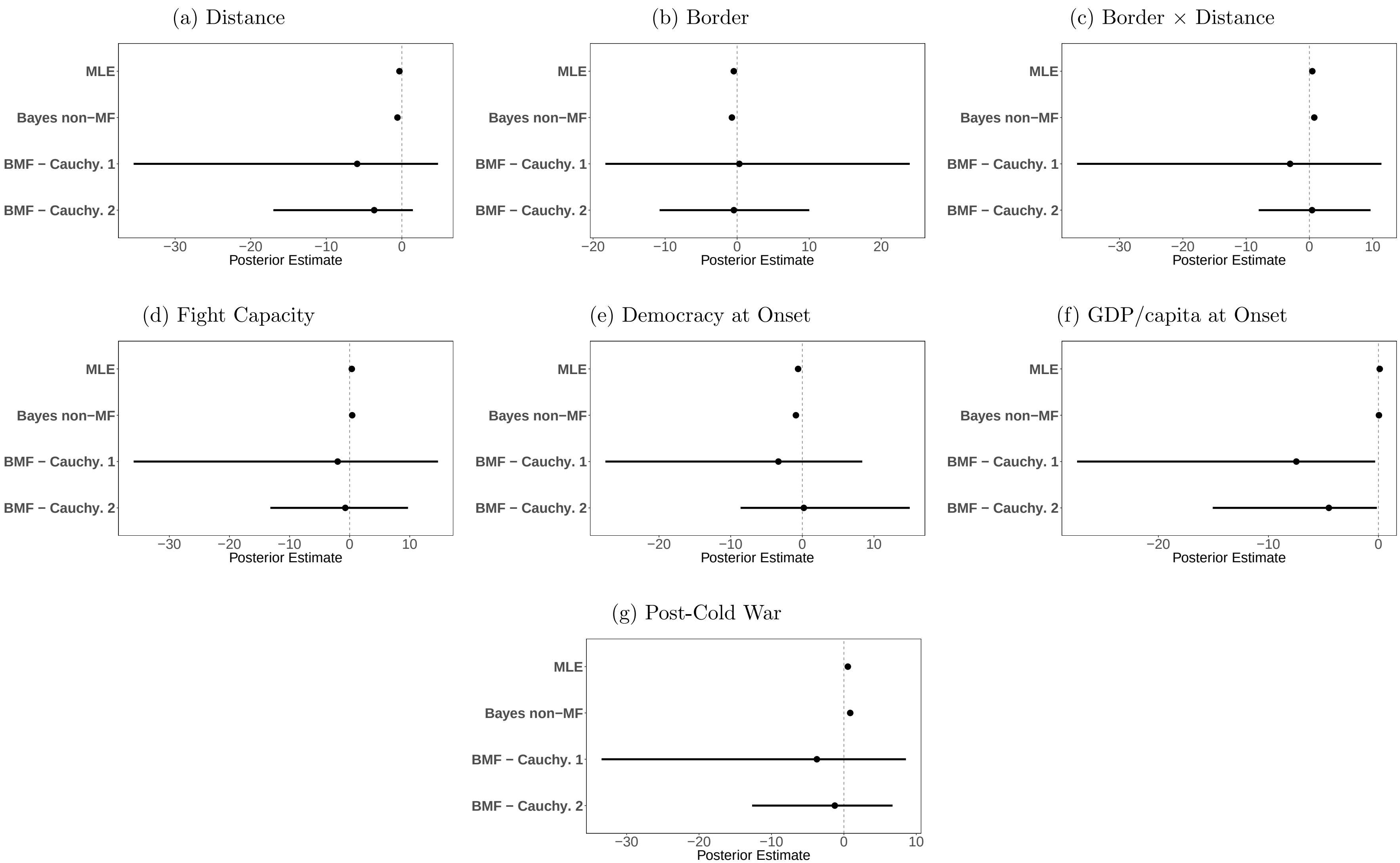


Figure A.33: Dot-Whisker Plots for Cauchy-Prior β Covariates in MLE and Bayesian (MF) Weibull Models



of the average daily per-capita caloric consumption (the inverse of basic needs satisfaction)" (RBS 2007, 687)—is a clear indicator of regressive socio-economic distribution. Higher levels of deprivation threaten democratic regime survival and increase the prospects of a political breakdown (i.e., regime failure) when the annual level of economic development per capita within a given country is moderately high. This is because in democracies with at least moderate income levels—that is, democracies with per capita income of $2,300 and beyond—citizens have sufficient socio-economic capacity to mobilize against incumbents and credibly threaten regime survival when deprivation of their basic needs increases (RBS 2007, 687).

RBS evaluate this moderated effect by interacting their main explanatory variable, *basic needs deprivation*, with *GDP per capita* (*logged*) to see how this interaction impacts their outcome variable, *democratic survival* (that is, democratic regime duration) (RBS 2007, 691) using a standard MLE Weibull model. In addition to this interaction term and its individual components, RBS also include the variables *economic growth*, the dummy variable *Presidentialism* for presidential democracies, *effective number of parties*, *religious fractionalization*, *ethnic fractionalization*, *past attempts at democracy*, and the level of *trade openness*. RBS (2007) find that *basic needs deprivation* × *GDP per capita* (*logged*) reliably increases the hazard of democratic regime failure in their MLE Weibull model, which supports their theoretical prediction. With respect to substantive effects, they find that increasing *basic needs deprivation* from 1 SD below to 1 SD above its mean in democracies reliably increases the hazard of democratic regime "breakdown" (or equivalently decreases democratic regime duration) when per capita income in their sample of democracies reaches $2,300 (RBS 2007, 692). They also find that *economic growth* and *trade openness* reliably decreases the hazard of democratic regime breakdown while higher *religious fractionalization* substantially and reliably increases the hazard of democratic "failure" (RBS 2007, 692).

Despite the substantive value of RBS's findings, the criteria used to code failure—that is, democratic regime breakdown—likely means that observed democratic failures within their data are contaminated with latent misclassified failure cases. To see why, first note that to be included in RBS' analysis of democratic regime (survival), a country must meet Dahl's (1976) threshold of "polyarchy" (that is, a political system that is both competitive and inclusive) and Linz and Stephan's (1996) criteria of "stateness". Countries that cease to meet any of these standards are thus coded as "failed" (i.e., given a score of one) and exit the dataset. However, the decision of what constitutes an important enough decline in inclusiveness, stateness, and competitiveness as to justify a democratic regime being coded as "failed" is *inherently subjective*. Such criteria are prone to misinterpretation or subjective judgments about the date of democratic regime breakdown, which thus inadvertently leads to the misclassification of observed event-failures. Indeed, owing to their subjective criteria for identifying democratic regime breakdown, we find for example that RBS code breakdown of democracy in Thailand in 1976 and Sri Lanka in 1983. Yet secondary sources such as the political regimes dataset in Polity IV and primary sources (listed in Table A.17) show that democracy (as per Dahl's polyarchy criteria used by RBS) persisted in Thailand beyond 1971 and also in Sri Lanka well beyond 1983 (in fact, into the 1990s as well). These two examples are hardly unique. In fact, numerous additional examples of recorded democratic-failure years listed in Table A.16 have been misidentified in the RBS (2007) data, suggesting that their observed democratic-failures are indeed contaminated with misclassified failure cases.

Given that contaminated misclassified failure cases in survival data generates econometric challenges, we estimate our Bayesian MF Weibull model (using the slice sampling MCMC algorithm) on RBS' survival data to statistically account for the possibility that the observed democratic regime failure in these data include misclassified failure cases. We replicate the main specification in column 2 of Table 1 in the RBS paper (see RBS 2007, 692 that tests the effect of *basic needs deprivation* × *GDP per capita* (logged) and other controls on their *democratic survival* outcome variable), by first estimating (i) the standard MLE Weibull model and then (ii) three different specifications of our Bayesian MF Weibull model on the RBS (2007) data using our slice sampling (MCMC) algorithm and the Mutivariate Normal prior (our results remain robust when we use the multivariate Cauchy Prior). To this end, we specified the Bayesian MF Weibull model's hyperparameters as follows: $a = 1$, $b = 1$, $S_\beta = I_{p1}$, $S_\gamma = I_{p2}$, $\nu_\beta = p1$ and $\nu_\gamma = p2$. The results from the Bayesian MF Weibull models are based on a set of

Table A.16: Democratic Regimes Examples in RBS (2007) that Survived Beyond "Failed" Date

Country	Event Description
Ghana	Recorded as Failed in 1972 in RBS (2007) but democracy survived (as per "competitiveness" criteria) in Ghana beyond this date according to secondary sources[1] and primary sources.[2]
Madagascar	Recorded as Failed in 1971 in RBS (2007). Yet democracy survived (as per "competitiveness" criteria) in Madagascar beyond this date according to secondary sources[1] and primary sources.[2]
Nigeria	Recorded as Failed in 1966 in RBS (2007). However, democracy survived (as per "competitiveness" criteria) in Nigeria well beyond this date (till the early 1970s) as per secondary sources[1] and primary sources.[2]
Malaysia	Recorded as Failed in 1969 in RBS (2007). But secondary sources[1] and primary sources[2] indicate that Malaysia made a transition to authoritarian rule only from the mid-1970s onwards which implies that it survived as a democracy well beyond this recorded date in RBS.
Peru	Recorded as Failed in 1992 in RBS (2007). Yet primary sources[2] have shown unambiguously that in terms of "inclusiveness" and "competitiveness," Peru continued as a democracy till the mid-1990s.
Uruguay	Recorded as Failed in 1973 in RBS (2007) but democracy survived (as per "inclusiveness" and "competitiveness" criteria) in Uruguay beyond this date according to secondary sources[1] and primary sources.[2]
Philippines	Recorded as Failed in 1972 in RBS (2007) but democracy survived (as per "inclusiveness" and "competitiveness" criteria) in Philippines beyond this date according to secondary sources[1] and primary sources.[2]
Suriname	Recorded as Failed in 1989 in RBS (2007). Yet democracy survived (as per "inclusiveness" and "competitiveness" criteria) in Madagascar beyond this date as according to secondary sources[1] and primary sources.[2]

Notes: [1]These secondary sources include the Varieties of Democracy (V-Dem) Database by Coppedge *et al* (2016), Cheibub *et al* (2010) political regimes dataset, and the Polity IV database. [2]For a list of these primary sources see Table A.17.

100,000 iterations with 10,000 burn-in scans and a thinning of 100.

To start, the standard MLE Weibull specification that we estimate for the RBS (2007) application includes all the same variables reported in column 2 of Table 1 in RBS (2007) which influence the *democratic survival* outcome variable (e.g., *basic needs deprivation* $\times$ *GDP per capita* (*logged*), the individual components of this interaction term, *economic growth*, *Presidentialism*, and so on). Next, we estimate the first (i.e. baseline) Bayesian MF Weibull specification in which the survival ($\mathbf{X}$) stage of this specification also includes all the same variables reported in column 2 (Table 1) in RBS (2007, 691). The misclassification stage of this baseline specification only includes an intercept.

The survival stage of the second Bayesian MF Weibull specification also includes all of the covariates reported in RBS (2007, 691). But the misclassification (**Z**) stage of this specification includes the following covariates that may affect the likelihood that some cases of democratic breakdown have been misclassified as terminated even when they had (possibly) not failed. We first include the dummy variable for Presidential democracies (labeled *Presidentialism*) in the misclassification stage. To understand why, note that extensive debates exist about whether Presidential regimes are "inclusive" and "competitive" and the extent to which these regimes are inclusive and competitive [7] – two key criteria in Dahl's polyarchy concept that RBS use to code when democratic regimes breakdown in their sample. Because the extent of inclusiveness and competitiveness in Presidential regimes are ambiguous, it may be difficult for researchers to accurately identify if and when breakdown of Presidential democracies occur using the criteria that RBS (2007) employ. This increases the possibility of misidentification of breakdown of Presidential democracies. Hence, we anticipate that *Presidentialism* will be positively associated with the probability of misclassified failure.

Next, we add *GDP per capita* (logged) and *economic growth* to the misclassification stage in the second Bayesian MF Weibull specification. The rationale for doing so is as follows. Specifically, studies have shown that the frequency of democratic breakdown is rare in democracies with a relatively high per capita income level of \$6,055 (1985 PPP USD) and beyond,[8] during periods of economic growth in democracies and in states characterized by high levels of trade openness (Przeworski and Limongi 1997; Boix 2003). Misclassifying or misidentifying democratic breakdowns under the aforementioned conditions is therefore less likely given the relative low frequency of democratic regime failure in the context of these conditions. Thus, we anticipate that logged *GDP per capita*, *economic growth*, and *trade openness* will each be negatively associated with the probability of misclassified failure. Moreover, unlike presidential

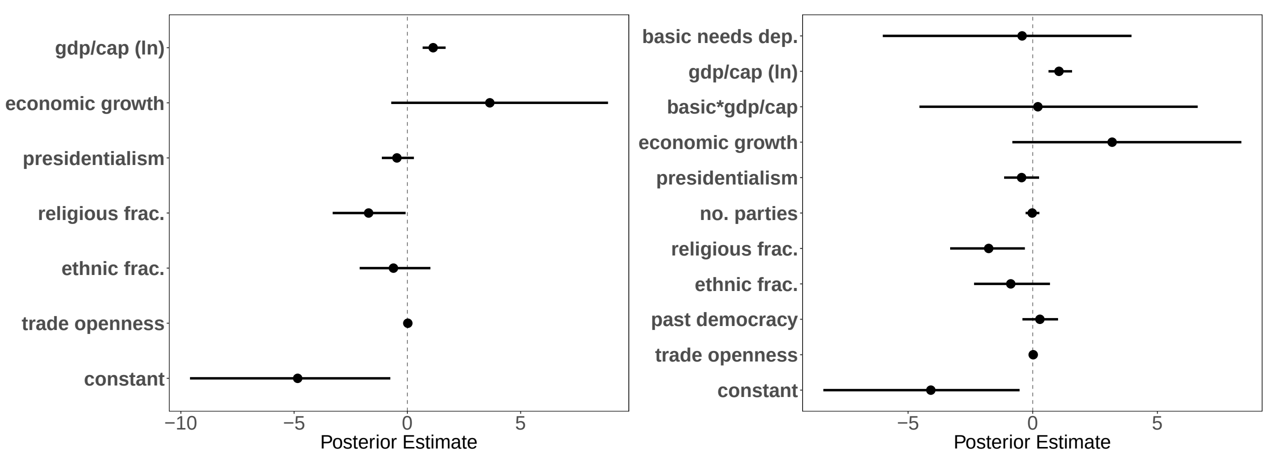
[7] For this see e.g., Linz (1994); Samuel and Shugart (2010).
[8] Przeworski and Limongi 1997, 165.

democracies, scholars have suggested that the extent of formally institutionalized inclusiveness in democracies with higher levels of ethnic or religious fractionalization is well defined and clear (Munck 2009; Teorell 2010). It is thus easier to accurately record if and when democratic regime breakdown occurs in more fractionalized societies as per the criteria used by RBS (2007). Hence the influence of *ethnic* and *religious fractionalization* on the probability of misclassified failure is likely to be negative. Finally, the third Bayesian MF Weibull specification for this empirical application includes all the covariates from the RBS (2007) specification within both the survival and misclassification stage of the Bayesian MF Weibull specification.

We turn to first discuss the misclassification stage and then the survival stage results from the Bayesian MF Weibull models that are estimated on the RBS (2007) democratic survival data. The misclassification stage results are presented via the following illustrations derived from the Bayesian MF models results: dot-whisker plots (Figure A.34) that illustrates each misclassification stage covariate's posterior mean estimate with its 95% Bayesian Credible Intervals (hereafter BCI) and the first difference in misclassification probabilities from the Bayesian MF model's misclassification stage ($\mathbf{Z}$) covariates (Figure A.35)

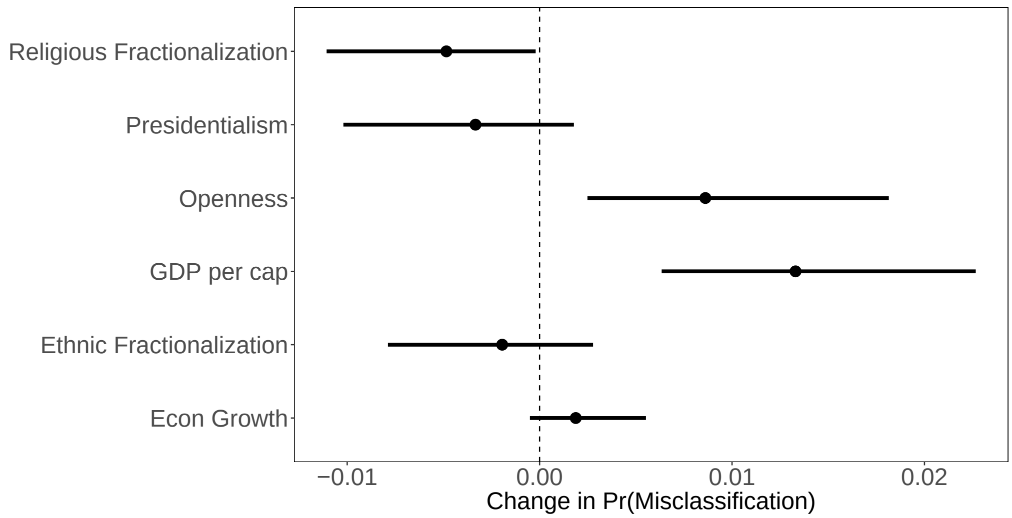
Figure A.34: Dot-Whisker Plots for γ Covariates in Bayesian MF Weibull Models

(a) Reenock *et al* (2007) Specification 2 (b) Reenock *et al* (2007) Specification 3



The dot-whisker plots from the posterior mean estimates for *ethnic* and *religious fractionalization* and the 95% BCI of these mean estimates in the misclassification stage illustrated in Figure A.34 statistically supports our claim that each of these two covariates will be negatively

Figure A.35: Change in Probability of Misclassification for Reenock *et al.* (2007)



Note: Black dot represents predicted change in the duration of democracy as the row variable changes from 1 SD below the mean to 1 SD above the mean for continuous variables and 0-1 change for dichotomous variables while holding all other variables at their mean or mode. Whiskers indicate the 95% credible interval. Results come from Reenock *et al* (2007) specification 3.

associated with the probability of misclassified failure. More specifically, the first difference in misclassification probabilities derived from the second Bayesian MF Weibull specification applied to the RBS (2007) data, which is also illustrated above in Figure A.35, shows that increasing *ethnic* and *religious fractionalization* from 1 SD below to 1 SD above their respective mean[9] decreases the probability of misclassified failure by approximately (i) 0.31% for *religious fractionalization* and (ii) 0.2% in the case of *ethnic fractionalization.* The 95% BCI of the substantive effect of *religious fractionalization* excludes zero, while the 95% BCI of *ethnic fractionalization* includes zero. The former result confirms our intuition that it is less likely that regime failure in democracies with high levels of religious fractionalization will be misclassified, and moreover, this result is reliable.

Contrary to our expectations, we find from the misclassification stage of the Bayesian MF Weibull specifications that *GDP per capita* (logged), *economic growth* and *trade openness* are each positively associated with the probability of misclassified failure. The first difference in misclassification probabilities with 95% BCI illustrated in Figure A.35 shows that the statisti-

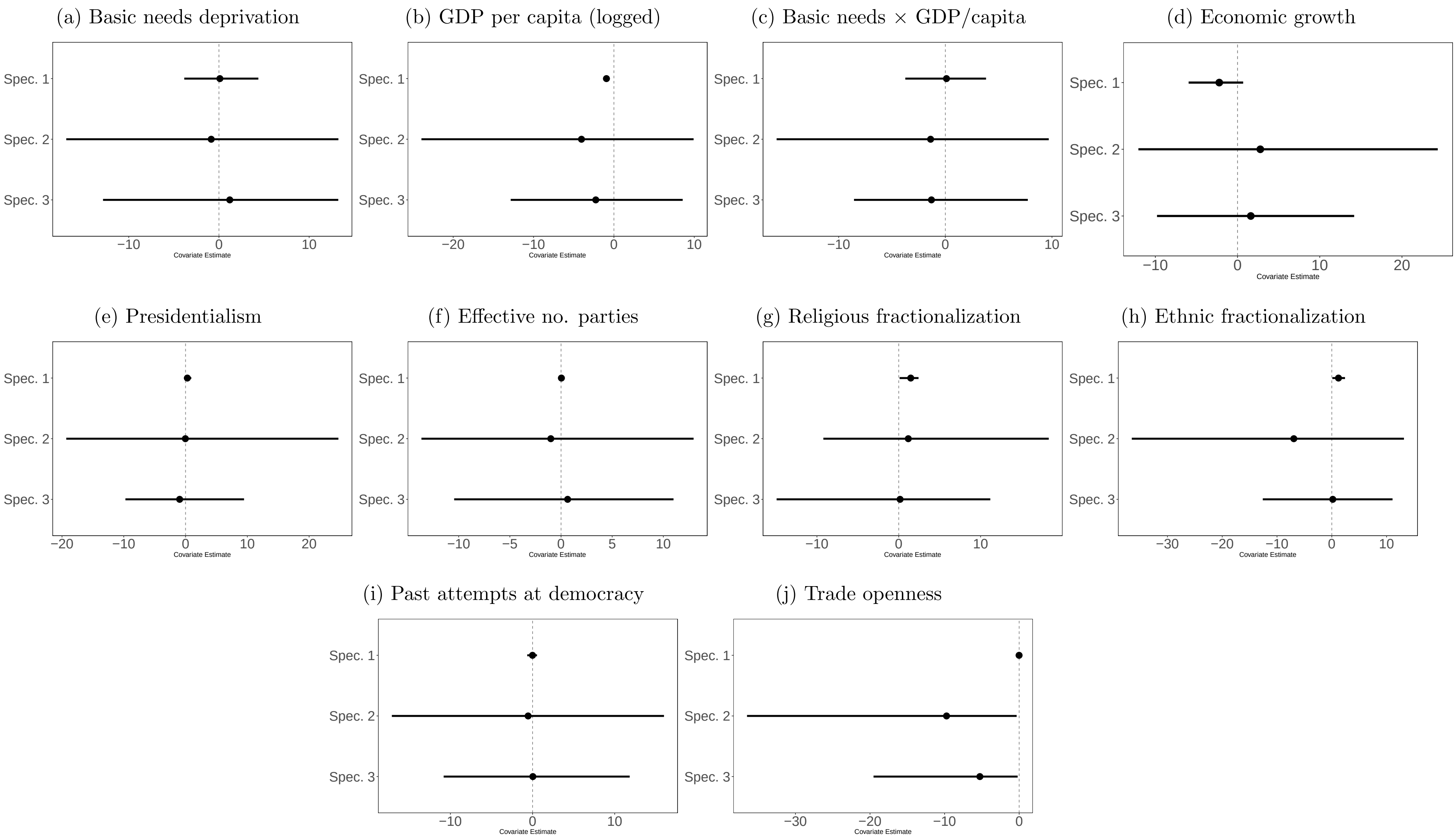
[9]While holding other covariates at their respective mean or mode in the sample.

cally positive association between two of the three aforementioned covariates (*GDP per capita* (logged) and *trade openness*) and the probability of misclassified failure is indeed reliable and substantial. The posterior mean estimate for *Presidentialism* in the misclassification stages of the main Bayesian MF Weibull specification (of interest) is -0.462 and -0.446. But the 95% BCI of the mean estimate and substantive effect of *Presidentialism* (see figure A.35) in the misclassification stage always includes zero, which indicates that the association between this dummy variable and the probability of misclassified failure is unreliable.

We next briefly report the results from the survival stage of the Bayesian MF Weibull specification(s) that are estimated on the RBS (2007) data. To this end, we find that the coefficient estimates from the standard MLE Weibull model applied to the RBS (2007) exactly mirror the results that is reported by RBS (2007) on page 689 of their published paper (we thus do not report the MLE Weibull's coefficient estimates to save space). In particular, the negative effect of *basic needs deprivation* $\times$ *GDP per capita* (logged) in the standard MLE Weibull specification is reliable. It thus supports the finding by RBS (2007) that this interaction term reliably increases the hazard of democratic regime failure or equivalently decreases the duration of democratic regimes. Next, consider the three rows in each of the dot-whisker plot in Figure A.36 that illustrate the posterior mean estimates (and their respective 95% BCI) of the survival stage covariates from the three main Bayesian MF Weibull specifications estimated on the RBS (2007) data. In sharp contrast to the results reported by RBS (2007) in their MLE Weibull model, the dot-whisker plots in Figure A.36c suggests that the *basic needs deprivation* $\times$ *GDP per capita* (logged) interaction term decreases the hazard of democratic regime failure or, in other words, increases the duration of democracies even though the 95% BCI of this mean estimate consistently includes zero. This result is exactly the opposite of what RBS (2007) find in their standard MLE Weibull model. It also indicates that once we statistically account for misclassified democratic regime failure cases, we find that there is a negative–rather than positive (as RBS 2007 find)– association between basic needs deprivation and the hazard of democratic regime failure once per capita income increases beyond its mean in the sample.

Although the key interaction term's result in the Bayesian MF Weibull model's survival

Figure A.36: Dot-Whisker Plots for β Covariates in Bayesian MF Weibull Models



stage is substantially different from those reported by RBS (2007), the rationale underlying this "contrarian" result is intuitive and is as follows. Specifically, democracies with moderately high levels of per capita income and beyond are on average more accountable to their citizens and also have greater material capacity for addressing crises such as food (a key "basic need") shortages that result from exogenous shocks (e.g., Sen 1982; Lindert 2004). Hence when basic needs deprivation occur in relatively higher income democracies, governments in these democracies are more likely to successfully resolve such deprivation crises. This will serve to reinforce the citizens' faith in the democratic political process in these countries which in turn helps to increase the prospects for survival and duration of democracies, as shown empirically by our Bayesian MF Weibull models.

Other key results from the survival stage of all the Bayesian MF Weibull specifications applied to the RBS (2007) data also vary dramatically from those reported by RBS (2007). For instance, RBS (2007) suggest and find that *economic growth* reliably decreases the hazard of democratic regime failure. The posterior mean estimate and the 95% BCI for *economic growth* across the Bayesian MF Weibull's survival stage specifications, however, reveal that the association between this variable and the hazard of democratic regime failure is weak, inconsistent and fragile (see Figure A.36d). RBS (2007) report that *religious fractionalization* reliably increases the hazard of democratic regime failure. In contrast, the posterior mean estimate and the 95% BCI for *religious fractionalization* in the Bayesian MF Weibull model's survival stage specifications is also weak and inconsistent, therein suggesting that the association between *religious fractionalization* and the hazard of democratic regime failure is unreliable. Lastly, standard convergence diagnostic checks for the parameters in each Bayesian MF Weibull specification estimated for the RBS (2007) data also suggests that the Markov chain has reached a steady state in each case. Indeed, autocorrelation plots (available on request) from the Bayesian MF Weibull's misclassification and survival stage specifications reveals not only good mixing and rapid convergence but also indicates that there is no high degree of autocorrelation for the posterior samples. Further, none of the misclassification and survival stage parameters produced significant z-scores from the Geweke diagnostic test applied to the Bayeisan MF

Weibull models estimated on the RBS (2007) data. This indicates that there is *no evidence against convergence* based on the Geweke diagnostic and that the Markov chain has successfully converged to the desired posterior.

Table A.17: Sources for Identifying Misclassified Democratic Failure Cases in RBS (2007) Data

Countries in sample by region	Region and country-specific sources
South and South-East Asia Bangladesh; Cambodia; India; Indonesia; Malaysia; Myanmar; Nepal; Pakistan; Philippines; Sri Lanka; Suriname; Thailand	Human Rights Watch. *Descent into Chaos: Thailand's 2010 Protests and the Government Crackdown.* New York and Bangkok: Human Rights Watch, 2011. Brass Paul (ed). *Routledge handbook of South Asian politics: India, Pakistan, Bangladesh, Sri Lanka, and Nepal.* Routledge, New York. Hussain . 2008. Politics of alliances in Pakistan, unpublished PhD thesis submitted to the National Institute of Pakistan Studies, Quaid-i-Azam University, Islamabad, Pakistan. Kochanek SA. 2000. "Governance, patronage politics, and democratic transition in Bangladesh." *Asian Survey* 40(3): 530550 Kershaw, Roger. 2001. *Monarchy in South-East Asia: The Faces of Tradition in Transition.* London and NewYork: Routledge. Case, William F. 2002. *Politics in Southeast Asia: Democracy or Less.* London and New York: Routledge Curzon. Connors, Michael K. 2011, "Ambivalent about Human Rights: Thai Democracy," in Thomas W. D. Davis and Brian Galligan (eds), *Human Rights in Asia*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 10322 Freedman, Amy L. 2007, "Consolidation or Withering Away of Democracy? Political Changes in Thailand and Indonesia." *Asian Affairs: An American Review* vol. 33, no. 4 (Winter), pp. 195-216. Frolic, Michael B. 2001, "Transitions to Democracy after the Cold War," in Amitav Acharya, B. Michael Frolic, and Richard Stubbs (eds), *Democracy, Human Rights and Civil Society in South East Asia*, Toronto, Canada: University of TorontoYork University Joint Centre for Asia Pacific Studies, pp. 21–35.
Latin America Colombia; Ecuador; El Salvador; Dominican Republic; Guatemala; Grenada; Honduras; Mexico; Nicaragua; Paraguay; Peru; Suriname; Uruguay; Venezuela	Mettenheim, Kurt., and James Malloy. 1998. *Deepening Democracy in Latin America.* Pittsburgh: University of Pittsburgh Press. Millet, Richard. 2009. "Democratic Consolidation in Latin America?" In *Latin American Democracy : Emerging reality or endangered species?* eds, Richard L. Millet, Jennifer S. Holmes, Orlando J. Prez. New York: Routledge Pinkney, Robert. 2003. *Democracy in the Third World.* London: Lynne Rienner Publishers. Hagopian, Frances. and Scott Mainwaring., 2005. The Third Wave of *Democratization: Advances and Setbacks.* New York: Cambridge University Press Hagopian, Mainwaring, and Daniel Brinks. 2008. "Political Regimes in Latin America, 1900-2007" `http://kellogg.nd.edu/scottmainwaring/Political_Regimes.pdf` Rector, John. 2003. The History of Chile. New York: Palgrave Macmillan Schedler, Andreas. 2001. "Measuring Democratic Consolidation." Studies in *Comparative International Development* 36 (1): 66-92. Blake, Charles. 2005. *Politics in Latin America.* Boston: Houghton Mifflin Company. Buxton Julia., and Nicola Philips. 1999. *Case Studies in Latin America Political Economy.* Manchester: Manchester University Press.

Countries in sample by region	Region and country-specific sources
Africa Benin; Botswana; Chad; Congo; Gambia; Ghana; Madagascar; Mali; Mozambique; Namibia; Niger; Nigeria; Sierra Leone; South Africa; Sudan; Tanzania; Uganda; Zambia	Bogaards, M. 2004. "Counting parties and identifying dominant party system in Africa." *European Journal of Political Research.* 43: 173-197. Bratton, M., Walle, N. van de. 1997. *Democratic experiments in Africa.* Cambridge: Cambridge University Press. Osei, A. 2012. Party-Voter Linkage in Africa, Party Research in Africa: Findings and Problems. Ellis, S. 1994. "Democracy and Human Rights in Africa" in Rob Van Berg UlbeBosma (eds) *Poverty and Development: Historical Dimension of Development, Change and Conflict in the South*, Ministry of Foreign Affairs. The Hague. Pp. 115-124. Bratton, M., Houessou, R. 2014. *Demand for democracy is rising in Africa, but most political leaders fail to deliver.* Afrobarometer Policy Paper No. 11. Bratton, M., Mattes, R., Gyimah-Boadi, E. 2005. *Public opinion, democracy, and market reform in Africa.* New York: Cambridge University Press. Gyimah-Boadi, E. 2015. "Africa's waning democratic commitment." *Journal of Democracy*, 26(1), 101–113. LeBas, A. 2014. "The perils of power sharing." *Journal of Democracy.* 25(2), 52–66 Wing, S. 2008. *Constructing democracy in transitioning societies in Africa: Constitutionalism and deliberation in Mali.* New York: Palgrave Macmillan.
Central and Eastern Europe; Former USSR Austria; Czechoslovakia; Estonia; Germany; Hungary Latvia; Lithuania; Poland; Romania; Russia (USSR); Slovakia; Ukraine	Linz, Stepan, and Richard Gunther. 1995."Democratic Transitions and Consolidation in Southern Europe, with Reflections on Latin America and Eastern Europe." In *The Politics of Democratic Consolidation: Southern Europe in Comparative Perspective*, eds Richard Gunther, Nikiforos Diamandouros, Hans-Jurgen Puhel. Baltimore: John Hopkins University Press Plasser, Fritz and Ulram, Peter A. 1994. "Monitoring Democratic Consolidation: Political Trust and System Support in East-Central Europe", Paper for the XVI th World Congress of the International Political Science Association, Berlin. Banac, Ivo. 2014. "Twenty-Five Years after the Fall of the Berlin Wall. East European Politics and Societies and Cultures," *Special Issue on the Post-1989 Developments in the Region*, 28 (2): 653-657. Bogaards, Matthijs. 2009. "How to classify hybrid regimes? Defective democracy and electoral authoritarianism." *Democratization*, 16 (2): 399-423. Coman, Ramona and Tomini, Luca 2014. "A Comparative Perspective on the State of Democracy in Central and Eastern Europe." *Europe-Asia Studies* 66 (6): 853-858 Dzihic, Vedran. 2014. "Grey zones between democracy and authoritarianism: Re-thinking the current state of democracy in Eastern and South Eastern Europe" in Wiersma et al., eds. 2014. *Problems of Representative Democracy in Europe*, Amsterdam: Foundation for European Progressive Studies: 21-30. Papadopoulos, Yannis. 2013. *Democracy in Crisis? Politics, Governance and Policy.* Houndmills: Palgrave-Macmillan, 296.

Countries in sample by region	Region and country-specific sources
Central and Eastern Europe; Former USSR (Cont.) Austria; Czechoslovakia; Estonia; Germany; Hungary Latvia; Lithuania; Poland; Romania; Russia (USSR); Slovakia; Ukraine	Pappas, Takis. 2014. Populist Democracies: Post-Authoritarian Greece and Post-Communist Hungary. *Government and Opposition* 49 (1): 1-25. Golosov, G. 2008. "Electoral authoritarianism in Russia." *Pro et Contra*, JanuaryFebruary. Rose R, W. Mishler, N. Munro. 2006. *Russia transformed: developing popular support for a new regime.* Cambridge University Press, Cambridge, UK (2006) Rose R, N. Munro, W. Mishler. 2004. "Resigned acceptance of an incomplete democracy: Russia's political equilibrium." *Post-Soviet Affairs*, 20(3):195–218

References

Boix, Carles. 2003. *Democracy and Redistribution*. New York, NY: Cambridge University Press.

Box-Steffensmeier, Janet M. and Chrisopher Zorn. 1999. "Modeling Heterogeneity in Duration Models." Paper Presented at the 1999 Summer Meeting of The Political Methodology Society.

Buhaug, Halvard, Scott Gates and Päivi Lujala. 2009. "Geography, Rebel Capability, and the Duration of Civil Conflict." *Journal of Conflict Resolution* 53(4):544–569.

Dahl, Robert Alan. 1976. *Politics, economics, and welfare*. New Brunswick: Transaction Publishers.

Lindert, Peter H. 2004. *Growing public: Volume 1, The Story: Social Spending and Economic Growth Since the Eighteenth Century*. Vol. 1 Cambridge: Cambridge University Press.

Linz, Juan J. 1994. *The Failure of Presidential Democracy: Comparative Perspectives*. Vol. 1 Baltimore, MD: The Johns Hopkins University Press pp. 3–87.

Linz, Juan J and Alfred C Stepan. 1996. "Toward consolidated democracies." *Journal of democracy* 7(2):14–33.

Mahani, Alireza S. and Mansour T.A Sharabiani. 2016. *Package 'BayesMixSurv': Bayesian Mixture Survival Models using Additive Mixture-of-Weibull Hazards, with Lasso Shrinkage and Stratification*. R package version 0.9.1.

Munck, Gerardo L. 2009. *Measuring Democracy: A Bridge Between Scholarship and Politics*. Baltimore, MD: The Johns Hopkins Press Press.

Neal, Radford M. 2003. "Slice Sampling." *Annals of statistics* pp. 705–741.

Przeworski, Adam and Fernando Limongi. 1997. "Modernization: Theories and facts." *World Politics* 49(2):155–183.

Reenock, Christopher, Michael Bernhard and David Sobek. 2007. "Regressive socioeconomic distribution and democratic survival." *International Studies Quarterly* 51(3):677–699.

Samuel, David and Matthew S. Shugart. 2010. *Presidents, Parties, and Prime Ministers: How the Separation of Powers affects Party Organization and Behavior*. Cambridge: Cambridge University press.

Sen, Amartya. 1982. *Poverty and Famines – An Essay on Entitlement and Deprivation*. Oxford University Press.

Teorell, Jan. 2010. *Determinants of Democratization: Explaining Regime Change in the World, 1972–2006*. Cambridge: Cambridge University Press.