

Supplement2.R

PL39

Sat Sep 02 15:35:27 2017

```
# Effect of parameters and covariates using four different definitions of "failed" hearing screen.  
# 484 infants failed a first test but did not have documented follow up. We categorized them  
# four different ways to determine how sensitive the model was to these definitions.  
# Some parameters, such as population density and the random effect for block group, were not  
# included in the models presented in the manuscript, but were evaluated here.
```

```
setwd("~/Project Data/Hearing study")  
load("~/Project Data/Hearing study/workspace.8.21.RData")
```

```
library(brms)
```

```
## Loading required package: Rcpp
```

```
## Loading required package: ggplot2
```

```
## Loading 'brms' package (version 1.8.0). Useful instructions
```

```
## can be found by typing help('brms'). A more detailed introduction
```

```
## to the package is available through vignette('brms_overview').
```

```
library(ggplot2)
```

```
library(viridis)
```

```
## Loading required package: viridisLite
```

```
library(bayesplot)
```

```
## This is bayesplot version 1.2.0
```

```
##### RANDOMIZE THOSE WITH NO FOLLOW UP TO PASS VS FAIL AT RATE OF ~25% FAILURE
```

```
##### * This reflects actual rate of fail vs pass on follow up among those who were followed up
```

```
# Rand_unadj <- brm(As_Fail ~ s(long, lat) data=data, family=bernoulli, chains=6, iter=3000,  
# control= list(adapt_delta=0.99))
```

```
# Rand_full <- brm(As_Fail ~ s(long, lat) + BirthWeight + GestAge + Gender + NICU + STI + Parity + Race +  
# ADI + PopDens + (1|bgfips), data=data, family=bernoulli, chains=6, iter=3000,  
# control= list(adapt_delta=0.99))
```

```
##### COUNT AS FAIL IF FAILED FIRST TEST AND NO FOLLOW UP
```

```
##### * This is the definition of "fail" used for most analyses in this study
```

```
# Fail_unadj <- brm(As_Fail ~ s(long, lat) data=data, family=bernoulli, chains=6, iter=3000,  
# control= list(adapt_delta=0.99))
```

```
# As_Fail <- brm(As_Fail ~ s(long, lat) + BirthWeight + GestAge + Gender + NICU + STI + Parity + Race +  
# ADI + PopDens + (1|bgfips), data=data, family=bernoulli, chains=6, iter=3000,  
# control= list(adapt_delta=0.99))
```

```
##### COUNT AS PASS IF NO FOLLOW UP, EVEN IF FAILED FIRST TEST
```

```
# Pass_unadj <- brm(As_Fail ~ s(long, lat) data=data, family=bernoulli, chains=6, iter=3000,  
# control= list(adapt_delta=0.99))
```

```
# As_Pass <- brm(As_Fail ~ s(long, lat) + BirthWeight + GestAge + Gender + NICU + STI + Parity + Race +
```

```

#           ADI + PopDens + (1/bgfips), data=data, family=bernoulli, chains=6, iter=3000,
#           control= list(adapt_delta=0.99))

#### EXCLUDE INFANTS WHO DID NOT HAVE A SECOND TEST AFTER FAILING FIRST
# data2<-subset(data, pass_fail < 2) ## select out / exclude infants who failed test 1 but no follow up
# Rand_unadj <- brm(As_Fail ~ s(long, lat) data=data, family=bernoulli, chains=6, iter=3000,
#                 control= list(adapt_delta=0.99))
# Exclude <- brm(As_Fail ~ s(long, lat) + BirthWeight + GestAge + Gender + NICU + STI + Parity + Race +
#               ADI + PopDens + (1/bgfips), data=data, family=bernoulli, chains=6, iter=3000,
#               control= list(adapt_delta=0.99))

### Variable names in model summaries:
# ga_std = gestational age (standardized / centered)
# bw_std = birth weight (standardized / centered)
# gender = gender (0 = female, 1 = male)
# nicu = NICU admission (0 = no, 1 = yes)
# binrace = (binary) race/ethnicity (0 = nonhispanic white, 1 = minority)
# parity_std = maternal parity (standardized / centered)
# sti = maternal STI (0 = no, 1 = yes)
# adi_std = ADI (area deprivation index, standardized / centered)
# pop_std = population density (standardized / centered)
# slonglat1 and 2 = smoothed long/lat pairs

#### Predict distribution of hearing loss with different definitions
Rand_fit<-as.data.frame(fitted(Rand_unadj, newdata=grid))
fit1<-fitted(Rand_unadj, newdata=grid, summary=FALSE)
m1<-mean(fit1)
p1<-colMeans(fit1>m1)
Rand_table<-cbind(grid[,1:2], Rand_fit, p1)
Randplot<-ggplot(Rand_table) +
  ggtitle("Randomized") +
  geom_raster(aes(long, lat, fill=Estimate)) +
  scale_fill_viridis(option="B") +
  geom_contour(aes(x=long, y=lat, z=p1), breaks=c(0.025), color="Blue", cex=1.2, lty=1) +
  geom_contour(aes(x=long, y=lat, z=p1), breaks=c(0.975), color="Red", cex=1.2, lty=1) +
  theme_bw()

Fail_fit<-as.data.frame(fitted(Fail_unadj, newdata=grid))
fit2<-fitted(Fail_unadj, newdata=grid, summary=FALSE)
m2<-mean(fit2)
p2<-colMeans(fit2>m2)
Fail_table<-cbind(grid[,1:2], Fail_fit, p2)
Failplot<-ggplot(Fail_table) +
  ggtitle("As Fail") +
  geom_raster(aes(long, lat, fill=Estimate)) +
  scale_fill_viridis(option="B") +
  geom_contour(aes(x=long, y=lat, z=p2), breaks=c(0.025), color="Blue", cex=1.2, lty=1) +
  geom_contour(aes(x=long, y=lat, z=p2), breaks=c(0.975), color="Red", cex=1.2, lty=1) +
  theme_bw()

Pass_fit<-as.data.frame(fitted(Pass_unadj, newdata=grid))
fit3<-fitted(Pass_unadj, newdata=grid, summary=FALSE)

```

```

m3<-mean(fit3)
p3<-colMeans(fit3>m3)
Pass_table<-cbind(grid[,1:2], Pass_fit, p3)
Passplot<-ggplot(Pass_table) +
  ggtitle("As Pass") +
  geom_raster(aes(long, lat, fill=Estimate)) +
  scale_fill_viridis(option="B") +
  geom_contour(aes(x=long, y=lat, z=p3), breaks=c(0.025), color="Blue", cex=1.2, lty=1) +
  geom_contour(aes(x=long, y=lat, z=p3), breaks=c(0.975), color="Red", cex=1.2, lty=1) +
  theme_bw()

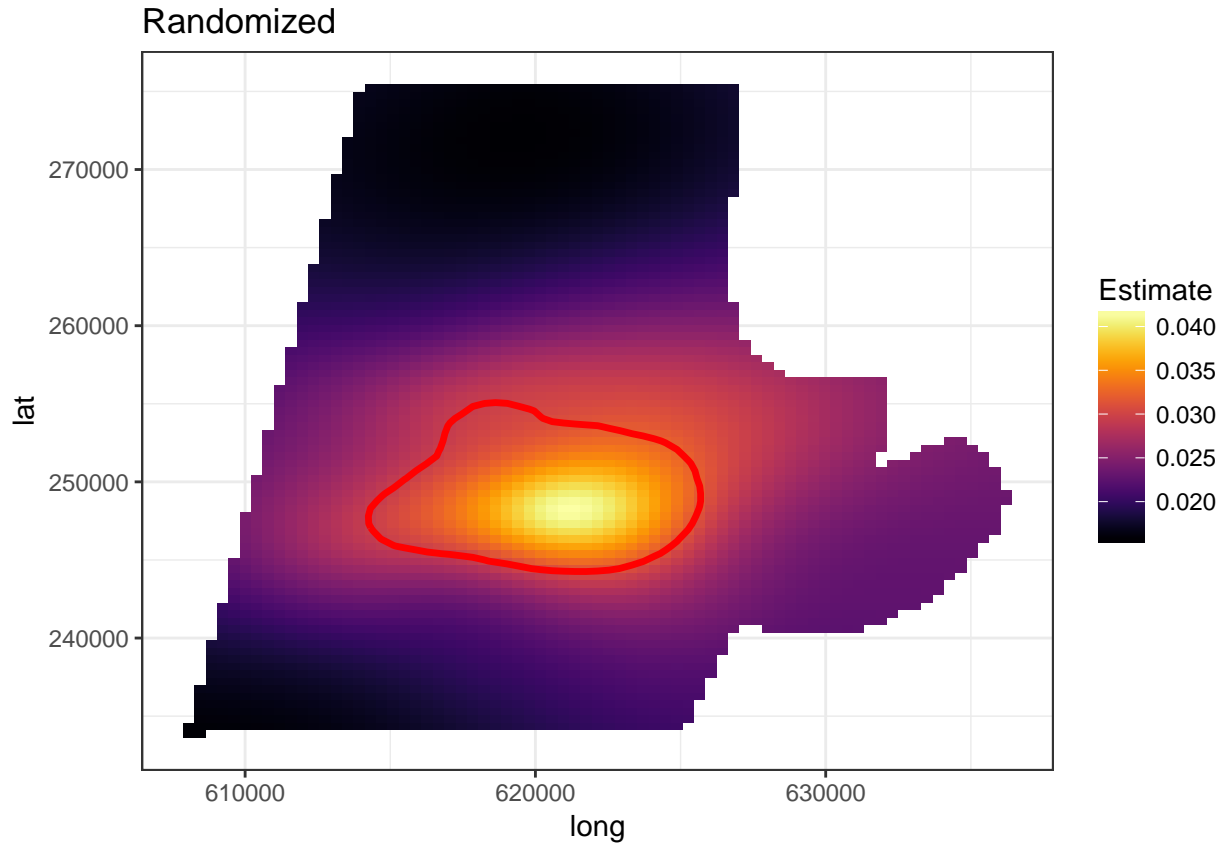
Ex_fit<-as.data.frame(fitted(Ex_unadj, newdata=grid))
fit4<-fitted(Ex_unadj, newdata=grid, summary=FALSE)
m4<-mean(fit4)
p4<-colMeans(fit4>m4)
Ex_table<-cbind(grid[,1:2], Ex_fit, p4)
Explot<-ggplot(Ex_table) +
  ggtitle("Exclude") +
  geom_raster(aes(long, lat, fill=Estimate)) +
  scale_fill_viridis(option="B") +
  geom_contour(aes(x=long, y=lat, z=p4), breaks=c(0.025), color="Blue", cex=1.2, lty=1) +
  geom_contour(aes(x=long, y=lat, z=p4), breaks=c(0.975), color="Red", cex=1.2, lty=1) +
  theme_bw()

## Estimates differ, because Pass and Exclude have 2/3 fewer "failures, so overall odds of hearing
## loss are lower, but spatial distribution and significant covariates do not substantially change.

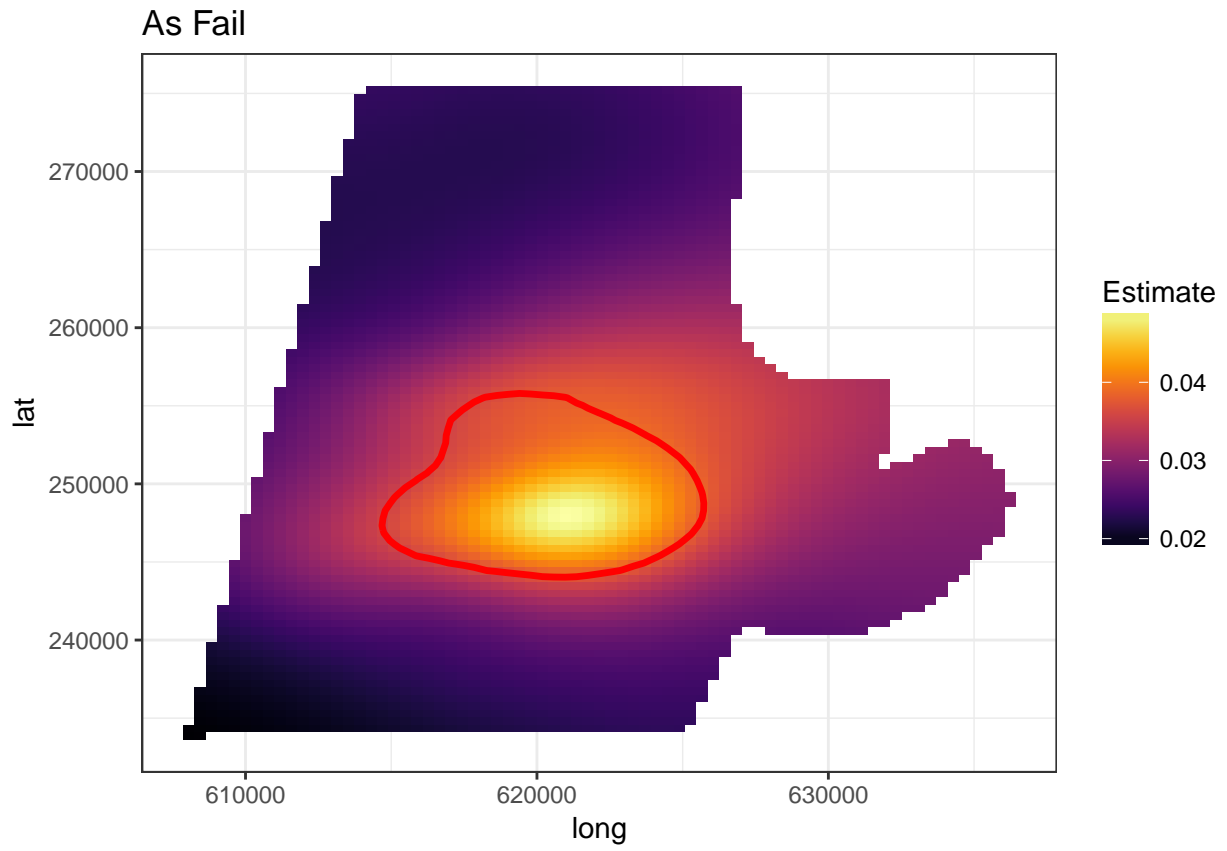
Randplot; Failplot; Passplot; Explot

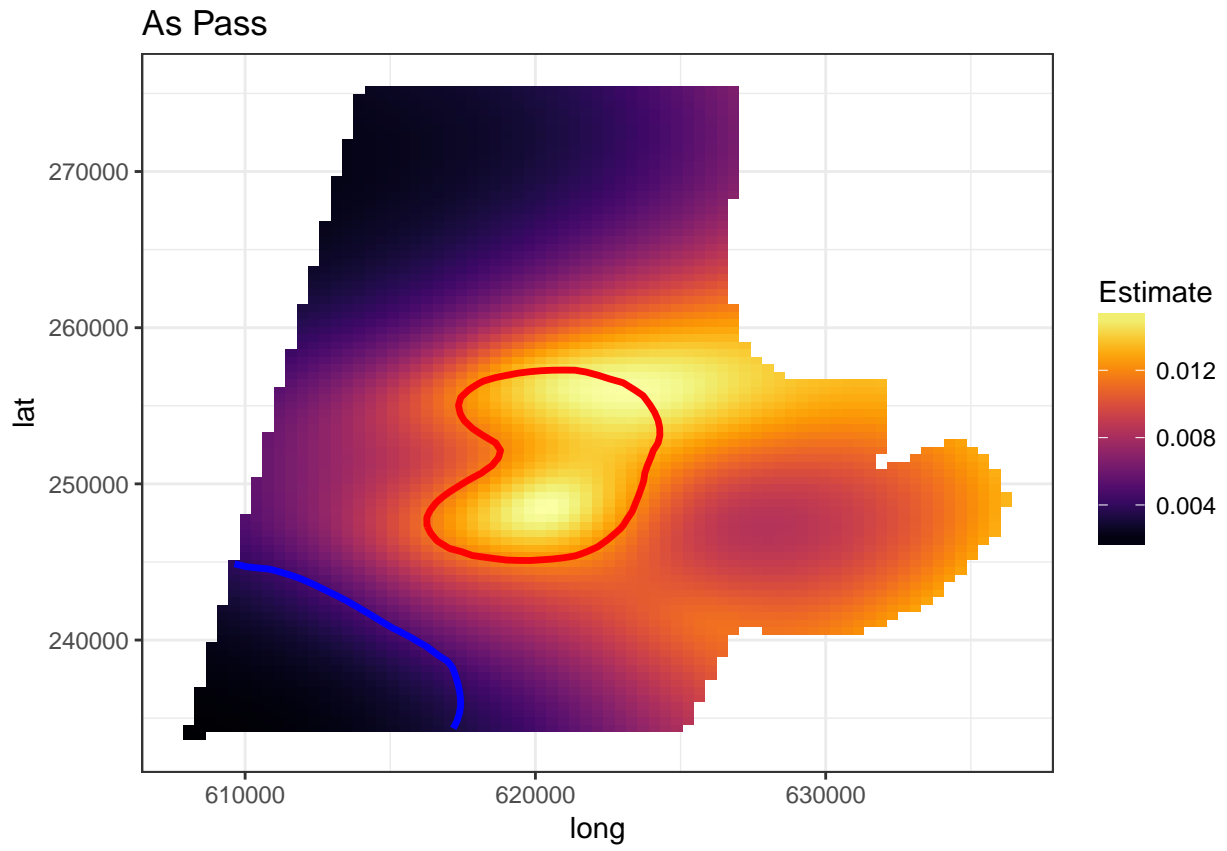
## Warning: Not possible to generate contour data

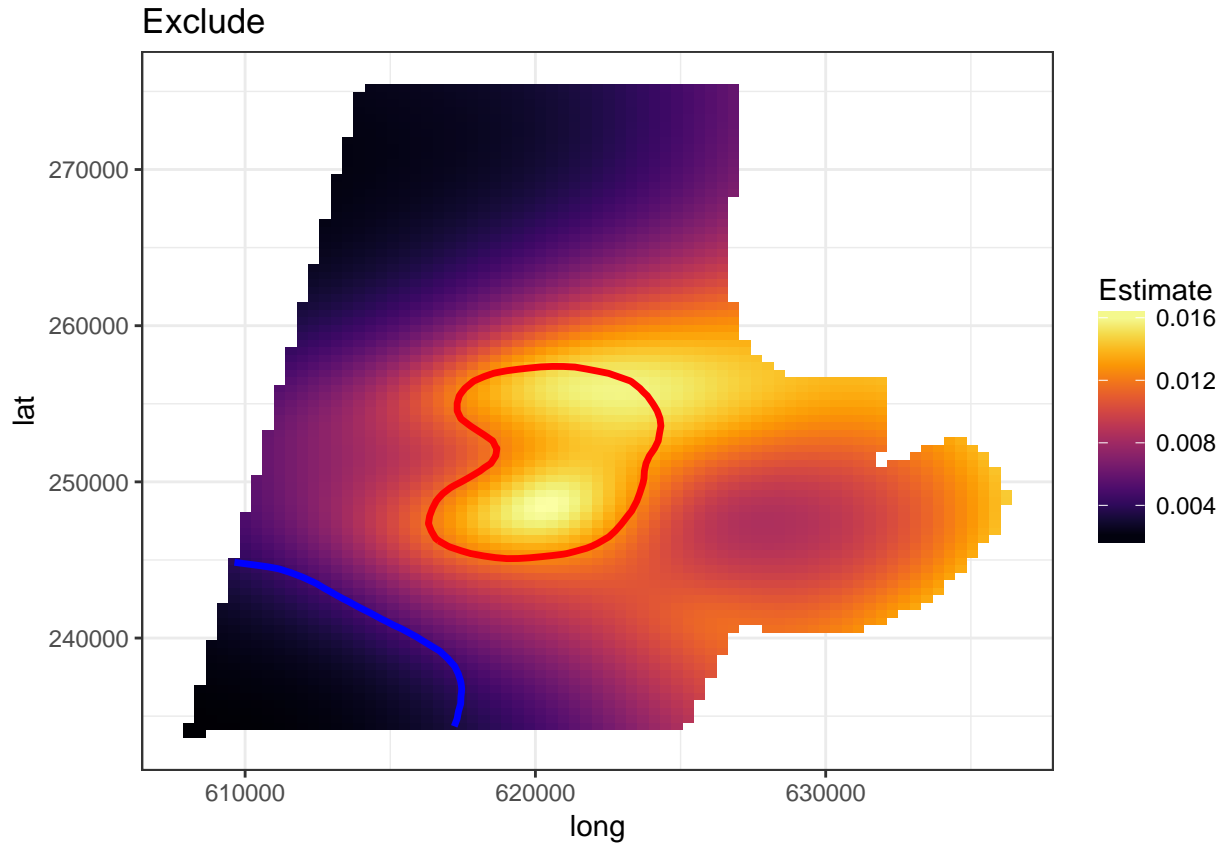
```



Warning: Not possible to generate contour data







```
#### Model summaries
# 1 - Full model summaries
# 2 - Histogram of posterior distribution of parameters
# A parameter that does not contain 0 is significant. In these fully adjusted models
# low gestational age and minority race are always significant predictors of hearing loss.
# Maternal STI is a significant predictor in the "As fail" and the "Randomized" models.
# It is not significant in the "As Pass" and the "Exclude" models; however, higher maternal parity
# is significant in these cases.
```

```
summary(Rand_full)
```

```
## Family: bernoulli(logit)
## Formula: Random ~ s(long, lat) + bw_std + ga_std + gender + nicu + sti + parity_std + binrace + adi_
## Data: data (Number of observations: 19348)
## Samples: 8 chains, each with iter = 3000; warmup = 1500; thin = 1;
##           total post-warmup samples = 12000
## ICs: LOO = Not computed; WAIC = Not computed
##
## Smooth Terms:
##           Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## sds(slonglat_1)    0.49     0.36    0.02    1.35     4305    1
##
## Group-Level Effects:
## ~bgfips (Number of levels: 152)
##           Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## sd(Intercept)    0.16     0.09    0.01    0.33     2706    1
```

```
##
## Population-Level Effects:
##      Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## Intercept      -4.34      0.13   -4.61   -4.09     12000     1
## bw_std          0.04      0.05   -0.06    0.14     12000     1
## ga_std          -0.20      0.05   -0.30   -0.11     12000     1
## gender1         0.09      0.09   -0.08    0.26     12000     1
## nicu1           0.12      0.33   -0.59    0.73     12000     1
## sti1            0.29      0.09    0.11    0.48     12000     1
## parity_std      0.02      0.04   -0.06    0.10     12000     1
## binrace1        0.82      0.13    0.56    1.09     12000     1
## adi_std         0.04      0.06   -0.09    0.16     12000     1
## pop_std        -0.03      0.06   -0.15    0.09     12000     1
## slonglat_1      0.01      0.10   -0.22    0.20      9605     1
## slonglat_2      0.00      0.09   -0.17    0.18      9601     1
##
## Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
## is a crude measure of effective sample size, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
summary(Fail_full)
```

```
## Family: bernoulli(logit)
## Formula: As_Fail ~ s(long, lat) + bw_std + ga_std + gender + nicu + sti + parity_std + binrace + adi
## Data: data (Number of observations: 19348)
## Samples: 6 chains, each with iter = 3000; warmup = 1500; thin = 1;
##          total post-warmup samples = 9000
##          ICs: LOO = Not computed; WAIC = Not computed
##
## Smooth Terms:
##      Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## sds(slonglat_1)  0.38      0.29    0.02    1.12     3308     1
##
## Group-Level Effects:
## ~bgfips (Number of levels: 152)
##      Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## sd(Intercept)  0.11      0.07    0.01    0.27     2337     1
##
## Population-Level Effects:
##      Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## Intercept      -4.11      0.12   -4.34   -3.88     9000     1
## bw_std          0.05      0.05   -0.05    0.14     9000     1
## ga_std          -0.21      0.04   -0.29   -0.12     9000     1
## gender1         0.12      0.08   -0.04    0.28     9000     1
## nicu1           0.37      0.27   -0.19    0.88     9000     1
## sti1            0.24      0.08    0.07    0.41     9000     1
## parity_std      0.01      0.04   -0.07    0.08     9000     1
## binrace1        0.78      0.12    0.55    1.01     9000     1
## adi_std         0.04      0.06   -0.07    0.15     9000     1
## pop_std         0.00      0.05   -0.10    0.11     9000     1
## slonglat_1      0.03      0.09   -0.15    0.20      5700     1
## slonglat_2     -0.03      0.08   -0.17    0.13      6305     1
##
## Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
## is a crude measure of effective sample size, and Rhat is the potential
```



```
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
summary(Pass_full)
```

```
## Family: bernoulli(logit)
## Formula: As_Pass ~ s(long, lat) + bw_std + ga_std + gender + nicu + sti + parity_std + binrace + adi
## Data: data (Number of observations: 19348)
## Samples: 6 chains, each with iter = 3000; warmup = 1500; thin = 1;
##           total post-warmup samples = 9000
##           ICs: LOO = Not computed; WAIC = Not computed
##
## Smooth Terms:
##           Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## sds(slonglat_1)    1.13      0.85    0.04    3.2      2402    1
##
## Group-Level Effects:
## ~bgfips (Number of levels: 152)
##           Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## sd(Intercept)    0.15      0.11    0.01    0.39      3529    1
##
## Population-Level Effects:
##           Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## Intercept        -5.47      0.23   -5.95   -5.03      9000    1
## bw_std            -0.09      0.09   -0.27    0.08      9000    1
## ga_std            -0.20      0.08   -0.36   -0.04      9000    1
## gender1           -0.01      0.15   -0.29    0.29      9000    1
## nicu1             0.15      0.54   -1.05    1.10      9000    1
## sti1              0.08      0.16   -0.23    0.40      9000    1
## parity_std        0.15      0.06    0.02    0.27      9000    1
## binrace1          0.84      0.24    0.37    1.33      9000    1
## adi_std           0.04      0.10   -0.17    0.24      9000    1
## pop_std           0.05      0.10   -0.15    0.23      9000    1
## slonglat_1        0.00      0.25   -0.60    0.40      4514    1
## slonglat_2       -0.23      0.19   -0.67    0.10      4896    1
##
## Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
## is a crude measure of effective sample size, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
summary(Ex_full)
```

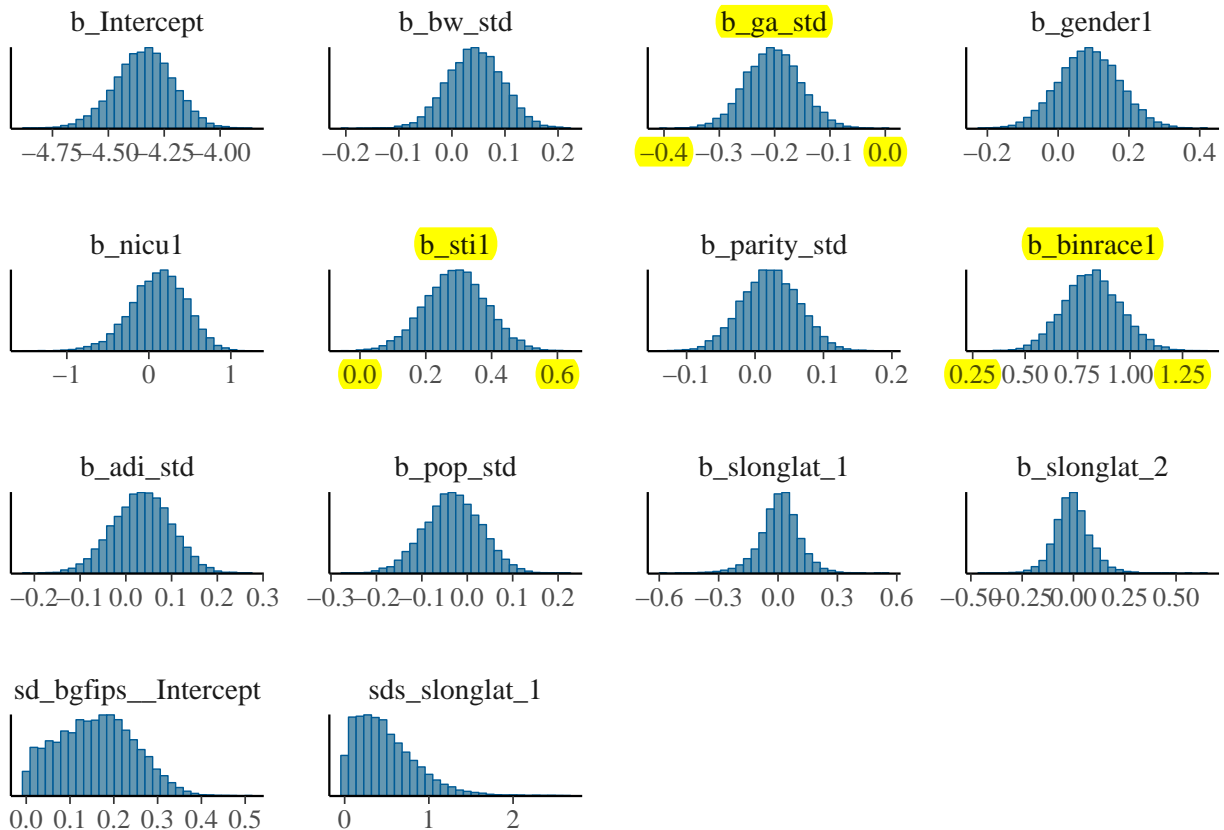
```
## Warning: There were 3 divergent transitions after warmup. Increasing adapt_delta above 0.99 may help
## See http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup
## Family: bernoulli(logit)
## Formula: As_Fail ~ s(long, lat) + bw_std + ga_std + gender + nicu + sti + parity_std + binrace + adi
## Data: data2 (Number of observations: 18864)
## Samples: 6 chains, each with iter = 3000; warmup = 1500; thin = 1;
##           total post-warmup samples = 9000
##           ICs: LOO = Not computed; WAIC = Not computed
##
## Smooth Terms:
##           Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## sds(slonglat_1)    0.83      0.64    0.04    2.36      2094    1
##
## Group-Level Effects:
```

```

## ~bgfips (Number of levels: 152)
##           Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## sd(Intercept)    0.15    0.11    0.01    0.4    2839    1
##
## Population-Level Effects:
##           Estimate Est.Error 1-95% CI u-95% CI Eff.Sample Rhat
## Intercept      -5.47    0.23   -5.94   -5.03    9000    1
## bw_std         -0.09    0.09   -0.26    0.08    9000    1
## ga_std         -0.20    0.08   -0.36   -0.04    9000    1
## gender1         0.00    0.15   -0.30    0.28    9000    1
## nicu1          0.16    0.55   -1.04    1.13    9000    1
## stil           0.10    0.16   -0.21    0.40    9000    1
## parity_std      0.15    0.06    0.02    0.27    9000    1
## binrace1        0.86    0.24    0.40    1.34    9000    1
## adi_std         0.04    0.10   -0.17    0.24    9000    1
## pop_std         0.05    0.10   -0.15    0.23    9000    1
## slonglat_1     -0.23    0.19   -0.68    0.09    4171    1
## slonglat_2      0.01    0.25   -0.59    0.39    2786    1
##
## Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
## is a crude measure of effective sample size, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
##### Marginal effects plots
stanplot(Rand_full, type="hist")

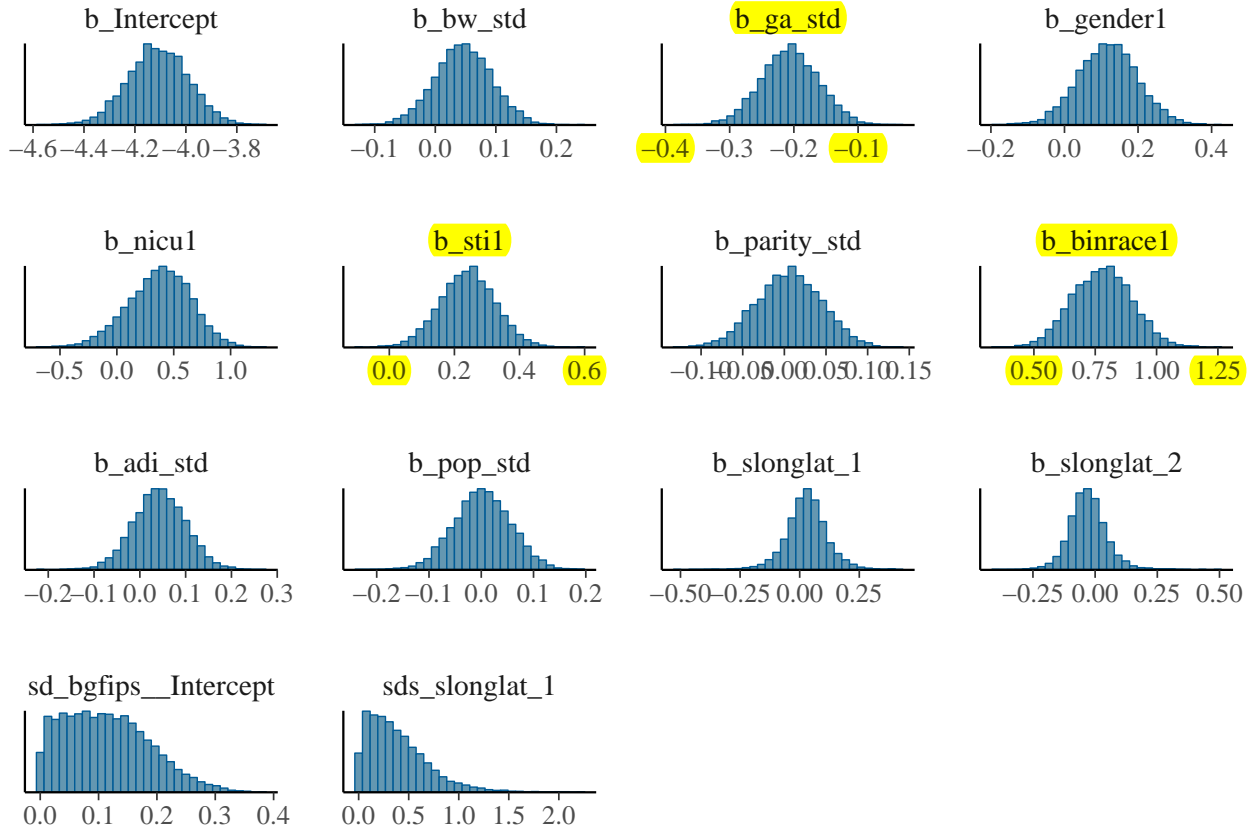
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

```



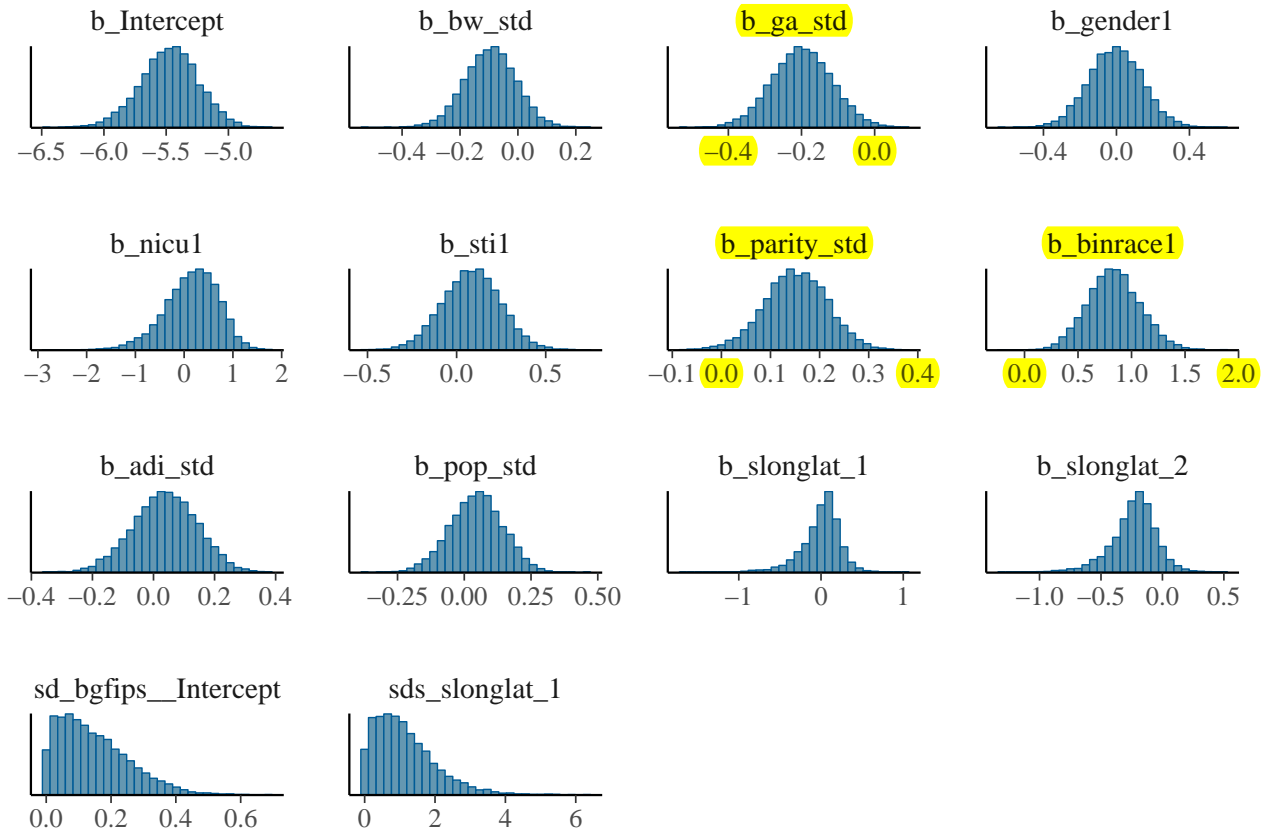
```
stanplot(Fail_full, type="hist")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
stanplot(Pass_full, type="hist")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
stanplot(Ex_full, type="hist")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

