Problem set 6: Bayesian OLS

Ancilla Marie Inocencio

Q1) OLS

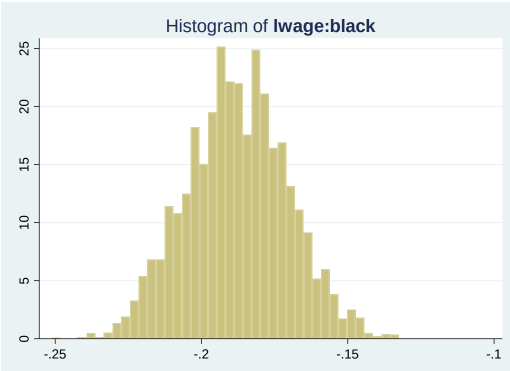
```
. reg lwage educ exper smsa black south

      Source |       SS           df       MS      Number of obs   =     3,010
-------------+----------------------------------   F(5, 3004)      =    232.21
       Model |  165.205654         5  33.0411308   Prob > F        =    0.0000
    Residual |  427.435957     3,004  .142288934   R-squared       =    0.2788
-------------+----------------------------------   Adj R-squared   =    0.2776
       Total |  592.641611     3,009  .196956335   Root MSE        =    .37721

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    .073807   .0035336    20.89   0.000     .0668784    .0807356
       exper |   .0393134   .0021955    17.91   0.000     .0350085    .0436183
        smsa |   .1647411   .0156919    10.50   0.000     .1339732     .195509
       black |  -.1882225   .0177678   -10.59   0.000    -.2230607   -.1533843
       south |  -.1290528   .0152285    -8.47   0.000    -.1589122   -.0991935
       _cons |   4.913331   .0631212    77.84   0.000     4.789566    5.037096
------------------------------------------------------------------------------
```

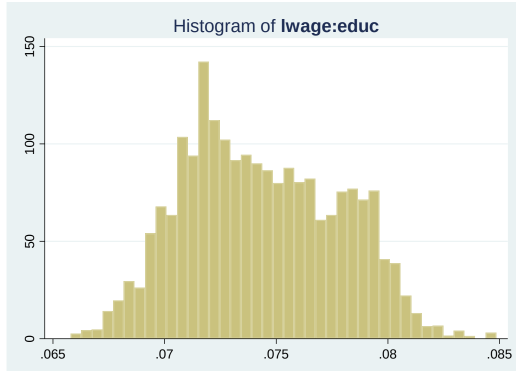
Q2) Bayes using Metropolis-Hastings, use flat prior, plot histograms

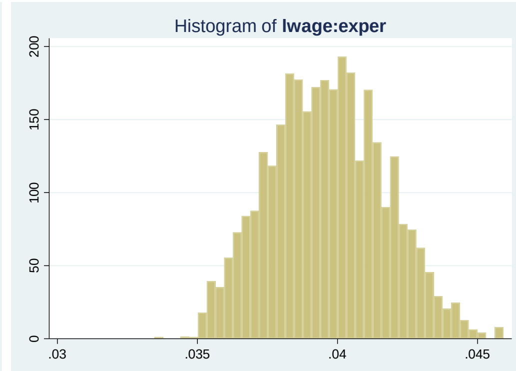
a) Bayes Metropolis- Hastings algorithm, using flat prior for all

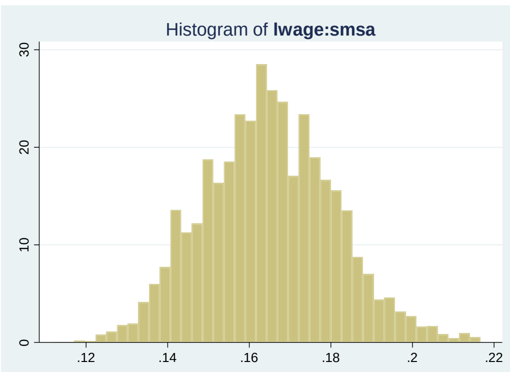
```
. bayesmh lwage educ exper smsa black south, likelihood(normal({var})) prior({lwage: _cons educ exper smsa black south}, flat) prior({var}, jeffreys)
```

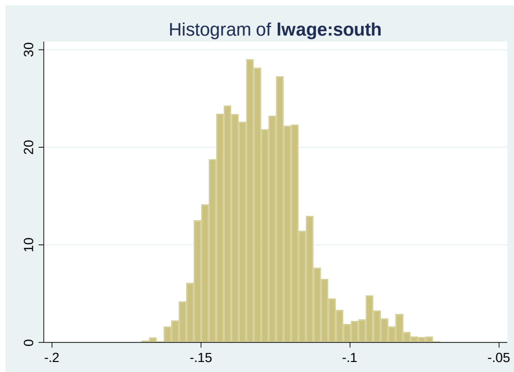
```
Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
Likelihood:
  lwage ~ normal(xb_lwage,{var})

Priors:
  {lwage:_cons educ exper smsa black south} ~ 1 (flat)                    (1)
                                      {var} ~ jeffreys
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_lwage.

Bayesian normal regression                       MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling         Burn-in          =      2,500
                                                 MCMC sample size =     10,000
                                                 Number of obs    =      3,010
                                                 Acceptance rate  =      .2392
                                                 Efficiency:  min =    .001266
                                                              avg =    .005404
Log marginal likelihood = -1360.3484                          max =     .01671

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |      Mean   Std. Dev.     MCSE     Median  [95% Cred. Interval]
-------------+----------------------------------------------------------------
lwage        |
        educ |  .0743258   .0034661   .000904   .0740133   .0683067   .0806296
       exper |   .039635   .0020627   .000434   .0396168   .0357716   .0437261
        smsa |  .1647284   .0160211    .00193   .1645684   .1353478   .1966974
       black | -.1878744   .0173699   .002068  -.1880804  -.2207895  -.1541442
       south | -.1292823   .0156651   .003353  -.1309543  -.1542596  -.0897909
       _cons |  4.903147   .0634373   .017832   4.915855   4.772623   4.999295
-------------+----------------------------------------------------------------
         var |  .1422125   .0036658   .000284   .1421001   .1349064   .1492788
------------------------------------------------------------------------------
Note: There is a high autocorrelation after 500 lags.
```

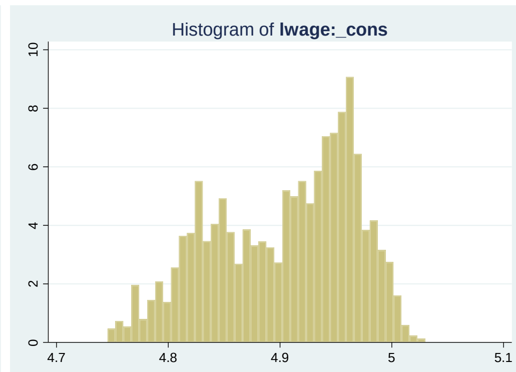

Histogram of lwage:black
0
5
10
15
20
25
-.25
-.2
-.15
-.1

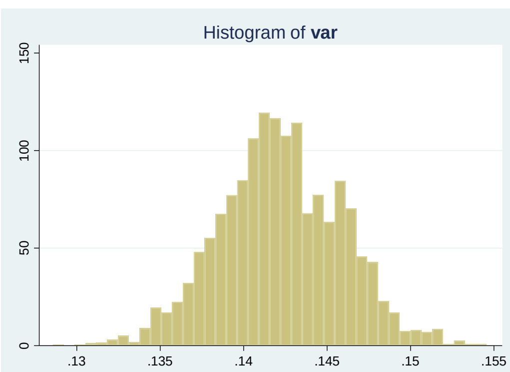

Histogram of lwage:educ
0
50
100
150
.065
.07
.075
.08
.085


Histogram of lwage:exper
0
50
100
150
200
.03
.035
.04
.045


Histogram of lwage:smsa
0
10
20
30
.12
.14
.16
.18
.2
.22


Histogram of lwage:south
0
10
20
30
-.2
-.15
-.1
-.05


Histogram of lwage:_cons
0
2
4
6
8
10
4.7
4.8
4.9
5
5.1

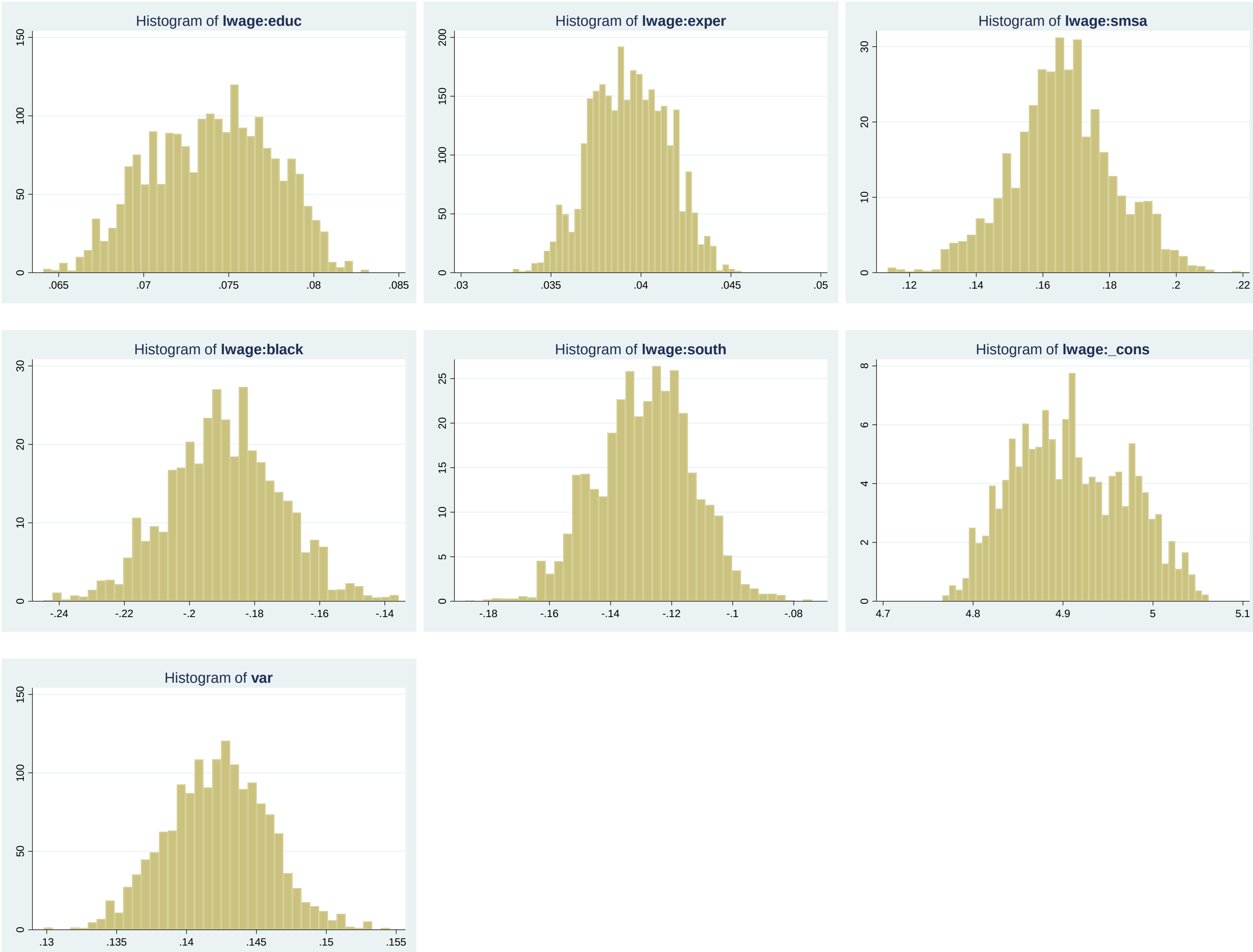

Histogram of var
0
50
100
150
.13
.135
.14
.145
.15
.155

b) Using prior from data provided (see do file for how variance was computed)

```
. bayesmh lwage educ exper smsa black south, likelihood(normal({var})) prior({lwage: _cons exper smsa black south}, flat) prior({lwage: educ},normal(0.06,1.97)) prior({var},
> jeffreys)

Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
Likelihood:
  lwage ~ normal(xb_lwage,{var})

Priors:
  {lwage:_cons exper smsa black south} ~ 1 (flat)                          (1)
                        {lwage:educ} ~ normal(0.06,1.97)                   (1)
                               {var} ~ jeffreys
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_lwage.


Bayesian normal regression                    MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling      Burn-in          =      2,500
                                              MCMC sample size =     10,000
                                              Number of obs    =      3,010
                                              Acceptance rate  =      .1611
                                              Efficiency:  min =    .001562
                                                           avg =     .01878
Log marginal likelihood = -1361.7424                       max =     .04299
```

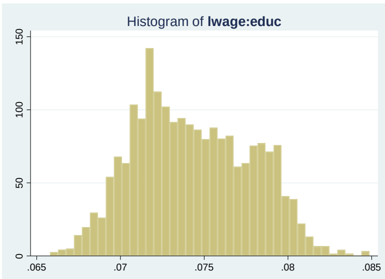
	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred.	Interval]
lwage						
educ	.0740197	.0035785	.00083	.074168	.0670933	.0803022
exper	.0393839	.0021328	.000468	.0393931	.0353249	.0433556
smsa	.166106	.0154831	.000816	.1658339	.1344929	.1964983
black	-.1891804	.017245	.001689	-.1893191	-.2233642	-.1556895
south	-.128783	.015629	.000754	-.1280511	-.1594032	-.0996628
_cons	4.908498	.0627174	.015867	4.904118	4.801841	5.029297
var	.1422329	.0036017	.000188	.1423463	.1351181	.1493051

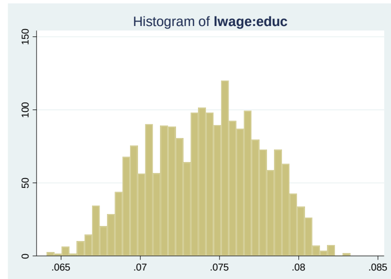
Note: There is a <u>high autocorrelation</u> after 500 lags.

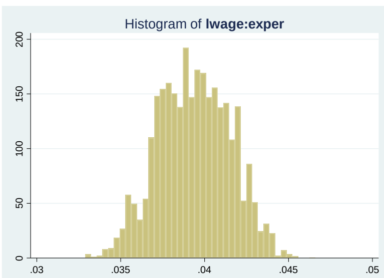

Histogram of lwage:educ
0
50
100
150
.065
.07
.075
.08
.085
Histogram of lwage:exper
0
50
100
150
200
.03
.035
.04
.045
.05
Histogram of lwage:smsa
0
10
20
30
.12
.14
.16
.18
.2
.22
Histogram of lwage:black
0
10
20
30
-.24
-.22
-.2
-.18
-.16
-.14
Histogram of lwage:south
0
5
10
15
20
25
-.18
-.16
-.14
-.12
-.1
-.08
Histogram of lwage:_cons
0
2
4
6
8
4.7
4.8
4.9
5
5.1
Histogram of var
0
50
100
150
.13
.135
.14
.145
.15
.155

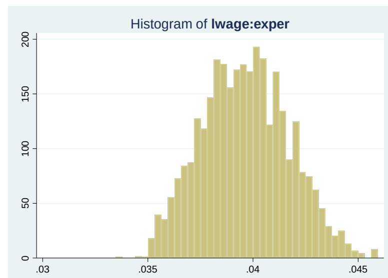
3) Doing a side-by-side comparison, prior distribution for education improved, centering around values closer to the mean. As a model, using prior from another study improved average efficiency (from 0.0054 to 0.01878) and at lower acceptance rate (from 0.239 to 0.161).

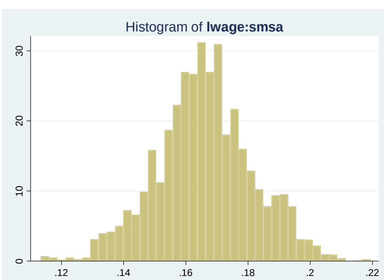
Flat prior for all | **Using prior from another study**

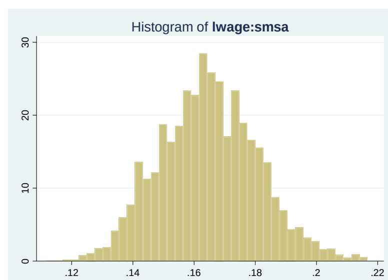












Flat prior for all

Using prior from another study