MCMC Diagnostics - IFLS data

Sarah Teichman

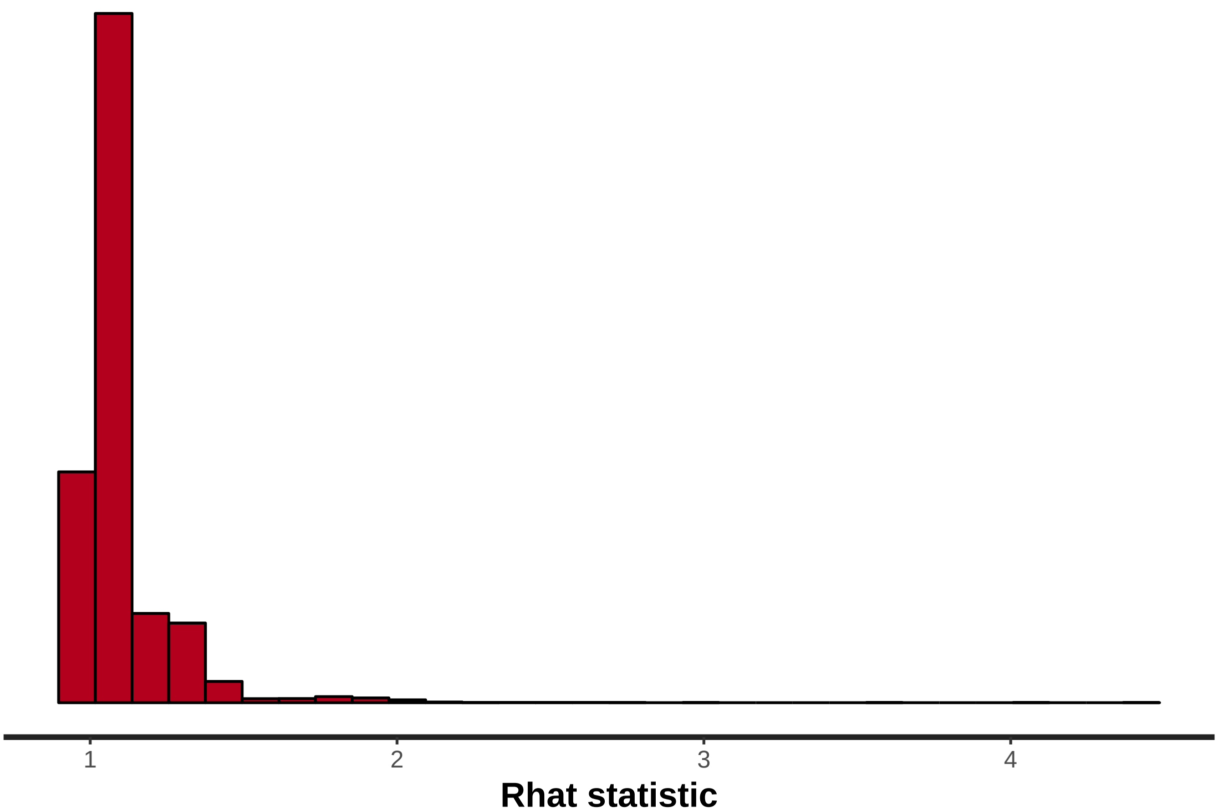
04/18/2020

```
K <- 7
Ti <- 3
N <- 1973
```

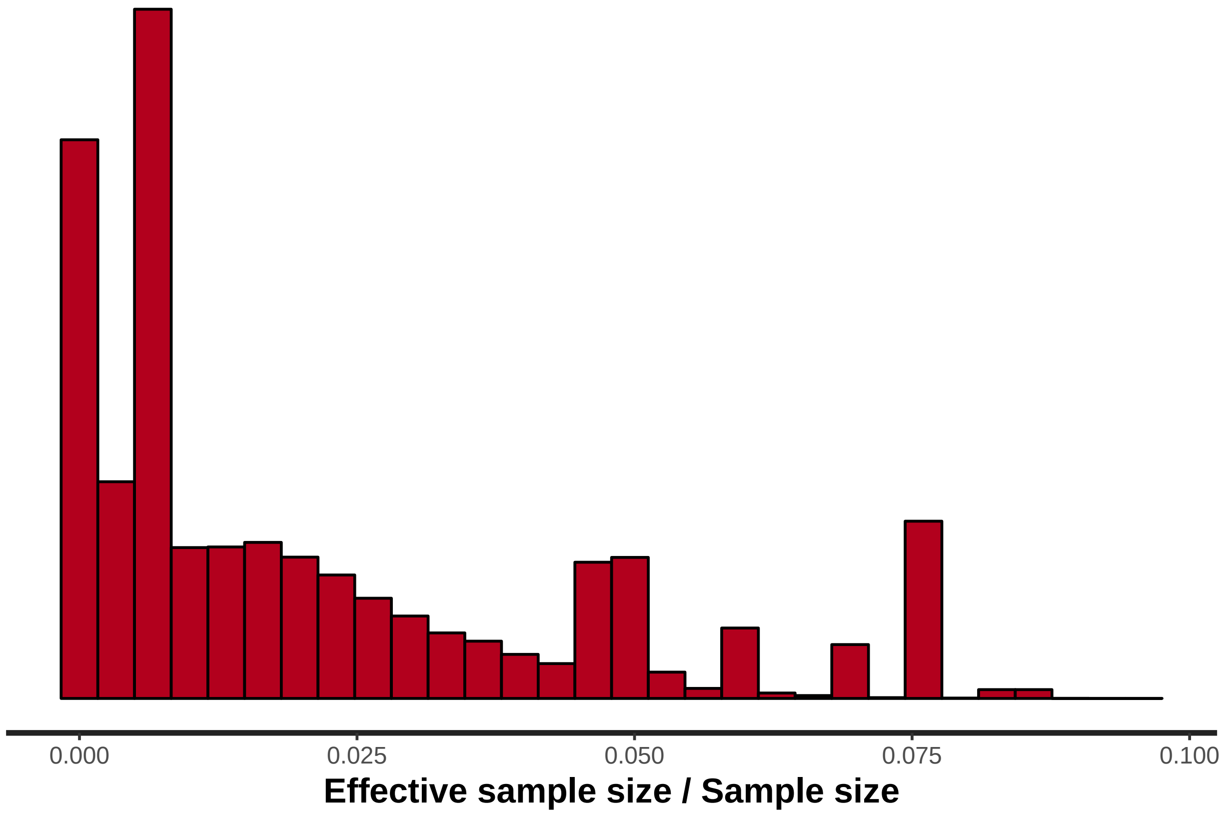
General MCMC diagnostic plots

Overall model diagnostics from rstan package.

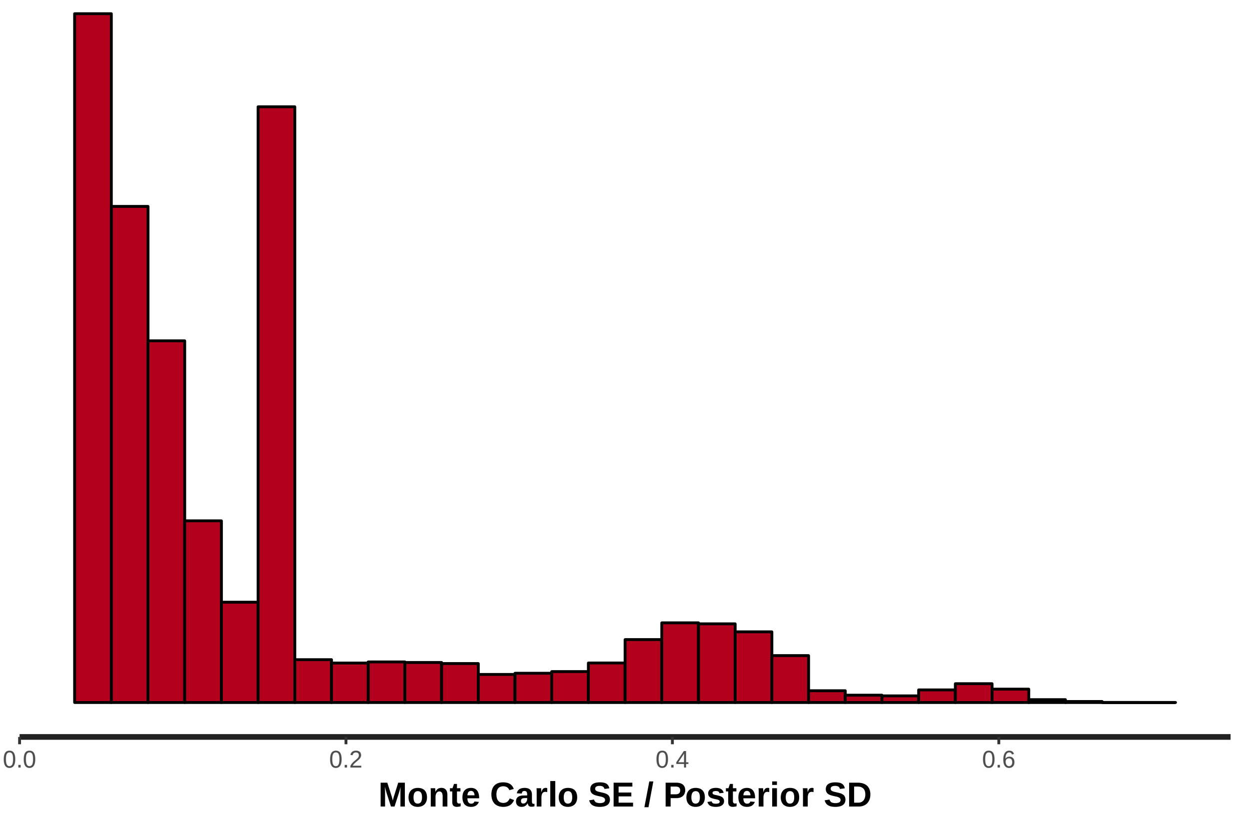
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




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




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



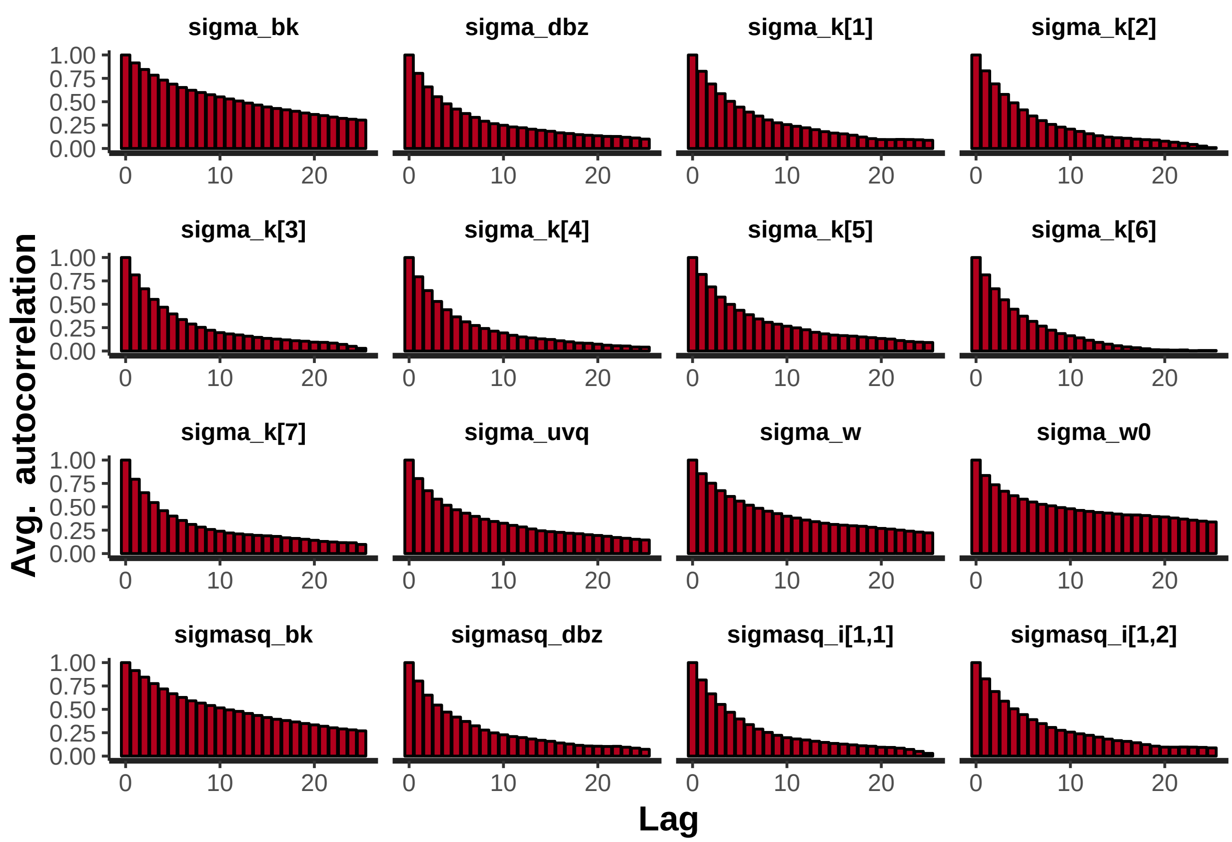
Individual Parameter Diagnostics

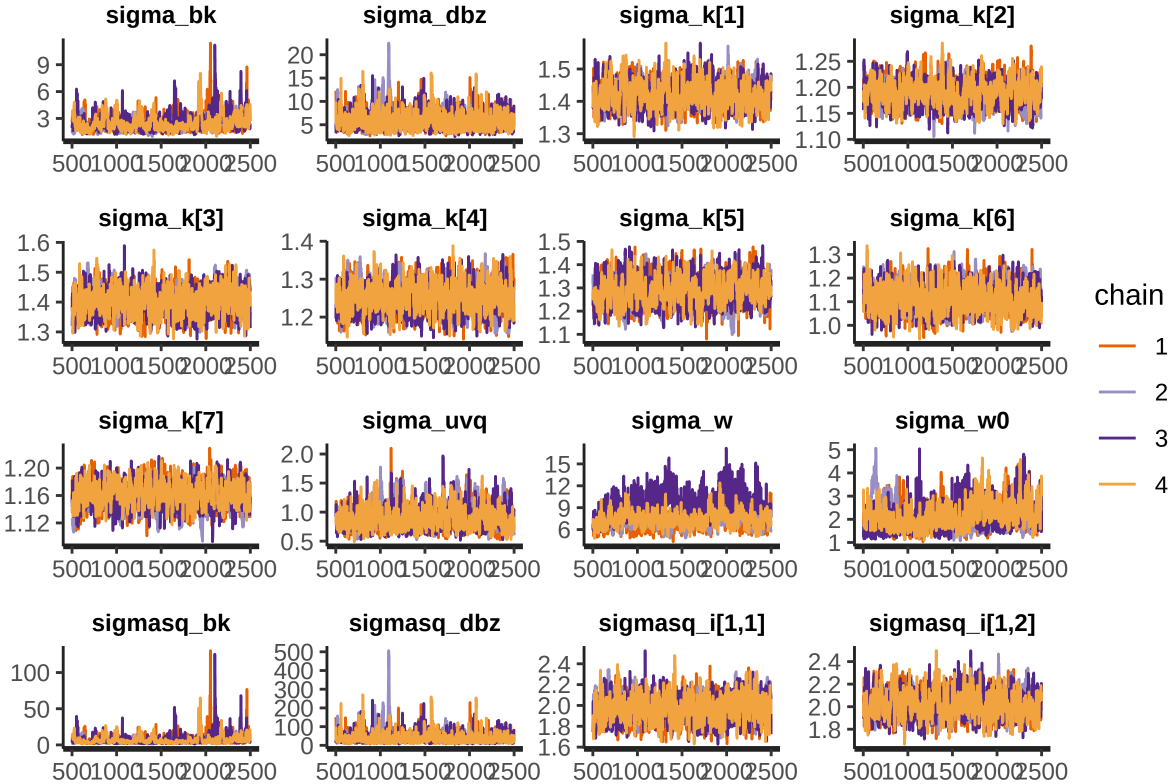
Individual parameter plots. Autocorrelation and trace plots for individual parameters, and histograms of posterior medians for group parameters.

```
get_single_plots <- function(fit, param) {
  print(fit_summ[param,c(1,2,3,5,6,7,9,10)])
  print(stan_ac(fit, pars = param))
  print(rstan::traceplot(fit, pars = param))
}
get_aggreg_plots <- function(fit, param, trim = F, trim_amount) {
  ind <- grep(paste0("^",param), rownames(as.data.frame(summary(fit)$summary)))
  medians <- data.frame(avg = as.data.frame(summary(fit)$summary)$`50%`[ind])
  print(paste0("Summary statistics for posterior medians of ",param))
  print(summary(medians))
  title <- paste0("Posterior Medians of ",param)
  print(ggplot(medians, aes(x = avg)) + geom_histogram(bins = 60) + ggtitle(title))
  if (trim == T) {
    lim <- quantile(abs(medians$avg), probs = trim_amount)
    meds_trim <- medians %>% filter(abs(medians$avg) < lim)
    print(ggplot(meds_trim, aes(x = avg)) + geom_histogram(bins = 60) +
            ggtitle(paste0(title, " Without Extreme ",100*(1-trim_amount),"%")))
  }
}
plot_fit <- function(fit) {
  get_single_plots(fit, sigma_params)
  get_single_plots(fit, beta_k)
  get_single_plots(fit, other_1d)
  get_single_plots(fit, u)
  get_single_plots(fit, v)
  get_single_plots(fit, q)
  get_aggreg_plots(fit, "w")
  get_aggreg_plots(fit, "z")
  get_aggreg_plots(fit, "p")
  get_aggreg_plots(fit, "eta", trim = T, trim_amount = .60)
  get_aggreg_plots(fit, "lambda", trim = T, trim_amount = .60)
  get_aggreg_plots(fit, "kappa", trim = T, trim_amount = .60)
}
plot_fit(fit)
```

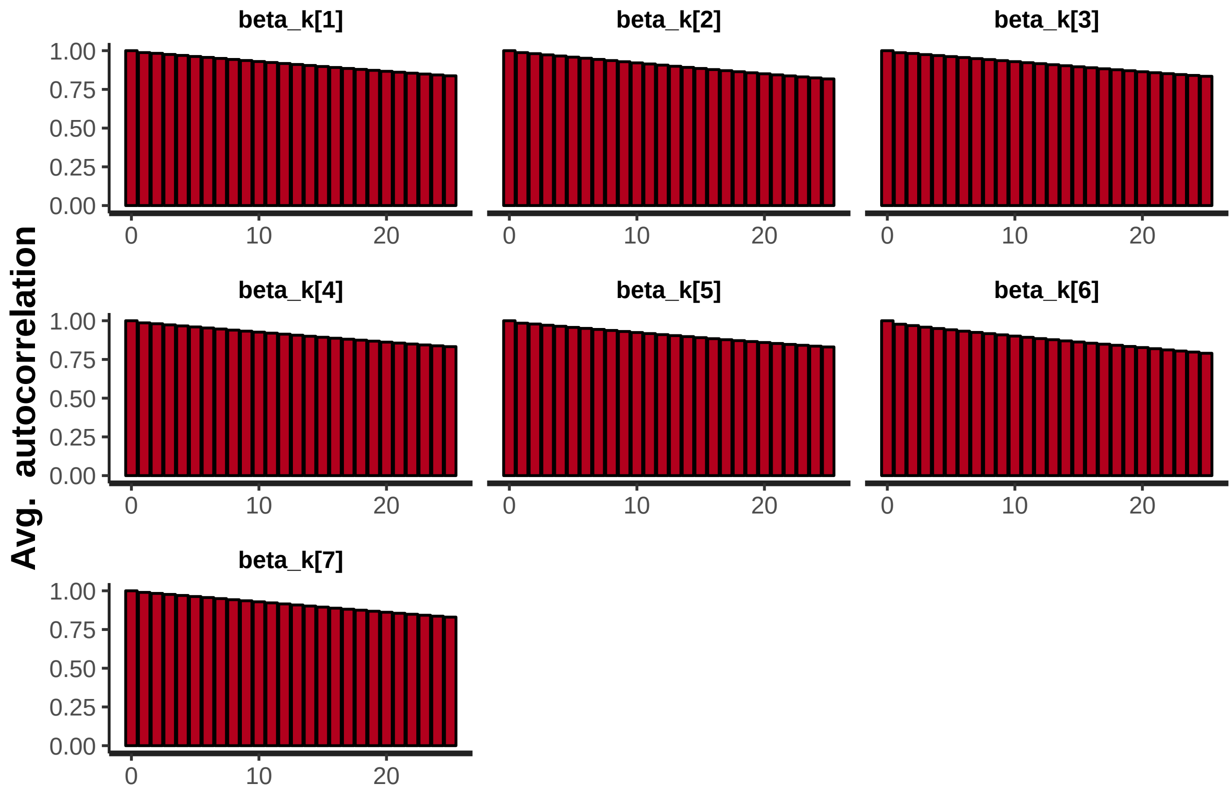
```
##                   mean      se_mean         sd       25%       50%
## sigma_bk     2.4841203 0.0753791904 0.90648334 1.8868803 2.2607473
## sigma_dbz    5.9963166 0.1246422210 1.98077944 4.6213343 5.5890537
## sigma_k[1]   1.4216260 0.0017643752 0.03864715 1.3950498 1.4194742
## sigma_k[2]   1.1882158 0.0009775429 0.02385794 1.1714872 1.1876302
## sigma_k[3]   1.4006206 0.0018058998 0.04238896 1.3705480 1.3993499
## sigma_k[4]   1.2498360 0.0015091294 0.03721624 1.2243974 1.2482965
## sigma_k[5]   1.2839141 0.0029316366 0.05919025 1.2418558 1.2827409
## sigma_k[6]   1.1182711 0.0021540259 0.05584899 1.0792843 1.1155865
## sigma_k[7]   1.1604961 0.0009041459 0.01766865 1.1491095 1.1608273
## sigma_uvq    0.9114963 0.0132959854 0.19330807 0.7739901 0.8813445
## sigma_w      7.7544651 0.7305075905 1.61250156 6.6564168 7.4107379
## sigma_w0     2.0969797 0.0622360624 0.54786232 1.7013136 2.0017151
## sigmasq_bk   6.9924631 0.4835888517 6.57876897 3.5603171 5.1109784
```

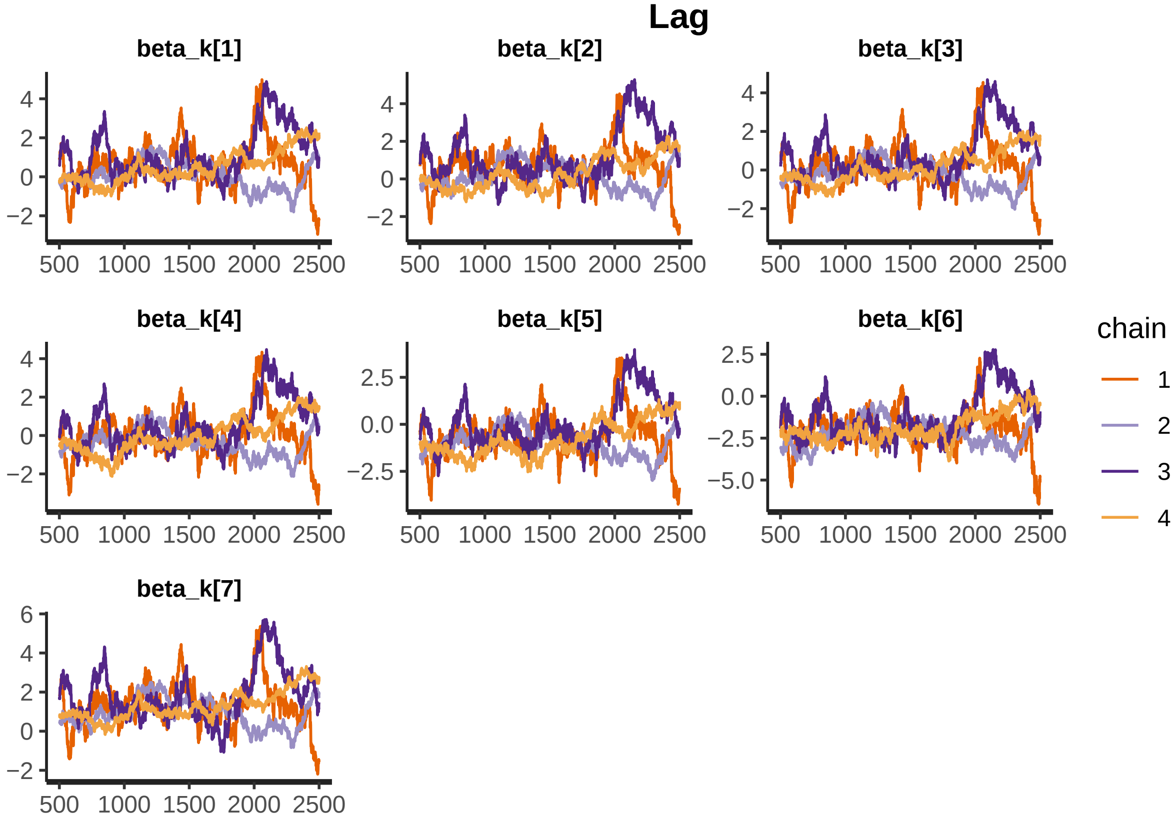
```
## sigmasq_dbz    39.8788090 2.1139701562 31.39568350 21.3567307 31.2375209
## sigmasq_i[1,1]  1.9635347 0.0050962181  0.11909019  1.8784019  1.9581802
## sigmasq_i[1,2]  2.0225138 0.0050488515  0.11029786  1.9461641  2.0149070
##                       75%      n_eff      Rhat
## sigma_bk         2.854211 144.616124 1.043225
## sigma_dbz        6.888012 252.546806 1.025210
## sigma_k[1]       1.447135 479.792180 1.014581
## sigma_k[2]       1.204197 595.654267 1.018949
## sigma_k[3]       1.429136 550.957696 1.006785
## sigma_k[4]       1.274738 608.151769 1.001997
## sigma_k[5]       1.323492 407.643140 1.009212
## sigma_k[6]       1.156003 672.246823 1.017023
## sigma_k[7]       1.172862 381.882606 1.014347
## sigma_uvq        1.018358 211.377564 1.016084
## sigma_w          8.501731   4.872488 1.356265
## sigma_w0         2.375399  77.492317 1.064715
## sigmasq_bk       8.146523 185.070292 1.035210
## sigmasq_dbz     47.444708 220.567835 1.028101
## sigmasq_i[1,1]   2.042429 546.079644 1.006833
## sigmasq_i[1,2]   2.094199 477.253371 1.014647
```





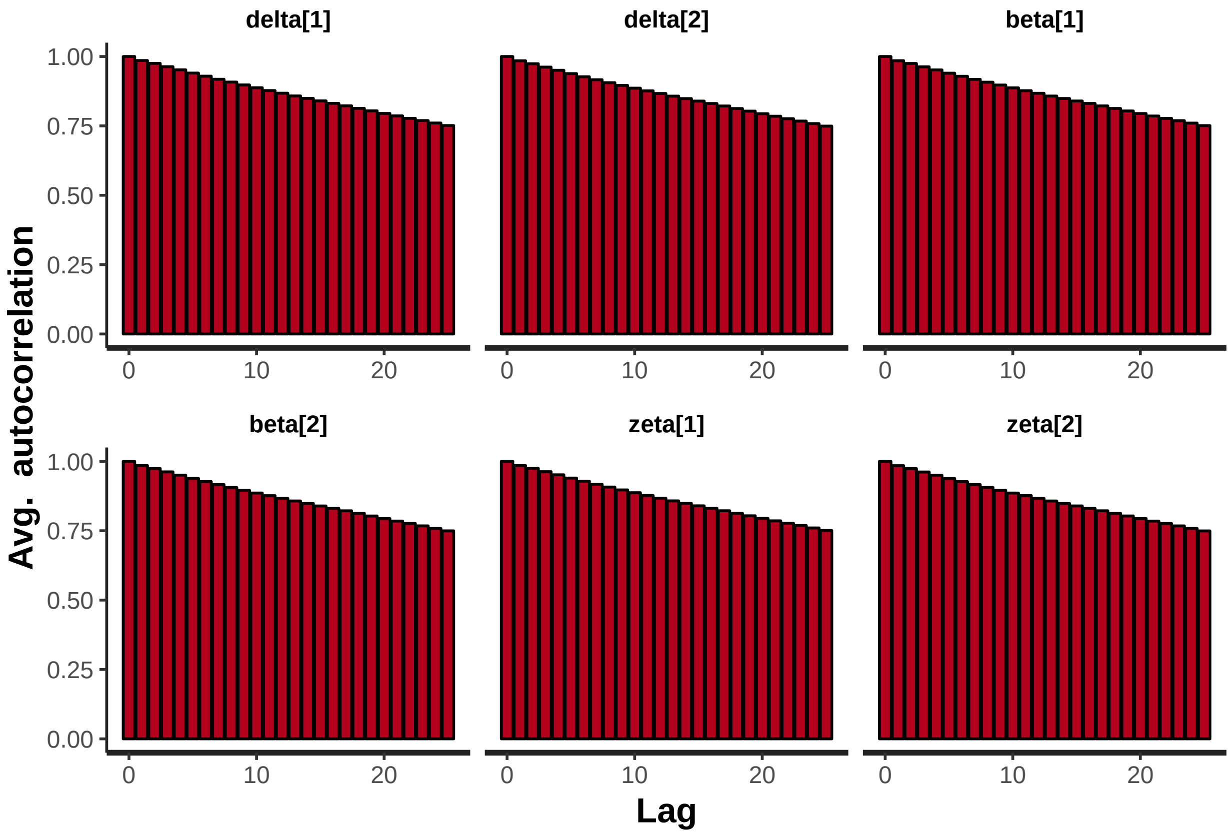

```
##                   mean   se_mean       sd         25%         50%
## beta_k[1]  0.60663102 0.2254949 1.080864 -0.05078121  0.43940966
## beta_k[2]  0.58117211 0.2459223 1.115348 -0.15784815  0.46051203
## beta_k[3]  0.26731812 0.2210444 1.072858 -0.37135223  0.09079881
## beta_k[4]  0.06272898 0.2230754 1.077985 -0.58765952 -0.11643326
## beta_k[5] -0.58534963 0.2350387 1.117956 -1.27800386 -0.72926633
## beta_k[6] -1.79848967 0.2247256 1.120884 -2.47404523 -1.96678688
## beta_k[7]  1.36332987 0.1885540 1.063390  0.74432101  1.21543420
##                    75%    n_eff     Rhat
## beta_k[1]  1.10741816 22.97568 1.191252
## beta_k[2]  1.11529528 20.56956 1.195851
## beta_k[3]  0.74845753 23.55730 1.187421
## beta_k[4]  0.59671414 23.35189 1.207362
## beta_k[5] -0.04239262 22.62409 1.205953
## beta_k[6] -1.25279438 24.87804 1.147036
## beta_k[7]  1.88370991 31.80631 1.112549
```

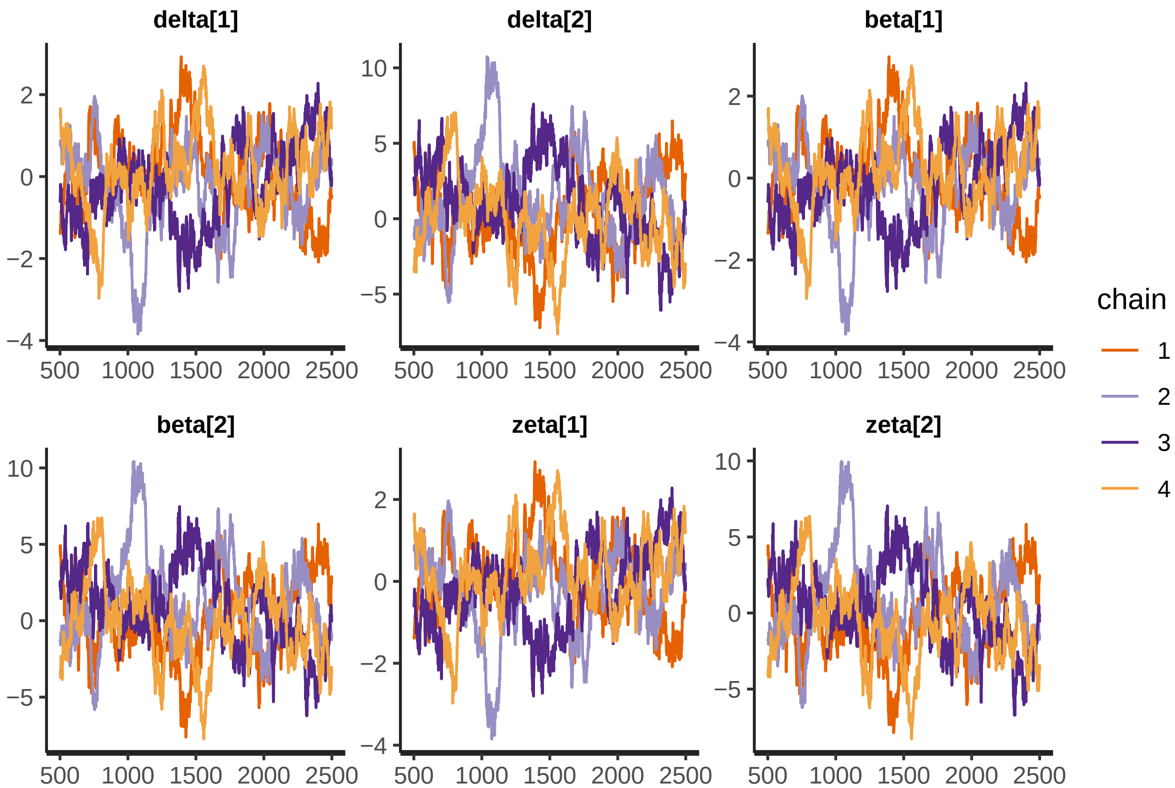




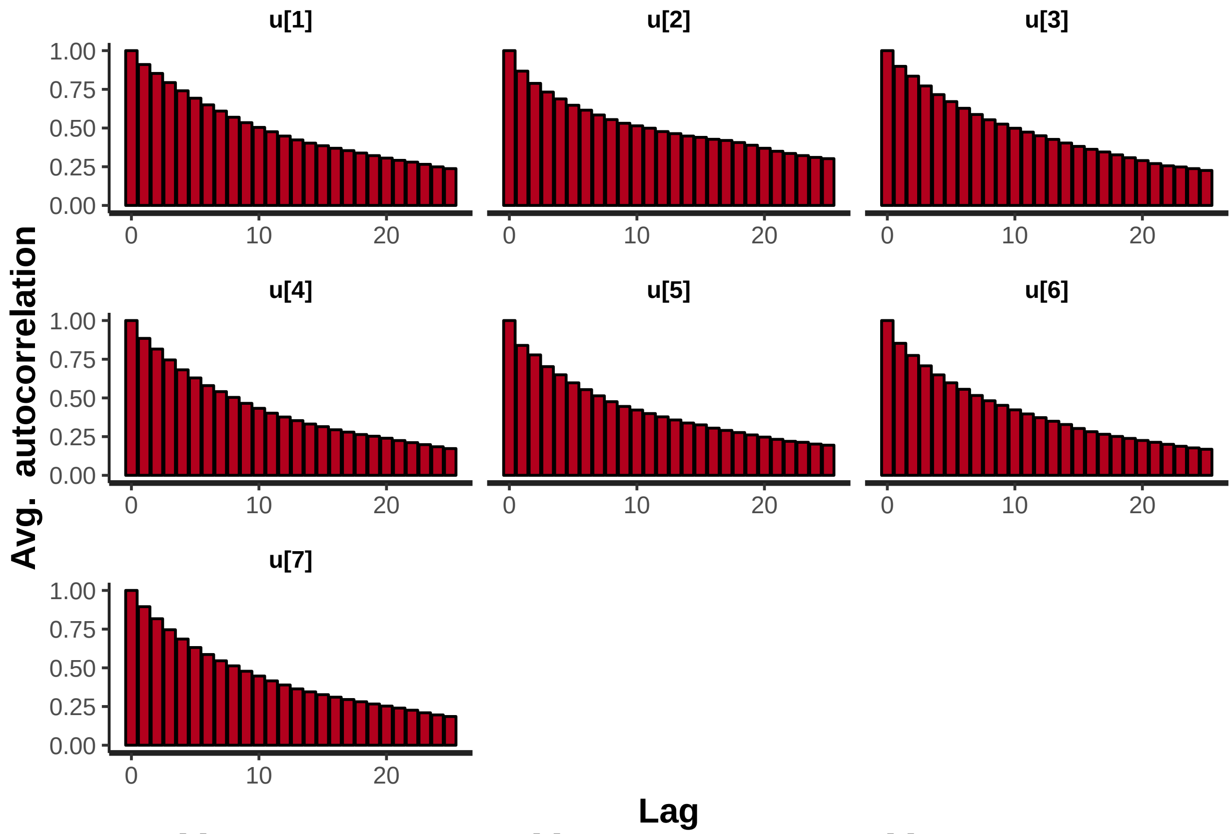

```
##                  mean   se_mean       sd        25%          50%       75%
## delta[1] -0.08947690 0.1517079 1.006301 -0.7054156 -0.041265468 0.5532885
## delta[2]  0.71835777 0.4069848 2.721174 -1.0089058  0.584376865 2.4086653
```

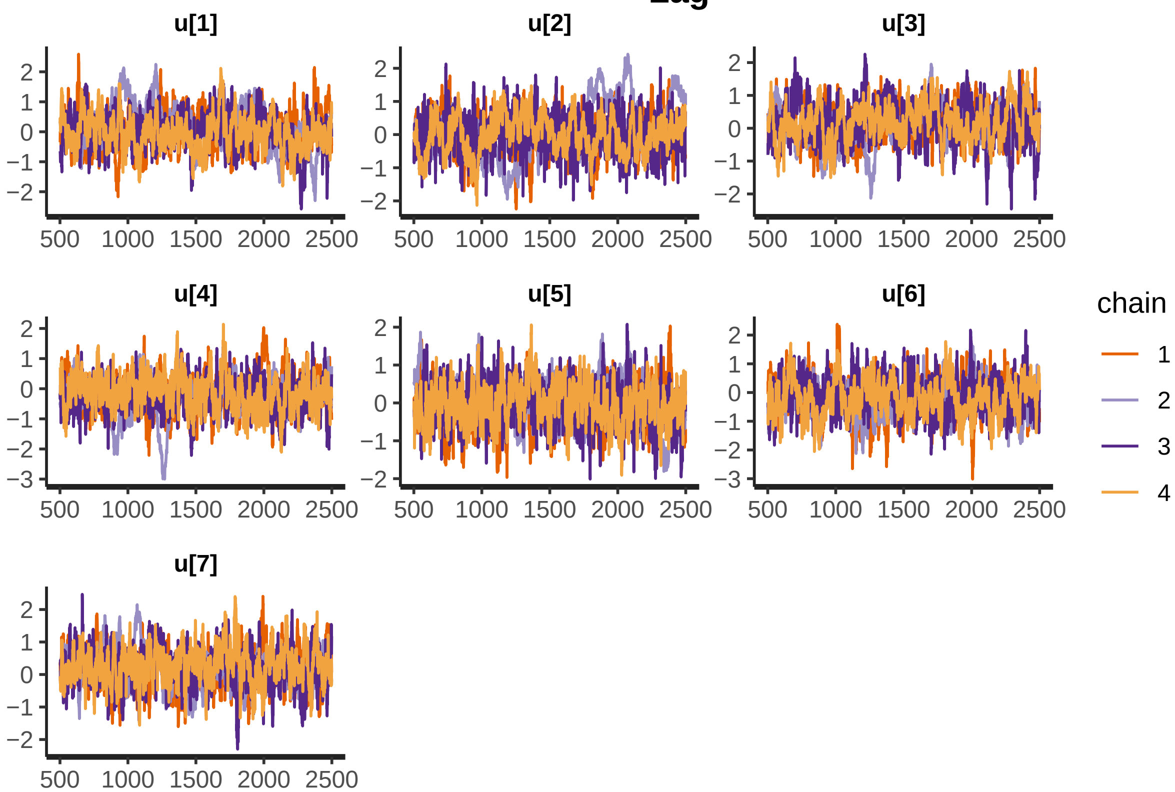
```
## beta[1]  -0.05073736 0.1517269 1.006597 -0.6671742 -0.002754029 0.5902232
## beta[2]   0.52070089 0.4072621 2.720567 -1.2099843  0.406408887 2.2047922
## zeta[1]  -0.08688189 0.1517785 1.006848 -0.7025336 -0.038609637 0.5545863
## zeta[2]   0.06429218 0.4074975 2.722354 -1.6564398 -0.056506268 1.7556404
##              n_eff     Rhat
## delta[1] 43.99863 1.095290
## delta[2] 44.70501 1.086566
## beta[1]  44.01354 1.095268
## beta[2]  44.62424 1.086925
## zeta[1]  44.00554 1.095218
## zeta[2]  44.63127 1.086730
```





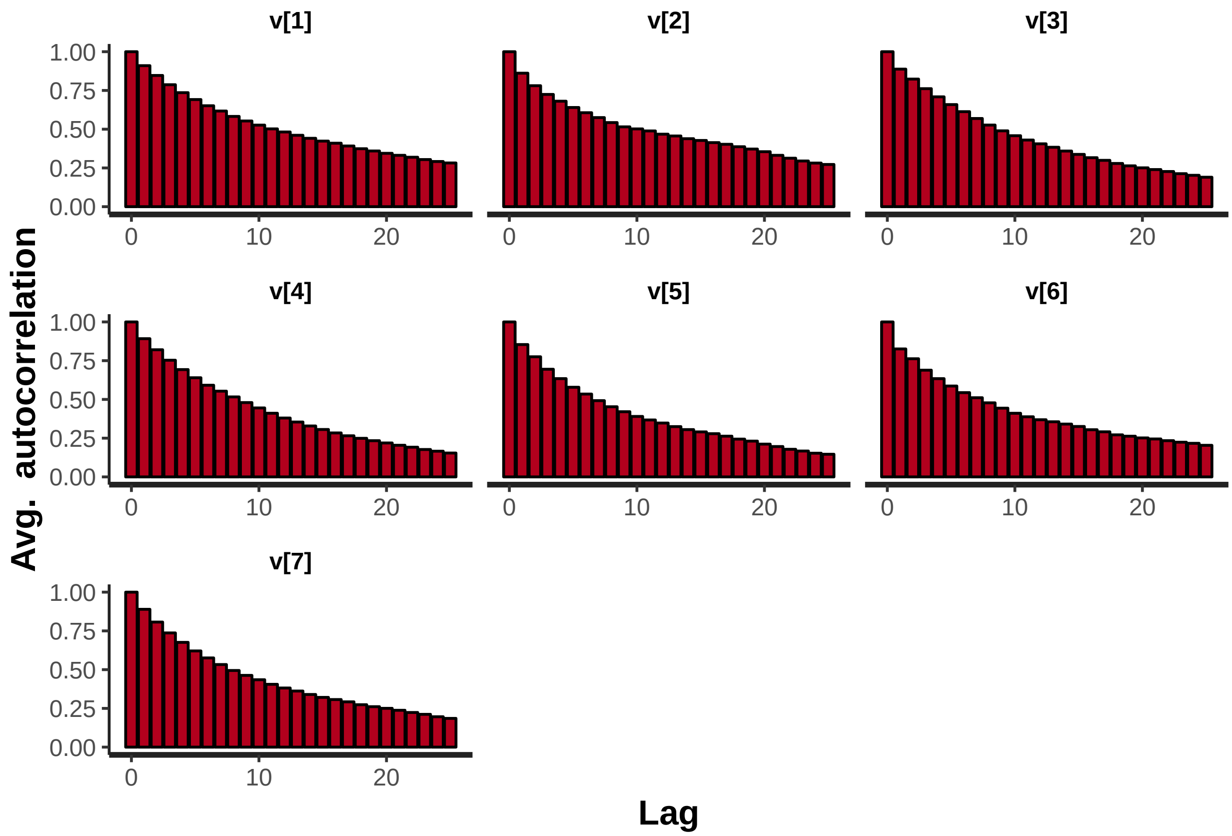

```
##              mean    se_mean        sd        25%         50%       75%
## u[1]  0.05894687 0.06391204 0.6344626 -0.3374245  0.05063563 0.4462591
## u[2]  0.04292392 0.08625491 0.6649740 -0.3840884  0.03223284 0.4475425
## u[3]  0.15229012 0.04423077 0.5878612 -0.2006122  0.16153999 0.5394425
## u[4] -0.20204250 0.04062363 0.6128957 -0.5664323 -0.19091926 0.1907404
## u[5] -0.02953950 0.04572176 0.5851756 -0.4151800 -0.02113300 0.3618943
## u[6] -0.19155551 0.03775252 0.6389671 -0.5997081 -0.19059230 0.2251249
## u[7]  0.18113800 0.04118032 0.6023605 -0.2189656  0.18249555 0.5679449
##          n_eff     Rhat
## u[1]  98.54772 1.036585
## u[2]  59.43492 1.091858
## u[3] 176.64475 1.039305
## u[4] 227.62273 1.023596
## u[5] 163.80474 1.023313
## u[6] 286.46071 1.006339
## u[7] 213.96052 1.015821
```

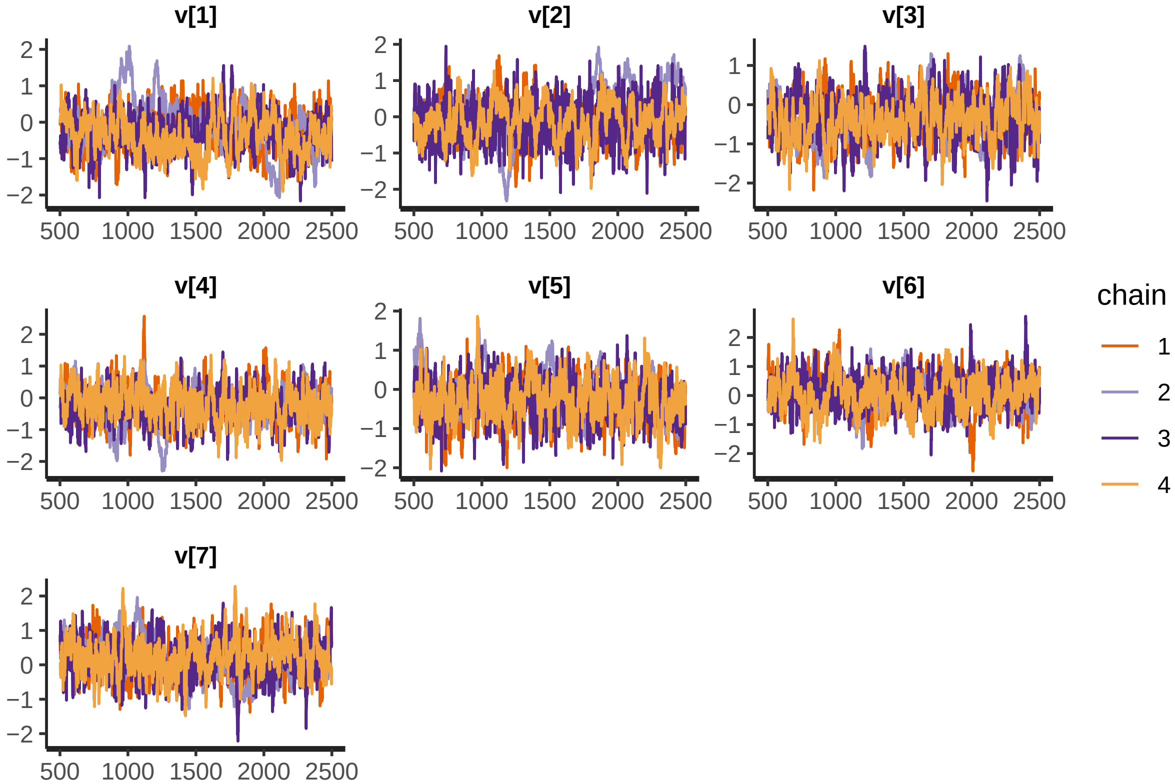




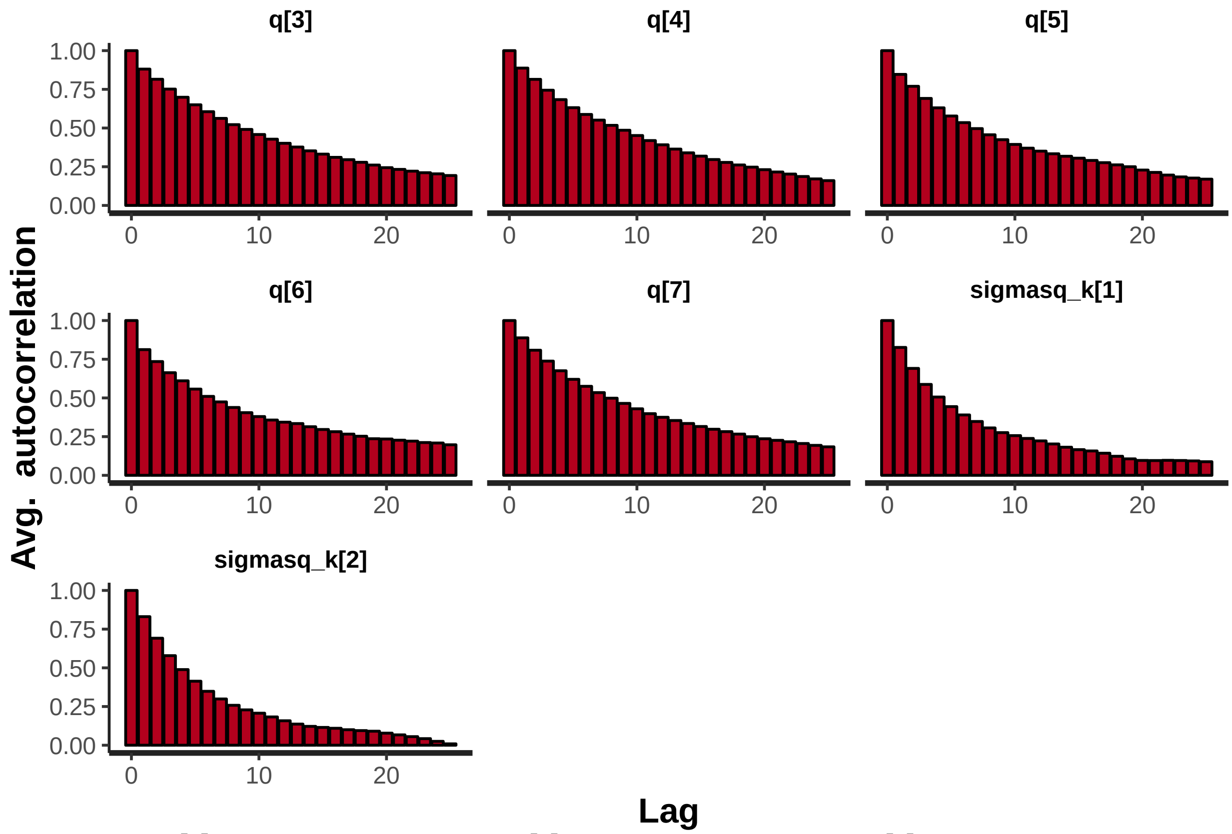

```
##              mean    se_mean        sd        25%         50%          75%
## v[1] -0.23986093 0.07173808 0.5860987 -0.6242038 -0.25901152  0.126465879
## v[2] -0.07468972 0.07551231 0.6041355 -0.4572723 -0.08588357  0.316170433
```

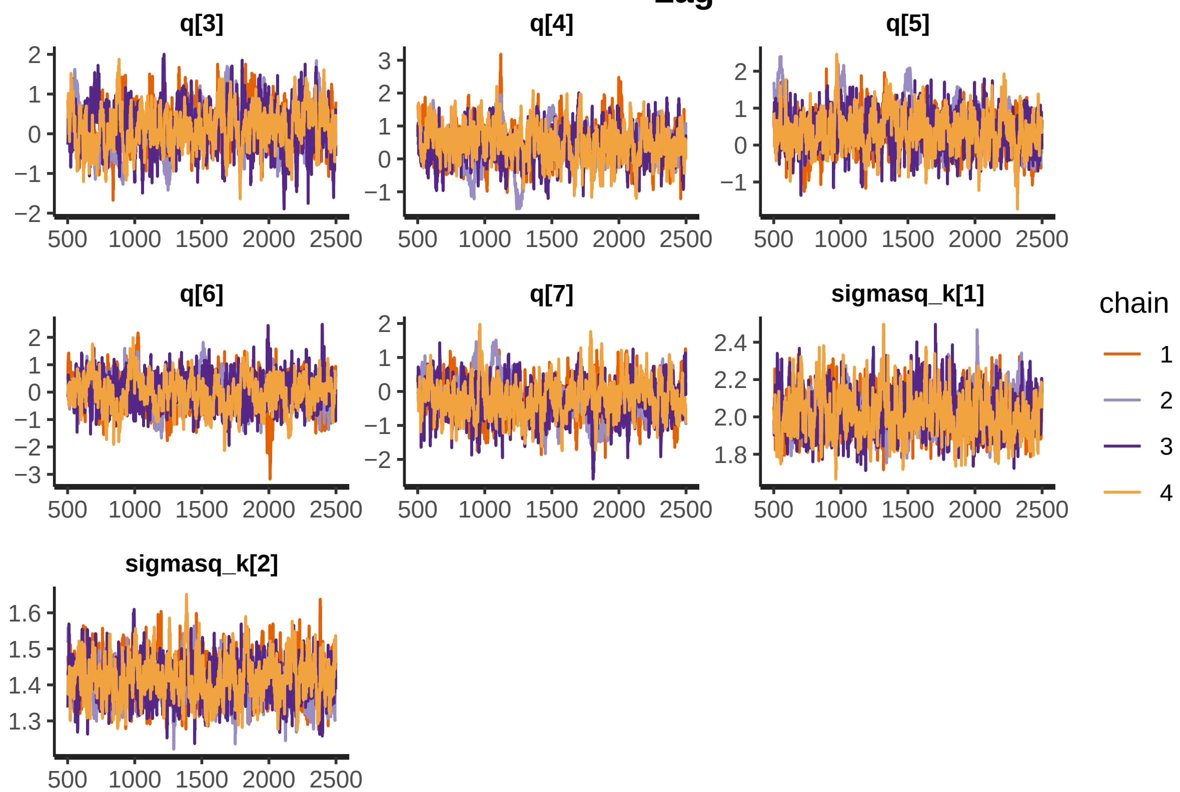
```
## v[3] -0.37858539 0.03676934 0.5504277 -0.7617850 -0.38226727 -0.009004447
## v[4] -0.23645453 0.03698296 0.5584609 -0.5837295 -0.23594863  0.120856120
## v[5] -0.25748499 0.03572365 0.5325335 -0.6292446 -0.26075410  0.083665229
## v[6]  0.08476178 0.03378280 0.5743423 -0.2871496  0.08592130  0.455135921
## v[7]  0.16256767 0.04150428 0.5476275 -0.2046696  0.15269990  0.517674398
##          n_eff     Rhat
## v[1]  66.74859 1.070905
## v[2]  64.00784 1.084411
## v[3] 224.09338 1.019112
## v[4] 228.02483 1.017355
## v[5] 222.21953 1.033142
## v[6] 289.03492 1.003979
## v[7] 174.09437 1.028799
```





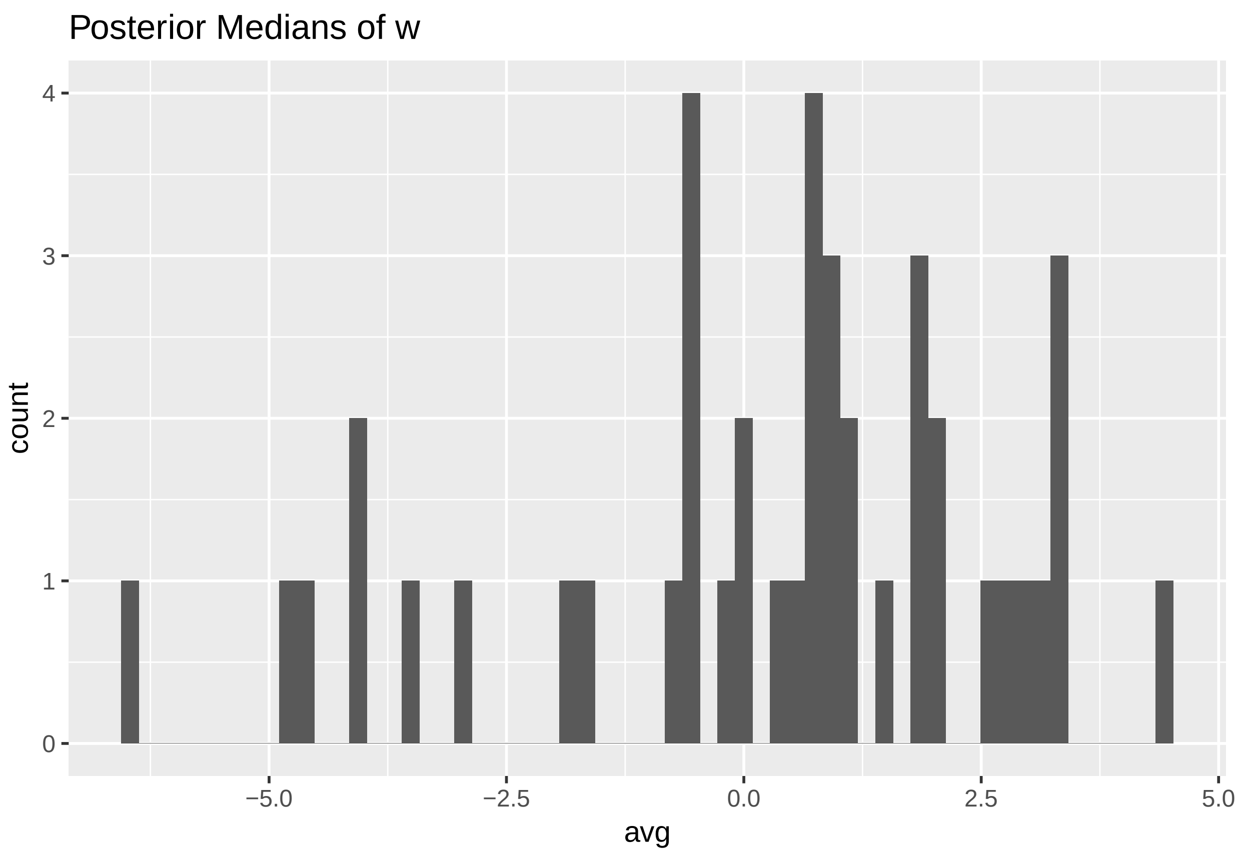

```
##                        mean      se_mean         sd         25%          50%
## q[3]           0.144512737 0.035571801 0.54802677 -0.22824400  0.146241688
## q[4]           0.447312174 0.035826799 0.55716215  0.10348836  0.443676579
## q[5]           0.389460631 0.038823956 0.54470674  0.01345793  0.372069192
## q[6]          -0.005286645 0.033831512 0.58895465 -0.37309908  0.007936894
## q[7]          -0.272254155 0.041797755 0.55097433 -0.63614955 -0.294573410
## sigmasq_k[1]   2.022513760 0.005048851 0.11029786  1.94616406  2.014906993
## sigmasq_k[2]   1.412425932 0.002325979 0.05679265  1.37238235  1.410465466
##                      75%    n_eff     Rhat
## q[3]          0.49805279 237.3514 1.018217
## q[4]          0.80124032 241.8506 1.020097
## q[5]          0.73476114 196.8457 1.031508
## q[6]          0.36748087 303.0546 1.005651
## q[7]          0.07859929 173.7628 1.028888
## sigmasq_k[1] 2.09419855 477.2534 1.014647
## sigmasq_k[2] 1.45009082 596.1737 1.018905
```



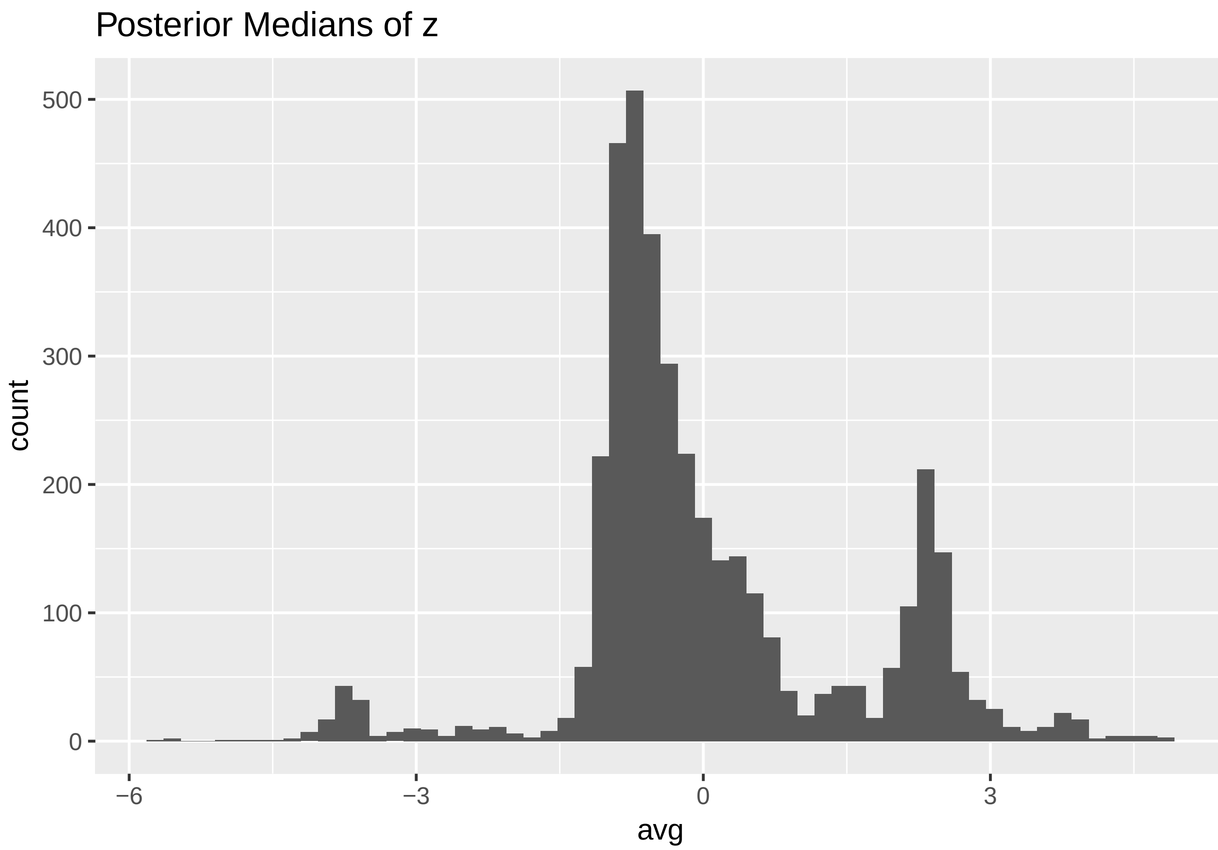



```
## [1] "Summary statistics for posterior medians of w"
##        avg
##  Min.   :-6.5442
```

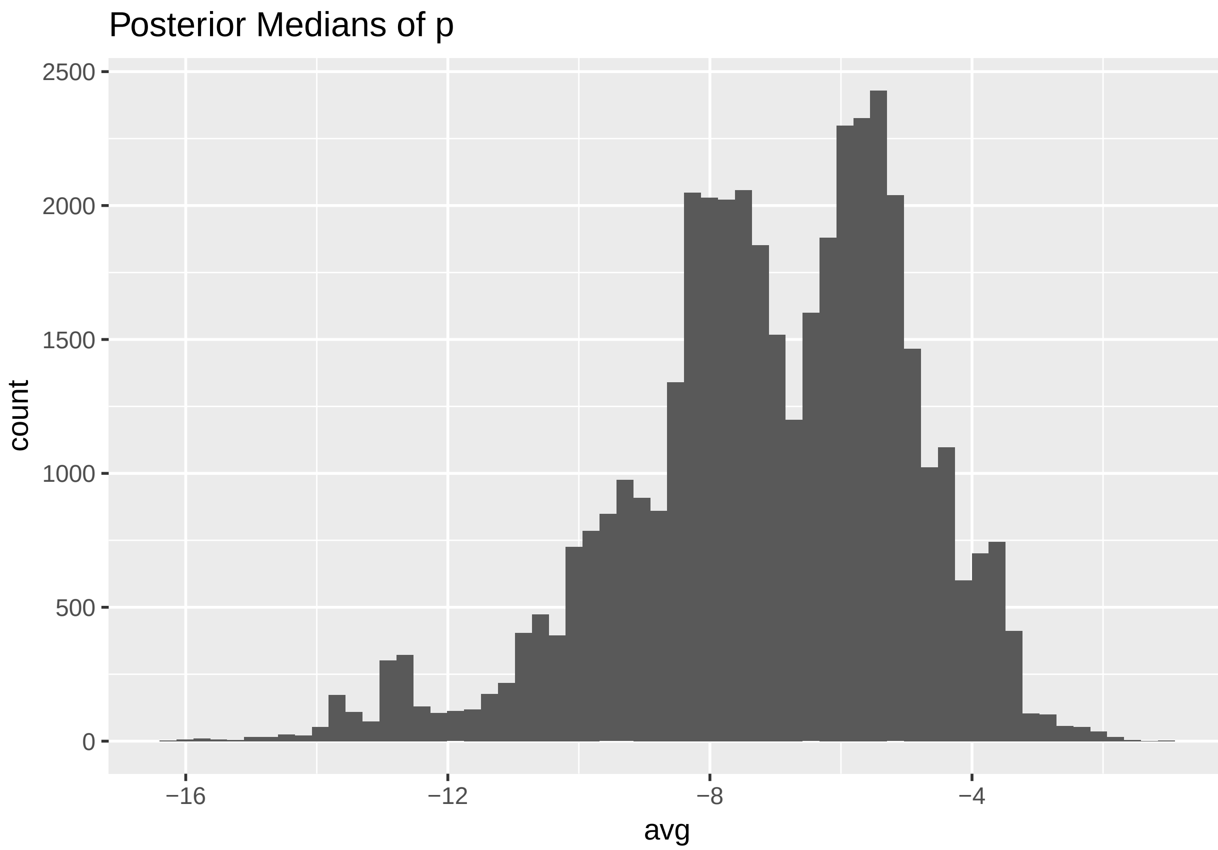
```
##  1st Qu.:-0.6335
##  Median : 0.7512
##  Mean   : 0.1932
##  3rd Qu.: 1.8451
##  Max.   : 4.3520
```



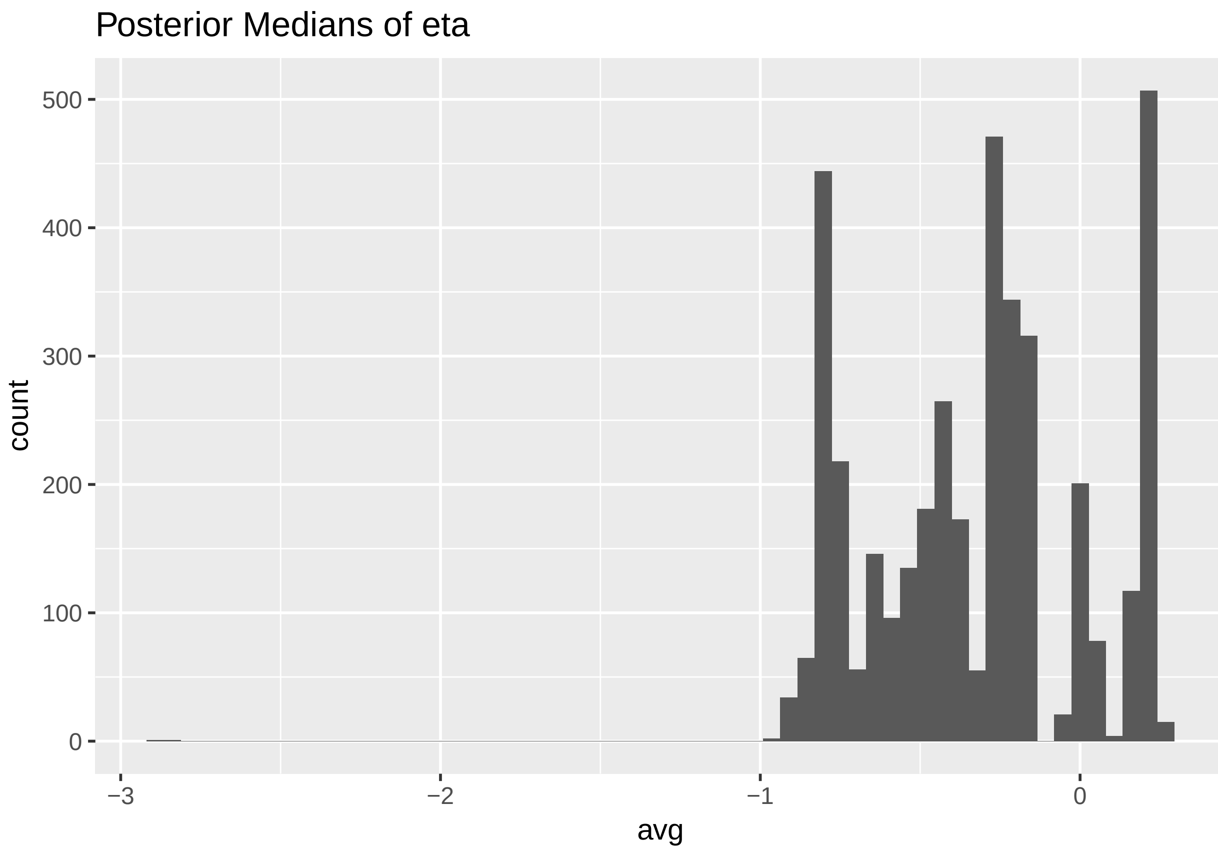

```
## [1] "Summary statistics for posterior medians of z"
##       avg
##  Min.   :-5.69091
##  1st Qu.:-0.79560
##  Median :-0.38091
##  Mean   : 0.05076
##  3rd Qu.: 0.64107
##  Max.   : 4.86801
```

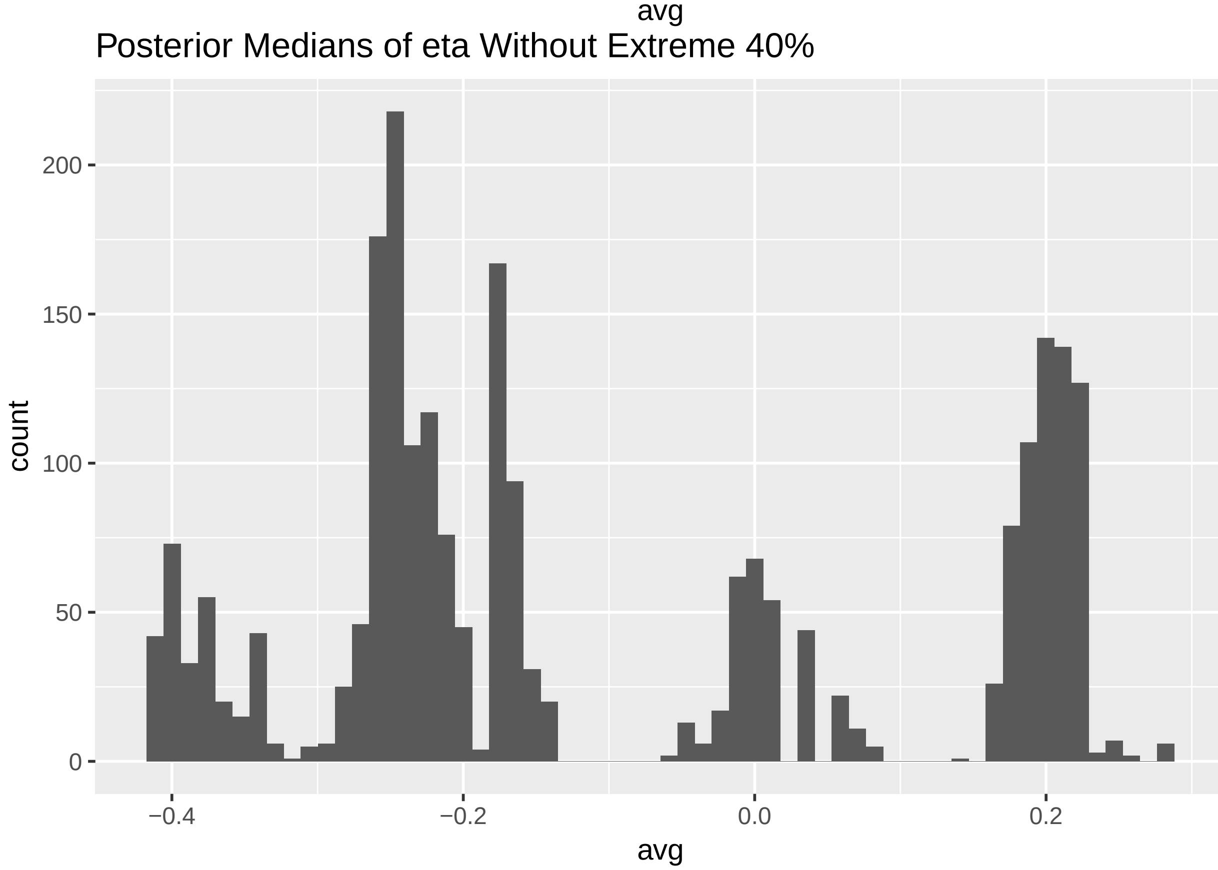



```
## [1] "Summary statistics for posterior medians of p"
##       avg
##  Min.   :-16.2050
##  1st Qu.: -8.3017
##  Median : -6.9538
##  Mean   : -7.0925
##  3rd Qu.: -5.4959
##  Max.   : -0.9664
```

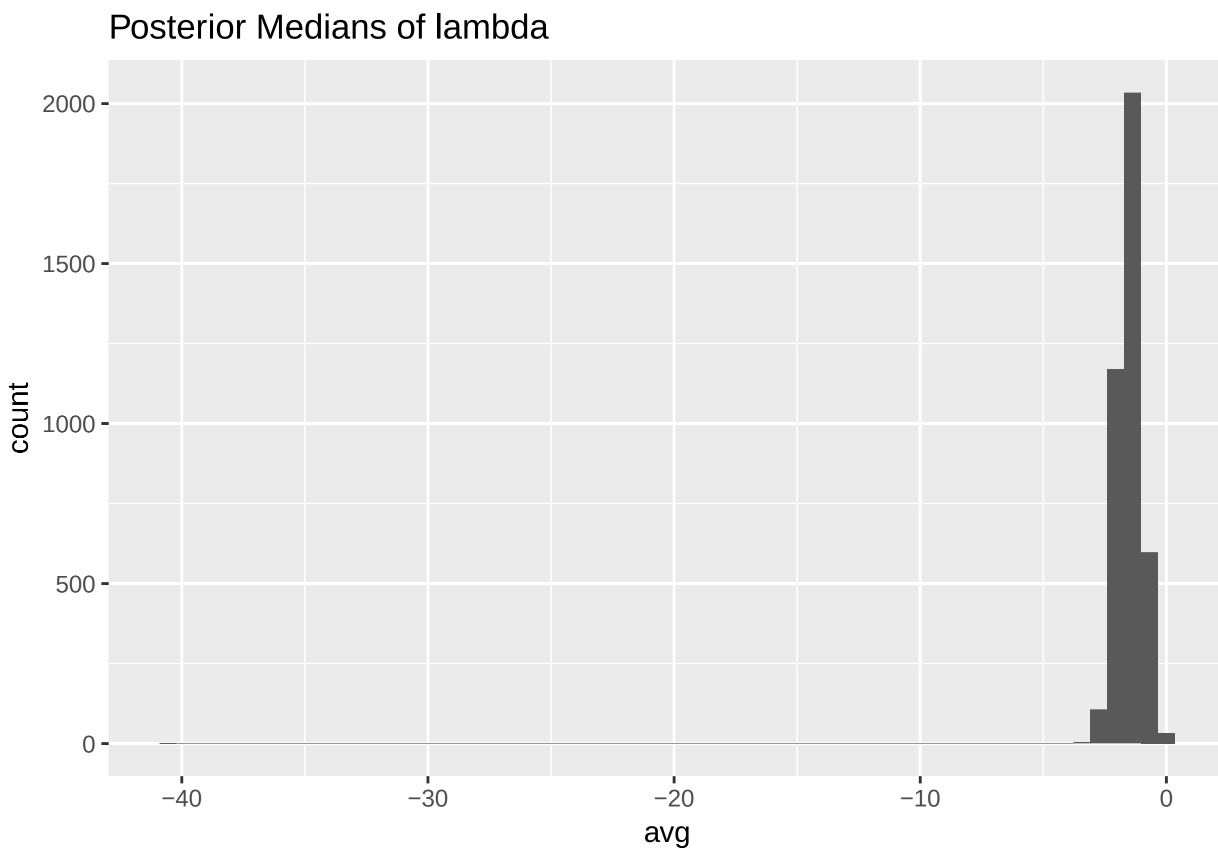



```
## [1] "Summary statistics for posterior medians of eta"
##       avg
##  Min.   :-2.8793
##  1st Qu.:-0.6049
##  Median :-0.2624
##  Mean   :-0.3248
##  3rd Qu.:-0.1571
##  Max.   : 0.2802
```

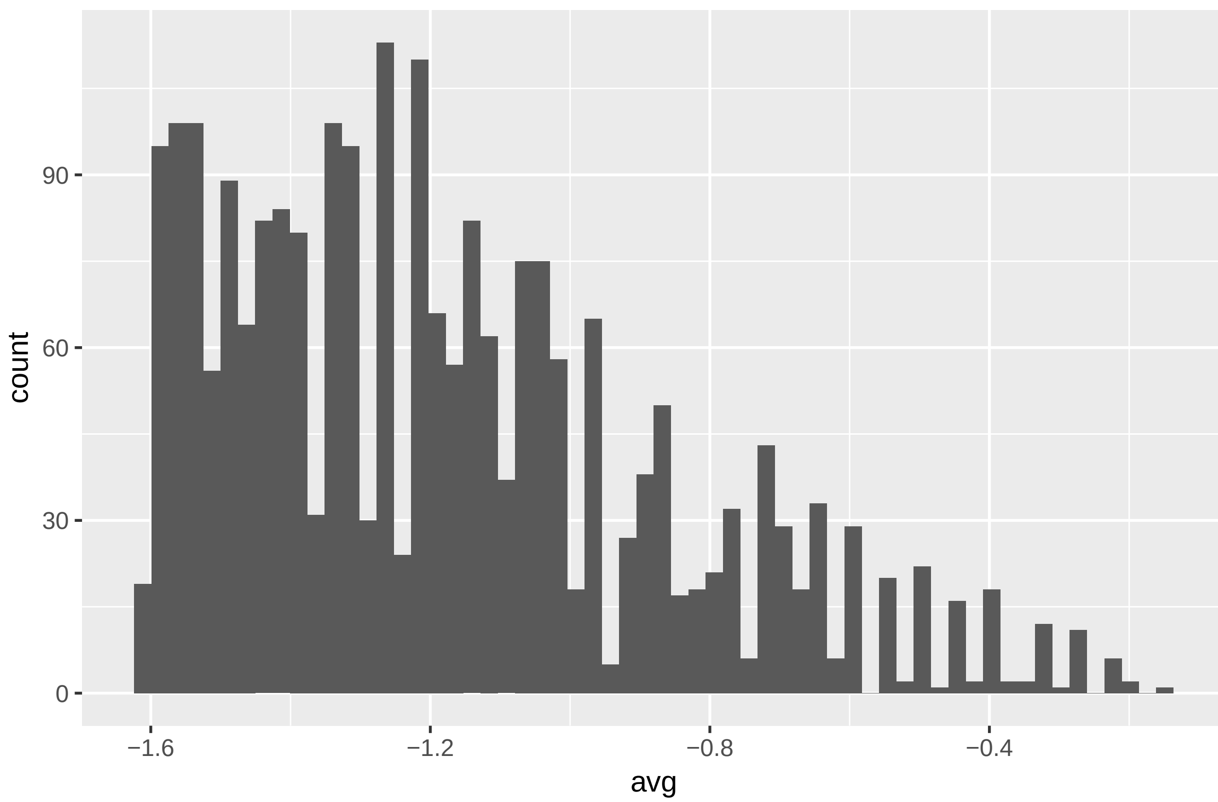

Posterior Medians of eta
count
500
400
300
200
100
0
−3
−2
−1
0
avg


Posterior Medians of eta Without Extreme 40%
count
200
150
100
50
0
−0.4
−0.2
0.0
0.2
avg

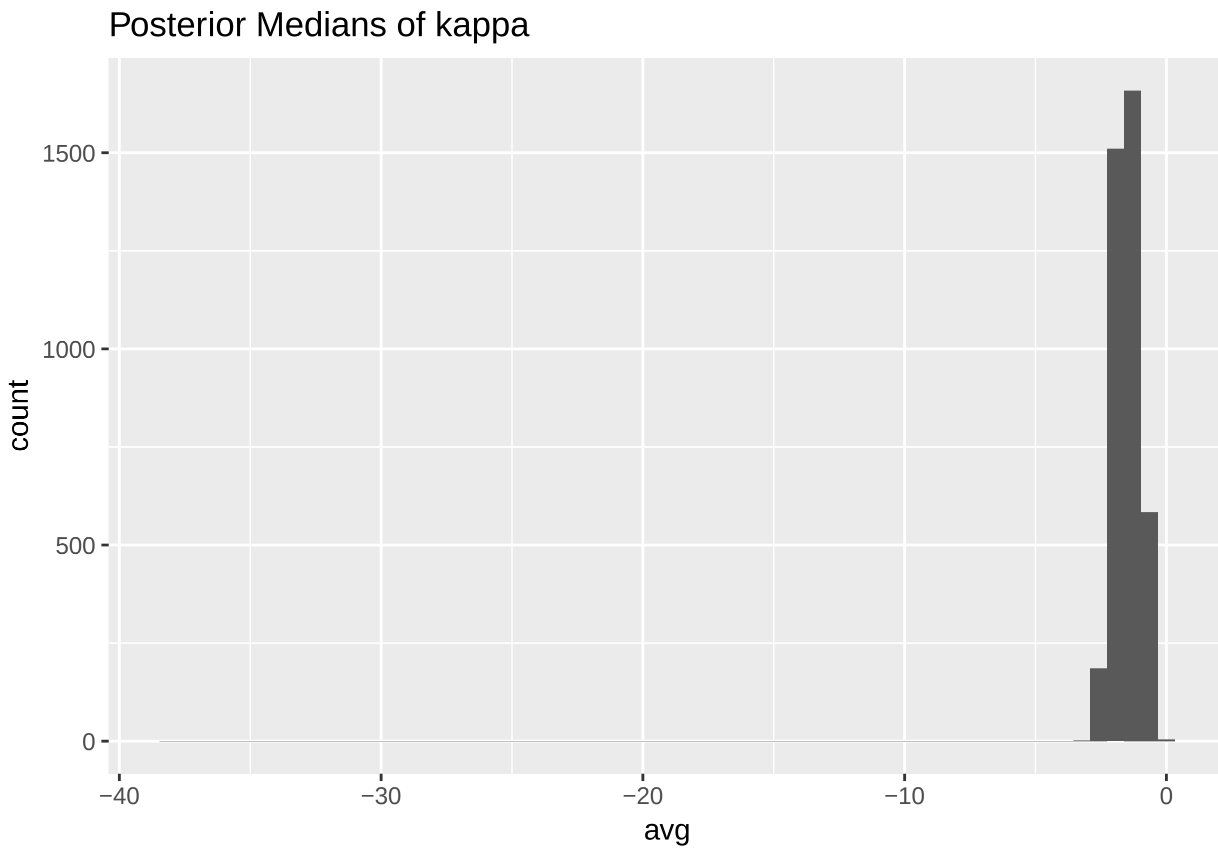
```
## [1] "Summary statistics for posterior medians of lambda"
##       avg
##  Min.   :-40.717
##  1st Qu.: -1.824
##  Median : -1.488
##  Mean   : -1.514
##  3rd Qu.: -1.168
##  Max.   : -0.155
```

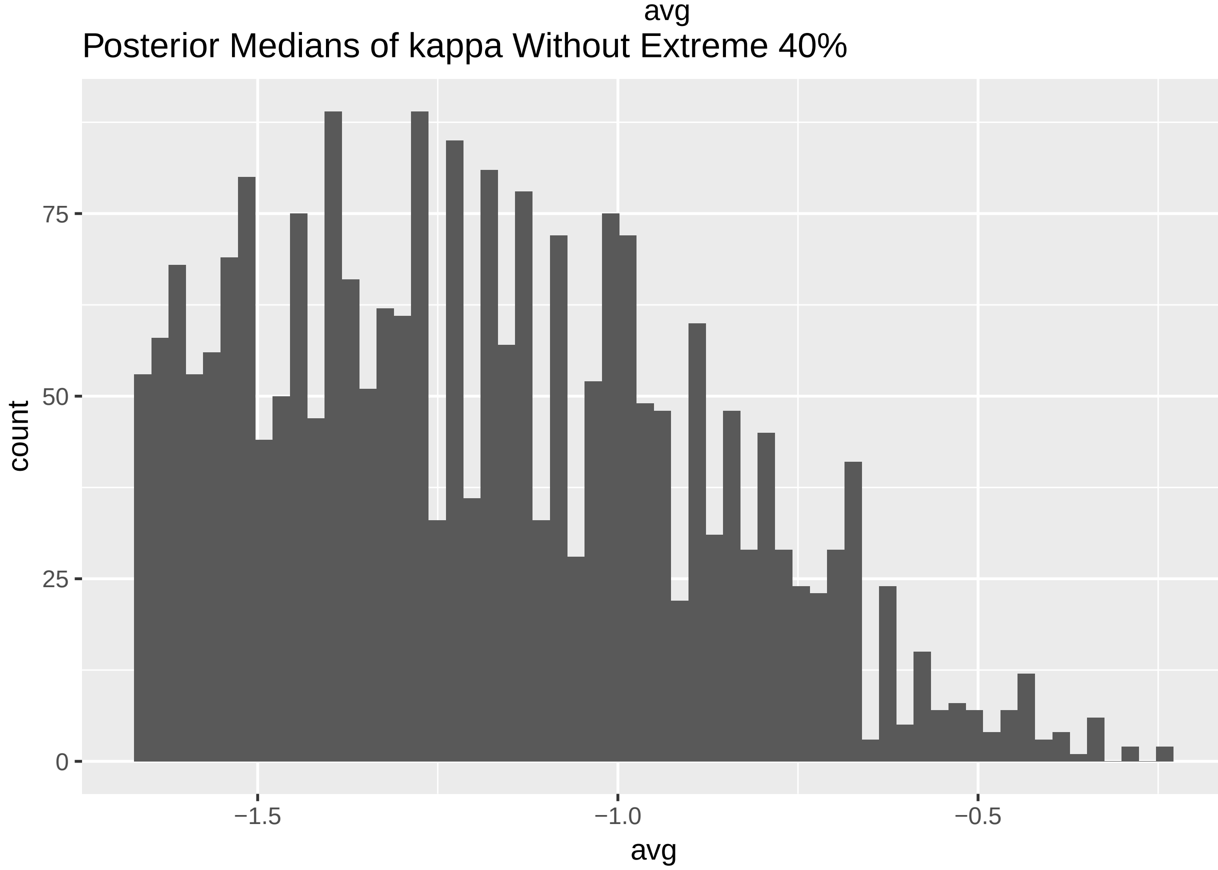


Posterior Medians of lambda Without Extreme 40%




```
## [1] "Summary statistics for posterior medians of kappa"
##       avg
##  Min.   :-38.3910
##  1st Qu.: -1.8758
##  Median : -1.5206
##  Mean   : -1.5280
##  3rd Qu.: -1.1386
##  Max.   : -0.2473
```


Posterior Medians of kappa
count
1500
1000
500
0
−40
−30
−20
−10
0
avg


Posterior Medians of kappa Without Extreme 40%
count
75
50
25
0
−1.5
−1.0
−0.5
avg

Identifying Parameters with Large Rhats

```
summary(fit_summ$Rhat)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.9999  1.0189  1.0503  1.1002  1.0957  4.4684
```

```
big_Rhat <- fit_summ$Rhat > 5
big_Rhat_dat <- fit_summ[big_Rhat,c(1,2,10)]
big_Rhat_dat
```

```
## [1] mean    se_mean Rhat
## <0 rows> (or 0-length row.names)
```