Homework2

Question1

a. Set true values for the model parameters. Generate data from the model and conduct Bayesian analysis on the basis of 10 replication.

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \\ y_{i5} \\ y_{i6} \\ y_{i7} \\ y_{i8} \\ y_{i9} \end{bmatrix} = \begin{bmatrix} \mu_1 & a_1 \\ \mu_2 & a_2 \\ \mu_3 & a_3 \\ \mu_4 & a_4 \\ \mu_5 & a_5 \\ \mu_6 & a_6 \\ \mu_7 & a_7 \\ \mu_8 & a_8 \\ \mu_9 & a_9 \end{bmatrix} \begin{bmatrix} 1 \\ c_i \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & \lambda_{62} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda_{83} \\ 0 & 0 & \lambda_{93} \end{bmatrix} \begin{bmatrix} \eta_i \\ \xi_{i1} \\ \xi_{i2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \\ \varepsilon_{i4} \\ \varepsilon_{i5} \\ \varepsilon_{i6} \\ \varepsilon_{i7} \\ \varepsilon_{i8} \\ \varepsilon_{i9} \end{bmatrix}$$

$$\eta_i = bd_i + \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \end{bmatrix} \begin{bmatrix} \xi_{i1} \\ \xi_{i2} \\ \xi_{i1}\xi_{i2} \\ \xi_{i2}^2 \end{bmatrix} + \delta_i$$

The true values of parameters set for this question are listed as follow, and 10 data sets are generated based on the true parameters.

```
mu <- c(3.0, 1.5, 2.0, 1.0, 2.5, 1.8, 3.2, 2.3, 2.8)
a <- c(0.8, 0.6, 0.7, 0.5, 0.9, 0.8, 1.0, 0.9, 0.7)
lambda <- c(0.6, 0.7, 0.4, 0.5, 0.8, 0.6) ##lambda_21, 31, 52, 62, 83, 93
b <- 1.2
gamma <- c(0.4, 0.6, 0.2, 0.3)
```

b. Demonstrate how to check convergence of the model.

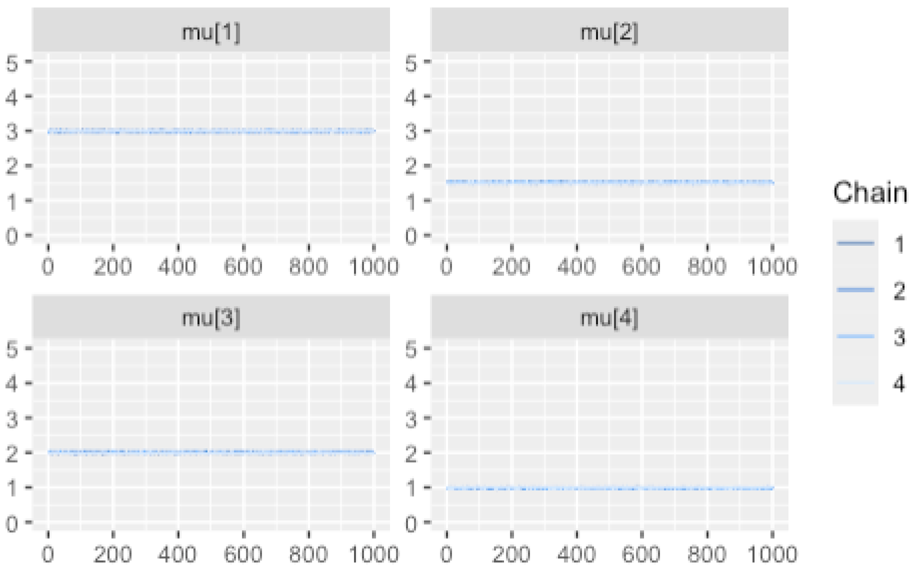
1. Check the Rhat of the 10 replications. If Rhat is close to 1, then the model converges well, otherwise it does not converge. In the following 10 replications, all estimations converge

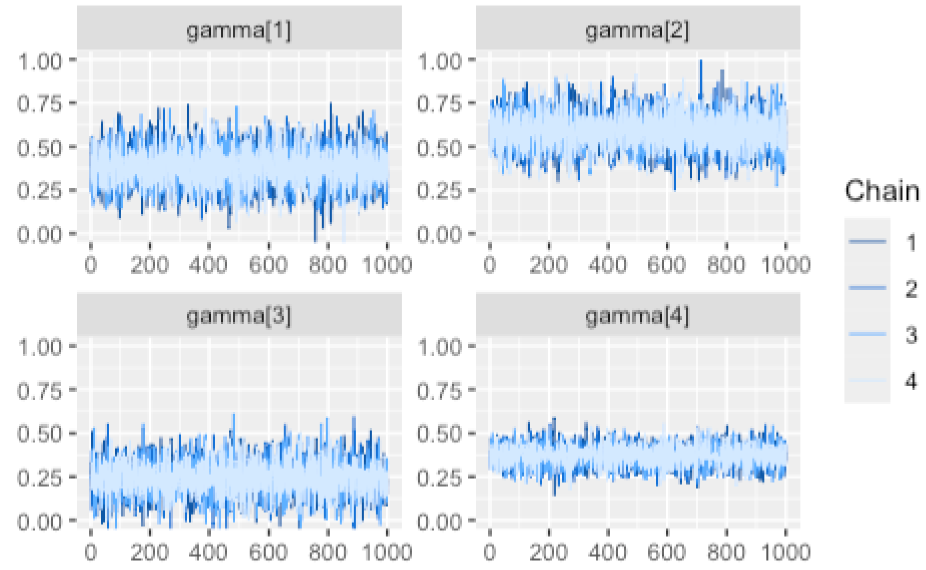
```
[1] "Rhat values:"
       mu[1]        mu[2]        mu[3]        mu[4] 
   1.0003066    0.9991804    0.9994207    0.9998072 
       mu[5]        mu[6]        mu[7]        mu[8] 
   0.9993043    0.9998817    1.0000675    0.9996022 
       mu[9]            b         a[1]         a[2] 
   1.0006596    1.0003084    0.9999338    0.9998860 
        a[3]         a[4]         a[5]         a[6] 
   0.9997470    0.9993103    0.9991942    0.9997591 
        a[7]         a[8]         a[9]     gamma[1] 
   0.9993286    0.9994857    0.9997709    0.9992572 
    gamma[2]     gamma[3]     gamma[4]    lambda[1] 
   0.9994981    0.9993308    1.0003887    0.9995764 
   lambda[2]    lambda[3]    lambda[4]    lambda[5] 
   1.0006462    0.9992140    0.9998694    0.9995630 
   lambda[6]    sigma_eta sigma_eps[1] sigma_eps[2] 
   1.0006977    0.9997413    1.0007863    0.9993388 
sigma_eps[3] sigma_eps[4] sigma_eps[5] sigma_eps[6] 
   0.9997111    0.9995392    0.9994385    1.0002183 
sigma_eps[7] sigma_eps[8] sigma_eps[9]         lp__ 
   0.9997375    0.9998503    0.9997121    1.0002380 
```

with Rhat close to 1.

2. Check the estimation process plot of chains. If different chains meet together as the iteration number grows, the model converges.

   Below are two figures depict the μ and γ estimation process of 4 chains. We can see that the model converges.





c. Use Bias and RMSE to summarize the estimation results.

The mean of posterior means, estimation bias and RMSE are listed below.

```
Mean of Posterior Means:
      mu[1]        mu[2]        mu[3]        mu[4]        mu[5]        mu[6] 
  2.99599887   1.51269347   1.99812651   0.98542591   2.51365282   1.79848541 
      mu[7]        mu[8]        mu[9]         a[1]         a[2]         a[3] 
  3.20854801   2.30006561   2.79283132   0.79435570   0.59340306   0.70077899 
       a[4]         a[5]         a[6]         a[7]         a[8]         a[9] 
  0.49612216   0.90651647   0.78699512   1.02011745   0.89459609   0.71504154 
  lambda[1]    lambda[2]    lambda[3]    lambda[4]    lambda[5]    lambda[6] 
  0.59111642   0.69938689   0.38321140   0.50732136   0.79826989   0.60951828 
          b     gamma[1]     gamma[2]     gamma[3]     gamma[4]    sigma_eta 
  1.05448711   0.38383597   0.58993826   0.24260778   0.36911443   1.00995737 
sigma_eps[1] sigma_eps[2] sigma_eps[3] sigma_eps[4] sigma_eps[5] sigma_eps[6] 
  0.09531023   0.10658906   0.10431425   0.10464561   0.10505293   0.10105482 
sigma_eps[7] sigma_eps[8] sigma_eps[9] 
  0.10106576   0.09477811   0.10250181 

Bias:
      mu_bias          <NA>          <NA>          <NA>          <NA>          <NA> 
-4.001129e-03  1.269347e-02 -1.873486e-03 -1.457409e-02  1.365282e-02 -1.514590e-03 
         <NA>          <NA>          <NA>        a_bias          <NA>          <NA> 
 8.548006e-03  6.561482e-05 -7.168682e-03 -5.644302e-03 -6.596941e-03  7.789859e-04 
         <NA>          <NA>          <NA>          <NA>          <NA>          <NA> 
-3.877837e-03  6.516474e-03 -1.300488e-02  2.011745e-02 -5.403908e-03  1.504154e-02 
  lambda_bias          <NA>          <NA>          <NA>          <NA>          <NA> 
-8.883575e-03 -6.131139e-04 -1.678860e-02  7.321359e-03 -1.730107e-03  9.518277e-03 
       b_bias    gamma_bias          <NA>          <NA>          <NA> 
-1.455129e-01 -1.616403e-02 -1.006174e-02  4.260778e-02  6.911443e-02 
```

```
RMSE:
         mu           a      lambda           b       gamma 
0.008858162 0.010342874 0.009204564 0.145512894 0.041697547 
```

d. Show your prior inputs and check whether the Bayesian analysis is sensitive to the inputs.

My prior inputs are as follows:

```
mu ~ normal(0, 5);
a ~ normal(0, 5);
lambda ~ normal(0, 5);
b ~ normal(0, 5);
gamma ~ normal(0, 5);
sigma_eta ~ cauchy(0, 5);
sigma_eps ~ cauchy(0, 5);
```

To check whether the Bayesian analysis is sensitive to the inputs, we modify the prior inputs to:

```
mu ~ normal(10, 5);
a ~ normal(10, 5);
lambda ~ normal(10, 5);
b ~ normal(10, 5);
gamma ~ normal(10, 5);
sigma_eta ~ cauchy(10, 5);
sigma_eps ~ cauchy(10, 5);
```

The model with new priori is in `HW3m2.stan` file. The metrics are listed below:

```
[1] "Rhat values:"
                   [,1]
mu[1]         0.9997265
mu[2]         0.9995139
mu[3]         0.9999071
mu[4]         0.9995851
mu[5]         0.9998282
mu[6]         1.0000129
mu[7]         1.0000968
mu[8]         0.9995017
mu[9]         1.0000093
b             0.9997704
a[1]          0.9994559
a[2]          0.9995017
a[3]          0.9993746
a[4]          0.9993503
a[5]          0.9998931
a[6]          0.9992300
a[7]          0.9993889
a[8]          0.9998872
a[9]          0.9994931
gamma[1]      0.9995667
gamma[2]      0.9998601
gamma[3]      0.9993839
gamma[4]      0.9994438
lambda[1]     0.9996634
lambda[2]     0.9991330
lambda[3]     0.9998159
lambda[4]     0.9997074
lambda[5]     0.9993279
lambda[6]     0.9996253
sigma_eta     0.9999720
sigma_eps[1]  0.9992682
sigma_eps[2]  0.9996428
sigma_eps[3]  0.9995294
sigma_eps[4]  1.0003774
sigma_eps[5]  0.9999163
sigma_eps[6]  0.9997111
sigma_eps[7]  0.9992756
sigma_eps[8]  0.9996577
sigma_eps[9]  1.0003259
lp__          1.0019579
```

```
Mean of Posterior Means:
       mu[1]         mu[2]         mu[3]         mu[4]         mu[5]         mu[6] 
  2.99654666    1.51280996    1.99842654    0.98537470    2.51347397    1.79854475 
       mu[7]         mu[8]         mu[9]          a[1]          a[2]          a[3] 
  3.20882184    2.30029141    2.79294753    0.79448627    0.59358515    0.70094235 
        a[4]          a[5]          a[6]          a[7]          a[8]          a[9] 
  0.49633075    0.90624001    0.78711292    1.02006092    0.89490626    0.71481500 
   lambda[1]     lambda[2]     lambda[3]     lambda[4]     lambda[5]     lambda[6] 
  0.59111281    0.69928764    0.38326455    0.50756424    0.79827083    0.60955667 
           b      gamma[1]      gamma[2]      gamma[3]      gamma[4]     sigma_eta 
  1.05924152    0.38484313    0.59069815    0.24759010    0.37119787    1.00905763 
sigma_eps[1]  sigma_eps[2]  sigma_eps[3]  sigma_eps[4]  sigma_eps[5]  sigma_eps[6] 
  0.09545608    0.10662235    0.10446003    0.10474832    0.10532052    0.10124413 
sigma_eps[7]  sigma_eps[8]  sigma_eps[9] 
  0.10105891    0.09480047    0.10237054 

Bias:
      mu_bias          <NA>          <NA>          <NA>          <NA>          <NA> 
-0.0034533367  0.0128099576 -0.0015734569 -0.0146252963  0.0134739736 -0.0014552481 
         <NA>          <NA>          <NA>        a_bias          <NA>          <NA> 
 0.0088218376  0.0002914073 -0.0070524667 -0.0055137291 -0.0064148548  0.0009423534 
         <NA>          <NA>          <NA>          <NA>          <NA>          <NA> 
-0.0036692514  0.0062400104 -0.0128870846  0.0200609169 -0.0050937370  0.0148149961 
  lambda_bias          <NA>          <NA>          <NA>          <NA>          <NA> 
-0.0088871914 -0.0007123627 -0.0167354532  0.0075642429 -0.0017291679  0.0095566716 
       b_bias    gamma_bias          <NA>          <NA>          <NA> 
-0.1407584817 -0.0151568739 -0.0093018546  0.0475900995  0.0711978740 


RMSE:
         mu           a      lambda           b       gamma 
0.008842501 0.010213218 0.009229507 0.140758482 0.043732740 
```

From the Rhat, bias, and RMSE of the estimation, we can conclude that the convergence and estimation precision are not affected by the different prior setting.

Question2

a. Compare the non-linear SEM in Q1 with its linear SEM counterpart.

The new non-linear model is in `HW3m3.stan` file. After we construct the linear SEM model, the metric is listed below:

```
Mean of Posterior Means:
      mu[1]        mu[2]        mu[3]        mu[4]        mu[5]        mu[6]        mu[7]        mu[8]        mu[9]         a[1]         a[2] 
 2.99633912   1.51260998   1.99823250   0.98546270   2.51360168   1.79817508   3.20876026   2.30019409   2.79304622   0.79401865   0.59348622 
       a[3]         a[4]         a[5]         a[6]         a[7]         a[8]         a[9]    lambda[1]    lambda[2]    lambda[3]    lambda[4] 
 0.70135169   0.49607844   0.90606369   0.78704690   1.01968174   0.89478929   0.71507625   0.59116664   0.69925810   0.38303366   0.50735144 
  lambda[5]    lambda[6]            b     gamma[1]     gamma[2]     gamma[3]     gamma[4]    sigma_eta sigma_eps[1] sigma_eps[2] sigma_eps[3] 
 0.79816521   0.60963181   1.07227272   0.42154744   0.56402264  -0.08881248  -0.02170741   1.21852128   0.09551678   0.10644530   0.10443322 
sigma_eps[4] sigma_eps[5] sigma_eps[6] sigma_eps[7] sigma_eps[8] sigma_eps[9] 
 0.10452039   0.10508377   0.10109931   0.10117860   0.09475222   0.10231530 

Bias:
     mu_bias          <NA>          <NA>          <NA>          <NA>          <NA>          <NA>          <NA>          <NA>        a_bias 
-0.0036608815  0.0126099771 -0.0017675044 -0.0145372983  0.0136016828 -0.0018249183  0.0087602595  0.0001940936 -0.0069537775 -0.0059813514 
        <NA>          <NA>          <NA>          <NA>          <NA>          <NA>          <NA>          <NA>   lambda_bias          <NA> 
-0.0065137824  0.0013516861 -0.0039215553  0.0060636891 -0.0129531027  0.0196817373 -0.0052107147  0.0150762520 -0.0088333620 -0.0007419049 
        <NA>          <NA>          <NA>          <NA>        b_bias    gamma_bias          <NA>          <NA>          <NA> 
-0.0169663425  0.0073514444 -0.0018347909  0.0096318146 -0.1277272822  0.0215474362 -0.0359773611 -0.2888124787 -0.3217074121 
RMSE:
         mu           a      lambda           b       gamma 
0.008821323 0.010229470 0.009279166 0.127727282 0.217179056 
```

Compare Bayes factor and DIC

The average DIC of the non-linear model is 7756.662, and the average DIC of the linear model is 7443.317. The average Bayes factor of non-linear vs. linear is 44.16.

This indicates that the Bayes factor and DIC are in favor of the original nonlinear model since the Bayes factor is pretty large.

b.Compare the non-linear SEM in Q1 with this new model.

The new non-linear model is in `HW3m4.stan` file. The Bayes factor is 10.2, and the average DIC of the new model is 7749.231. This indicates that the Bayes factor and the DIC are in favor of the original nonlinear model since the Bayes factor is large.

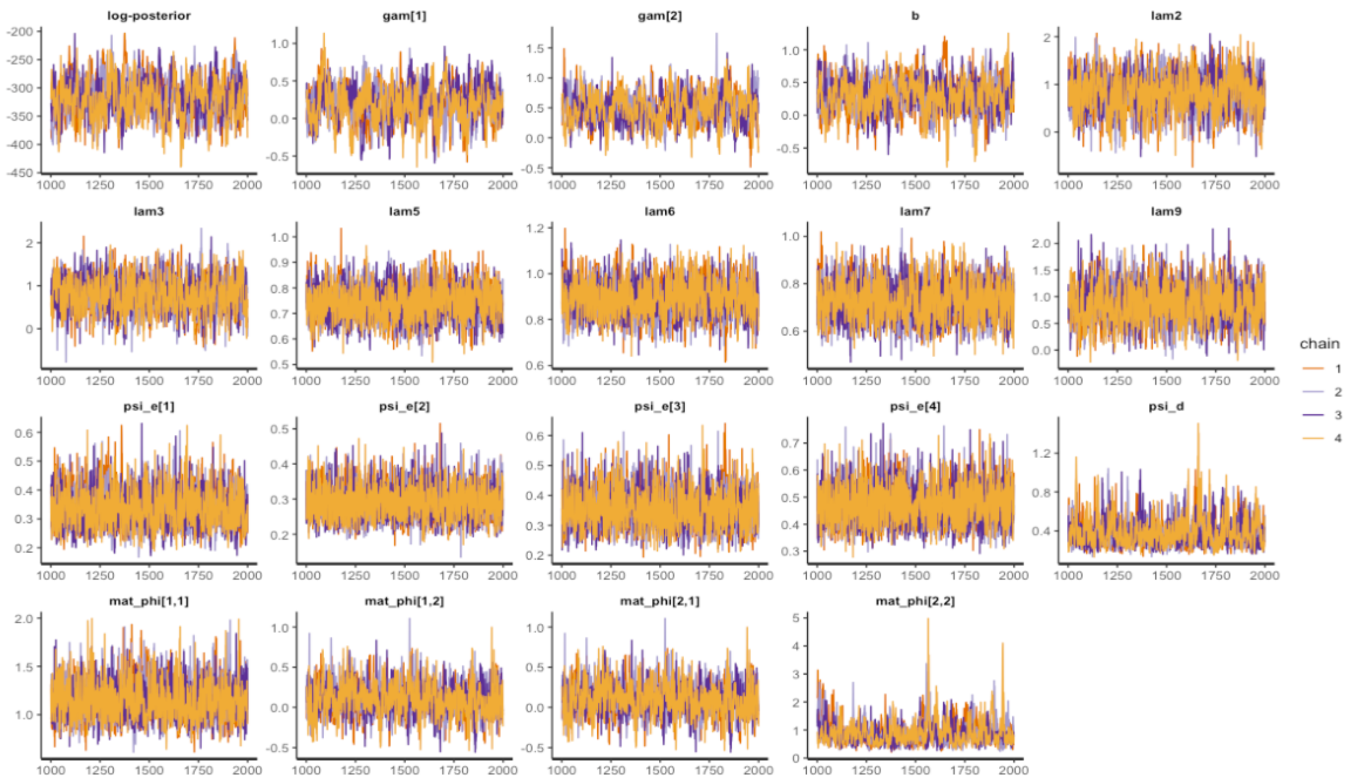
Question3

The true parameters are:

```
b <- 0.3
vec_lambda <- c(1, 0.8, 0.8, 1, 0.7, 0.9, 0.7, 1, 0.8)
vec_gamma <- c(0.2, 0.5)
mat_phi <- matrix(data = c(1, 0.2, 0.2, 0.81), ncol = 2)
vec_psi_eps <- c(1, 1, 1, 0.3, 0.3, 0.4, 0.4)
psi_delta <- 0.36
```

The posterior mean of parameters are:

gam.1	0.200777348	0.201739336	0.213630032	0.195693258	0.191822736	0.186883513	0.201788412	0.1947278005	0.205011814	0.210195752
gam.2	0.502626705	0.479065640	0.531059019	0.514092297	0.519701289	0.501389707	0.475886528	0.4923400927	0.512867401	0.507685193
b	0.290365838	0.318682542	0.311309871	0.283396489	0.303909810	0.302651564	0.302628083	0.3036206316	0.266221780	0.295658072
mu.1	0.009366186	0.018423184	-0.002927846	-0.014512741	-0.012414878	0.001485534	0.018452489	-0.0098526984	-0.013287401	-0.003924103
mu.2	-0.010753556	0.006709129	0.002733055	-0.019087272	-0.002872692	0.004786096	0.024069198	-0.0008994255	-0.010539082	-0.009683047
mu.3	-0.006969257	0.003419020	0.012213356	-0.008616189	-0.007421491	0.009896074	0.009045175	0.0220034205	-0.031211587	0.010365507
mu.4	-0.155243285	-0.107698909	-0.019726660	0.127493329	-0.046937627	0.150179467	0.050644458	0.1022414909	0.038689731	-0.023743312
mu.5	-0.067046723	-0.099741795	0.052475586	0.062478887	0.028056658	-0.001051156	0.086067665	0.1421168508	-0.002928196	0.017828701
mu.6	-0.016095730	-0.073744119	-0.026042053	0.231698907	0.032268864	0.018166750	-0.089778450	0.0200256463	-0.026504613	-0.022346890
mu.7	-0.039412794	0.092244297	0.035736154	0.141175365	-0.030912552	-0.191731083	0.080988757	-0.0726205035	0.059091415	-0.016590651
mu.8	0.116412330	-0.066737460	0.245261940	0.028093031	-0.202705989	0.168726334	-0.105523577	-0.0006431351	0.137596073	0.176192202
mu.9	-0.197494900	-0.239037934	-0.107843309	0.134418258	-0.197022605	0.173052856	-0.231885395	-0.2862671537	0.097899613	-0.170437877
lam2	0.797395391	0.779450866	0.809474265	0.802278700	0.819053451	0.808765246	0.799425125	0.7703715169	0.789084521	0.811161696
lam3	0.797608493	0.805914451	0.816510819	0.798994894	0.819770967	0.809607989	0.794473388	0.8172393038	0.814369412	0.817964004
lam5	0.620622815	0.713665038	0.723552341	0.722827299	0.654323230	0.824093962	0.692058070	0.7711985396	0.541195119	0.682135471
lam6	0.913199732	0.924092882	0.865203957	0.955903024	0.999390560	0.981075629	0.894557218	0.9507051926	0.850528012	0.907327958
lam7	0.713905206	0.732942740	0.645619631	0.820156169	0.647204886	0.711581490	0.694345816	0.7660435248	0.663972122	0.634583037
lam9	0.655700319	0.792721576	0.755276301	0.818260420	0.825536829	0.891215779	0.543345092	0.8384625375	0.621821509	0.510856248
psi_e.1	0.327525552	0.428874725	0.333647190	0.337754598	0.259770454	0.271778222	0.368247260	0.2722553027	0.336509161	0.313764477
psi_e.2	0.275463408	0.251546245	0.302590930	0.308736021	0.312695550	0.267603392	0.258232307	0.2642581425	0.297972186	0.268076743
psi_e.3	0.371738697	0.51022037	0.377620228	0.464834251	0.297985648	0.370055741	0.411042209	0.3351680001	0.358775984	0.370006795
psi_e.4	0.327079817	0.291300141	0.368351816	0.402113549	0.344436338	0.366272071	0.329032466	0.3336612252	0.371350390	0.332138987
psi_d	0.361043132	0.369050314	0.369897844	0.378773066	0.381865889	0.378994268	0.381892434	0.3689275574	0.371793938	0.367119190
mat_phi.1.1	1.093261175	0.853945672	1.032178378	0.891904019	0.920516235	0.997812878	1.055688874	1.0876881346	0.940597679	1.161981882
mat_phi.2.1	0.254880356	-0.155924289	-0.083434974	0.135746622	0.178086675	-0.134720861	0.326089245	0.3370082713	0.337485708	0.378542763
mat_phi.1.2	0.254880356	-0.155924289	-0.083434974	0.135746622	0.178086675	-0.134720861	0.326089245	0.3370082713	0.337485708	0.378542763
mat_phi.2.2	0.777128969	0.807006193	0.807672629	0.865393361	0.916519182	0.971010344	0.868510473	1.0302326353	0.890294156	0.950207235



The bias and RMSE of parameter estimations are:

	gam.1	gam.2	b	mu.1	mu.2	mu.3	mu.4	mu.5	mu.6	mu.7	mu.8	mu.9	lam2	lam3	lam5	lam6	lam7
BIAS	0.00023	0.00367	-0.00216	-0.00092	-0.00155	0.00127	0.01159	0.02183	0.00476	0.00580	0.04967	-0.10246	-0.00135	0.00925	-0.00543	0.02420	0.00304
RMSE	0.00779	0.01692	0.01431	0.01197	0.01161	0.01444	0.09601	0.07058	0.08432	0.09197	0.14478	0.19208	0.01444	0.01286	0.07465	0.05162	0.05618

	lam9	psi_e.1	psi_e.2	psi_e.3	psi_e.4	psi_d	mat_phi.1.1	mat_phi.2.1	mat_phi.1.2	mat_phi.2.2
BIAS	-0.07468	0.02501	-0.01928	-0.01326	-0.05343	0.01294	0.00356	-0.04262	-0.04262	0.0784
RMSE	0.14659	0.05412	0.02879	0.05990	0.06101	0.01456	0.09467	0.20301	0.20301	0.1093

From the bias and RMSE of the estimation, we can conclude that the model estimation is robust to different simulated data.