Modeling Customer Personality and Marketing Response:

A Bayesian Perspective

Yi-Hsuan Kuo, Felix Liu, Xuan Tao†

Content

† Sorted in alphabetical order by surname.

1. Introduction and Data Description

1.1 Background

Marketing activities are essential tools for businesses to engage with customers and increase revenue. However, customer responses to marketing efforts can vary significantly. It is therefore important to identify which demographic and behavioral characteristics best predict responsiveness to such campaigns.

This project employs a Bayesian hierarchical logistic regression model to explore the factors influencing customers' acceptance of marketing offers. By leveraging detailed demographic and transactional data, we aim to uncover both individual-level predictors and group-level patterns that explain variability in response behavior. These insights can help businesses refine and optimize their promotional strategies. We also examined several alternative prior specifications and conducted a straightforward comparison with frequentist models to assess robustness and relative performance.

1.2 Dataset Overview

The dataset used in this project is the **"Customer Personality Analysis dataset (2021)"**. Originally published on Kaggle. It contains detailed information for **2,240** customers, including demographics, purchasing behavior, and responses to marketing campaigns. **Key variables include**:

- **Age**: Customer's age in years
- **Income**: Yearly household income
- **Education**: Categorical variable with five levels (e.g., Graduation, PhD)
- **Marital_Status**: Categorical variable (e.g., Married, Single, Together)
- **Kidhome and Teenhome**: Number of children and teenagers in the household
- **Commodity types**: Amounts spent in various product categories
- **Response**: Binary outcome indicating if the customer accepted the most recent campaign offer

The dataset supports modeling both personal economic behavior (Income, Age, Spending) and group-level variation (Education, Marital_Status).

We select **Education** as the grouping factor due to its five well-distributed categories—Graduation (1,127), PhD (486), Master (370), 2n Cycle (203), and Basic (54)—suitable for estimating group-specific effects. Although ordinal, we model Education with exchangeable random intercepts, leaving ordinal modeling for future work. In contrast, **Marital_Status** includes rare or ambiguous labels (e.g., "YOLO", "Alone") and lacks a meaningful order, making it unsuitable for random effects modeling here.

2. Bayesian Model Specification

We specify a two-level hierarchical logistic regression. At the first level, individual-level predictors model each customer's probability of response. At the second level, we introduce a random intercept for Education, allowing baseline response rates to vary across education groups[1].

The likelihood for the observed data $y = \left(y_1, ..., y_N\right)$ given the parameters $\boldsymbol{\theta}$ is:

$$p(y|\boldsymbol{\theta}) = \prod_{i=1}^{N} Bernoulli\left(y_i|p_i\right)$$

where the probability p_i is modeled via the logit link:

$$logit\left(p_i\right) = \beta_0 + \sum_{k=1}^{4} \beta_k x_{ik} + \alpha_{g(i)}$$

where, $x_{i1}, x_{i2}, x_{i3}, x_{i4}$ represent Income, Age, Spending, and Children for observation i, respectively, and $g(i)$ is the Education group for observation i.

The set of parameters to be estimated is $\boldsymbol{\theta} = \left(\beta, \alpha, \sigma_\alpha\right)$, where:

- $\beta = \left(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4\right)$ is the vector of fixed effect coefficients.
- $\alpha = \left(\alpha_1, ..., \alpha_J\right)$ is the vector of random intercepts for the J Education groups.
- σ_α is the standard deviation of the random intercepts.

The prior distributions

- For fixed effects coefficients: $\pi\left(\beta_k\right) = N\left(\beta_k|0, 5^2\right), \quad for\ k = 0, ..., 4.$
- For random intercepts (hierarchical prior): $\pi\left(\alpha_j|\sigma_\alpha\right) = N\left(\alpha_j|0, \sigma_\alpha^2\right), \quad for\ j = 1, ..., J.$
- For the standard deviation of random intercepts (hyperprior): $\pi\left(\sigma_\alpha\right) = N^+\left(\sigma_\alpha|0, 5^2\right), \quad \sigma_\alpha > 0.$

The joint posterior distribution for the parameters $\boldsymbol{\theta}$ given the data y is:

[1] For the DAG, see Appendix B.2.1.

$$p(\boldsymbol{\theta}|y) \propto p(y|\theta)\,\pi(\beta)\,\pi\left(\alpha|\sigma_\alpha\right)\pi\left(\sigma_\alpha\right)$$

where $\pi(\beta) = \prod_{k=0}^{4} \pi\left(\beta_k\right)$ and $\pi\left(\alpha|\sigma_\alpha\right) = \prod_{j=1}^{J} \pi\left(\alpha_j|\sigma_\alpha\right)$.

Our goal is to characterize the posterior distributions of the fixed-effect coefficients β_k, the group-level intercepts α_j, and the hyperparameter σ_α.

3. Computational Approach

3.1 MCMC Settings

We used Markov Chain Monte Carlo (MCMC) sampling using the `brms` package in R. Four chains were run, each with 8,000 iterations, including 4,000 warm-up steps. This gave us 16,000 post-warmup samples in total. A higher `adapt_delta` value of 0.95 was used to reduce the chance of divergent transitions during sampling.

```
> summary(model)
 Family: bernoulli
  Links: mu = logit
Formula: Response ~ Income + Age + Spending + Children + (1 | Education)
   Data: df_train (Number of observations: 1772)
  Draws: 4 chains, each with iter = 8000; warmup = 4000; thin = 1;
         total post-warmup draws = 16000

Multilevel Hyperparameters:
~Education (Number of levels: 5)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.51      0.50     0.05     1.71 1.01      462      825

Regression Coefficients:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept    -1.10      0.47    -2.03    -0.18 1.01      764      712
Income       -0.00      0.00    -0.00    -0.00 1.00    13035    12422
Age          -0.01      0.01    -0.02     0.00 1.00     3826     6029
Spending      0.00      0.00     0.00     0.00 1.00    11776    11478
Children     -0.23      0.12    -0.46     0.00 1.00     3472     6101

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Figure 3.1: Model structure and summary statistics for fixed and group-level parameters[2]

[2] Due to rounding, some credible intervals (e.g., for Income) may appear as [-0.00, -0.00], but they actually represent small non-zero ranges. This is supported by the shape of the posterior histograms and the sensitivity analysis discussed in Section 5.

3.2 Convergence Diagnostics

We assessed convergence using several diagnostics:

- **Trace plots** of sd_Education__Intercept and all regression coefficients showed good mixing across chains, indicating convergence.
- **R-hat values** were all close to 1 (≤ 1.01), suggesting minimal between-chain variability.
- **Effective Sample Size (ESS)** was sufficiently high for all fixed effects and group-level parameters, indicating good sampling efficiency.
- **Monte Carlo Standard Error (MCSE)** was low, confirming that the posterior means are estimated precisely.

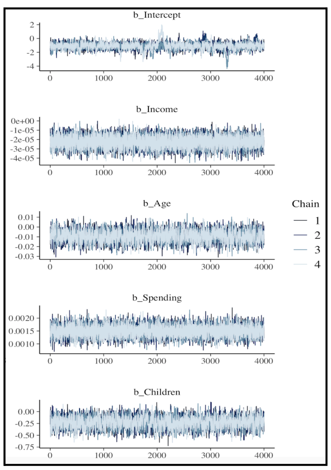


Figure 3.2a: Trace plots of regression coefficients show convergence for each predictor.

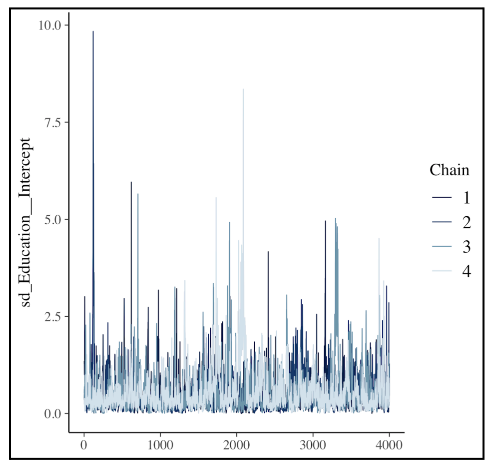


Figure 3.2b: Trace plots of sd_Education__Intercept shows convergence

3.3 Software Implementation

The model was implemented using the `brms` package in R, which interfaces with `Stan` to handle Bayesian inference. This allowed for flexible model specification and access to convergence diagnostics and predictive checks.

4. Results

4.1 Posterior Summaries

Posterior estimates for the regression coefficients suggest limited evidence of strong fixed-effect influences on the binary response variable.

- The coefficient for **Children** had a posterior mean of **-0.23**, with a 95% credible interval of **[-0.46, 0.00]**. The coefficient for **Age** had a posterior mean of **-0.01**, with a 95% credible interval of **[-0.02, 0.00].** Although the interval lies mostly below zero, it still includes 0, indicating a potentially negative but **not statistically conclusive** effect.
- The coefficients for **Income** and **Spending** were all near zero, with 95% credible intervals that not **included zero**, suggesting **no strong evidence** that these predictors individually affect the binary response.
- **Posterior histograms** in Figure 4.1 for all coefficients were approximately normal in shape, supporting stability of the posterior estimates
- **Overall**, the 95% credible intervals for all fixed-effect predictors included zero or were very close to including zero, indicating no strong evidence of statistically significant effects.

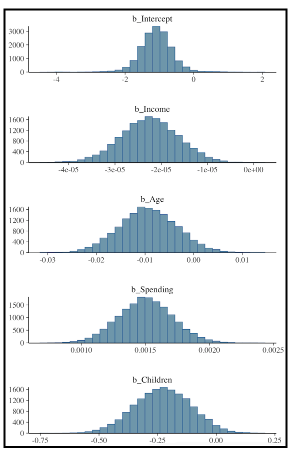


Figure 4.1: Posterior distributions of fixed-effect coefficients (Income, Age, Spending, Children)

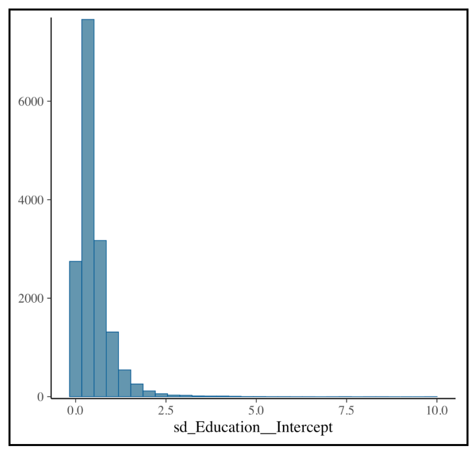


Figure 4.2: Posterior distribution of group-level standard deviation (sd_Education__Intercept)

4.2 Group-Level Effects

These results suggest that while individual predictors show limited explanatory power, group-level variation across education levels remains meaningful and warrants hierarchical modeling. The standard deviation of the random intercepts for Education was estimated with a mean of approximately 0.51, showing moderate variation across education levels. The credible interval excluded zero, supporting the inclusion of Education as a grouping factor.

4.3 Posterior Predictive Checks

We validated the model using two approaches:

Visual Checks: We used `pp_check(model)` to visually compare the distribution of predicted values (posterior predictive draws y_{rep}, with the actual observed data y). As shown in the density overlay plot, the predictive distribution closely matched the observed distribution, indicating that the model captured the overall data pattern well. This suggests that the model is capable of generating data similar to the real-world outcome.

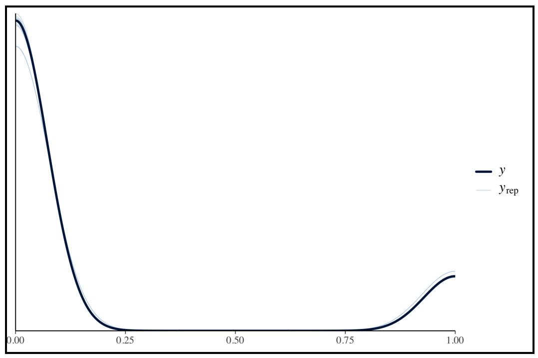


Figure 4.3: Overlay of the observed response distribution (y) and simulated responses from the model (y_{rep}), showing strong alignment and model fit.

Out-of-sample Prediction: We split the data into training and testing sets. The model was trained on the training set. Using a threshold of 0.5, we computed the predicted binary outcomes. The accuracy on the test set was 85.6%, indicating strong generalization performance. While the model's out-of-sample accuracy was high (85.6%), the contribution of each covariate appears uncertain, suggesting the need to explore alternative predictors, interaction terms, or non-linear effects in future work.

5. Sensitivity Analysis and Frequentist Comparison

5.1 Sensitivity Analysis

We evaluated how our key inferences and out-of-sample performance respond to a range of reasonable prior choices. All continuous predictors (Income, Age, Spending, Children) were standardized before fitting. In addition to the baseline weakly informative prior (Normal(0, 5^2) on coefficients;

Half-Normal(0, 5) on the random-intercept SD), we tested: a) Fixed-effect priors: weaker (Normal(0, 10^2)) vs. stronger (Normal(0, 2.5^2)); b) Random-effect SD priors: weaker (Half-Normal(0, 10)), stronger (Half-Normal(0, 1)), and a Half-Cauchy(0, 2.5) variant[3].

Across all six prior specifications[4], as shown in Figure 5.1, the posterior means for Income and Spending remained stably negative and positive, respectively, with 95 % credible intervals excluding zero. Age showed no credible effect under any prior, and Children retained a negative point estimate though its interval occasionally crossed zero. The estimate of σ_u varied modestly—from roughly 0.41 under the strongest shrinkage to 0.58 under the weakest—but its interval always excluded zero, confirming group-level heterogeneity by education. Importantly, the close overlap of means and intervals across all settings demonstrates that our inferences are robust to reasonable changes in prior choice.

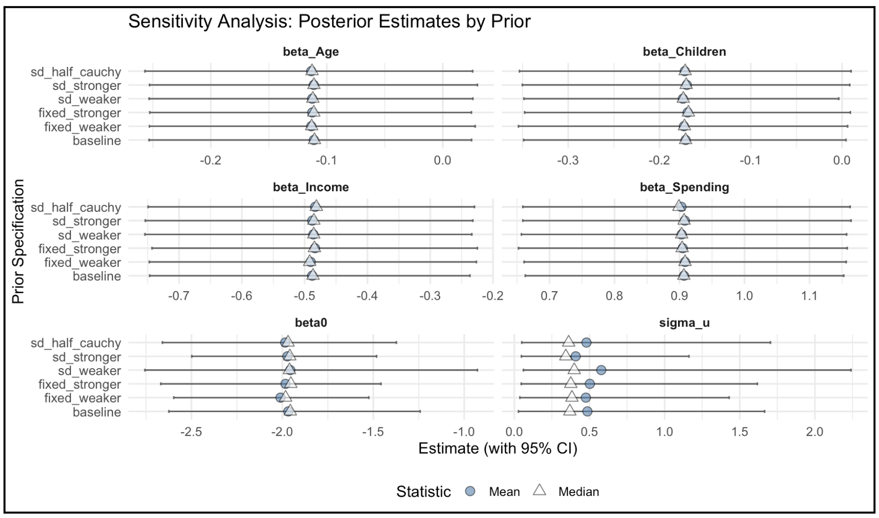


Figure 5.1 Sensitivity Analysis: Comparison of Posterior Estimates (standardized) by Prior[5]

Finally, predictive accuracy (≈ 85–85.4%) and AUC (≈ 0.72–0.73) on the test set changed by less than 0.005 across priors, demonstrating that our conclusions and classification performance are robust to these prior variations.

[3] See prior specifications and posterior summaries in Appendix B.5.1
[4] Sensitivity analysis was performed in JAGS (see BUGS code in Appendix A.5.1).
[5] See the full coefficient comparison table in Appendix B.5.2.

Table 5.1 Sensitivity Analysis: Model Predictive Performance (Test Set)

Prior Set	Accuracy	AUC
baseline	0.854	0.723
fixed_weaker	0.854	0.723
fixed_stronger	0.854	0.724
sd_weaker	0.851	0.724
sd_stronger	0.854	0.722
sd_half_cauchy	0.851	0.723

5.2 Frequentist Comparison

To compare the Bayesian model's performance, two frequentist benchmarks were also fitted: a standard logistic regression (GLM) treating Education as a fixed effect, and a mixed-effects logistic model (GLMER) with Education as a random intercept.

Table 5.2 presents a comparison of parameter estimates across GLM, GLMER, and Bayesian models. It confirms that all three models identify Income (approximately -0.50) and Spending (approximately 0.90) as their strongest predictors, with p-values less than 0.001 and 95% confidence intervals excluding zero. Differences in uncertainty emerge: Age never achieves significance, as all confidence intervals include zero, and Children is only marginally significant (frequentist p-values approximately 0.04 to 0.07; Bayesian confidence interval just touching zero).

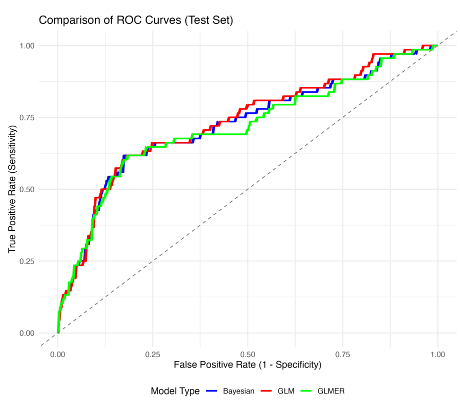
Furthermore, GLMER's point estimate for the Education standard deviation (0.165) provides no indication of uncertainty, while the Bayesian standard deviation (0.486 [0.028, 1.665]) reveals substantial between-group variability and accurately quantifies our posterior uncertainty, highlighting the Bayesian framework's advantage in assessing both effect size and confidence.

Table 5.2 Comparison of Parameter Estimates Across GLM, GLMER, and Bayesian Models

Parameter	GLM Estimate (SE)	GLMER Estimate (SE)	Bayesian Mean (Median)
Intercept	−2.1921 (0.2607)***	−1.9187 (0.1207)***	−1.967 (−1.956) [−2.624, −1.242]
Income	−0.5137 (0.1343)***	−0.4646 (0.1318)***	−0.488 (−0.487) [−0.747, −0.237]
Age	−0.1181 (0.0709)·	−0.1067 (0.0706)	−0.111 (−0.111) [−0.253, 0.025]
Spending	0.9032 (0.1281)***	0.8983 (0.1273)***	0.907 (0.907) [0.663, 1.153]
Children	−0.1864 (0.0906)*	−0.1623 (0.0898)·	−0.171 (−0.171) [−0.349, 0.004]
σ_u (Education SD)	—	0.1649	0.486 (0.369) [0.028, 1.665]

Note:
1) * p < 0.05; ** p < 0.01; *** p < 0.001; · p < 0.10;
2) Bayesian estimates show full posterior means, medians and 95 % credible intervals.

All three models exhibited nearly identical discrimination on the test set, with AUCs ranging from 0.714 to 0.732—GLM at 0.732, Bayesian at 0.723, and GLMER at 0.714—demonstrating that their ability to rank responders versus non-responders is effectively the same.



Model Predictive Performance Comparison (Test Set)

- Bayesian
  - Accuracy: 0.854[6]
  - AUC: 0.723
- GLM
  - Accuracy: 0.851
  - AUC: 0.732
- GLMER
  - Accuracy: 0.856
  - AUC: 0.714

Figure 5.2 Comparison of ROC and Predictive Performance (Bayesian vs GLM vs GLMER)

6. Discussion

We fit a Bayesian hierarchical logistic regression to predict customer response from standardized Income, Age, Spending, and Children, allowing Education to have its own random intercept. Weakly informative Normal and half-Normal priors provided regularization without overwhelming the data.

[6] The Bayesian model in this part was implemented in JAGS, whereas earlier analyses used the brms package, leading to minor differences in test-set accuracy (0.856 vs. 0.854) but essentially consistent results.

The Bayesian model revealed credible non-zero effects for standardized Income (negative) and Spending (positive), with their 95% credible intervals consistently excluding zero. The estimated standard deviation for the 'Education' random intercepts also robustly excluded zero, confirming significant response heterogeneity across educational backgrounds and validating the hierarchical approach. These results remained stable under various prior choices and matched the predictive accuracy of frequentist GLM and GLMER models. Although point estimates may differ slightly between Bayesian and frequentist methods, the Bayesian framework provides full posterior distributions, allowing us to account for uncertainty directly and to evaluate effect sizes more comprehensively than single-value estimates permit.

Despite these strengths, our analysis faces three key limitations. First, the simple linear specification omits interactions (such as Income × Age) and nonlinear effects that might uncover diminishing returns or synergies. Second, only Education was treated as a random intercept; small strata may remain underpowered, and other groupings (e.g. region or product category) could benefit from cross-classification or random slopes. Future work should therefore adopt more flexible and hierarchical extensions. Penalized splines or Gaussian process smooths could model nonlinear effects of Age and Spending, while explicitly including interactions would test for synergy. Allowing random slopes—such as varying the Spending effect by Education would yield more precise group-level estimates. These enhancements would deepen explanatory power without sacrificing the Bayesian framework's advantages in uncertainty quantification and regularization.

Contributions of Group Members

- Kuo, Yi-Hsuan:
  - Section 3 Computational Approach
  - Section 4 Results
- Liu, Felix:
  - Section 2 Bayesian Model Specification
  - Section 5 Sensitivity Analysis and Frequentist Comparison
  - Section 6 Discussion
- Tao, Xuan
  - Section 1 Introduction and Data Description
  - Section 2 Bayesian Model Specification

AI Acknowledgment

The initial draft of this report was written by the team. AI tools were used only to refine formulas and code formatting. all analysis and conclusions were made by the team members.

Appendix

A. Important R Code

A.3.1 Markov Chain Monte Carlo (MCMC) sampling

```
model <- brm(
  formula = Response ~ Income + Age + Spending + Children + (1 | Education),
  family = bernoulli(link = "logit"), # logistic regression
  data = df_train,
  chains = 4,
  iter = 8000, # original:2000
  warmup = 4000, # original:1000
  seed = 123,
  prior = c( set_prior("normal(0, 5)", class = "b"), # for Income, Age, Spending,
Children
    set_prior("normal(0, 5)", class = "Intercept"), # for intercept
    set_prior("normal(0, 5)", class = "sd") # foe education),
  control = list(adapt_delta = 0.95) # to control divergent transitions)
summary(model)
plot(model)
pp_check(model) # Posterior Predictive Check
```

A.5.1 Sensitive Analysis BUGS Models

```
// ---------------------------------------------------------------------
// model_baseline.bug
// Baseline priors:
// Betas: Normal(mean=0, sd=5)  => precision = 1/(5^2) = 0.04
// sigma_u: Half-Normal(loc=0, scale=5) => dnorm(0, 0.04)T(0,)
// ---------------------------------------------------------------------
model {
  # Likelihood: Bernoulli logistic regression model
  for (i in 1:N) {
    Response[i] ~ dbern(p[i])
    logit(p[i]) <- beta0 +
                    beta_Income * Income[i] +
                    beta_Age * Age[i] +
                    beta_Spending * Spending[i] +
                    beta_Children * Children[i] +
                    u[EducationID[i]]  # education_id from your R code
  }

  # Priors for fixed effects
  beta0 ~ dnorm(0, 0.04)          # Intercept
  beta_Income ~ dnorm(0, 0.04)
  beta_Age ~ dnorm(0, 0.04)
  beta_Spending ~ dnorm(0, 0.04)
  beta_Children ~ dnorm(0, 0.04)

  # Random intercepts by education level: u[j] ~ N(0, sigma_u^2)
  for (j in 1:J) {                # J is the number of education levels
    u[j] ~ dnorm(0, tau_u)
  }
```

```
  # Hyperprior for random effects precision
  tau_u <- pow(sigma_u, -2)       # Precision for random intercepts
  sigma_u ~ dnorm(0, 0.04) T(0,)  # Half-Normal prior for sigma_u (sd of u_j)
                                  # loc=0, scale=5 => precision = 1/(5^2) = 0.04
}
```text
// ------------------------------------------------------------------
// model_fixed_weaker.bug
// Fixed effects priors weaker:
// Betas: Normal(mean=0, sd=10) => precision = 1/(10^2) = 0.01
// sigma_u: Half-Normal(loc=0, scale=5) => dnorm(0, 0.04)T(0,)
// ------------------------------------------------------------------
model {
  # Likelihood
  for (i in 1:N) {
    Response[i] ~ dbern(p[i])
    logit(p[i]) <- beta0 +
                   beta_Income * Income[i] +
                   beta_Age * Age[i] +
                   beta_Spending * Spending[i] +
                   beta_Children * Children[i] +
                   u[EducationID[i]]
  }

  # Priors for fixed effects (weaker)
  beta0 ~ dnorm(0, 0.01)          # Intercept
  beta_Income ~ dnorm(0, 0.01)
  beta_Age ~ dnorm(0, 0.01)
  beta_Spending ~ dnorm(0, 0.01)
  beta_Children ~ dnorm(0, 0.01)

  # Random intercepts
  for (j in 1:J) {
    u[j] ~ dnorm(0, tau_u)
  }

  # Hyperprior for random effects (same as baseline)
  tau_u <- pow(sigma_u, -2)
  sigma_u ~ dnorm(0, 0.04) T(0,)
}
```text
// ------------------------------------------------------------------
// model_fixed_stronger.bug
// Fixed effects priors stronger:
// Betas: Normal(mean=0, sd=2.5) => precision = 1/(2.5^2) = 1/6.25 = 0.16
// sigma_u: Half-Normal(loc=0, scale=5) => dnorm(0, 0.04)T(0,)
// ------------------------------------------------------------------
model {
  # Likelihood
  for (i in 1:N) {
    Response[i] ~ dbern(p[i])
    logit(p[i]) <- beta0 +
                   beta_Income * Income[i] +
                   beta_Age * Age[i] +
                   beta_Spending * Spending[i] +
                   beta_Children * Children[i] +
                   u[EducationID[i]]
  }

  # Priors for fixed effects (stronger)
```

```
  beta0 ~ dnorm(0, 0.16)          # Intercept
  beta_Income ~ dnorm(0, 0.16)
  beta_Age ~ dnorm(0, 0.16)
  beta_Spending ~ dnorm(0, 0.16)
  beta_Children ~ dnorm(0, 0.16)

  # Random intercepts
  for (j in 1:J) {
    u[j] ~ dnorm(0, tau_u)
  }

  # Hyperprior for random effects (same as baseline)
  tau_u <- pow(sigma_u, -2)
  sigma_u ~ dnorm(0, 0.04) T(0,)
}
```text
// --------------------------------------------------------------------
// model_sd_weaker.bug
// Prior for sigma_u weaker:
// Betas: Normal(mean=0, sd=5)  => precision = 0.04
// sigma_u: Half-Normal(loc=0, scale=10) => dnorm(0, 0.01)T(0,)
// --------------------------------------------------------------------
model {
  # Likelihood
  for (i in 1:N) {
    Response[i] ~ dbern(p[i])
    logit(p[i]) <- beta0 +
                   beta_Income * Income[i] +
                   beta_Age * Age[i] +
                   beta_Spending * Spending[i] +
                   beta_Children * Children[i] +
                   u[EducationID[i]]
  }

  # Priors for fixed effects (same as baseline)
  beta0 ~ dnorm(0, 0.04)
  beta_Income ~ dnorm(0, 0.04)
  beta_Age ~ dnorm(0, 0.04)
  beta_Spending ~ dnorm(0, 0.04)
  beta_Children ~ dnorm(0, 0.04)

  # Random intercepts
  for (j in 1:J) {
    u[j] ~ dnorm(0, tau_u)
  }

  # Hyperprior for random effects (sigma_u prior weaker)
  tau_u <- pow(sigma_u, -2)
  sigma_u ~ dnorm(0, 0.01) T(0,) # Half-Normal, loc=0, scale=10
}
```text
// --------------------------------------------------------------------
// model_sd_stronger.bug
// Prior for sigma_u stronger:
// Betas: Normal(mean=0, sd=5)  => precision = 0.04
// sigma_u: Half-Normal(loc=0, scale=1) => dnorm(0, 1)T(0,)
// --------------------------------------------------------------------
model {
  # Likelihood
  for (i in 1:N) {
    Response[i] ~ dbern(p[i])
```

```
    logit(p[i]) <- beta0 +
                    beta_Income * Income[i] +
                    beta_Age * Age[i] +
                    beta_Spending * Spending[i] +
                    beta_Children * Children[i] +
                    u[EducationID[i]]
  }

  # Priors for fixed effects (same as baseline)
  beta0 ~ dnorm(0, 0.04)
  beta_Income ~ dnorm(0, 0.04)
  beta_Age ~ dnorm(0, 0.04)
  beta_Spending ~ dnorm(0, 0.04)
  beta_Children ~ dnorm(0, 0.04)

  # Random intercepts
  for (j in 1:J) {
    u[j] ~ dnorm(0, tau_u)
  }

  # Hyperprior for random effects (sigma_u prior stronger)
  tau_u <- pow(sigma_u, -2)
  sigma_u ~ dnorm(0, 1) T(0,)    # Half-Normal, loc=0, scale=1
}
```text
// -----------------------------------------------------------------
// model_sd_half_cauchy.bug
// Prior for sigma_u is Half-Cauchy:
// Betas: Normal(mean=0, sd=5)  => precision = 0.04
// sigma_u: Half-Cauchy(loc=0, scale=2.5) => dt(0, 1/(2.5^2), 1)T(0,) = dt(0, 0.16,
1)T(0,)
// -----------------------------------------------------------------
model {
  # Likelihood
  for (i in 1:N) {
    Response[i] ~ dbern(p[i])
    logit(p[i]) <- beta0 +
                    beta_Income * Income[i] +
                    beta_Age * Age[i] +
                    beta_Spending * Spending[i] +
                    beta_Children * Children[i] +
                    u[EducationID[i]]
  }

  # Priors for fixed effects (same as baseline)
  beta0 ~ dnorm(0, 0.04)
  beta_Income ~ dnorm(0, 0.04)
  beta_Age ~ dnorm(0, 0.04)
  beta_Spending ~ dnorm(0, 0.04)
  beta_Children ~ dnorm(0, 0.04)

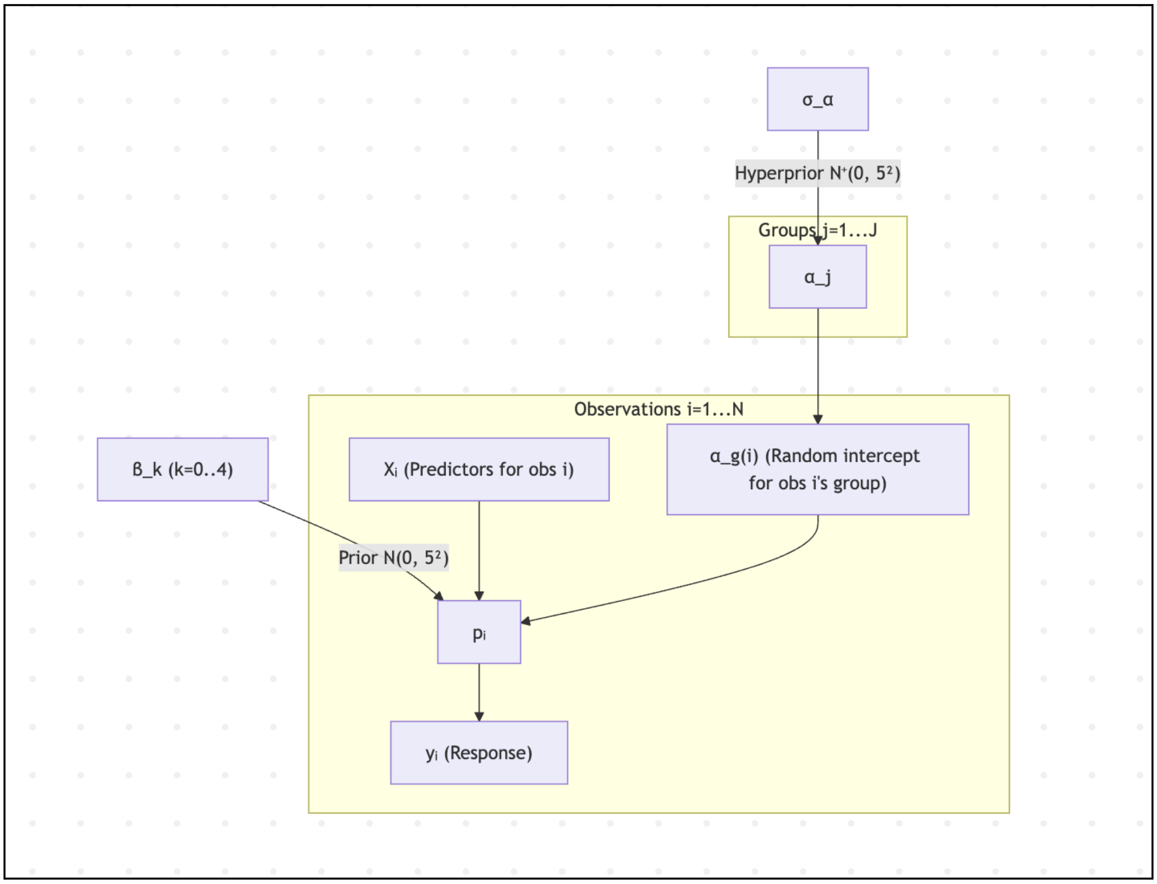
  # Random intercepts
  for (j in 1:J) {
    u[j] ~ dnorm(0, tau_u)
  }

  # Hyperprior for random effects (sigma_u prior is Half-Cauchy)
  tau_u <- pow(sigma_u, -2)
  sigma_u ~ dt(0, 0.16, 1) T(0,) # Half-Cauchy, loc=0, scale=2.5
                                 # (precision for t-dist = 1/scale^2, df=1 for
Cauchy)
```
```

```
}
```

B. Additional Figures or Tables

B.2.1 DAG of Bayesian Hierarchical Model



B.5.1 Overview of Prior Settings Used in Sensitivity Analysis

Prior Set Name	Prior for Fixed Effects (β)	Prior for Random Effect SD (σ_u)
baseline	Normal(0, sd=5)	Half-Normal(loc=0, scale=5)
fixed_weaker	Normal(0, sd=10)	Half-Normal(loc=0, scale=5)
fixed_stronger	Normal(0, sd=2.5)	Half-Normal(loc=0, scale=5)
sd_weaker	Normal(0, sd=5)	Half-Normal(loc=0, scale=10)
sd_stronger	Normal(0, sd=5)	Half-Normal(loc=0, scale=1)
sd_half_cauchy	Normal(0, sd=5)	Half-Cauchy(loc=0, scale=2.5)

Note:
1. baseline: Standard weakly informative priors.
2. fixed_weaker: Weaker (less informative) priors for fixed effects.
3. fixed_stronger: Stronger (more informative, more shrinkage towards 0) priors for fixed effects.
4. sd_weaker: Weaker prior for the random effect standard deviation σ_u.
5. sd_stronger: Stronger prior for the random effect standard deviation σ_u.
6. sd_half_cauchy: Alternative prior family (Half-Cauchy) for σ_u.

B.5.2 Sensitivity Analysis (rjags): Parameter Posterior Statistics

Prior Set	**Mean**	**Median**	**2.5% CI**	**97.5% CI**
β_0				
baseline	-1.967	-1.956	-2.624	-1.242
fixed_weaker	-2.01	-1.982	-2.597	-1.525
fixed_stronger	-1.983	-1.954	-2.668	-1.458
sd_weaker	-1.956	-1.961	-2.756	-0.927
sd_stronger	-1.972	-1.959	-2.498	-1.482
sd_half_cauchy	-1.984	-1.969	-2.66	-1.373
β_{Income}				
baseline	-0.488	-0.487	-0.747	-0.237
fixed_weaker	-0.49	-0.492	-0.747	-0.226
fixed_stronger	-0.484	-0.484	-0.743	-0.225
sd_weaker	-0.486	-0.486	-0.754	-0.234
sd_stronger	-0.488	-0.485	-0.754	-0.232
sd_half_cauchy	-0.483	-0.481	-0.749	-0.229
β_{Age}				
baseline	-0.111	-0.111	-0.253	0.025
fixed_weaker	-0.113	-0.113	-0.253	0.028
fixed_stronger	-0.113	-0.111	-0.253	0.025
sd_weaker	-0.113	-0.112	-0.253	0.026
sd_stronger	-0.111	-0.111	-0.253	0.03
sd_half_cauchy	-0.113	-0.113	-0.257	0.026
$\beta_{Spending}$				
baseline	0.907	0.907	0.663	1.153
fixed_weaker	0.909	0.909	0.661	1.156
fixed_stronger	0.905	0.904	0.652	1.158

sd_weaker	0.903	0.903	0.657	1.157
sd_stronger	0.908	0.907	0.659	1.164
sd_half_cauchy	0.903	0.899	0.659	1.162
$\beta_{Children}$				
baseline	-0.171	-0.171	-0.349	0.004
fixed_weaker	-0.173	-0.173	-0.355	0.006
fixed_stronger	-0.169	-0.168	-0.348	0.009
sd_weaker	-0.174	-0.174	-0.349	-0.004
sd_stronger	-0.17	-0.171	-0.35	0.008
sd_half_cauchy	-0.172	-0.172	-0.353	0.01
σ_u				
baseline	0.486	0.369	0.028	1.665
fixed_weaker	0.476	0.383	0.037	1.429
fixed_stronger	0.503	0.376	0.046	1.617
sd_weaker	0.579	0.399	0.061	2.24
sd_stronger	0.408	0.343	0.047	1.162
sd_half_cauchy	0.48	0.362	0.05	1.704