# MCMC Diagnostics - IFLS data
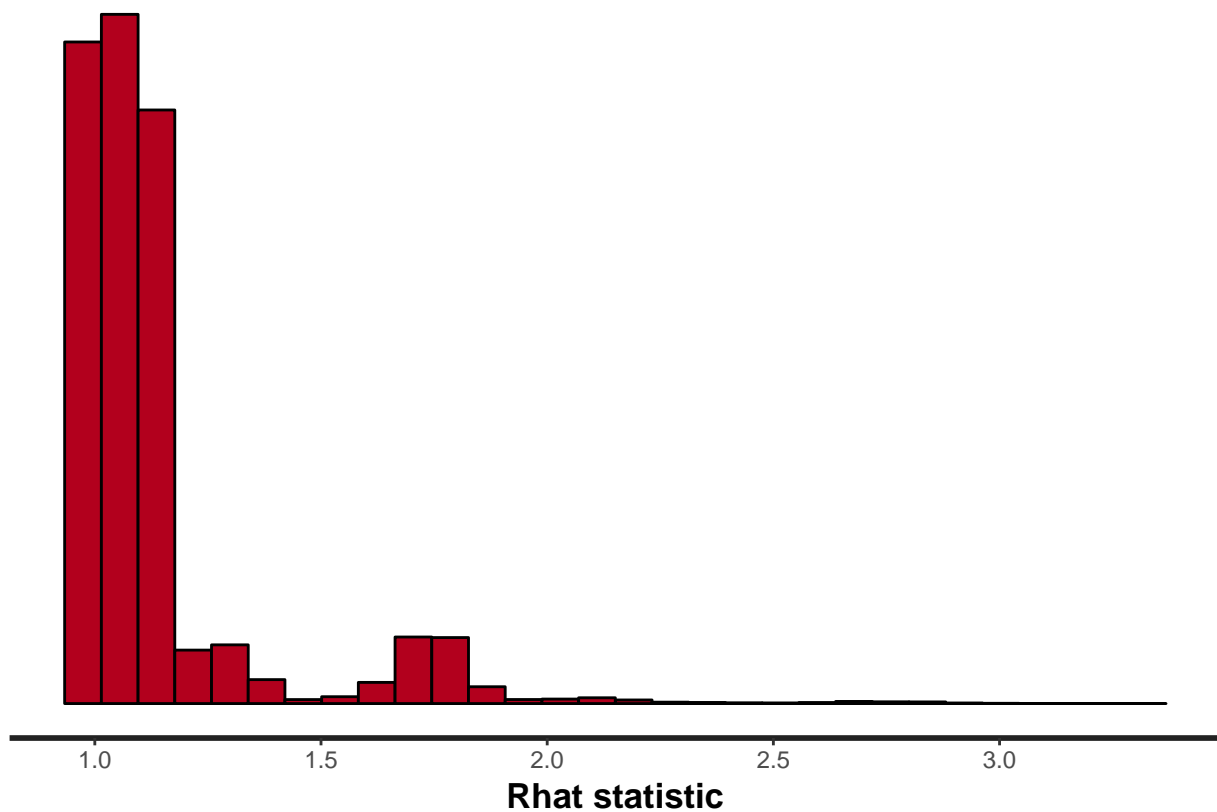
*Sarah Teichman*

*04/19/2020*

```
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`.
```

```
## Warning: Removed 1 rows containing non-finite values (stat_bin).
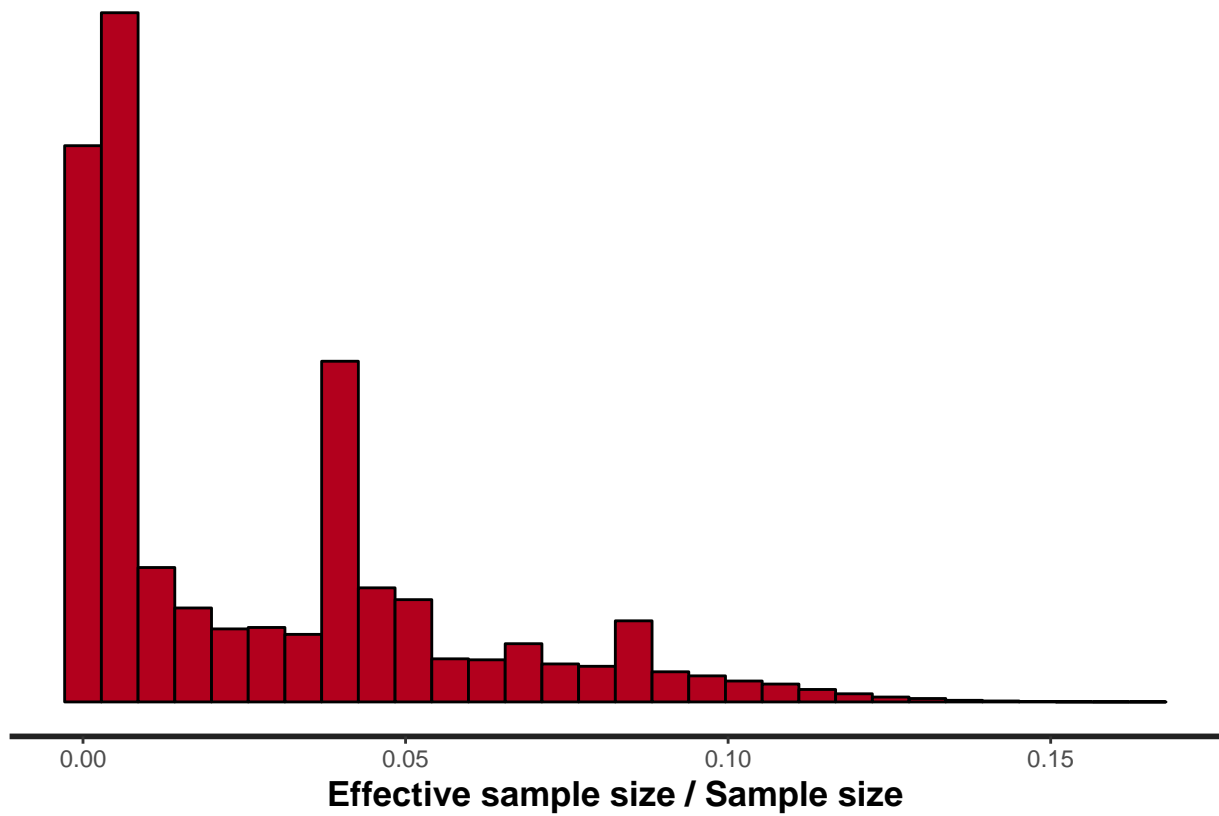```
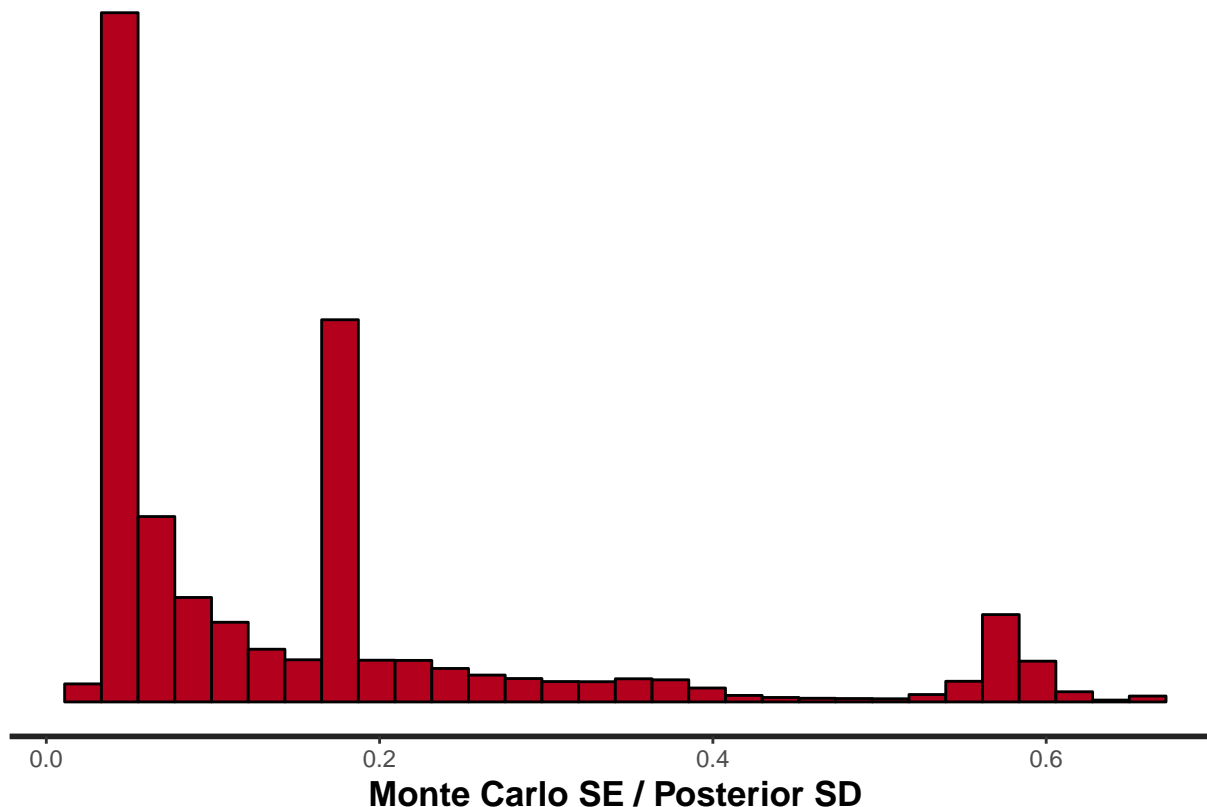


```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 1 rows containing non-finite values (stat_bin).
```

**Effective sample size / Sample size**

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

## Warning: Removed 1 rows containing non-finite values (stat_bin).
```

**Monte Carlo SE / Posterior SD**

## 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)
```

```r
  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       2.9379111 0.0548042319  0.98543641   2.2532452   2.7315117
## sigma_dbz      5.9618956 0.1531721374  2.10201450   4.6028339   5.5179874
## sigma_k[1]     1.4201093 0.0019071867  0.03762082   1.3945737   1.4177880
## sigma_k[2]     1.1903085 0.0012987731  0.02359614   1.1736554   1.1892611
## sigma_k[3]     1.3975246 0.0021592681  0.04145671   1.3692684   1.3961222
## sigma_k[4]     1.2485897 0.0013705865  0.03611541   1.2229964   1.2476776
## sigma_k[5]     1.2866408 0.0024396985  0.05782873   1.2457468   1.2854248
## sigma_k[6]     1.1166341 0.0031188263  0.05790421   1.0769499   1.1141701
## sigma_k[7]     1.1617929 0.0009474342  0.01737274   1.1499946   1.1615907
## sigma_uvq      0.8966181 0.0094373651  0.17621206   0.7718418   0.8742251
## sigma_w        7.8724395 0.5969345280  1.62320691   6.7586328   7.5695920
## sigma_w0       2.0459986 0.0675797532  0.49314753   1.6857170   1.9742228
## sigmasq_bk     9.6022854 0.4194648291  7.82248485   5.0771141   7.4611563
## sigmasq_dbz   39.9621113 2.6900754420 39.15072748  21.1860797  30.4481846
## sigmasq_i[1,1] 1.9547935 0.0060394995  0.11620463   1.8748960   1.9491571
## sigmasq_i[1,2] 2.0181254 0.0054346815  0.10724044   1.9448357   2.0101228
##                     75%       n_eff      Rhat
## sigma_bk        3.3987291 323.317354 1.017509
## sigma_dbz       6.7675884 188.326689 1.026977
## sigma_k[1]      1.4447160 389.107872 1.006147
## sigma_k[2]      1.2066920 330.076985 1.013662
## sigma_k[3]      1.4244575 368.617973 1.006945
## sigma_k[4]      1.2720098 694.340187 1.006795
## sigma_k[5]      1.3252109 561.843108 1.006017
## sigma_k[6]      1.1538031 344.697268 1.003386
## sigma_k[7]      1.1735012 336.231418 1.003518
## sigma_uvq       0.9985709 348.634014 1.007525
## sigma_w         8.6030632   7.394254 1.361811
## sigma_w0        2.3193463  53.250111 1.082235
## sigmasq_bk     11.5513599 347.774998 1.016537
## sigmasq_dbz    45.8002529 211.812094 1.023880
## sigmasq_i[1,1]  2.0290793 370.207297 1.006901
## sigmasq_i[1,2]  2.0872042 389.375862 1.006095
```
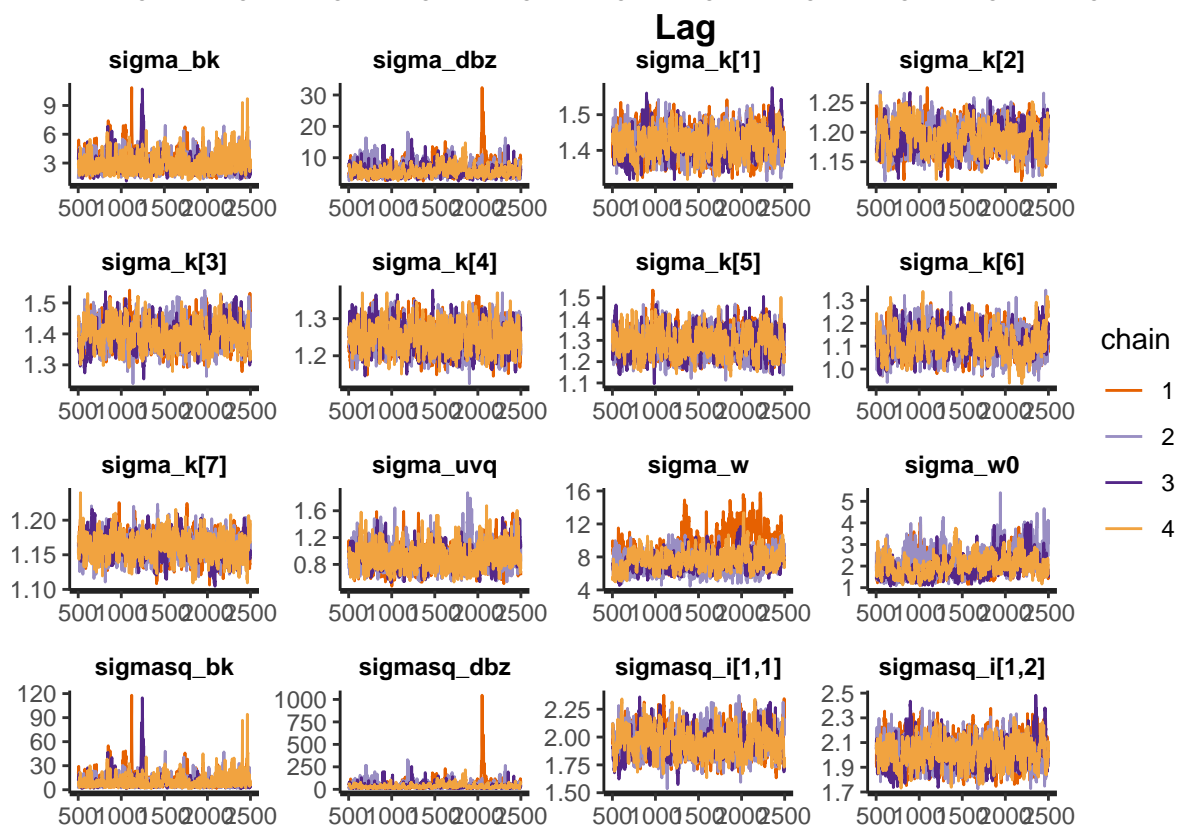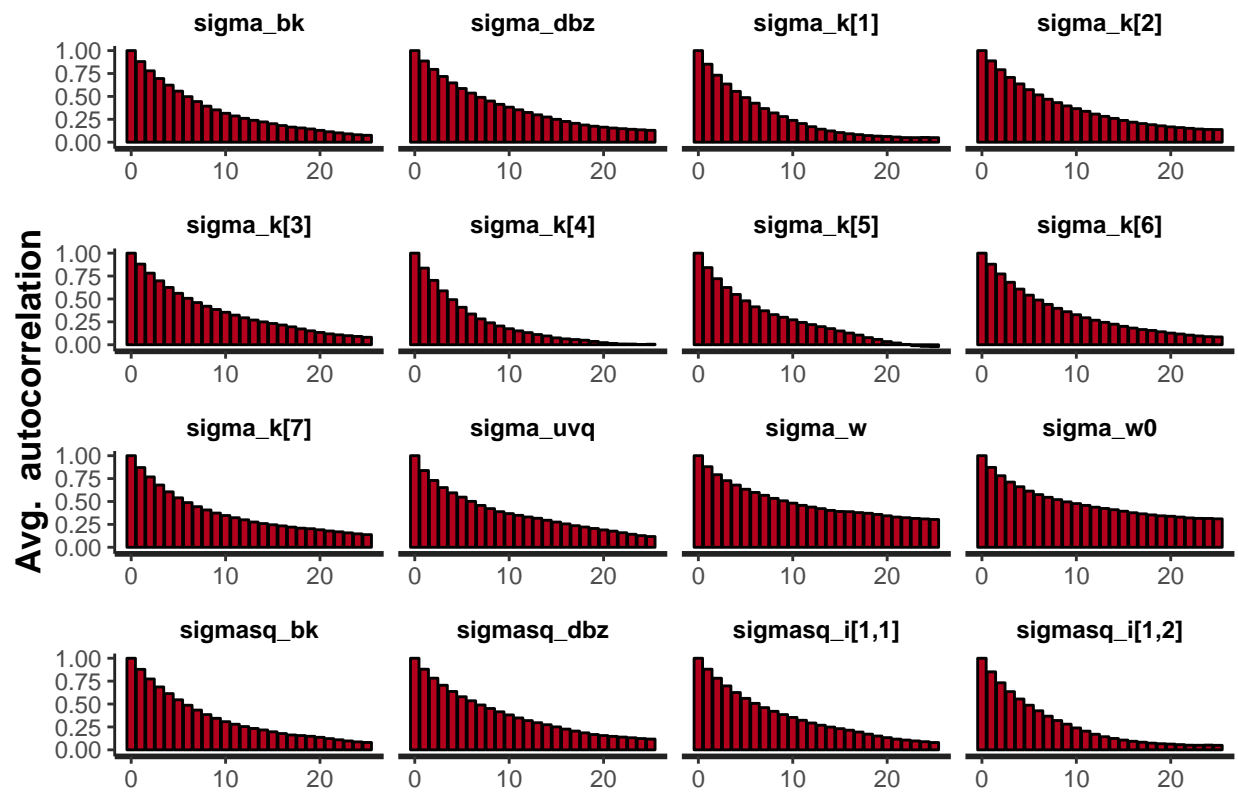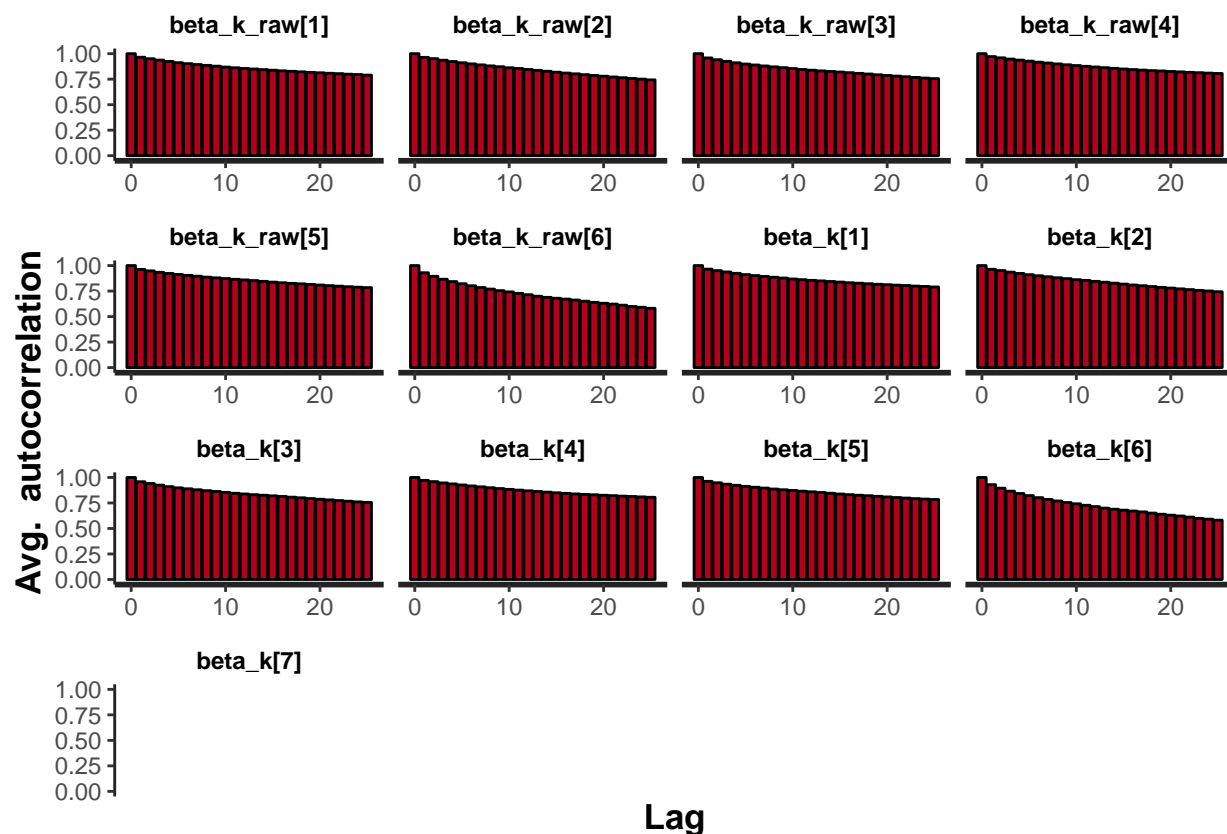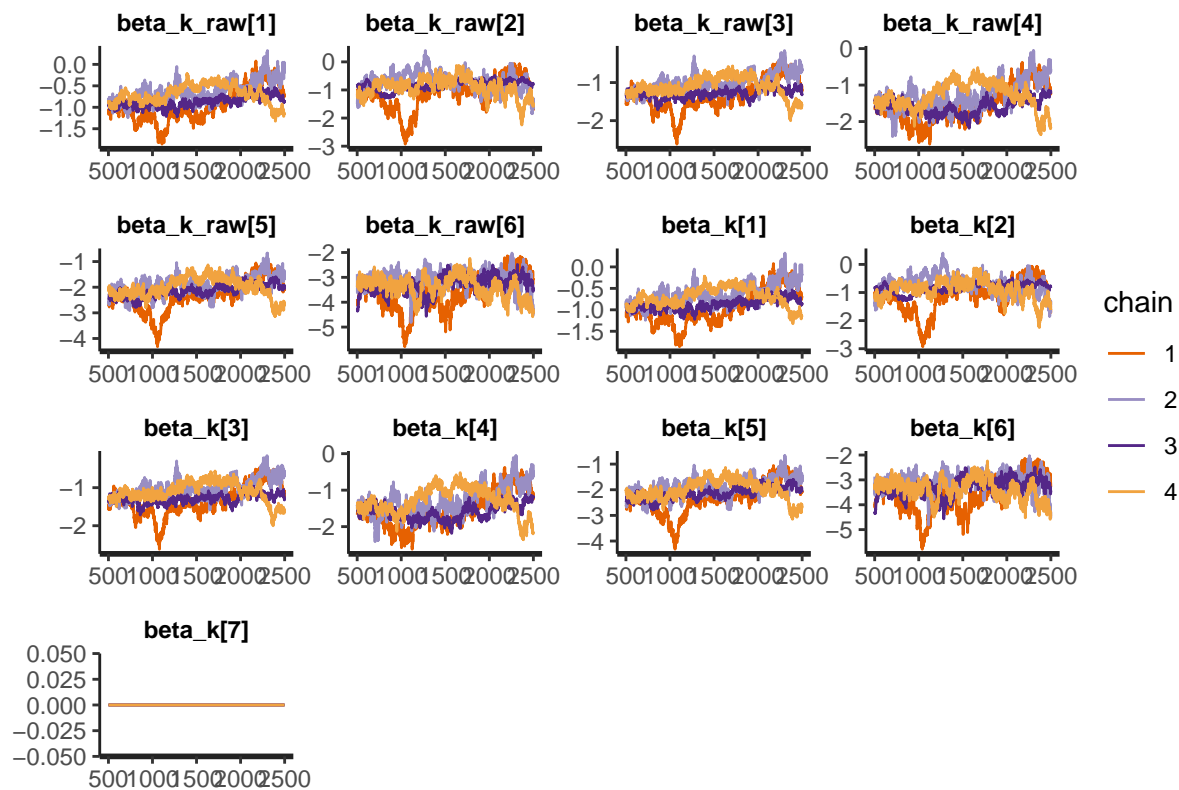
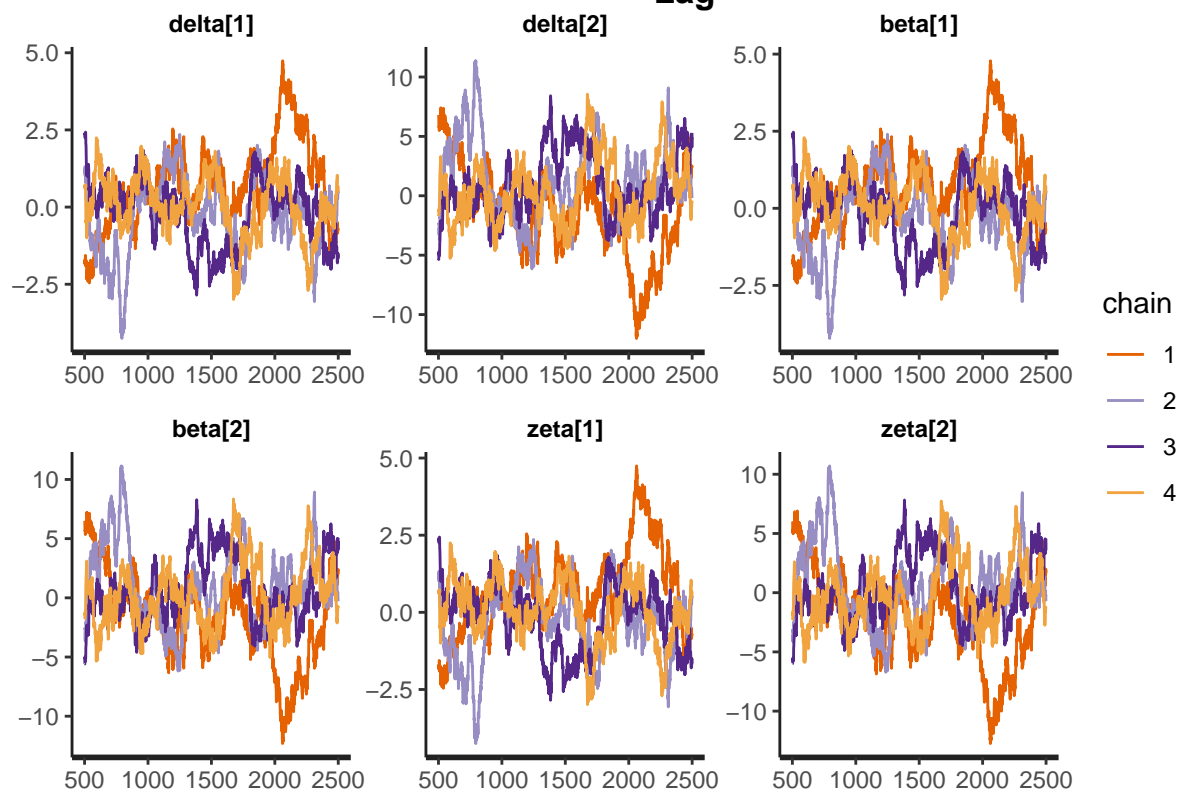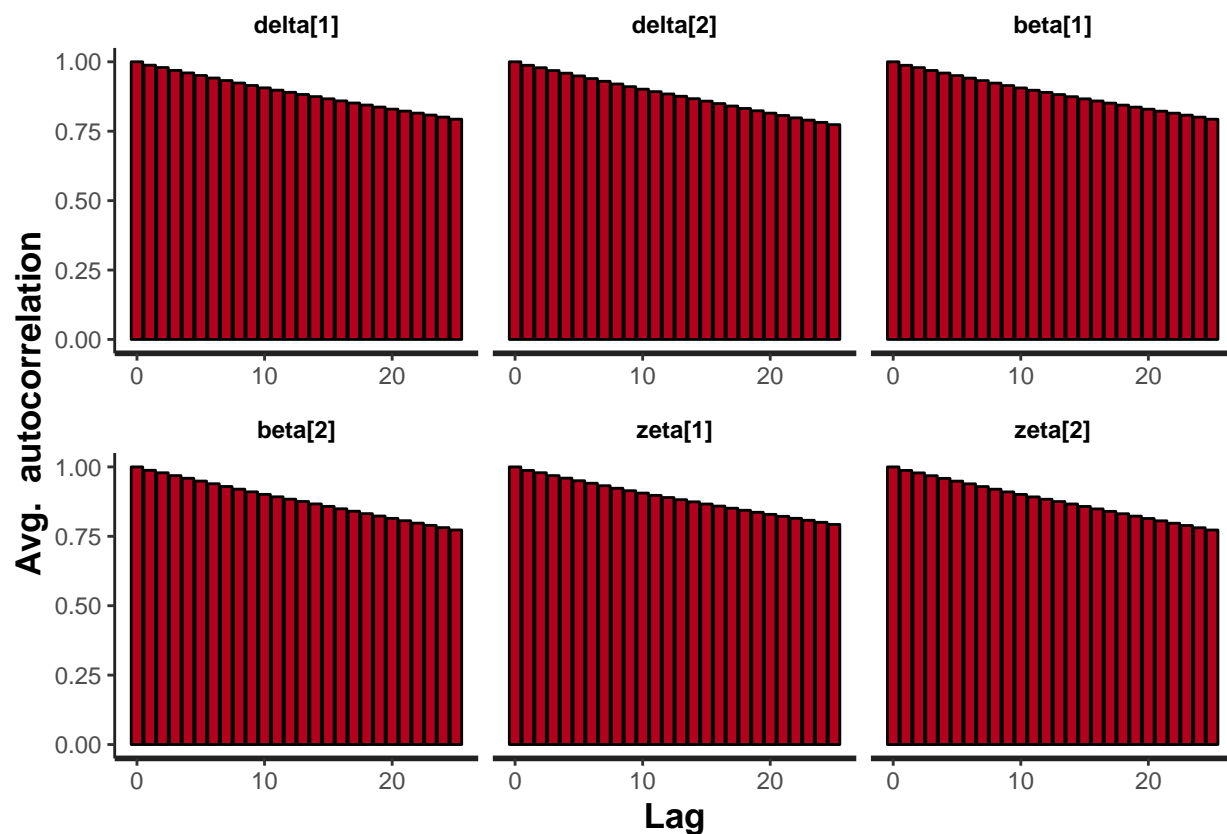| ## | | mean | se_mean | sd | 25% | 50% |
|---|---|---|---|---|---|---|
| ## | beta_k_raw[1] | -0.7841747 | 0.08613496 | 0.2953686 | -0.9573467 | -0.8023418 |
| ## | beta_k_raw[2] | -0.8837152 | 0.09922041 | 0.4039552 | -1.0523171 | -0.8267221 |

```
## beta_k_raw[3]  -1.1800951 0.08460307 0.3058298 -1.3476195 -1.2016308
## beta_k_raw[4]  -1.4250433 0.10374989 0.3955319 -1.6959977 -1.4678671
## beta_k_raw[5]  -2.0878480 0.11657333 0.4465316 -2.3197480 -2.0835746
## beta_k_raw[6]  -3.3506530 0.11878065 0.5058635 -3.6168634 -3.2786784
## beta_k[1]       -0.7841747 0.08613496 0.2953686 -0.9573467 -0.8023418
## beta_k[2]       -0.8837152 0.09922041 0.4039552 -1.0523171 -0.8267221
## beta_k[3]       -1.1800951 0.08460307 0.3058298 -1.3476195 -1.2016308
## beta_k[4]       -1.4250433 0.10374989 0.3955319 -1.6959977 -1.4678671
## beta_k[5]       -2.0878480 0.11657333 0.4465316 -2.3197480 -2.0835746
## beta_k[6]       -3.3506530 0.11878065 0.5058635 -3.6168634 -3.2786784
## beta_k[7]        0.0000000        NaN 0.0000000  0.0000000  0.0000000
##                      75%    n_eff     Rhat
## beta_k_raw[1]  -0.6035666 11.75898 1.424825
## beta_k_raw[2]  -0.6462663 16.57541 1.263956
## beta_k_raw[3]  -0.9965240 13.06735 1.315172
## beta_k_raw[4]  -1.1712609 14.53409 1.276879
## beta_k_raw[5]  -1.7889950 14.67257 1.315576
## beta_k_raw[6]  -3.0132480 18.13741 1.223897
## beta_k[1]       -0.6035666 11.75898 1.424825
## beta_k[2]       -0.6462663 16.57541 1.263956
## beta_k[3]       -0.9965240 13.06735 1.315172
## beta_k[4]       -1.1712609 14.53409 1.276879
## beta_k[5]       -1.7889950 14.67257 1.315576
## beta_k[6]       -3.0132480 18.13741 1.223897
## beta_k[7]        0.0000000      NaN      NaN
```

```
## Warning: Removed 104 rows containing non-finite values (stat_summary).
```

```
##                mean    se_mean       sd         25%         50%        75%
## delta[1]  0.05866516  0.2141304  1.220341  -0.6687869   0.10346451  0.8128439
## delta[2]  0.27142480  0.5683795  3.275683  -1.7938739   0.13780933  2.2599668
## beta[1]   0.09741983  0.2141576  1.220626  -0.6309216   0.14302669  0.8522835
## beta[2]   0.07129960  0.5683742  3.275464  -1.9916972  -0.06663923  2.0686005
## zeta[1]   0.06127766  0.2141302  1.220621  -0.6692158   0.10661212  0.8146261
## zeta[2]  -0.38199853  0.5684384  3.276158  -2.4479499  -0.50214945  1.6172228
##             n_eff      Rhat
## delta[1]  32.47926  1.126171
## delta[2]  33.21443  1.126973
## beta[1]   32.48613  1.126116
## beta[2]   33.21062  1.126901
## zeta[1]   32.49421  1.126137
## zeta[2]   33.21720  1.127209
```

```
##            mean      se_mean         sd           25%          50%           75%
## u[1] 0.10217182 0.03550160 0.5830579 -0.2925936  0.10338710 0.4970478
## u[2] 0.01736152 0.03645510 0.5807092 -0.3567947  0.01781394 0.3947438
```

```
## u[3]   0.23876946 0.02987063 0.5628365 -0.1195697  0.23193792 0.6091596
## u[4]  -0.21637117 0.03941035 0.5925109 -0.5992536 -0.22380531 0.1474855
## u[5]  -0.07794555 0.04149630 0.5878812 -0.4636175 -0.06572891 0.3204047
## u[6]  -0.06632961 0.03847205 0.6290430 -0.4768316 -0.07759662 0.3436157
## u[7]   0.20578478 0.04340685 0.5750972 -0.1685098  0.21148063 0.5825033
##          n_eff      Rhat
## u[1] 269.7289 1.016659
## u[2] 253.7470 1.007819
## u[3] 355.0387 1.005826
## u[4] 226.0331 1.015912
## u[5] 200.7060 1.004157
## u[6] 267.3438 1.007469
## u[7] 175.5359 1.025237
```
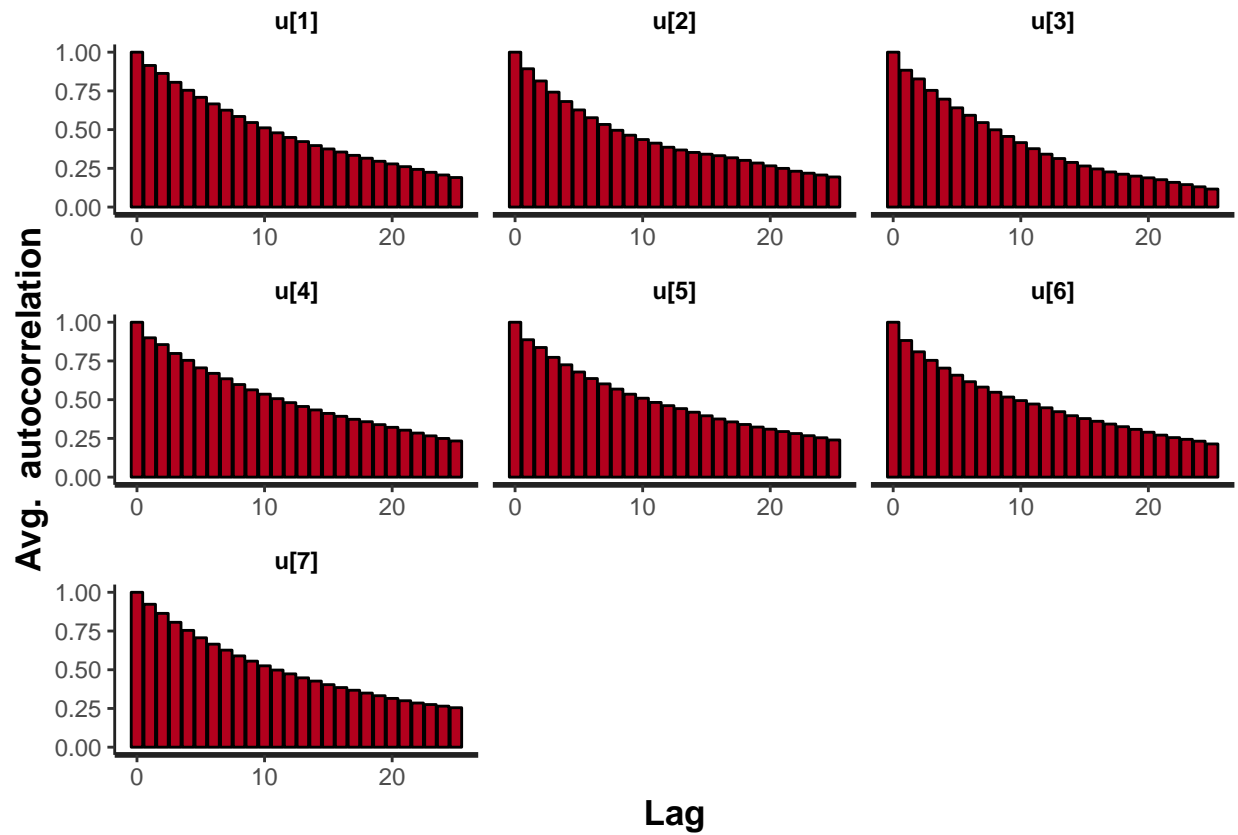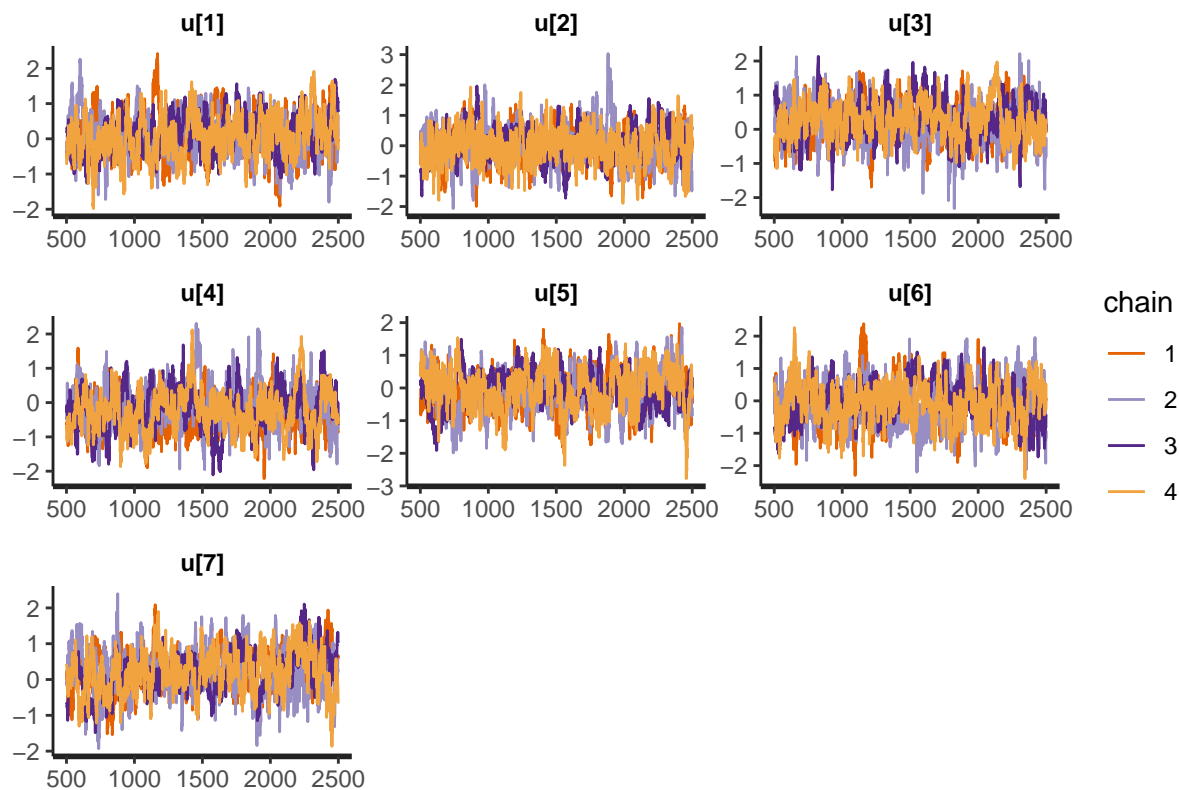
```
##               mean      se_mean         sd          25%          50%          75%
## v[1]  -0.2303948  0.03604165  0.5469551  -0.5792755  -0.2292528  0.12664494
## v[2]  -0.1310269  0.03467290  0.5625785  -0.5011073  -0.1405971  0.24299252
## v[3]  -0.3247641  0.03328886  0.5403205  -0.6406247  -0.3032770  0.02385453
## v[4]  -0.2834901  0.04041505  0.5698921  -0.6456070  -0.2857439  0.08471085
## v[5]  -0.3405938  0.03884276  0.5460112  -0.6962840  -0.3223014  0.01540685
## v[6]   0.1671997  0.04229167  0.5943235  -0.2218204   0.1440038  0.54386514
## v[7]   0.1562508  0.03485039  0.5114810  -0.1810607   0.1680546  0.50683002
##          n_eff       Rhat
## v[1]  230.3000  1.016268
## v[2]  263.2606  1.008399
## v[3]  263.4542  1.010663
## v[4]  198.8378  1.016530
## v[5]  197.5982  1.018512
## v[6]  197.4860  1.010228
## v[7]  215.3990  1.027077
```
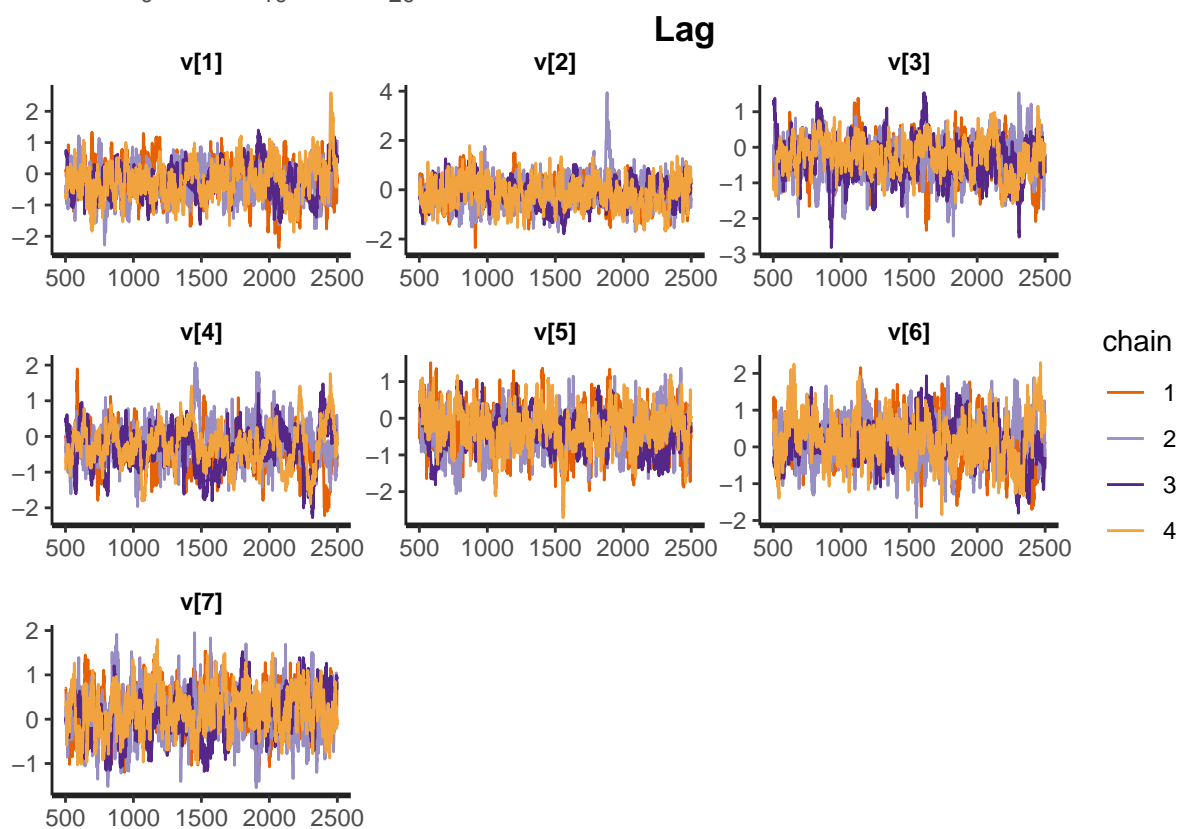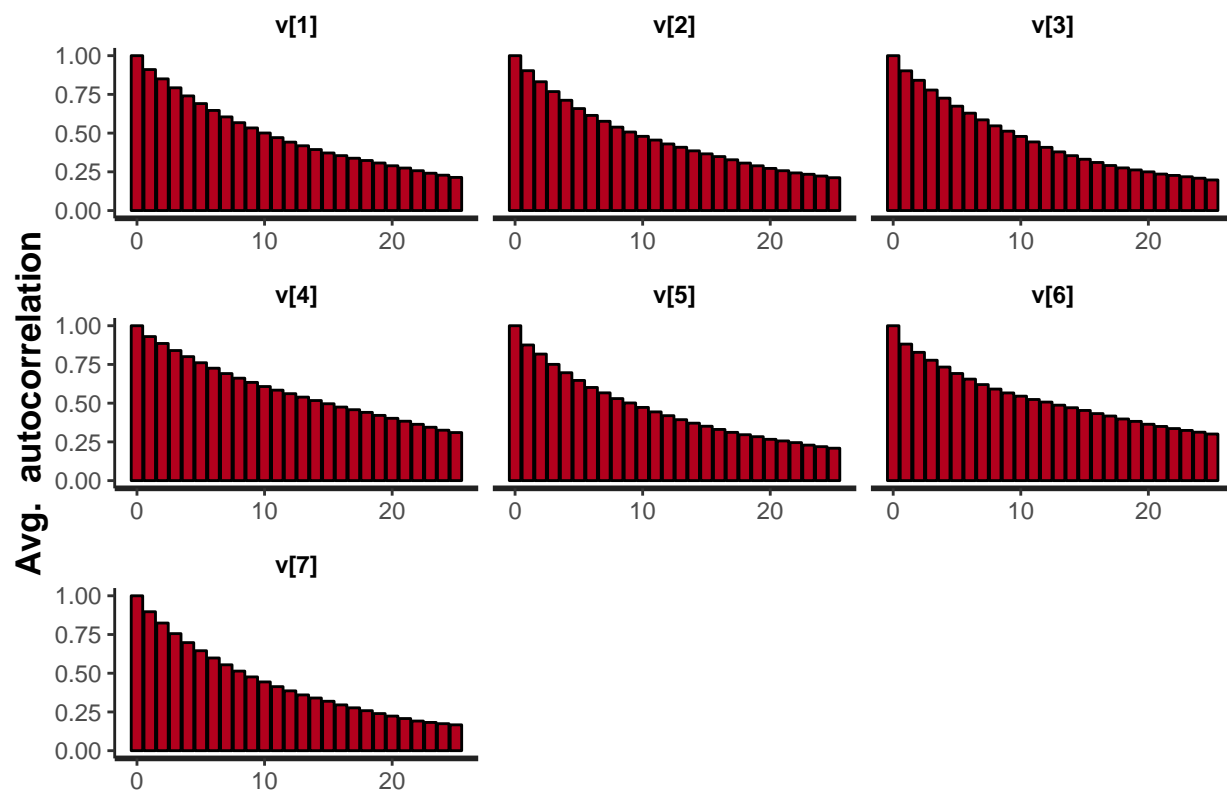
```
##            mean      se_mean         sd          25%          50%         75%
## q[1]  0.25798488  0.03591332  0.5465137  -0.10660388  0.25549161  0.61194541
## q[2]  0.14696916  0.03569596  0.5728526  -0.22498080  0.12847919  0.51221265
```

```
## q[3]   0.20153028 0.03346278 0.5385208 -0.12398815   0.22414043 0.53949828
## q[4]   0.40244560 0.04071184 0.5658380   0.03744034   0.39367184 0.76550634
## q[5]   0.30768418 0.03752371 0.5375002 -0.03478011   0.31597861 0.66442759
## q[6]   0.07927985 0.04034887 0.6014185 -0.31685238   0.06373473 0.44532134
## q[7]  -0.27730947 0.03719057 0.5156113 -0.62646540 -0.26350837 0.07867479
##          n_eff      Rhat
## q[1] 231.5746 1.017277
## q[2] 257.5417 1.008531
## q[3] 258.9888 1.011905
## q[4] 193.1714 1.018580
## q[5] 205.1850 1.019266
## q[6] 222.1727 1.010732
## q[7] 192.2115 1.027714
```
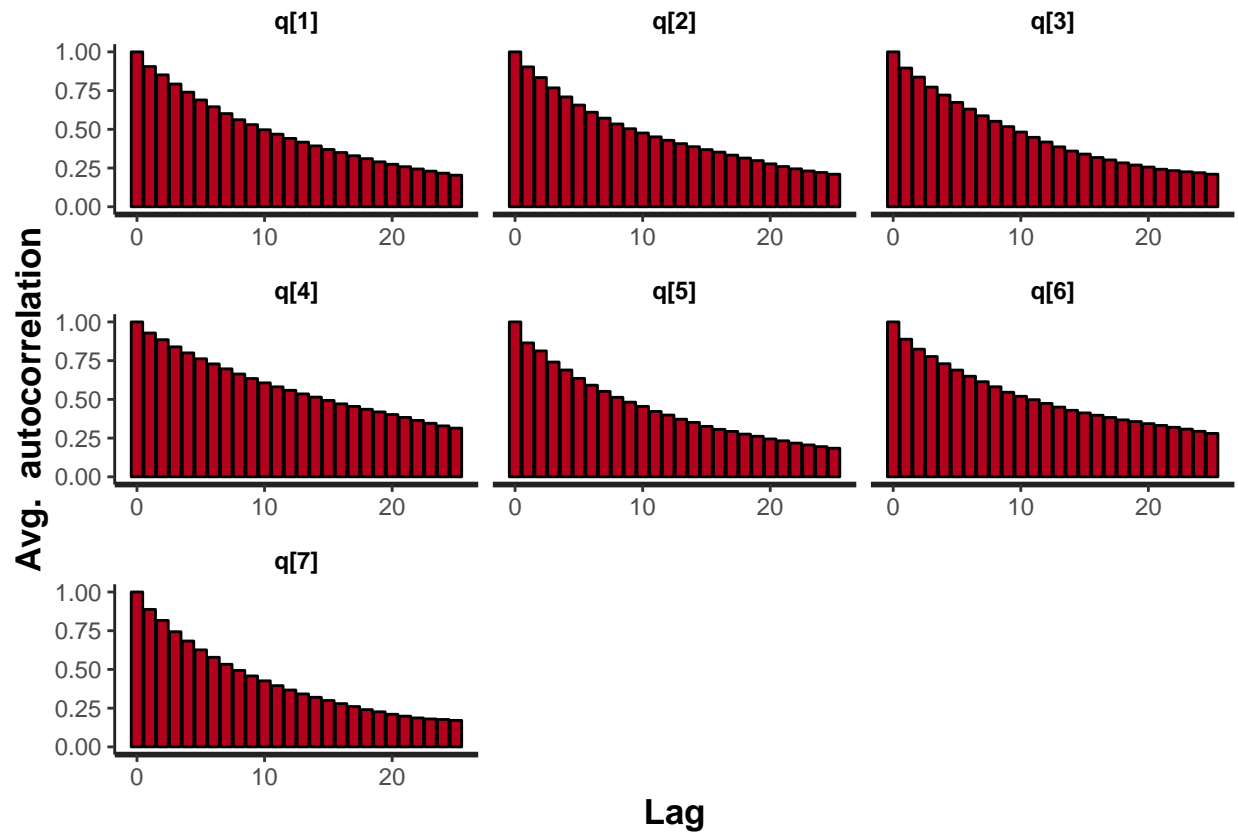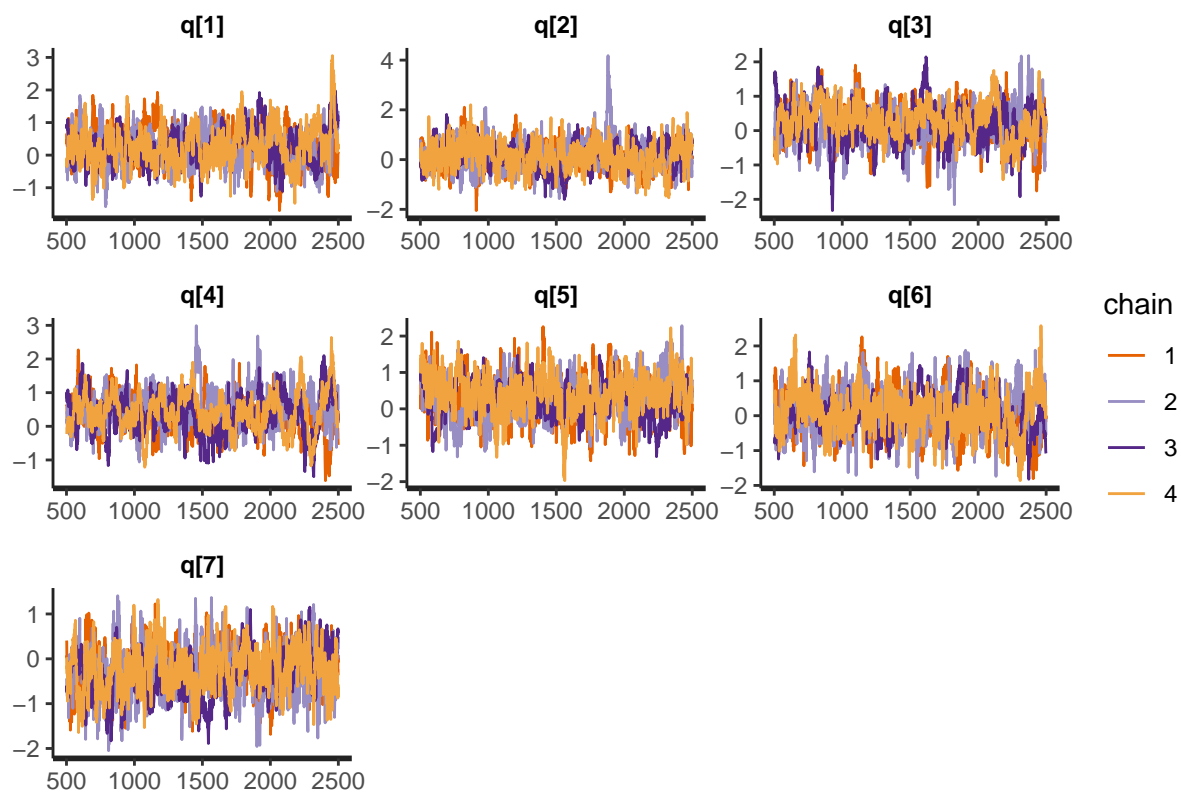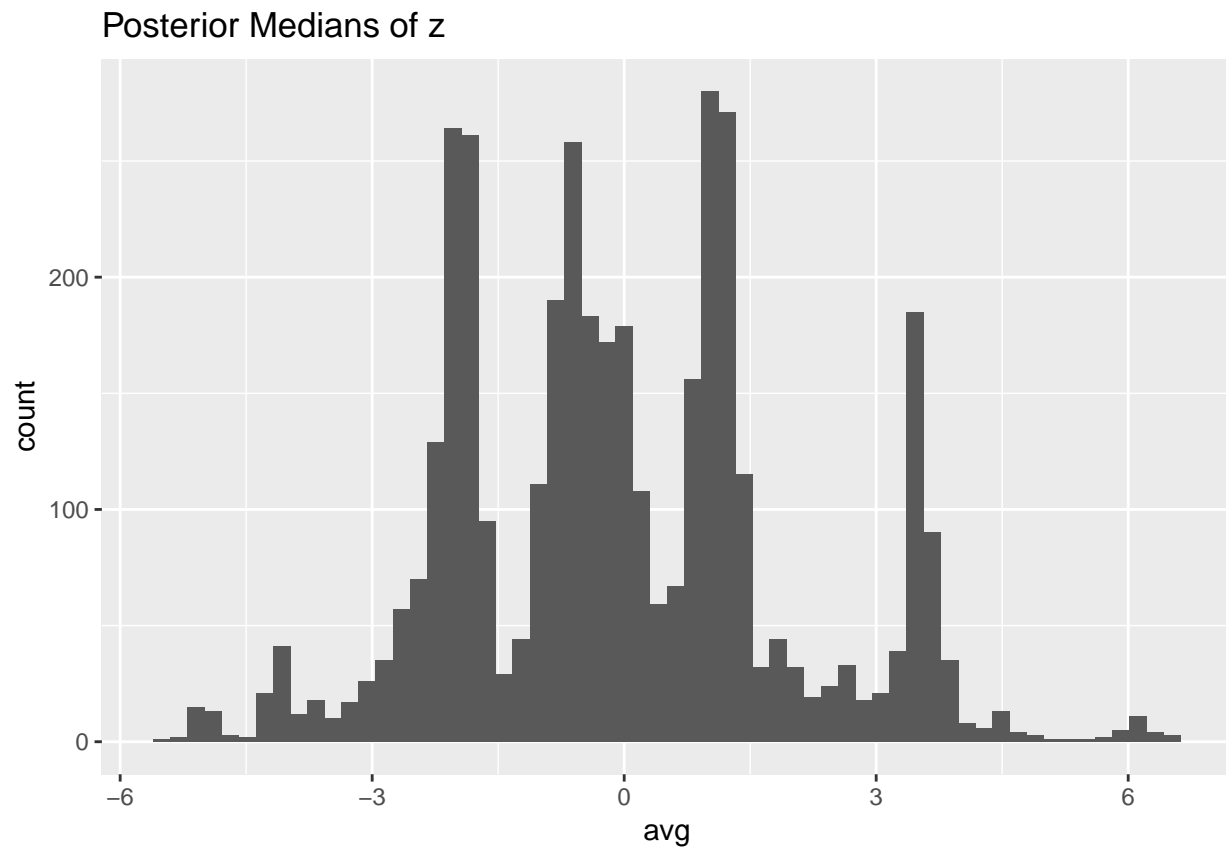
```
## [1] "Summary statistics for posterior medians of w"
##          avg
##  Min.    :-4.3700
##  1st Qu.:-0.6818
##  Median : 0.3614
##  Mean   : 1.1056
##  3rd Qu.: 1.9260
##  Max.   : 8.5503
```
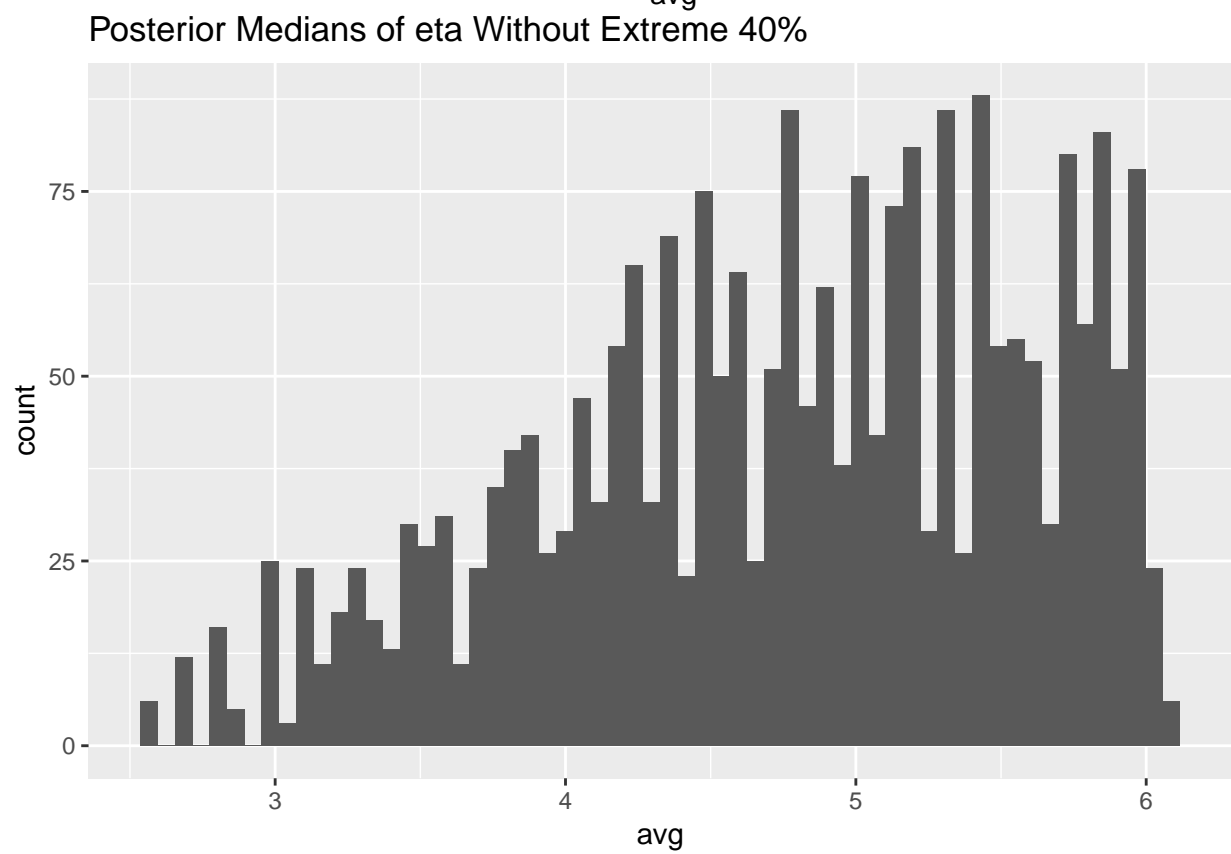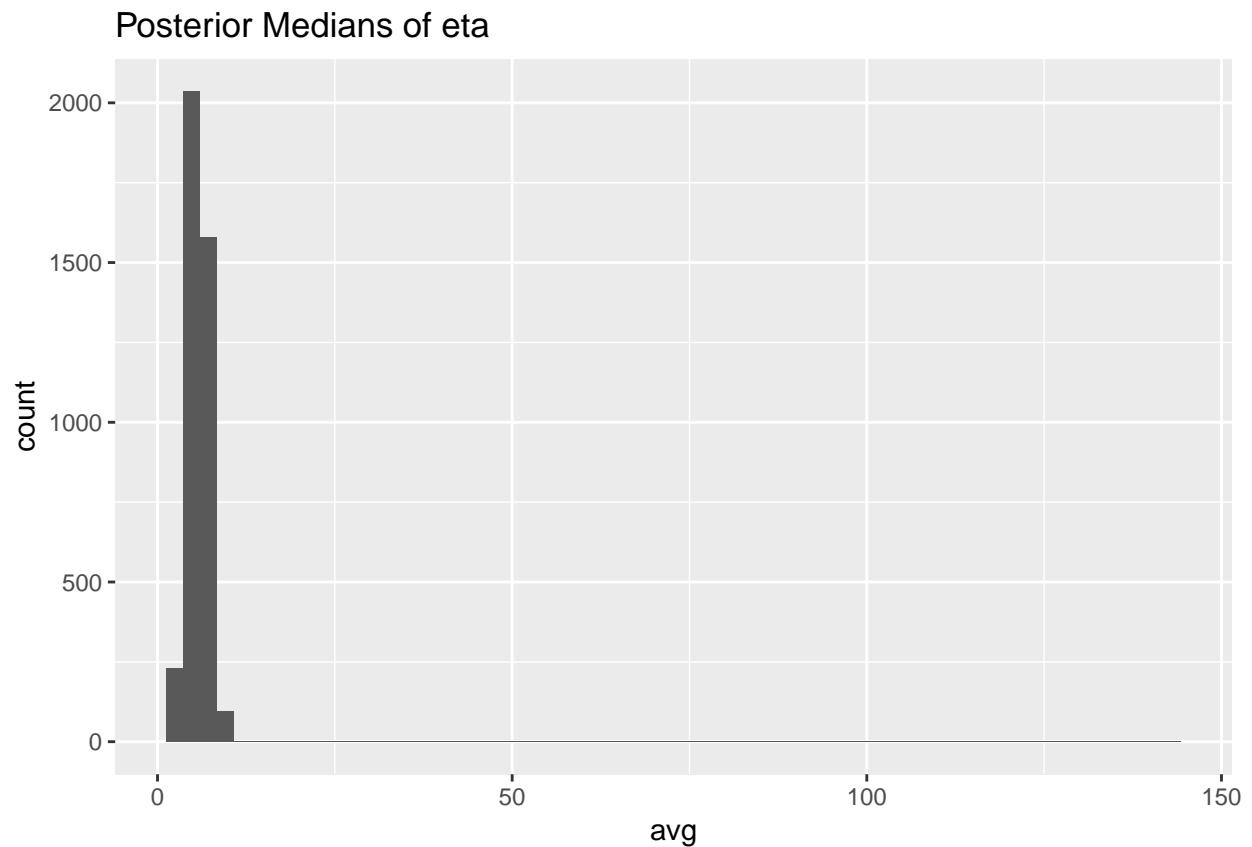
## Posterior Medians of w
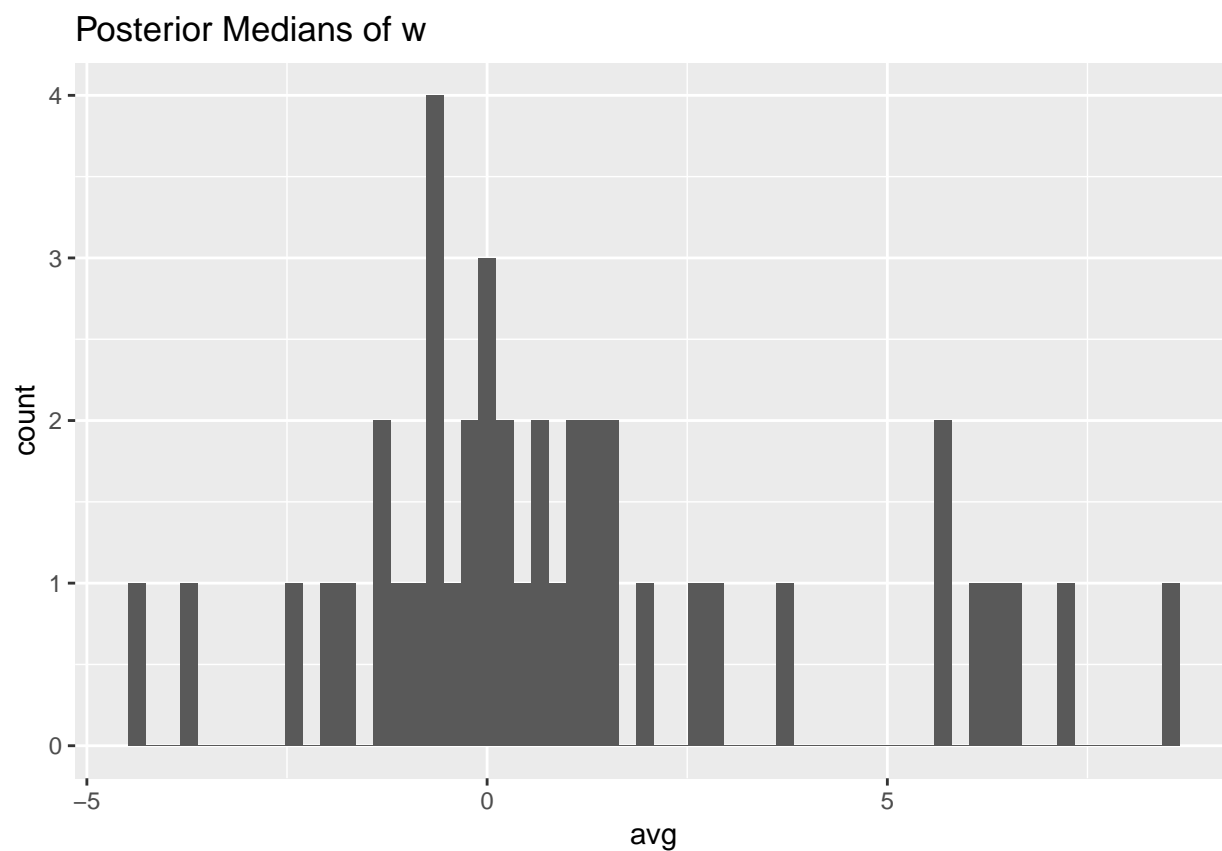


```
## [1] "Summary statistics for posterior medians of z"
##        avg
##  Min.   :-5.40736
##  1st Qu.:-1.74782
##  Median :-0.20634
##  Mean   :-0.04478
##  3rd Qu.: 1.14335
##  Max.   : 6.62060
```

## Posterior Medians of z



```
## [1] "Summary statistics for posterior medians of p"
##         avg
##  Min.   :-17.454
##  1st Qu.:-10.224
##  Median : -8.491
##  Mean   : -8.652
##  3rd Qu.: -6.877
##  Max.   : -2.472
```

## Posterior Medians of p
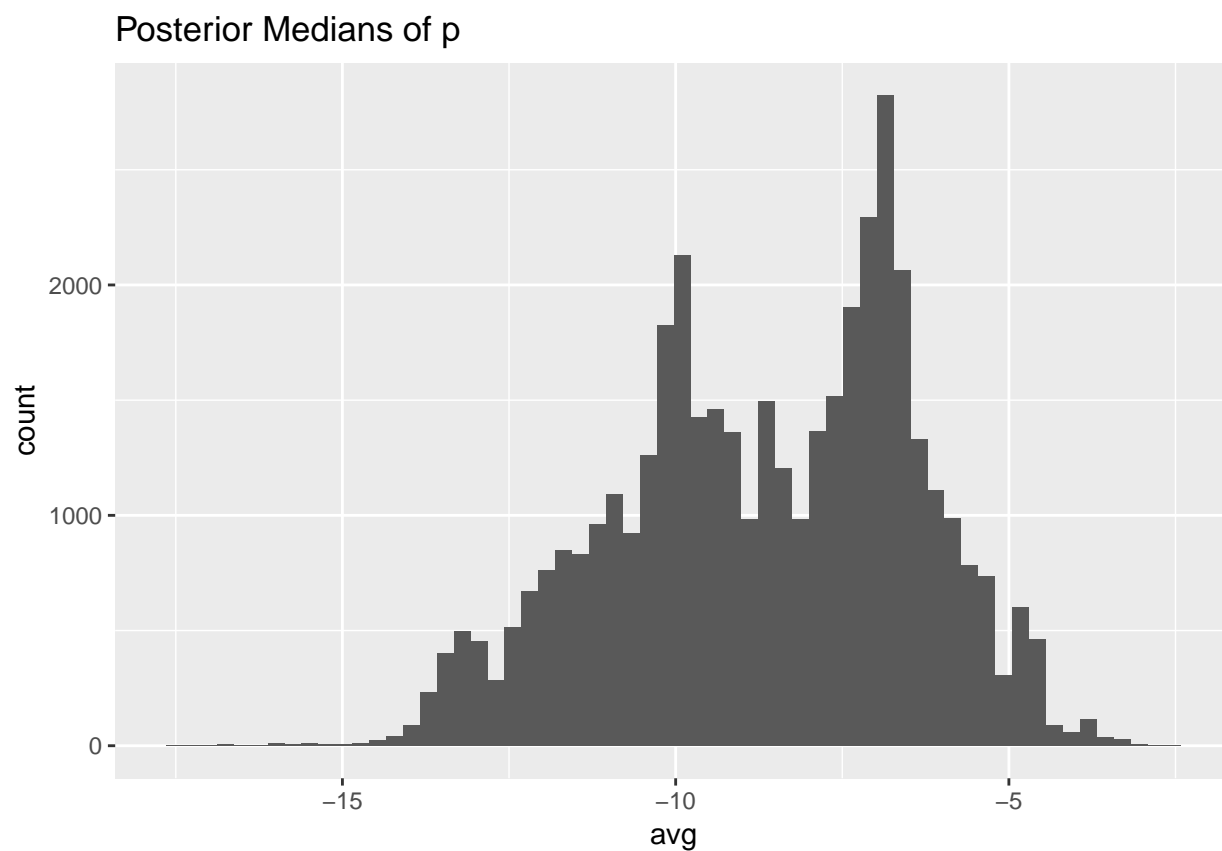


```
## [1] "Summary statistics for posterior medians of eta"
##         avg
##  Min.    :  2.542
##  1st Qu.:  4.656
##  Median :  5.659
##  Mean   :  5.735
##  3rd Qu.:  6.628
##  Max.    :143.303
```

Posterior Medians of eta

Posterior Medians of eta Without Extreme 40%
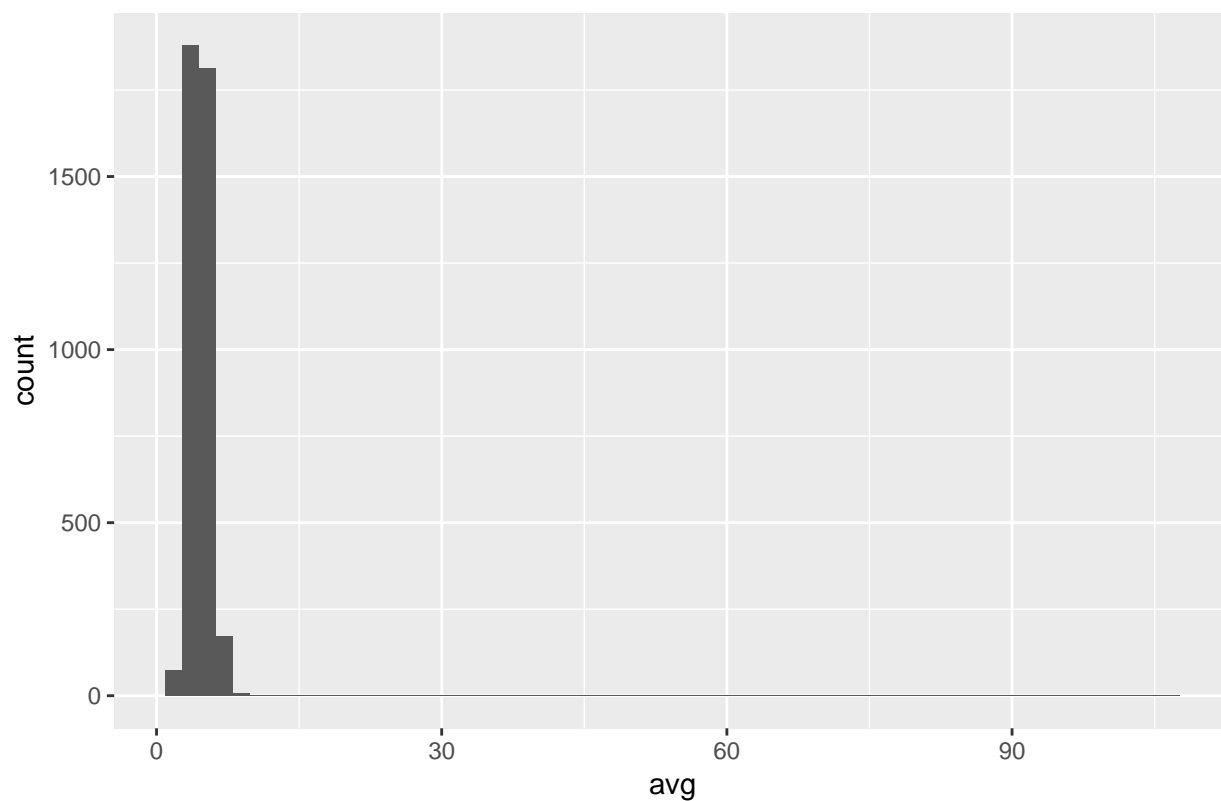
```
## [1] "Summary statistics for posterior medians of lambda"
##        avg
##  Min.   :  2.014
##  1st Qu.:  3.764
##  Median :  4.501
##  Mean   :  4.562
##  3rd Qu.:  5.233
##  Max.   :104.192
```

## Posterior Medians of lambda

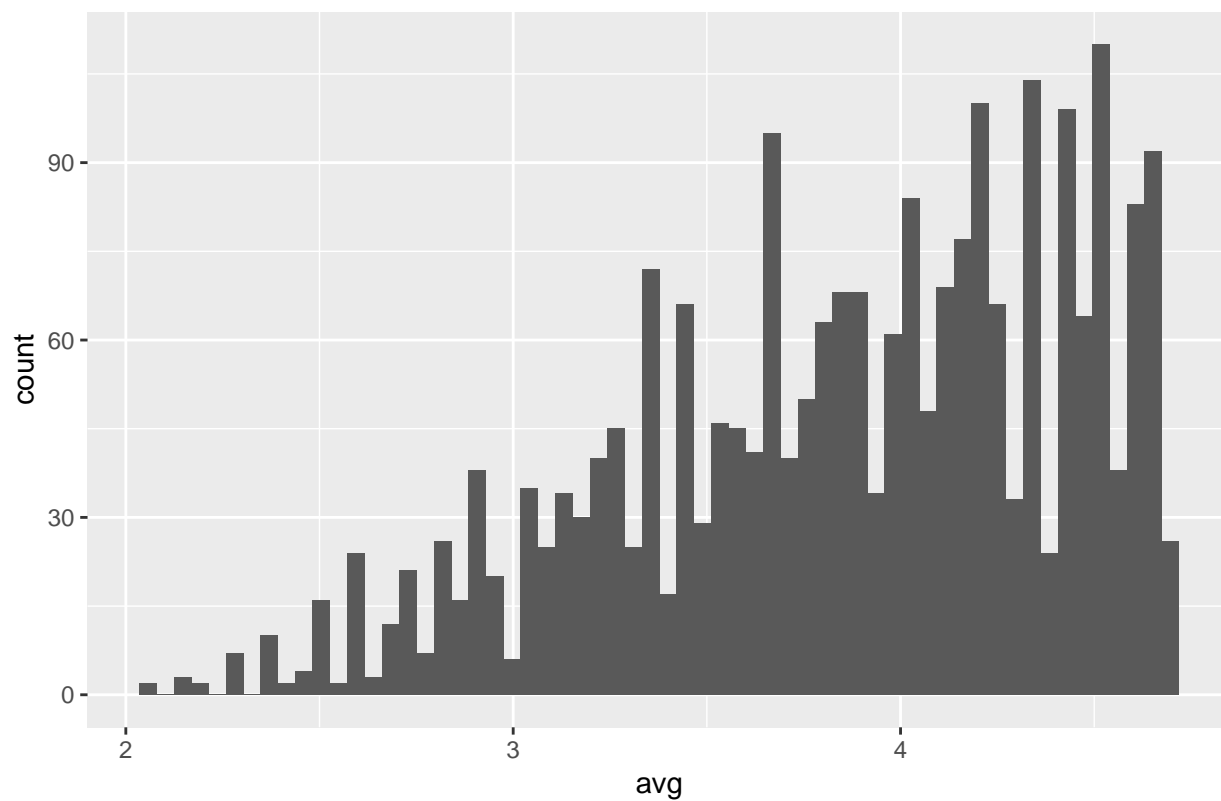## Posterior Medians of lambda Without Extreme 40%
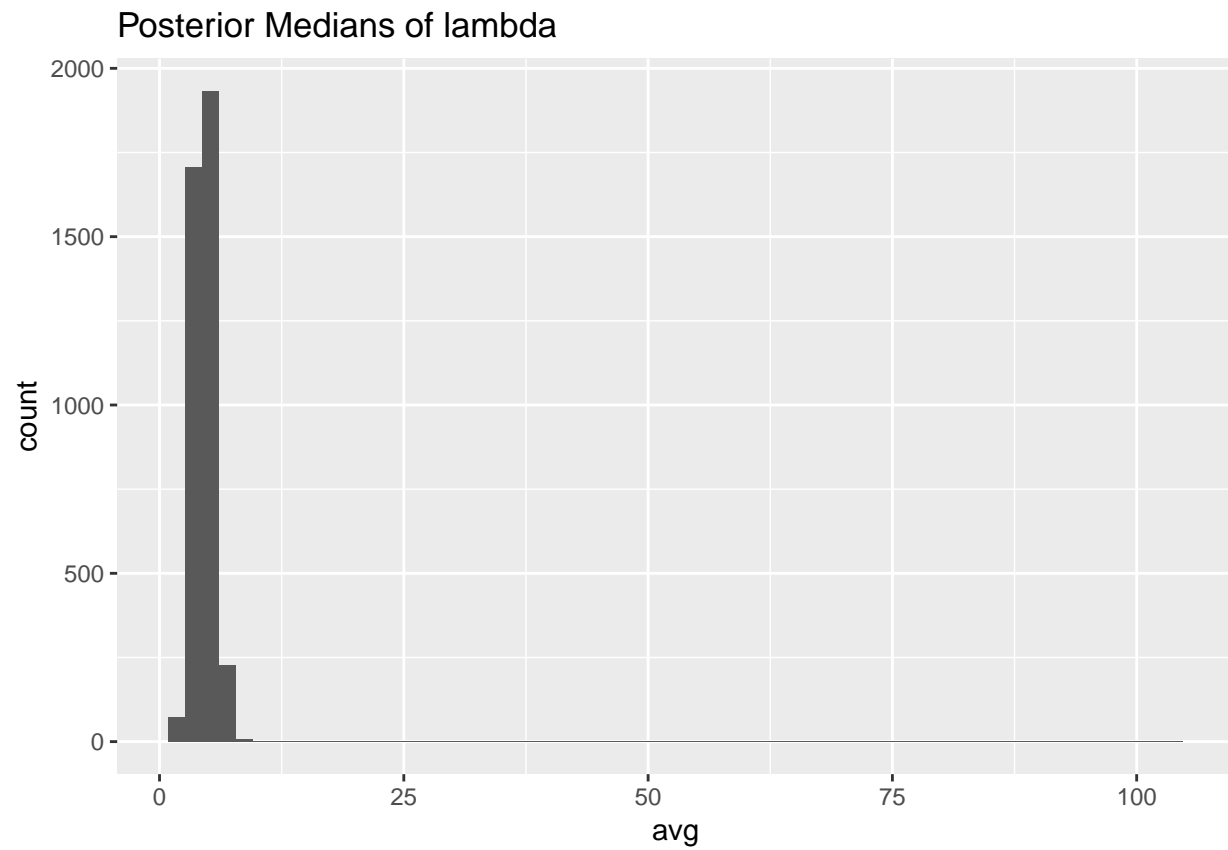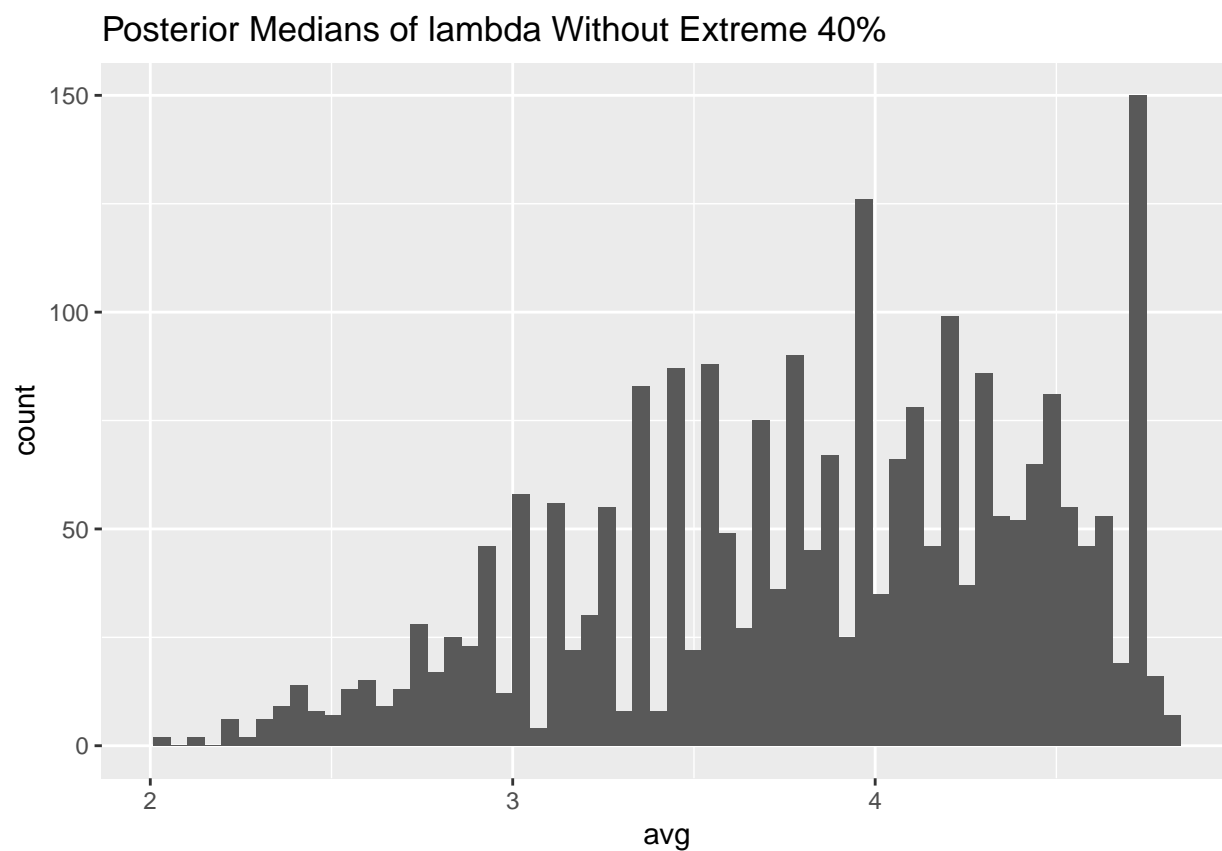


```
## [1] "Summary statistics for posterior medians of kappa"
##       avg
## Min.   : 2.075
## 1st Qu.: 3.794
## Median : 4.463
## Mean   : 4.551
## 3rd Qu.: 5.133
## Max.   :107.110
```

Posterior Medians of kappa



Posterior Medians of kappa Without Extreme 40%

## Identifying Parameters with Large Rhats

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

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.9999  1.0136  1.0493  1.1379  1.1263  3.3535       1
```

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

```
##     mean se_mean Rhat
## NA    NA      NA   NA
```