

# MCMC Diagnostics - IFLS data

*Sarah Teichman*

*04/17/2020*

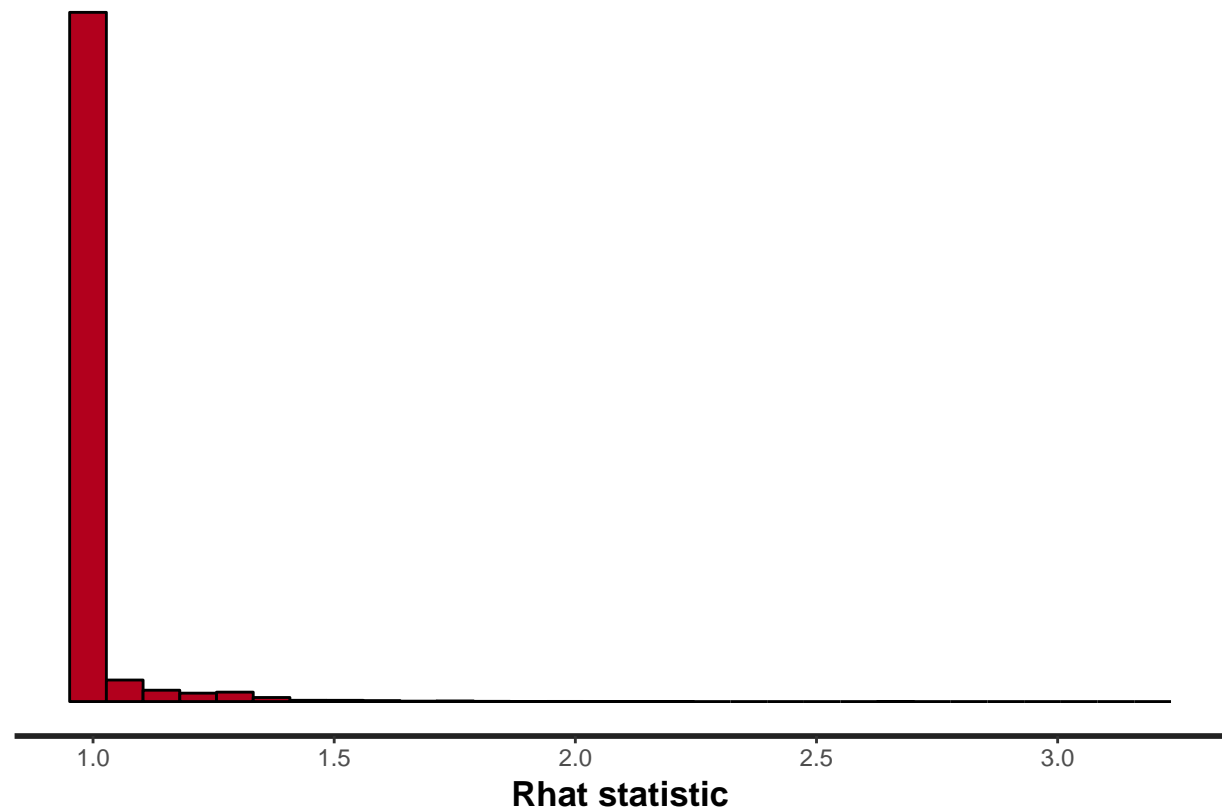
```
K <- 7  
Ti <- 3  
N <- 1973
```

## General MCMC diagnostic plots

Overall model diagnostics from rstan package.

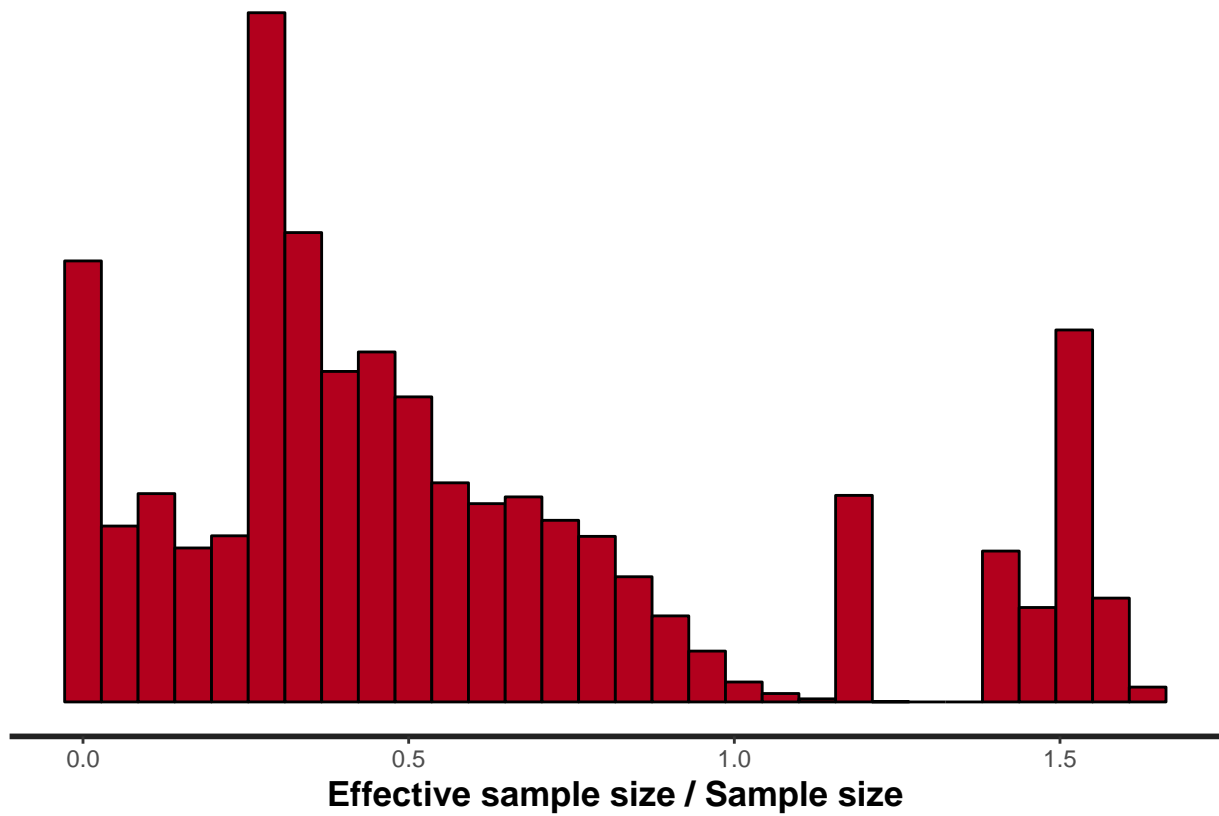
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

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



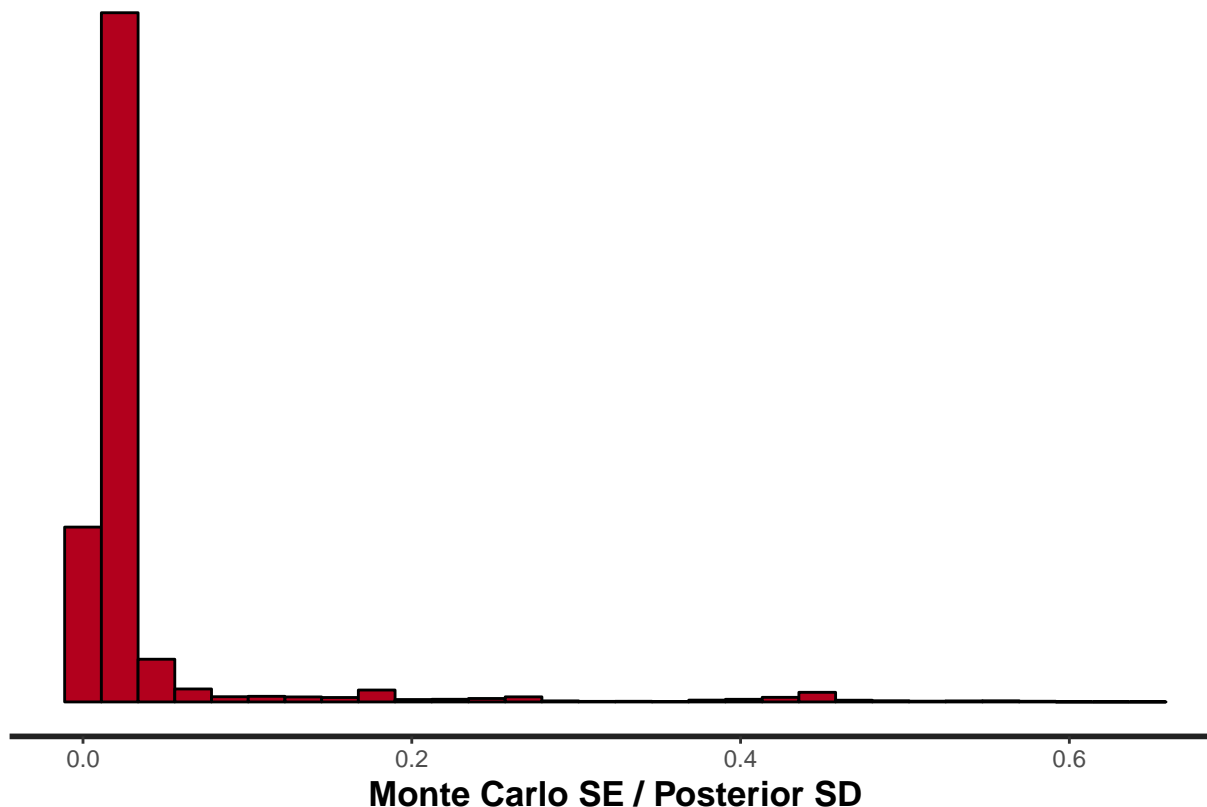
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```



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

```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```



### 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),"%")))
  }
}

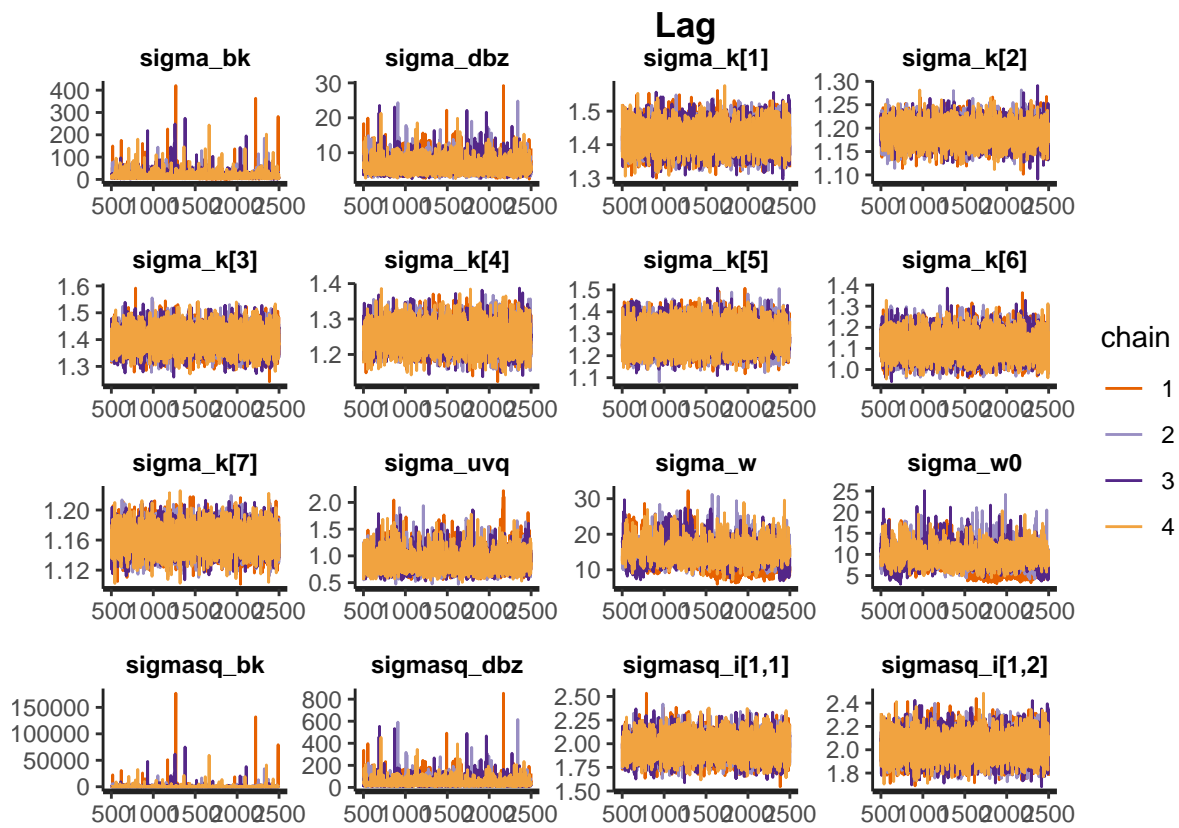
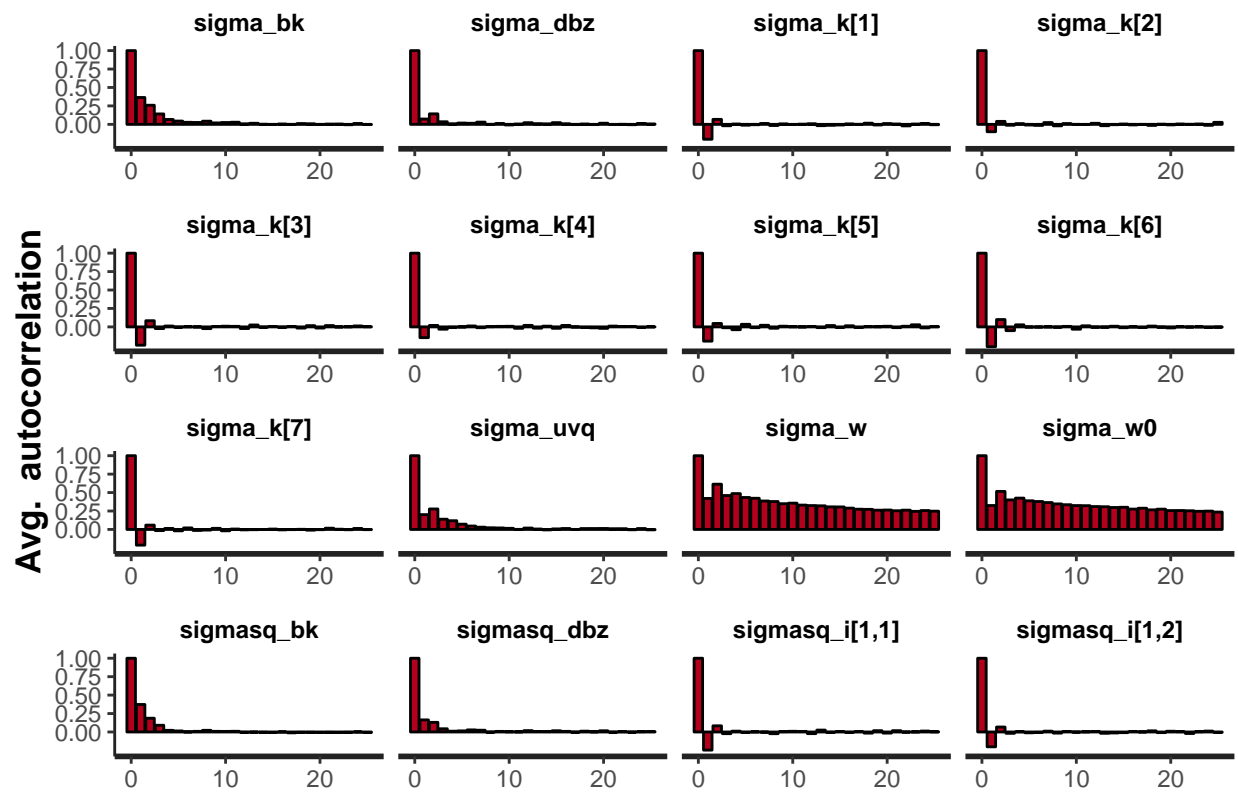
plot_fit <- function(fit) {
  get_single_plots(fit, sigma_params)
  # get_single_plots(fit, beta_k)
```

```

get_single_plots(fit, beta_0)
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)

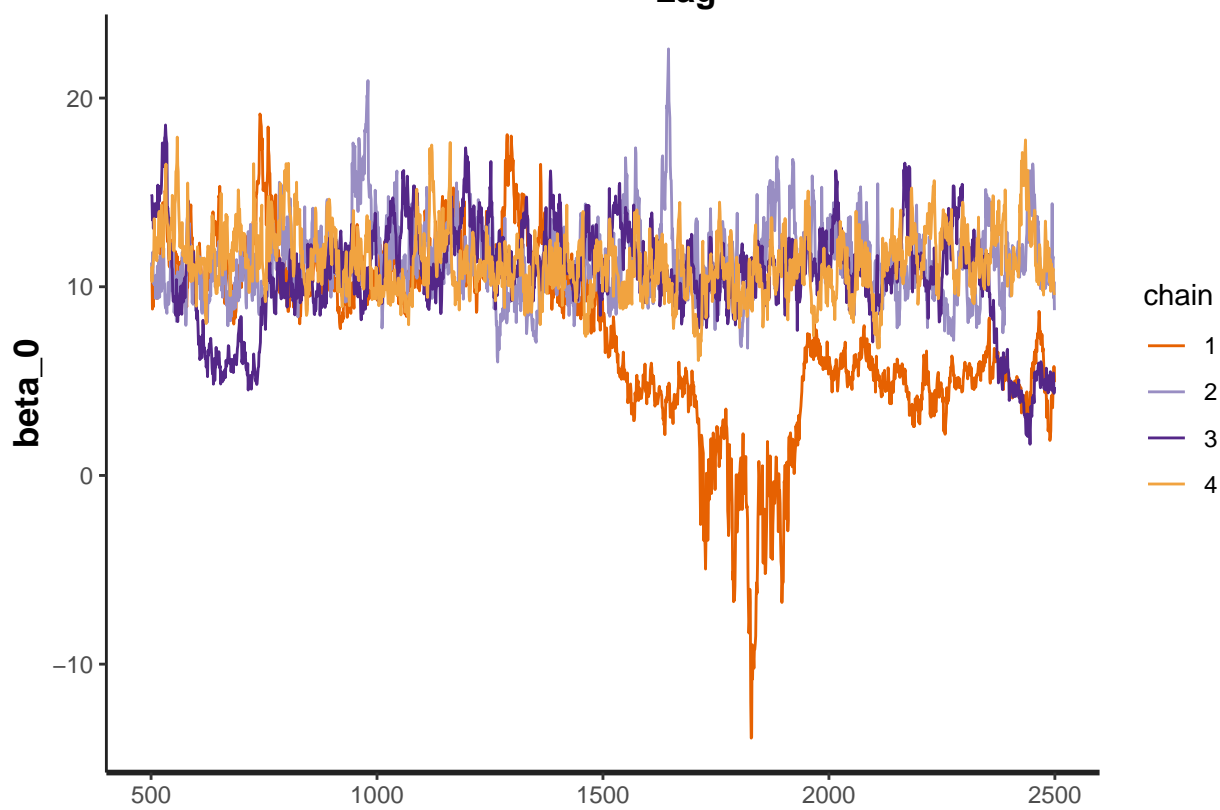
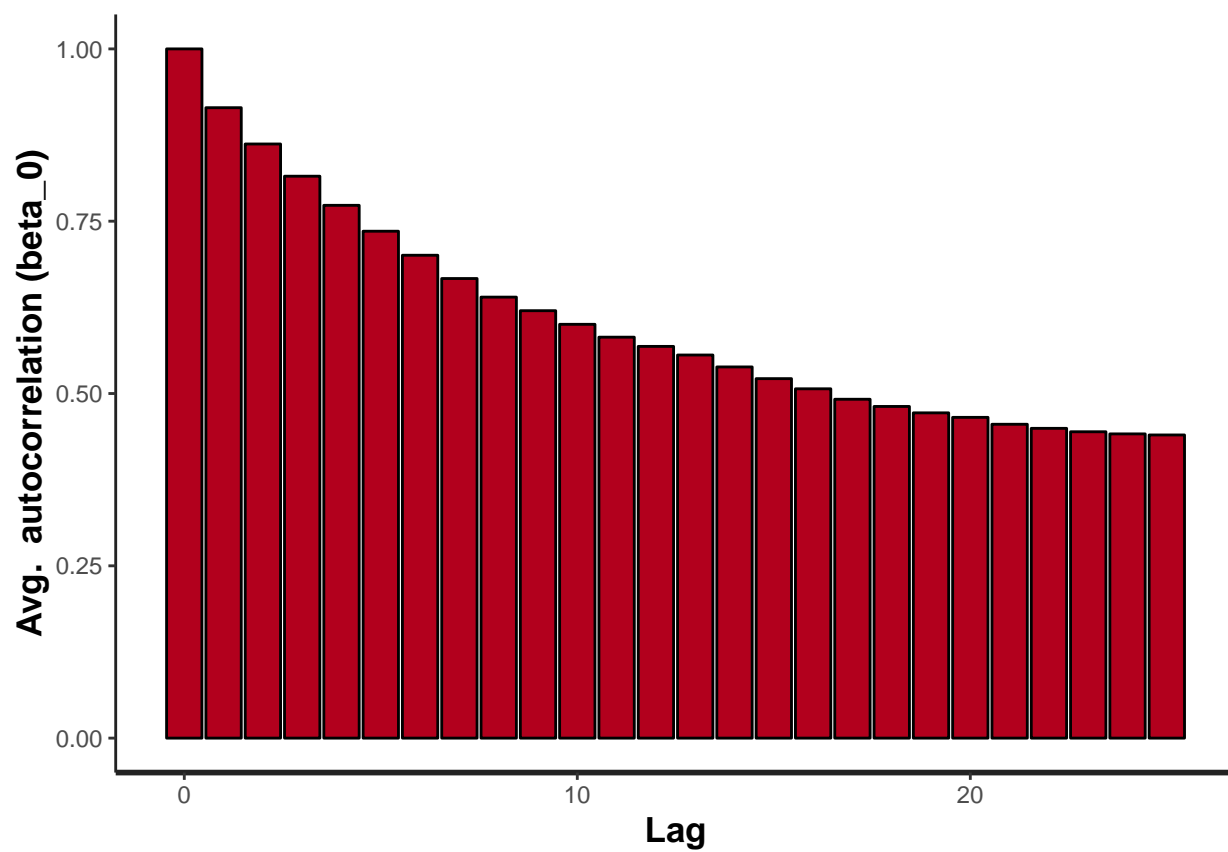
```

##		mean	se_mean	sd	25%	50%
##	sigma_bk	14.4813192	4.094575e-01	1.862502e+01	6.6972041	9.7696693
##	sigma_dbz	6.0914123	3.051015e-02	2.143320e+00	4.6377166	5.6438425
##	sigma_k[1]	1.4223674	3.588147e-04	3.848352e-02	1.3964981	1.4217275
##	sigma_k[2]	1.1893354	2.437528e-04	2.361346e-02	1.1733560	1.1889912
##	sigma_k[3]	1.3987752	3.747604e-04	4.193581e-02	1.3701924	1.3974290
##	sigma_k[4]	1.2503451	3.367239e-04	3.565065e-02	1.2256265	1.2496364
##	sigma_k[5]	1.2850727	5.341038e-04	5.797902e-02	1.2455559	1.2822851
##	sigma_k[6]	1.1132882	5.003429e-04	5.722053e-02	1.0738674	1.1111667
##	sigma_k[7]	1.1609815	1.566390e-04	1.719914e-02	1.1491937	1.1607915
##	sigma_uvq	0.9102344	3.705440e-03	1.888752e-01	0.7755486	0.8839208
##	sigma_w	14.4572192	6.189074e-01	3.529047e+00	11.9371727	14.1331008
##	sigma_w0	8.9281166	4.990989e-01	2.686637e+00	7.0645577	8.6658863
##	sigmasq_bk	556.5565607	7.271117e+01	3.979033e+03	44.8525422	95.4464385
##	sigmasq_dbz	41.6985480	5.638679e-01	3.743462e+01	21.5084162	31.8529587
##	sigmasq_i[1,1]	1.9583304	1.053369e-03	1.176294e-01	1.8774271	1.9528078
##	sigmasq_i[1,2]	2.0246098	1.025358e-03	1.096740e-01	1.9502070	2.0213090
##		75%	n_eff	Rhat		
##	sigma_bk	15.454028	2069.07264	1.0064615		
##	sigma_dbz	7.016465	4934.97699	0.9999258		
##	sigma_k[1]	1.448186	11502.94516	1.0000745		
##	sigma_k[2]	1.204919	9384.69463	0.9998642		
##	sigma_k[3]	1.426149	12521.68400	0.9996248		
##	sigma_k[4]	1.273895	11209.51992	0.9998352		
##	sigma_k[5]	1.322651	11783.93864	1.0000231		
##	sigma_k[6]	1.150165	13078.81172	0.9996835		
##	sigma_k[7]	1.172589	12056.29036	0.9997209		
##	sigma_uvq	1.013803	2598.18625	1.0006227		
##	sigma_w	16.561652	32.51348	1.1364444		
##	sigma_w0	10.498865	28.97642	1.1397399		
##	sigmasq_bk	238.826995	2994.69578	1.0008627		
##	sigmasq_dbz	49.230781	4407.49812	0.9998741		
##	sigmasq_i[1,1]	2.033900	12470.12862	0.9996256		
##	sigmasq_i[1,2]	2.097241	11440.78405	1.0000940		

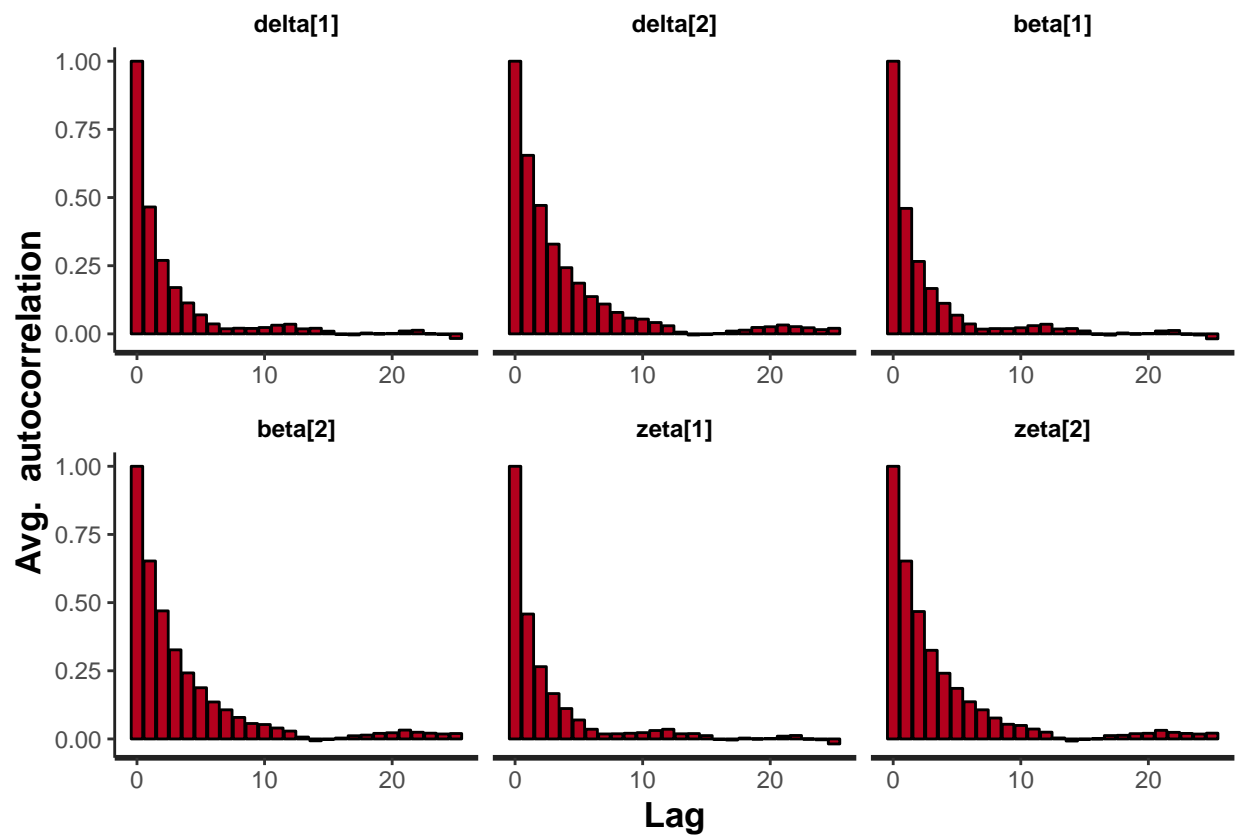


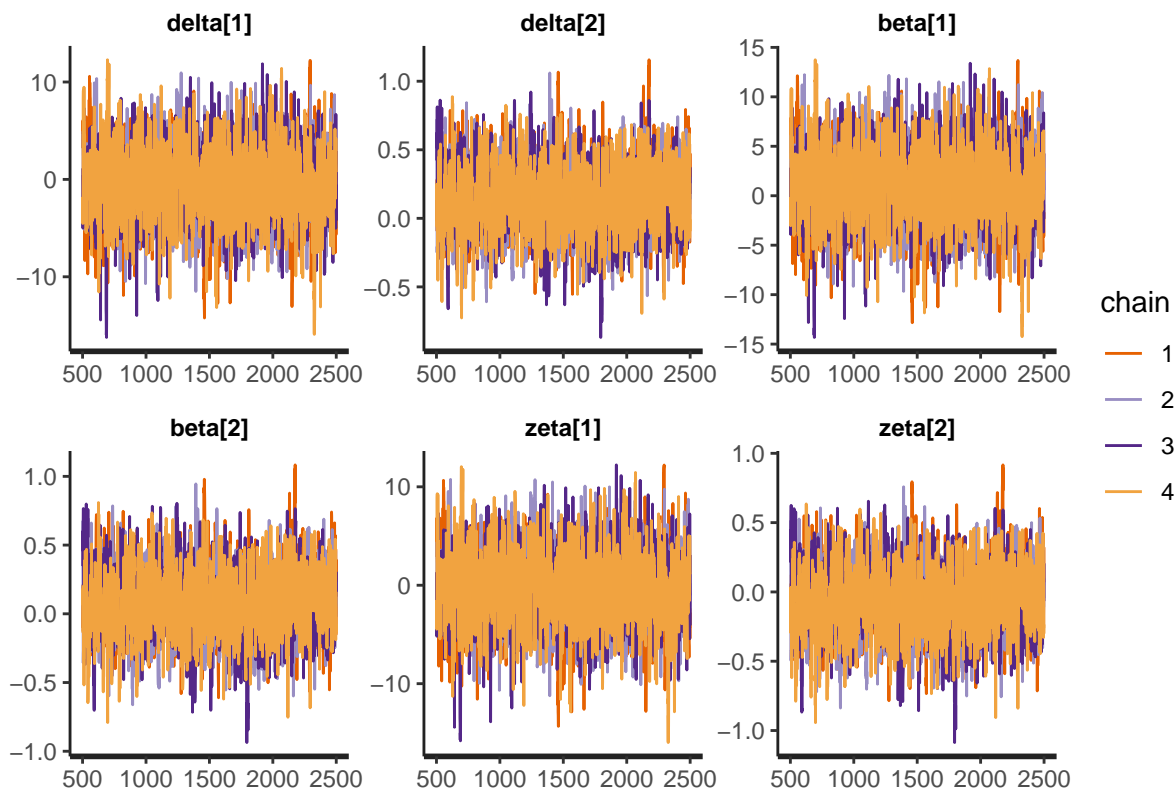
```
##          mean    se_mean      sd      25%      50%      75%    n_eff
## beta_0 10.19448  0.9550696  3.400682  9.247678 10.72088 12.10632 12.67833
##          Rhat
```

## beta\_0 1.517788



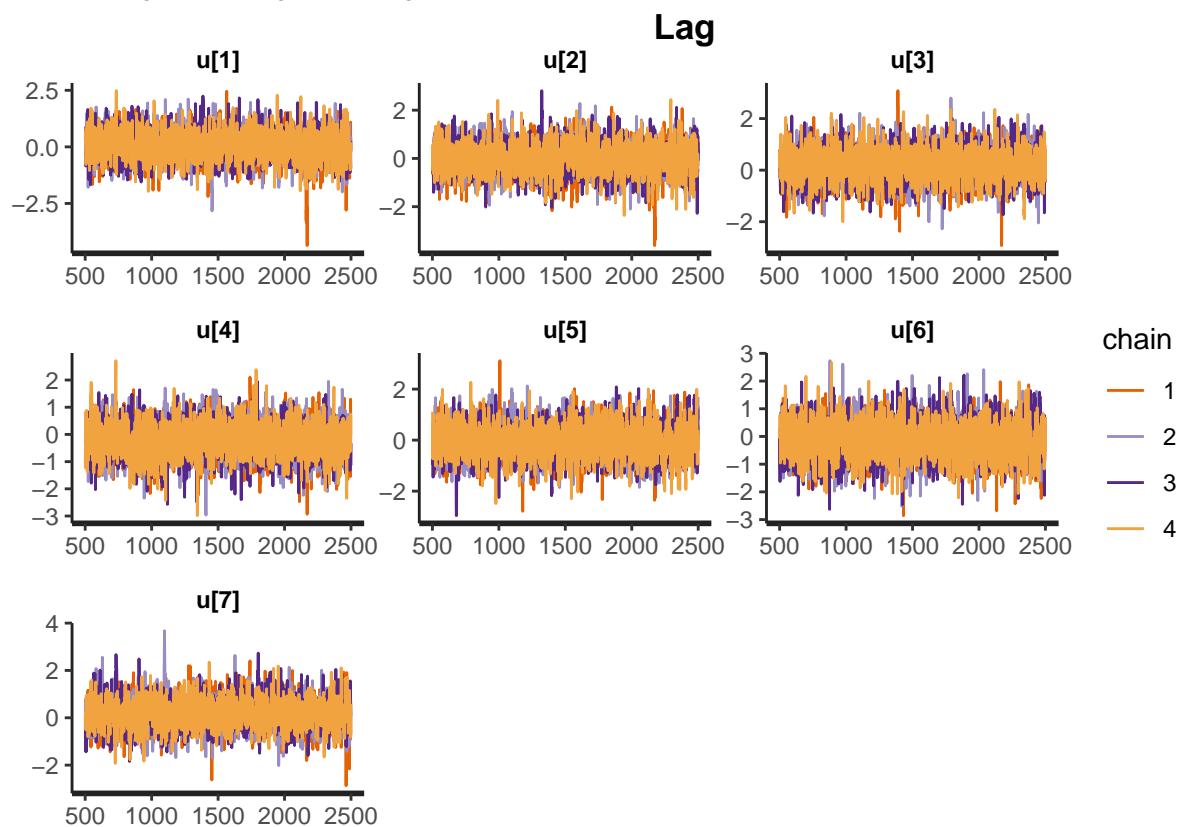
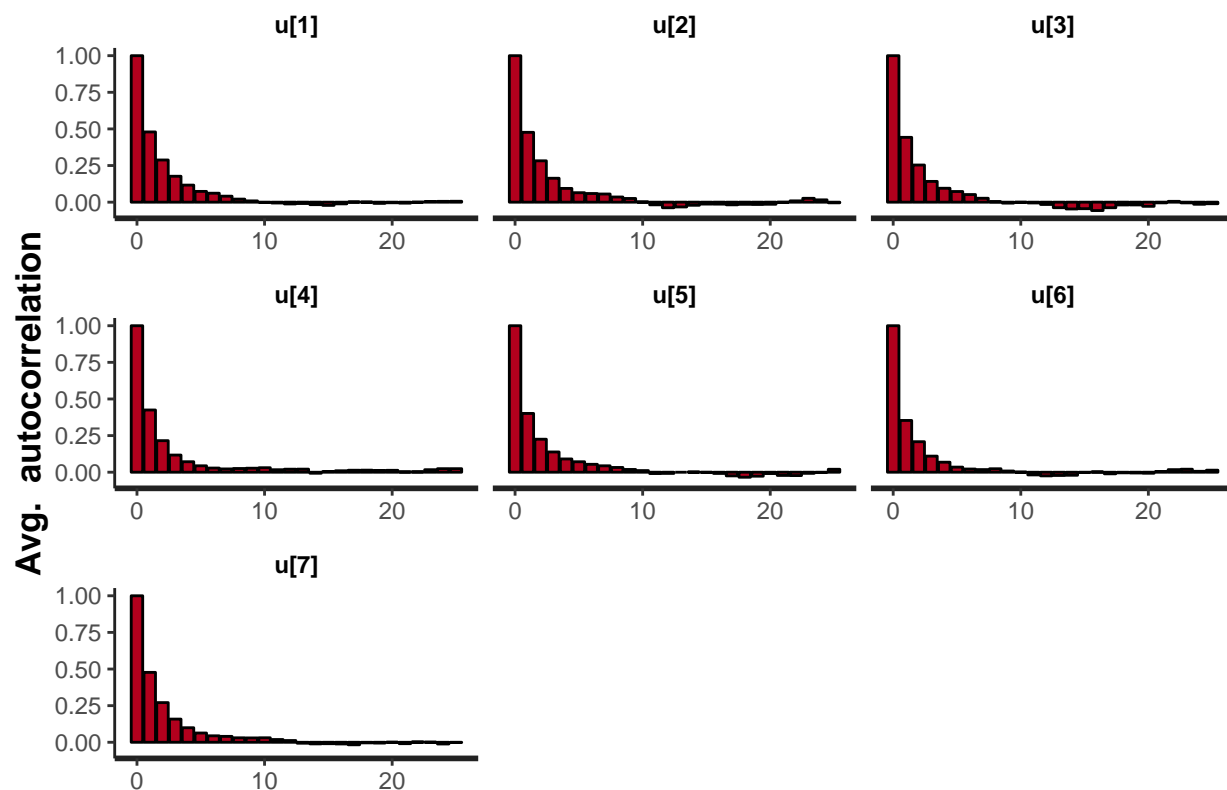
##		mean	se_mean	sd	25%	50%
##	delta[1]	-0.54124455	0.076417009	3.5574998	-2.761978670	-0.48853726
##	delta[2]	0.15650526	0.006971788	0.2402965	-0.001721143	0.15752341
##	beta[1]	0.95321962	0.075930745	3.5579784	-1.255663200	1.00300531
##	beta[2]	0.09166297	0.006904967	0.2390552	-0.065130255	0.09318813
##	zeta[1]	-0.44214125	0.076009063	3.5589116	-2.646725094	-0.38420387
##	zeta[2]	-0.05555315	0.006862859	0.2391714	-0.214879659	-0.05321322
##		75%	n_eff	Rhat		
##	delta[1]	1.6834004	2167.253	1.001419		
##	delta[2]	0.3132452	1187.973	1.007233		
##	beta[1]	3.1568963	2195.691	1.001376		
##	beta[2]	0.2485569	1198.597	1.007384		
##	zeta[1]	1.7831428	2192.319	1.001404		
##	zeta[2]	0.1021078	1214.530	1.007160		





##		mean	se_mean	sd	25%	50%	75%
##	u[1]	0.055076063	0.01361923	0.6372646	-0.3445164	0.049696667	0.4587086
##	u[2]	-0.009081963	0.01314849	0.6169057	-0.4090494	-0.008176074	0.3868941
##	u[3]	0.196587426	0.01228229	0.6122260	-0.1923362	0.198850739	0.5841634
##	u[4]	-0.212500035	0.01214545	0.6077615	-0.6036213	-0.206360717	0.1755529
##	u[5]	-0.051300339	0.01243640	0.6212712	-0.4561162	-0.045763822	0.3479618
##	u[6]	-0.146314779	0.01205475	0.6498731	-0.5659115	-0.152966030	0.2775148
##	u[7]	0.167552756	0.01304957	0.6139695	-0.2284199	0.154631499	0.5566034
##		n_eff	Rhat				
##	u[1]	2189.448	1.002057				
##	u[2]	2201.333	1.002860				
##	u[3]	2484.646	1.001715				
##	u[4]	2504.028	1.001143				
##	u[5]	2495.589	1.001515				
##	u[6]	2906.302	1.004522				
##	u[7]	2213.612	1.001427				



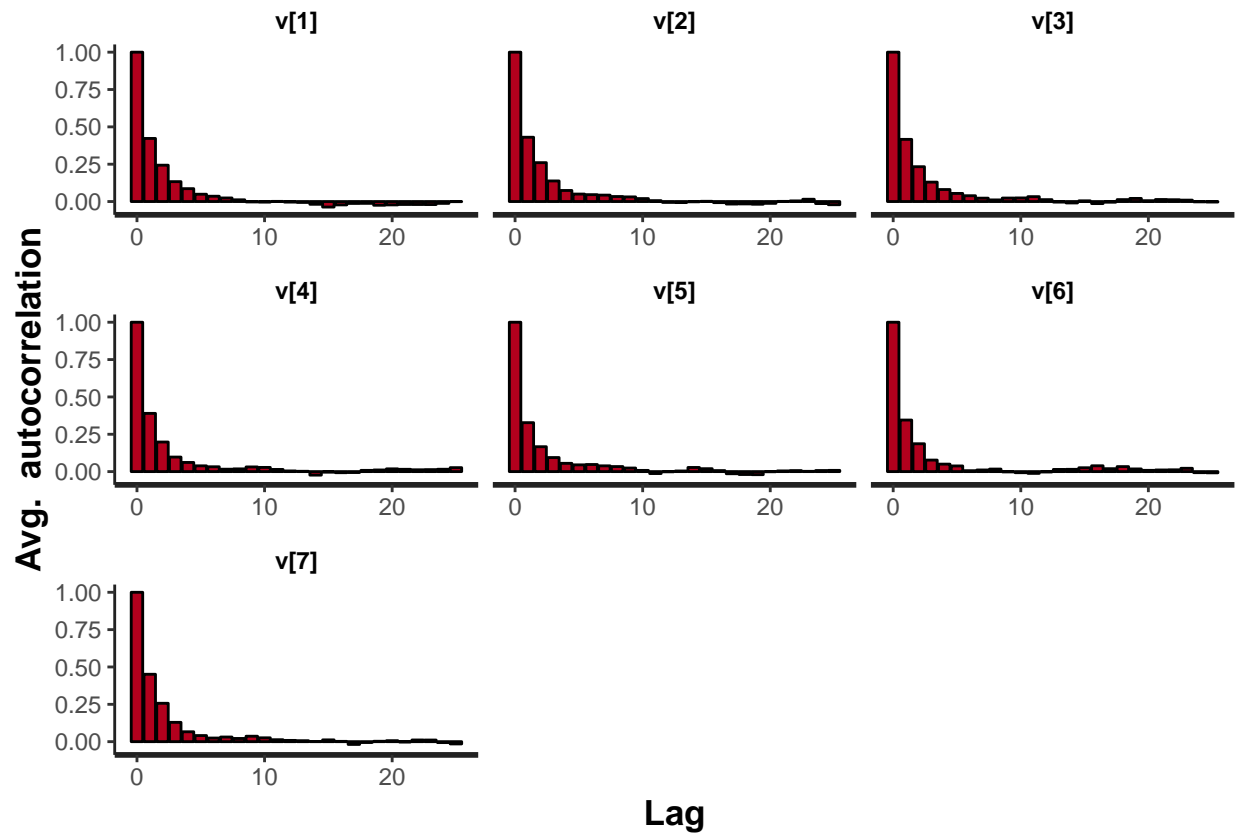


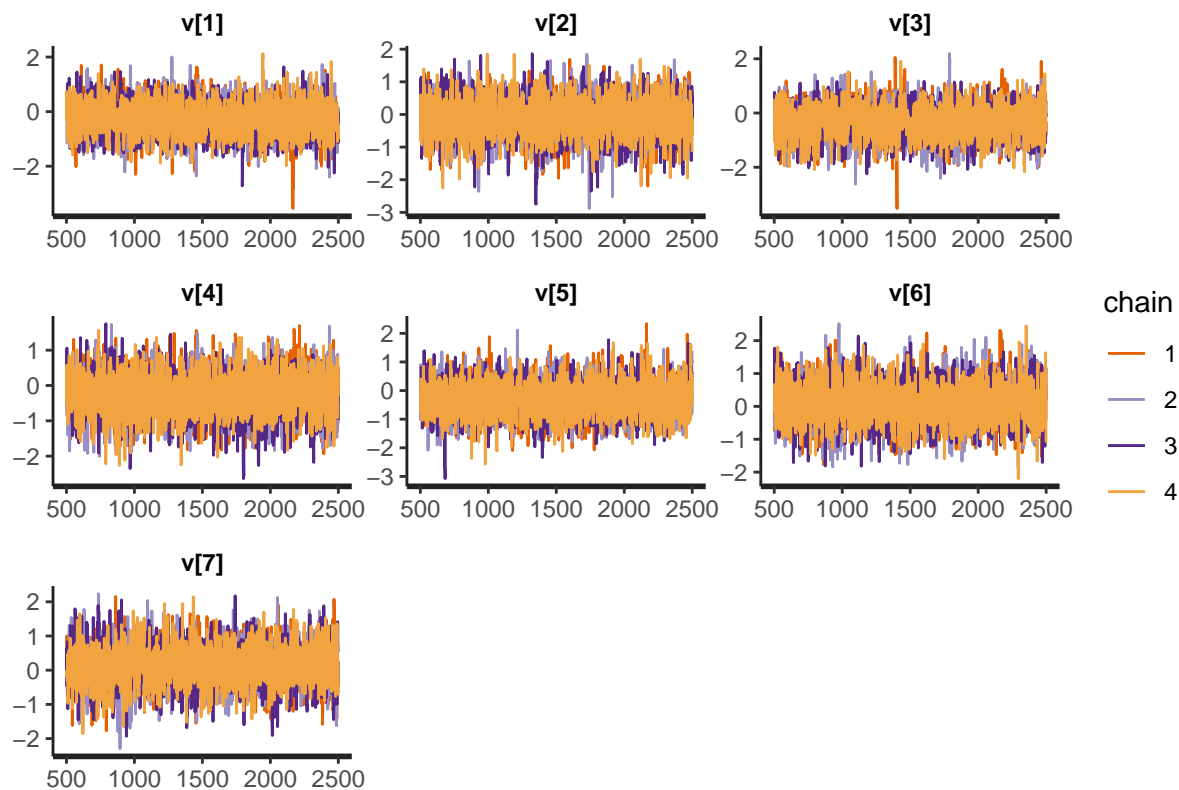
##	mean	se_mean	sd	25%	50%	75%
## v[1]	-0.2608380	0.01129872	0.5792088	-0.6232146	-0.2496927	0.100281238
## v[2]	-0.1418966	0.01160104	0.5689952	-0.4981460	-0.1310038	0.226428819

```

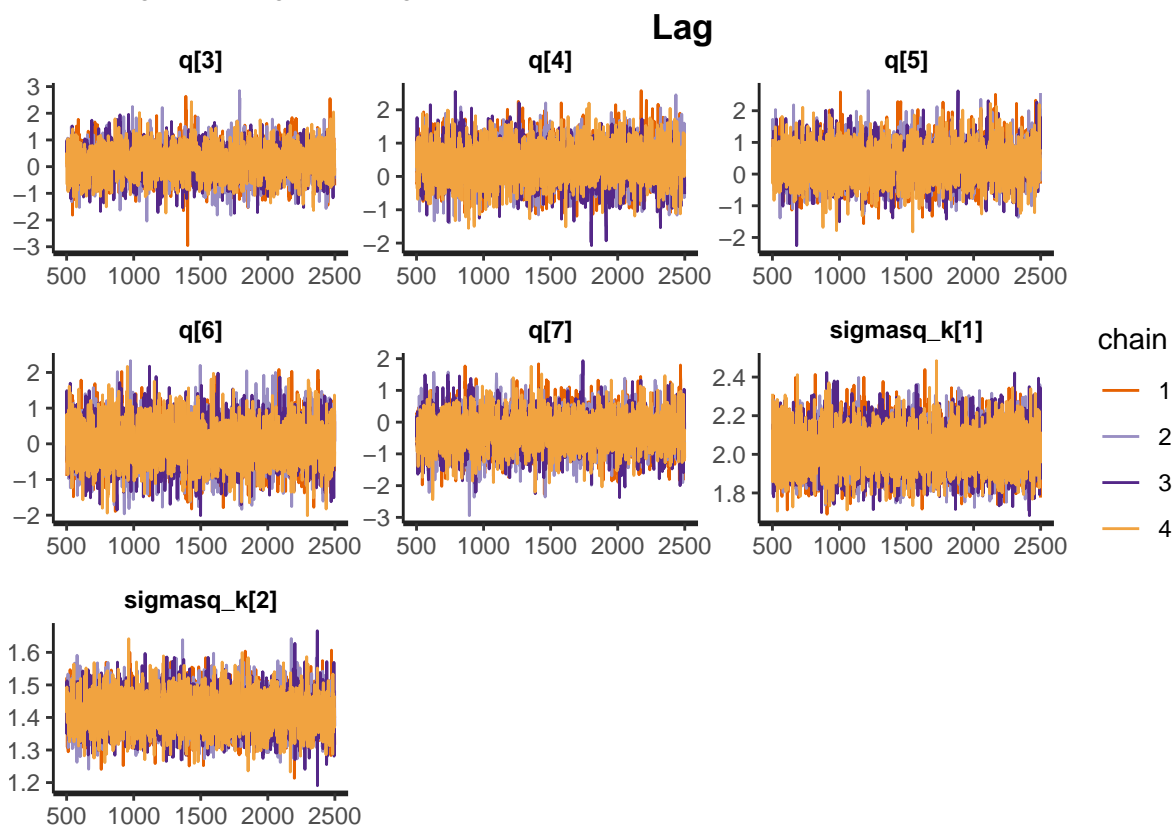
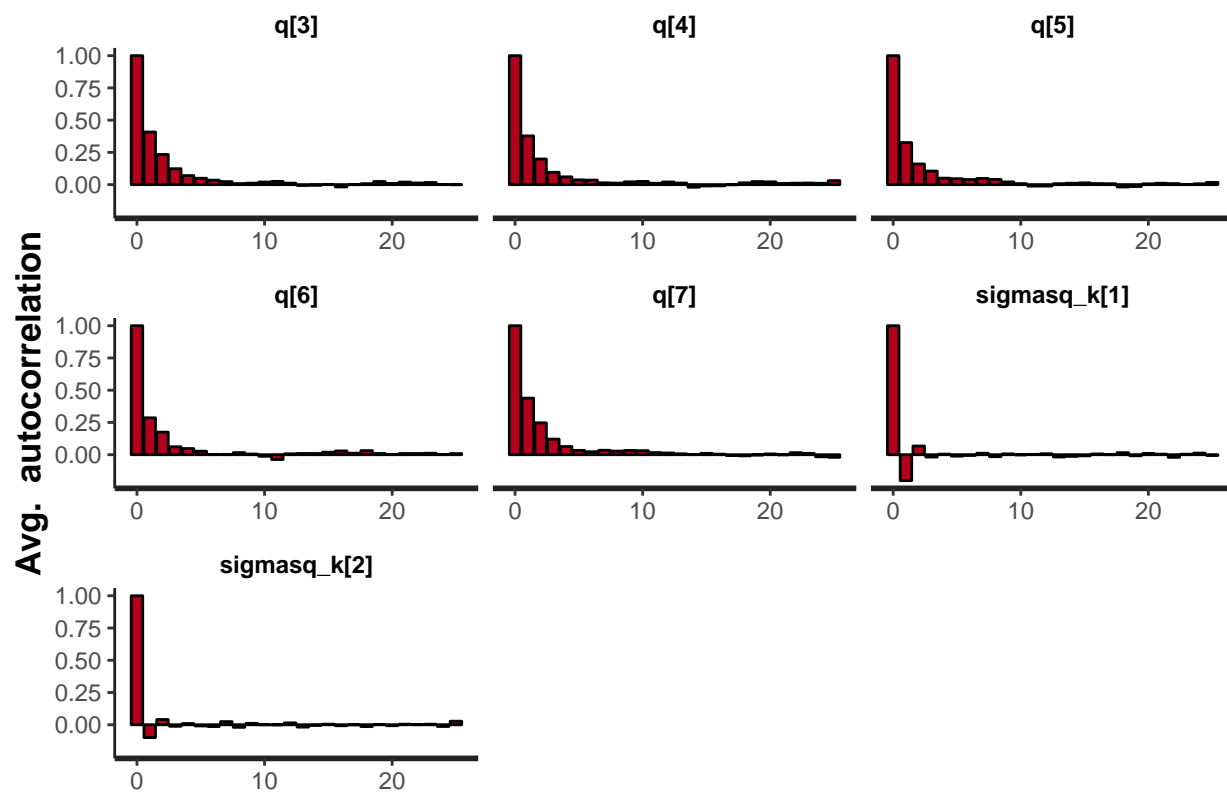
## v[3] -0.3526331 0.01122603 0.5652307 -0.7103787 -0.3492257 0.008732518
## v[4] -0.2645598 0.01104952 0.5635386 -0.6183355 -0.2562956 0.111397571
## v[5] -0.2958225 0.01036844 0.5632586 -0.6592896 -0.3043598 0.063933442
## v[6] 0.1089564 0.01008438 0.5759614 -0.2715071 0.1096889 0.485380135
## v[7] 0.1324663 0.01149226 0.5650412 -0.2314916 0.1320214 0.498668213
##      n_eff      Rhat
## v[1] 2627.917 1.001606
## v[2] 2405.601 1.002077
## v[3] 2535.122 1.000617
## v[4] 2601.121 1.006429
## v[5] 2951.136 1.000614
## v[6] 3262.035 1.000778
## v[7] 2417.406 1.000907

```



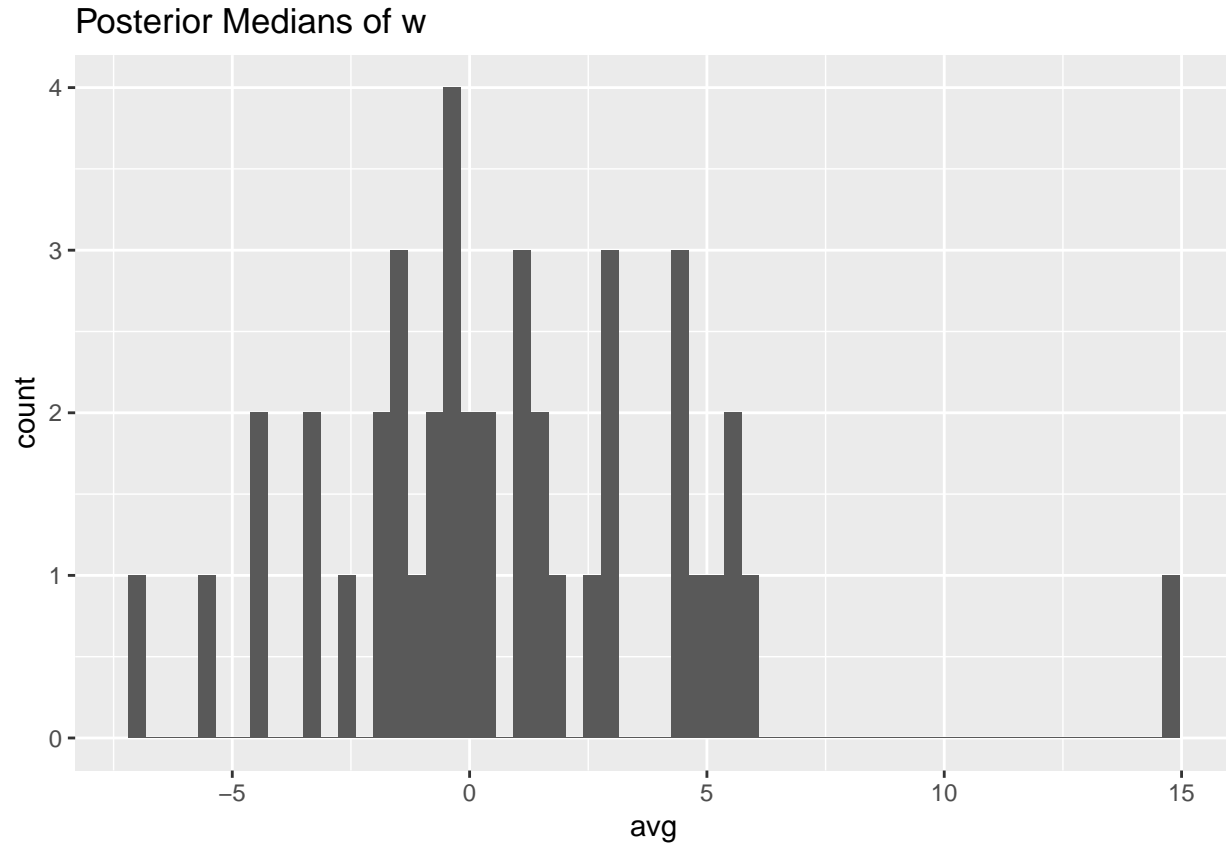


##		mean	se_mean	sd	25%	50%
##	q[3]	0.16956330	0.0110319519	0.56717598	-0.19309145	0.17599109
##	q[4]	0.41768884	0.0109319451	0.56253214	0.05693179	0.42279231
##	q[5]	0.35139701	0.0104119048	0.56557436	-0.01434797	0.34435449
##	q[6]	0.01684901	0.0097722515	0.58894465	-0.37323704	0.01252318
##	q[7]	-0.30458056	0.0113500339	0.56419677	-0.66923661	-0.30432573
##	sigma <sub>sq</sub> _k[1]	2.02460978	0.0010253585	0.10967399	1.95020702	2.02130898
##	sigma <sub>sq</sub> _k[2]	1.41507614	0.0005811733	0.05623213	1.37676433	1.41370003
##		75%	n_eff	Rhat		
##	q[3]	0.52665667	2643.206	1.0005014		
##	q[4]	0.78950071	2647.889	1.0062464		
##	q[5]	0.71257767	2950.659	1.0005779		
##	q[6]	0.40128437	3632.116	1.0011634		
##	q[7]	0.05875604	2470.967	1.0007789		
##	sigma <sub>sq</sub> _k[1]	2.09724127	11440.784	1.0000940		
##	sigma <sub>sq</sub> _k[2]	1.45183018	9361.765	0.9998693		

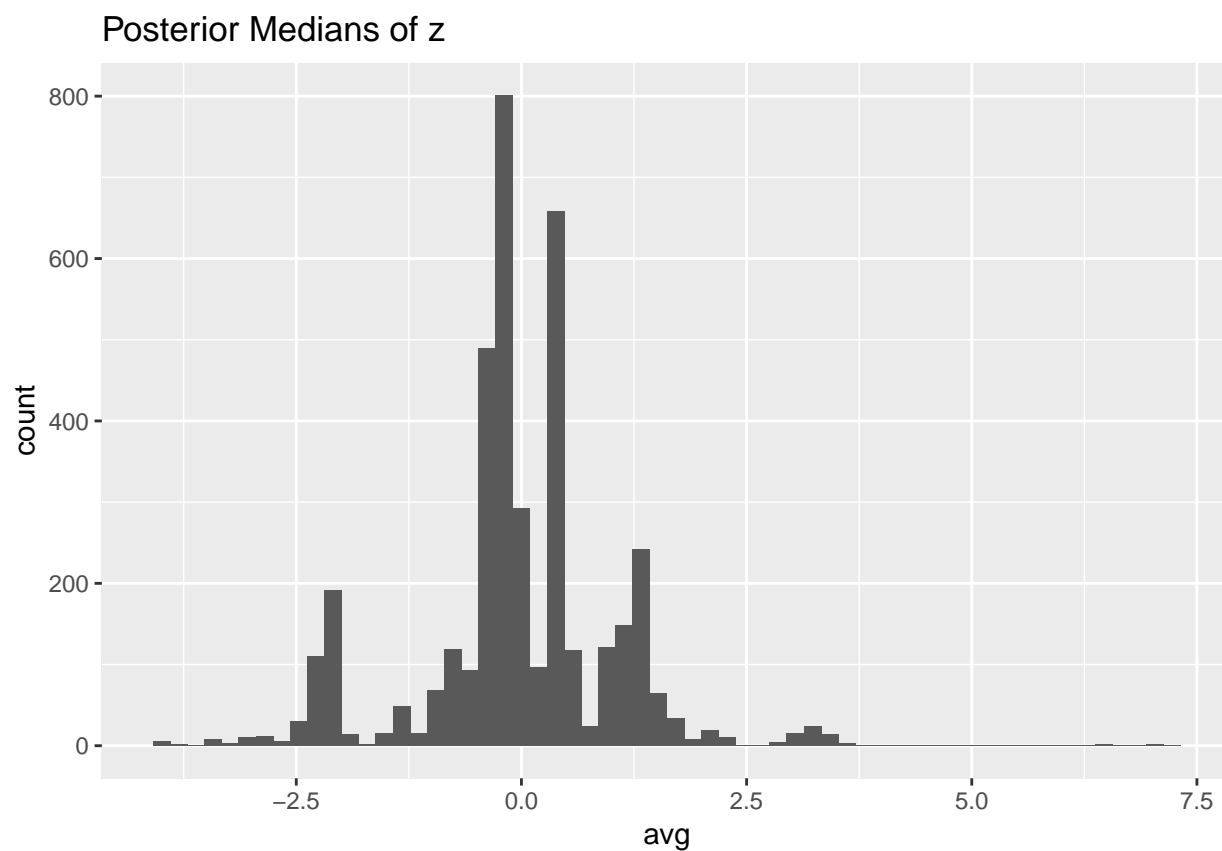


```
## [1] "Summary statistics for posterior medians of w"
##      avg
## Min.   :-7.1460
```

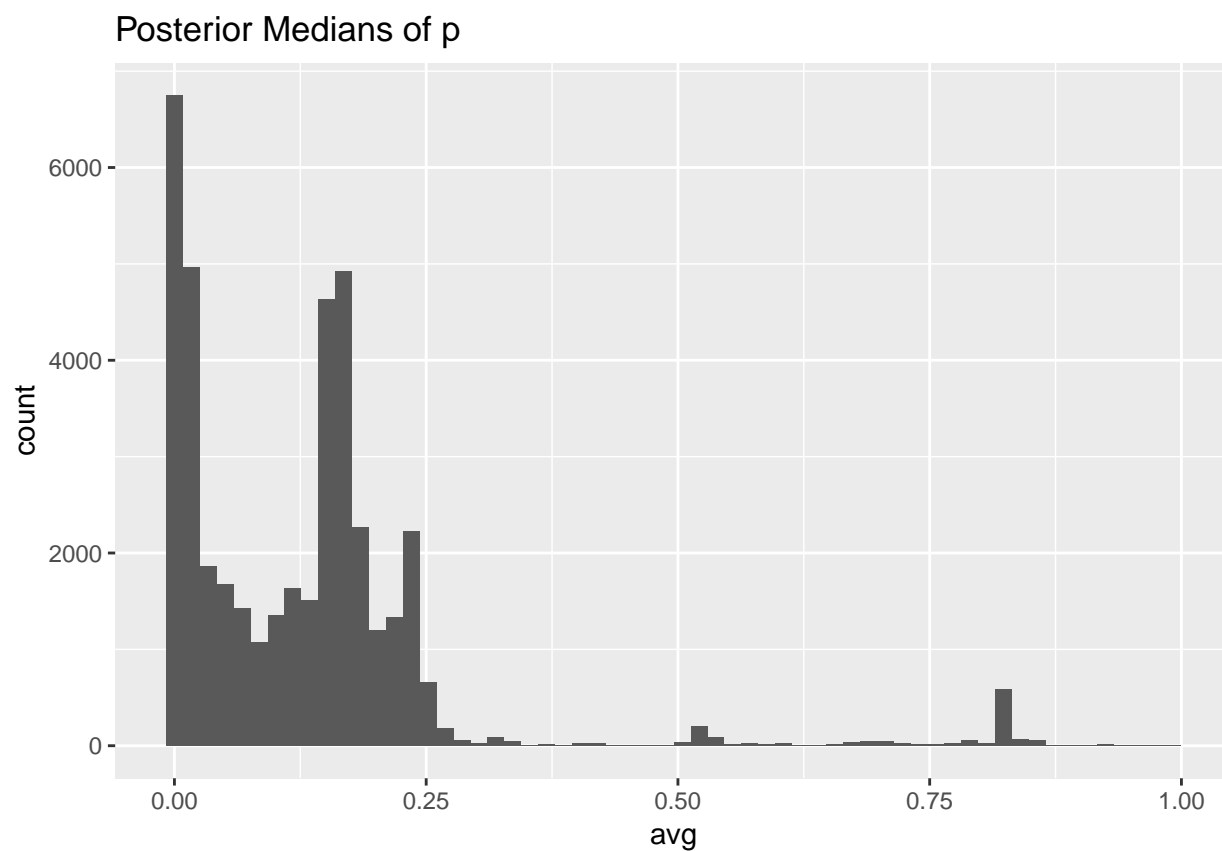
```
## 1st Qu.: -1.3850
## Median : 0.1449
## Mean   : 0.7491
## 3rd Qu.: 2.9871
## Max.   : 14.6435
```



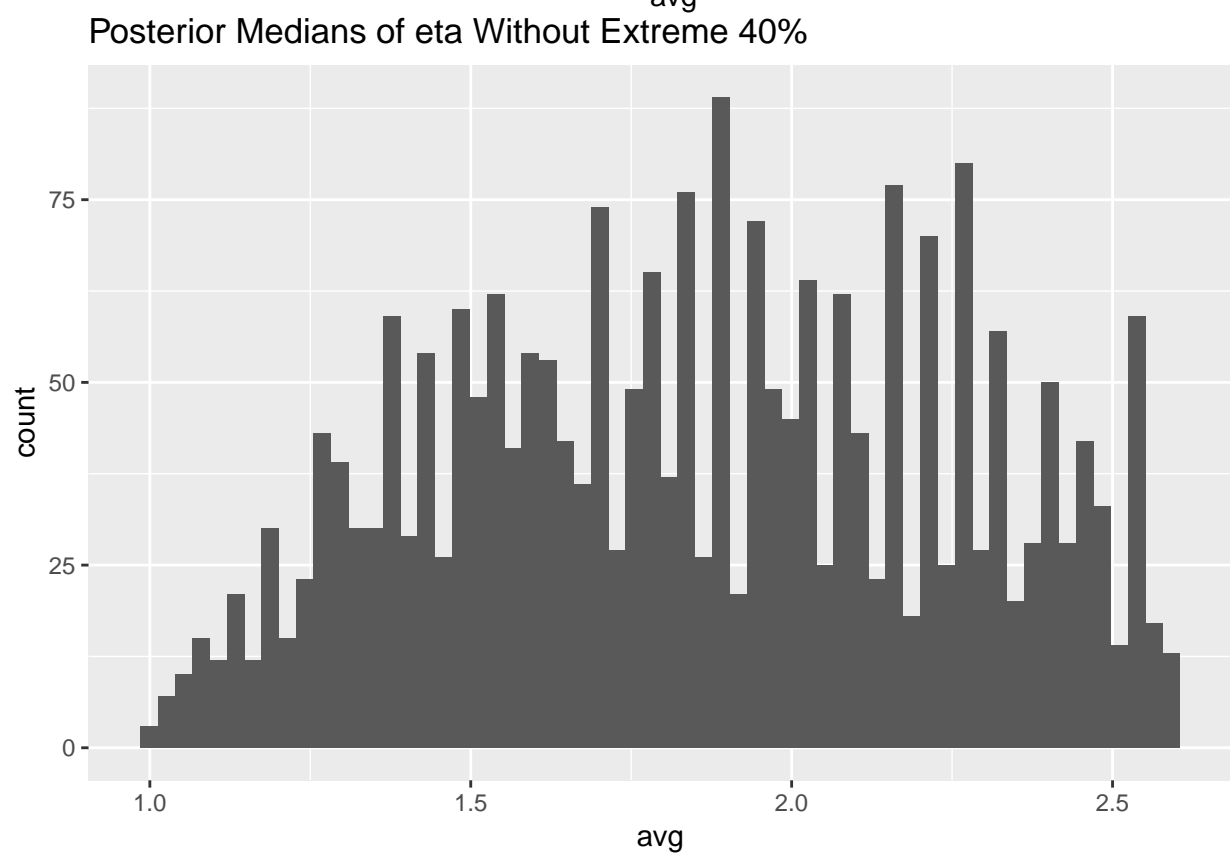
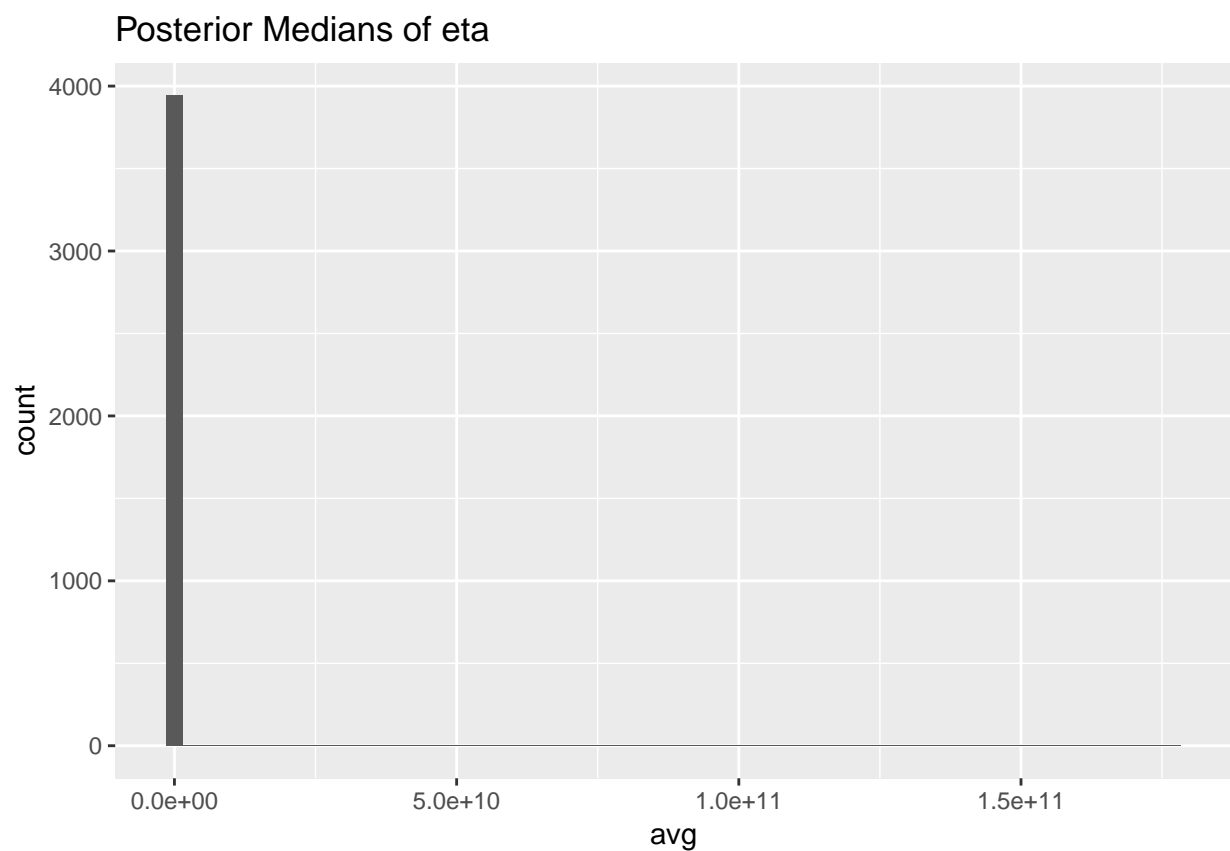
```
## [1] "Summary statistics for posterior medians of z"
##      avg
## Min.   : -4.00616
## 1st Qu.: -0.32516
## Median : -0.12500
## Mean    : -0.03138
## 3rd Qu.:  0.44105
## Max.    :  7.21230
```



```
## [1] "Summary statistics for posterior medians of p"
##      avg
##  Min.   :0.0000001
## 1st Qu.:0.0174563
##  Median :0.1257900
##   Mean  :0.1283479
## 3rd Qu.:0.1728865
##   Max.  :0.9915268
```

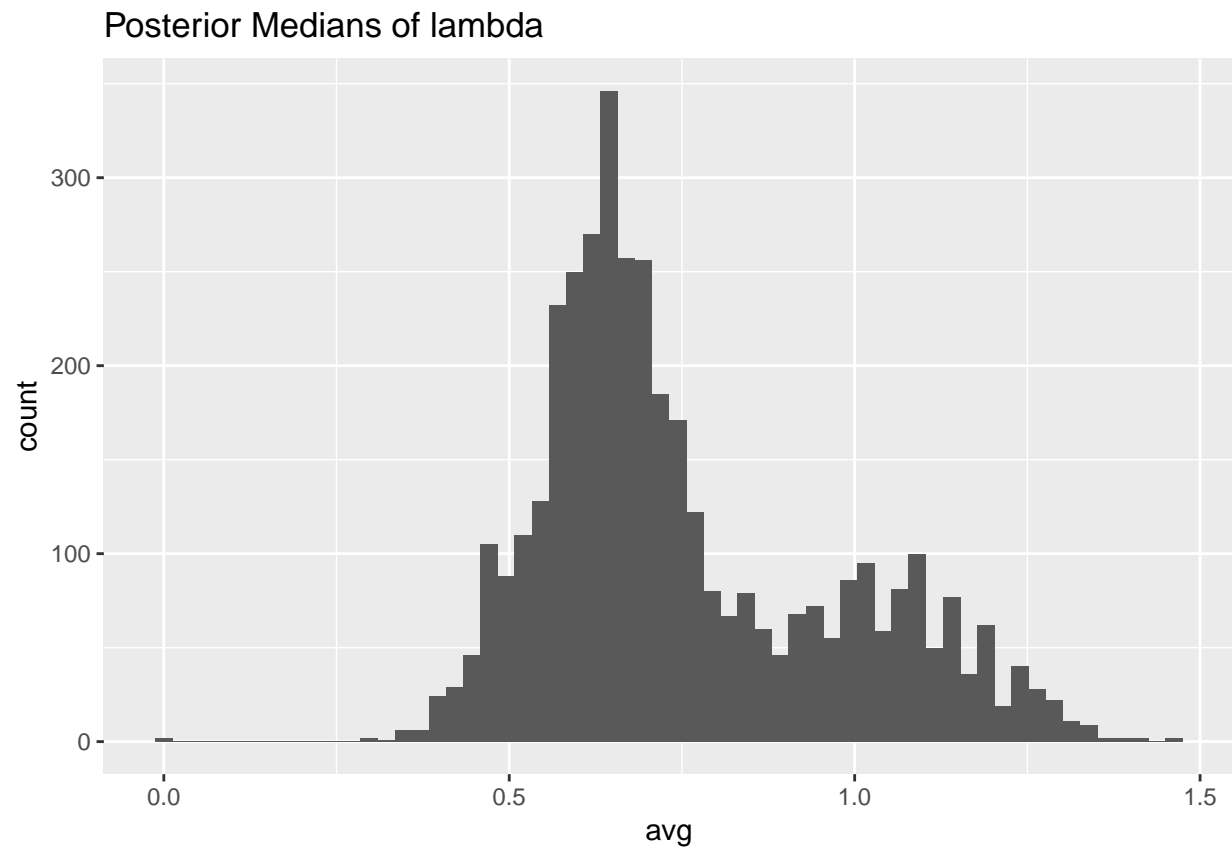


```
## [1] "Summary statistics for posterior medians of eta"
##      avg
##  Min.   :1.000e+00
## 1st Qu.:2.000e+00
##  Median :2.000e+00
##   Mean   :7.493e+07
## 3rd Qu.:3.000e+00
##   Max.   :1.769e+11
```

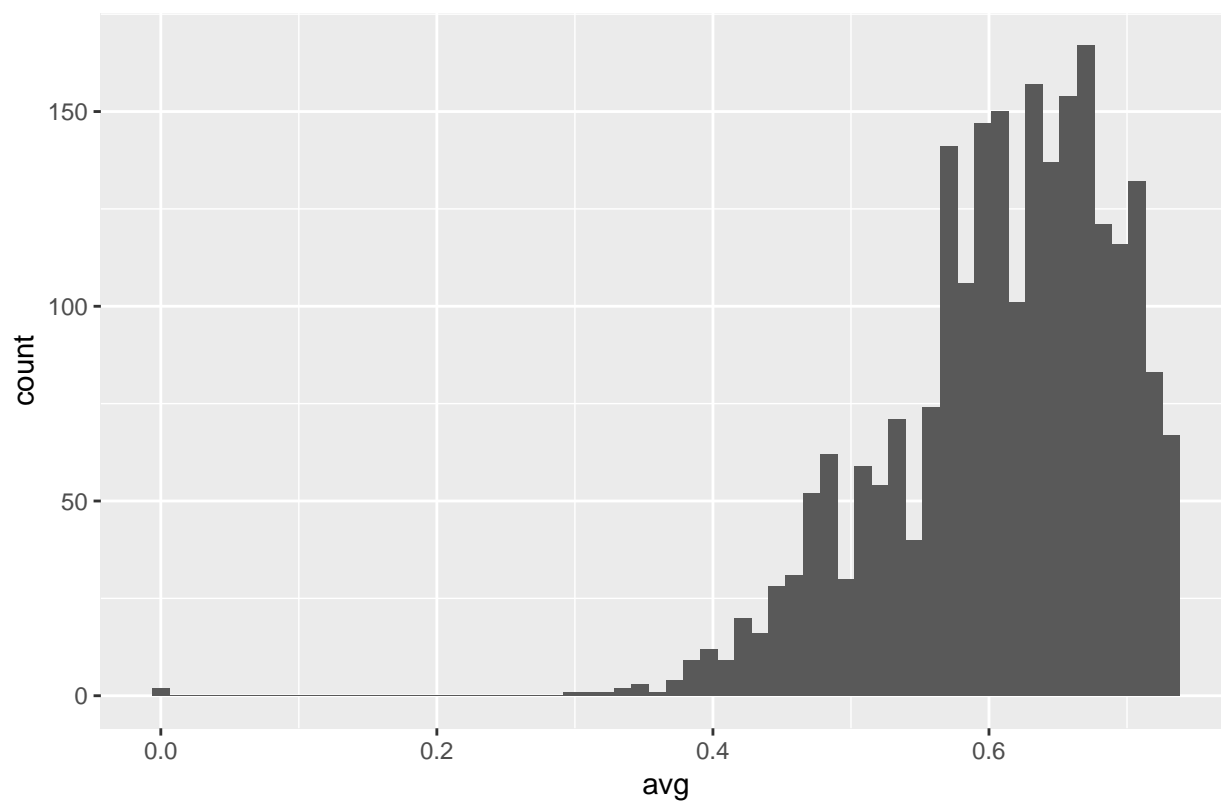




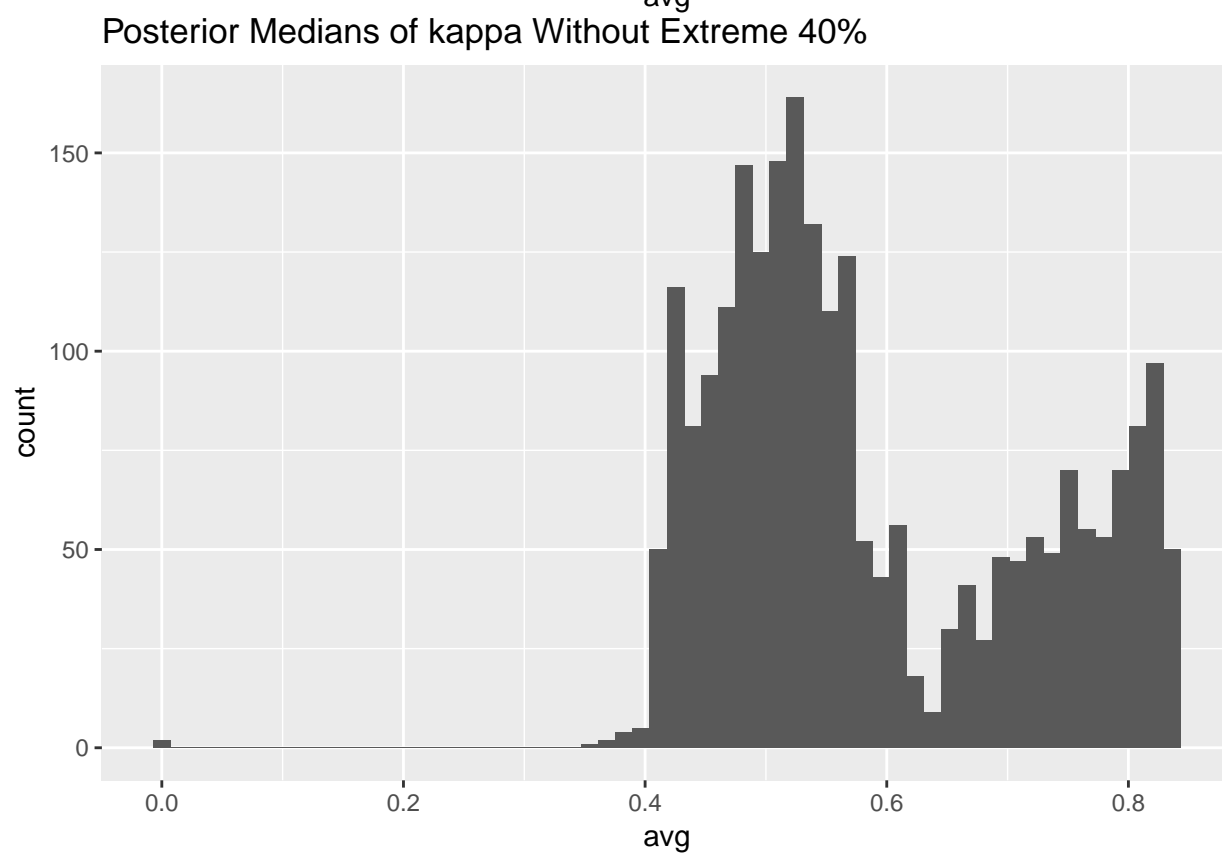
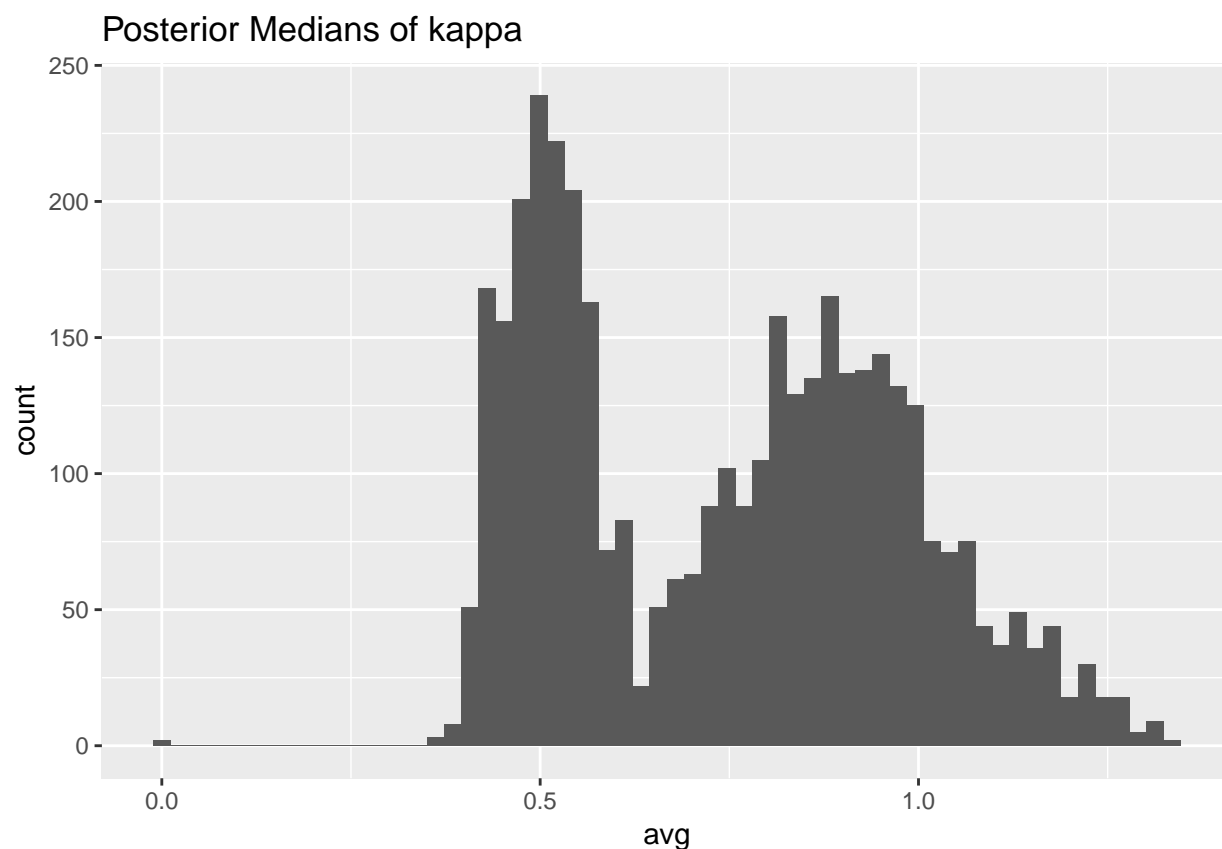
```
## [1] "Summary statistics for posterior medians of lambda"
##      avg
##  Min.   :0.000004
## 1st Qu.:0.603693
## Median :0.688819
## Mean   :0.754512
## 3rd Qu.:0.896007
##  Max.   :1.462525
```



Posterior Medians of lambda Without Extreme 40%



```
## [1] "Summary statistics for posterior medians of kappa"
##      avg
##  Min.   :0.0000434
## 1st Qu.:0.5245244
##  Median :0.7600591
##   Mean  :0.7479723
## 3rd Qu.:0.9305208
##   Max.  :1.3353141
```



## Identifying Parameters with Large Rhats

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

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
## 0.9995  1.0003  1.0010  1.0185  1.0046  3.2069     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
```