a.

First, run in its default setting i.e. correlation of x1 and x2 equal o 0

```
SigmaX <- c(1,  0,  0,
            0,  1,  0,
            0,  0,  1)
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs:

(Intercept)	X1	X2	X3
1.003911	2.000822	3.001982	4.001642

True standard errors across 1000 simulation runs:

(Intercept)	X1	X2	X3
0.1446771	0.1432049	0.1452139	0.1513109

Average estimated standard errors across 1000 simulation runs:

(Intercept)	X1	X2	X3
0.1442141	0.1453379	0.1458833	0.1456037

True t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
6.911945	13.966005	20.659178	26.435641

Average estimated t-stat across 1000 simulation runs"

(Intercept)	X1	X2	X3
6.961251	13.766690	20.577963	27.483105

1) Set the correlation of x1 and x2 to 0.5

```
SigmaX <- c(1,  0.5,  0,
            0.5,  1,  0,
            0,  0,  1)
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

(Intercept)	X1	X2	X3
1.004253	1.999523	3.003760	3.998139

True standard errors across 1000 simulation runs:

(Intercept)	X1	X2	X3
0.1353240	0.1612381	0.1725336	0.1439625

Average estimated standard errors across 1000 simulation runs:

(Intercept)	X1	X2	X3
0.1441633	0.1674769	0.1690186	0.1455700

True t-stat across 1000 simulation runs:

(Intercept)	X1	X2	X3
7.389673	12.404015	17.387919	27.785009

Average estimated t-stat across 1000 simulation runs:

(Intercept)	X1	X2	X3
6.966079	11.939099	17.771778	27.465395

2) Set the correlation of x1 and x2 to 0.9

```
SigmaX <- c(1,  0.9,  0,
            0.9,   1,   0,
            0,   0,  1)
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs:

(Intercept)	X1	X2	X3
1.000647	2.001827	2.992671	3.999894

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1406567	0.3353353	0.3425352	0.1485056

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1440609	0.3321339	0.3318834	0.1457313

True t-stat across 1000 simulation runs"

(Intercept)	X1	X2	X3
7.109508	5.964180	8.758224	26.935010

Average estimated t-stat across 1000 simulation runs"

(Intercept)	X1	X2	X3
6.946001	6.027169	9.017235	27.447056

3) Set the correlation of x1 and x2 to 0.99

```
SigmaX <- c(1,  0.99,  0,
            0.99,   1,   0,
            0,   0,  1)
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs:

(Intercept)	X1	X2	X3
0.9957557	1.9905524	3.0047330	3.9956545

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1426535	1.0319404	1.0339733	0.1471823

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1436530	1.0277199	1.0278200	0.1459758

True t-stat across 1000 simulation runs"

(Intercept)	X1	X2	X3
7.009992	1.938096	2.901429	27.177182

Average estimated t-stat across 1000 simulation runs"

(Intercept)	X1	X2	X3
6.931672	1.936863	2.923404	27.372031

Observation and Conclusion:

The finding is that 1) collinearity does not cause bias, for we can see that the true values of parameter, standard error and t-score, and their estimates, as correlation of x1 and x2 increases, are very close and about the same; 2) collinearity, however, does cause the estimate to be less precies, for we can see that, as correlation of x1 and x2 increases, standard error grows larger, indicating more uncertainty.

b.

set correlation of x1 and x2 to 1

```
SigmaX <- c(1,  1,  0,
            1,   1,   0,
            0,   0,  1)
```

Results: the value of x2's cofficients are NAs

True parameters: 1,2,3,4

Average LS estimate across 1000 simulation runs

(Intercept)	X1	X2	X3
1.004571	5.002623	NA	3.998549

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1354517	0.1500680	NA	0.1426182

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X3
0.1433162	0.1446791	0.1447385

True t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
7.382707	13.327291	NA	28.046920

Average estimated t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
7.00947	34.57737	NA	27.90019

Conclusion:
In a model with perfect multicollinearity, the regression coefficients are indeterminate and their standard errors are infinite. In other words, there is no unique solution to the linear algebra problem of OLS. There will be infinite number of equally good solutions, and there's no way to tell which one is better.

c.
run the regression in mcobv.r with x2 omitted

```
Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs:
  (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
    0.9916968     1.9809368     3.9970380

True standard errors across 1000 simulation runs
  (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
    0.3392692     0.3353871     0.3510692
Average estimated standard errors across 1000 simulation runs
  (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
    0.3349979     0.3363610     0.3381439

True t-stat across 1000 simulation runs
  (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
     2.947512      5.963259     11.393766
Average estimated t-stat across 1000 simulation runs
  (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
     2.960308      5.889317     11.820525
```

Conclusion:
when the correlation of x1 and x2 is 0, the omission of x2 does not cause bias--i.e. the estimates are about the same as the "true" value, but it makes the estimates of x1 and x3's coefficients less precise--i.e. the standard error is larger, compared with the original model in mcls.r (results printed out in quesiton a.)

d.
set the correlation of x1 and x2 to 0.9 in mcovb.r

```
SigmaX <- c(1,  0.9,  0,
            0.9,   1,   0,
            0,   0,  1)
```

```
Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs
  (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
    0.9995109     4.6940817     3.9929369
```

True standard errors across 1000 simulation runs
```
 (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
   0.1929092    0.2116235    0.1964159
```
Average estimated standard errors across 1000 simulation runs
```
 (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
   0.1946141    0.1973194    0.1977131
```

True t-stat across 1000 simulation runs
```
 (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
    5.183787     9.450747    20.364955
```
Average estimated t-stat across 1000 simulation runs
```
 (Intercept) X[, c(1, 3)]1 X[, c(1, 3)]2
     5.13586     23.78926     20.19561
```

Conclusion:
The results differ from what are found in question c.
When the correlation of x1 and x2 is 0.9, the omission of x2 causes bias of the estimates of x1's coefficients, the biased estimates are less precise (having larger standard error, when compared with the origninal complete model in mcls.r)

e.

keep the correlation of x1 and x2 at 0.9 and rewrite mcovb.r to omit x3

```
SigmaX <- c(1,  0.9,  0,
            0.9,   1,   0,
            0,   0,  1)
```

Rewritten part:
```
 res <- lm(y~X[,c(1,2)])
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs
```
 (Intercept) X[, c(1, 2)]1 X[, c(1, 2)]2
    1.019461     2.010988     2.999740
```

True standard errors across 1000 simulation runs
```
 (Intercept) X[, c(1, 2)]1 X[, c(1, 2)]2
   0.4270246    1.0210131    1.0333707
```
Average estimated standard errors across 1000 simulation runs
```
 (Intercept) X[, c(1, 2)]1 X[, c(1, 2)]2
   0.4285566    0.9903458    0.9914714
```

True t-stat across 1000 simulation runs
```
 (Intercept) X[, c(1, 2)]1 X[, c(1, 2)]2
    2.341786     1.958839     3.870828
```

Average estimated t-stat across 1000 simulation runs
(Intercept) X[, c(1, 2)]1 X[, c(1, 2)]2
2.378824 2.030592 3.025544

Conclusion:

When the correlation of x1 and x2 is kept at 0.9, the omission of x3 has causes no bias of the estimates of x1 and x2's coefficients, with the precision of the estimates is slightly improved (i.e. standard error smaller than the standard complete model)

f.

Findings:

Omitted variables cause bias of the coefficient estimates of included variables when they are correlated with the omitted variables. The precision of the estimates may also be affected, whenever variables are omitted.

Suggestion on how to deal with hihgly correlated covariates:

If these highly correlated covariates are necessary for a well specified model, the researcher should leave them in the model. While doing so causes imprecise estimates (large standard error), the benefit is that the estimates of the coefficients are still unbiased. Keeping the higly correlated covariates in the model is better than taking one of them out, which causes bias in the estimates.

g.

run mcselect.r

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs
(Intercept) selectX1 selectX2 selectX3
0.2936736 1.7963454 2.6895235 3.5887842

True standard errors across 1000 simulation runs
(Intercept) selectX1 selectX2 selectX3
0.3206626 0.2260434 0.2472244 0.2766393
Average estimated standard errors across 1000 simulation runs
(Intercept) selectX1 selectX2 selectX3
0.3155037 0.2222701 0.2427435 0.2711980

True t-stat across 1000 simulation runs
(Intercept) selectX1 selectX2 selectX3
3.118543 8.847857 12.134722 14.459263
Average estimated t-stat across 1000 simulation runs
(Intercept) selectX1 selectX2 selectX3
0.9308088 8.0818133 11.0796912 13.2330787

Conclusion:
selection on y causes both bias (less accurate estimates) and imprecision (larger standard error).

h.
Run mchet.r under its default setting
Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

(Intercept)	X1	X2	X3
1.008089	2.004764	3.002107	3.995437

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1417963	0.1471852	0.1425660	0.1477049

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1431662	0.1439025	0.1450042	0.1448510

True t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
7.05237	13.58832	21.04289	27.08102

Average estimated t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
7.04139	13.93140	20.70359	27.58308

Conclusion:
Homoskedastic means Equal Variances. According to this assumption, although different samples can come from populations with different means, they have about the same variance.

When Gamma1 = 0, y is still homoskedastic, as we can see that the standard errors of the estimates of the coefficients of x1, x2 and x3 are still about the same as the "true" values, meaning that the model meets the equal variances assumption.

Adding heteroskedasticity by increasing Gamma1 to 1
g <- c(log(2),1)

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

(Intercept)	X1	X2	X3
0.9941065	1.9935321	3.0051387	4.0037156

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1805888	0.2559456	0.1829800	0.1861140

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1834027	0.1841175	0.1846952	0.1852426

True t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
5.537443	7.814161	16.395236	21.492203

Average estimated t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
5.420349	10.827499	16.270801	21.613364

Conclusion:

When Gamma1 = 1, y is heteroskedastic, as we can see that now the standard errors of the estimate of the coefficients of x1 is different from the "true" value, meaning that the model does NOT meet the equal variances assumption.

The effect of heteroskedasticity here is that:

(1) The OLS estimators and regression predictions based on them remains unbiased and consistent. (i.e. true value and estimates are about the same)
(2) The OLS estimators are no longer the BLUE (Best Linear Unbiased Estimators) because they are no longer efficient, so the regression predictions will be inefficient too. (i.e. in this case, larger variance)
(3) Because of the inconsistency of the covariance matrix of the estimated regression coefficients, the tests of hypotheses, (i.e. t-test) are no longer valid. (i.e. in this case, smaller t-score)

i.

run mcautocor.r in default setting

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

(Intercept)	X1	X2	X3
1.000647	1.997953	2.999894	3.994637

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1406567	0.1479150	0.1485056	0.1445140

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1440609	0.1446835	0.1457313	0.1453965

True t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
7.109508	13.521281	20.201257	27.678979

Average estimated t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
6.946001	13.809126	20.585111	27.474098

Rerun with Rho = 0.5 and Rho*Xk = 0.5 for all k

```
rho <- 0.5
rhoX <- c(0.5, 0.5, 0.5)
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

(Intercept)	X1	X2	X3
0.9938191	1.9988496	3.0000771	3.9974895

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.2147806	0.1679558	0.1687629	0.1750445

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1611188	0.1454341	0.1463197	0.1456935

True t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
4.655913	11.907894	17.776421	22.851326

Average estimated t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
6.168239	13.744022	20.503582	27.437658

Rerun with Rho = 0.9 and Rho*Xk = 0.9 for all k

```
rho <- 0.9
rhoX <- c(0.9, 0.9, 0.9)
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

(Intercept)	X1	X2	X3
1.007825	2.003087	3.006584	4.007556

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.2765791	0.1745296	0.1784950	0.1860279

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.1948548	0.1456802	0.1464527	0.1465591

True t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
3.615602	11.459377	16.807196	21.502154

Average estimated t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
5.172184	13.749894	20.529380	27.344306

Conclusion:
Serial correlation causes NO bias in coefficient estimates, but it causes bias in estimates of standard error.

Experimenting further...
(1) serial correlation in y but not in X

```
rho <- 0.9
rhoX <- c(0, 0, 0)
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

```
(Intercept)        X1        X2        X3
   1.013455  2.009385  3.001280  3.995287
```

True standard errors across 1000 simulation runs

```
(Intercept)        X1        X2        X3
  0.2674998 0.1999162 0.1900501 0.1985933
```

Average estimated standard errors across 1000 simulation runs

```
(Intercept)        X1        X2        X3
  0.1914158 0.1923742 0.1937853 0.1936890
```

True t-stat across 1000 simulation runs

```
(Intercept)        X1        X2        X3
    3.73832  10.00419  15.78531  20.14166
```

Average estimated t-stat across 1000 simulation runsa

```
(Intercept)        X1        X2        X3
   5.294518 10.445191 15.487658 20.627328
```

(2) serial correlation in X but not in y

```
rho <- 0
rhoX <- c(0.9, 0.9, 0.9)
```

Results:
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

```
(Intercept)        X1        X2        X3
   0.995535  2.002749  3.001765  4.001038
```

True standard errors across 1000 simulation runs

```
(Intercept)        X1        X2        X3
  0.1440512 0.1091848 0.1123510 0.1140063
```

verage estimated standard errors across 1000 simulation runs

```
(Intercept)        X1        X2        X3
  0.1457337 0.1098399 0.1094366 0.1096701
```

True t-stat across 1000 simulation runs

```
(Intercept)         X1          X2          X3
  6.941975   18.317561   26.702040   35.085776
```

Average estimated t-stat across 1000 simulation runs

```
(Intercept)         X1          X2          X3
  6.831193   18.233349   27.429252   36.482501
```

Conclusion:

These two further experiments show that serial correlation does not cause bias in coefficient estimates, but it does affect the precision of estimates. When there is correlation in y but not in X, the standard error becomes larger than the "true" one. When there is correlation in X but not in y, the standard error becomes smaller than the "true" one. Imprecision then affects the judgment about whether one should reject the null hypothesis, i.e. standard error is related to the t-score.

j.

(I am answering the example question)

Increase the size of n by setting n = 1000 instead of 100
Then, test the effect of a larger dataset in different settings.
As a benchmark, the results of the original standard BLUE model are as follows:

True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs:

```
(Intercept)         X1          X2          X3
  1.003911    2.000822    3.001982    4.001642
```

True standard errors across 1000 simulation runs:

```
(Intercept)         X1          X2          X3
 0.1446771   0.1432049   0.1452139   0.1513109
```

Average estimated standard errors across 1000 simulation runs:

```
(Intercept)         X1          X2          X3
 0.1442141   0.1453379   0.1458833   0.1456037
```

True t-stat across 1000 simulation runs"

```
(Intercept)         X1          X2          X3
  6.911945   13.966005   20.659178   26.435641
```

Average estimated t-stat across 1000 simulation runs"

```
(Intercept)         X1          X2          X3
  6.961251   13.766690   20.577963   27.483105
```

(1) multicolinearity

Using mcls.r
Set the correlation of x1 and x2 to 0.9

```
SigmaX <- c(1,  0.9,  0,
           0.9,   1,   0,
           0,   0,  1)
```

Results: (When n = 1000):
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

(Intercept)	X1	X2	X3
0.9982659	1.9981138	3.0029378	4.0000838

True standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.04380777	0.10378126	0.10238086	0.04460450

Average estimated standard errors across 1000 simulation runs

(Intercept)	X1	X2	X3
0.04479301	0.10277685	0.10280891	0.04487235

True t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
22.82700	19.27130	29.30235	89.67706

Average estimated t-stat across 1000 simulation runs

(Intercept)	X1	X2	X3
22.28620	19.44128	29.20893	89.14362

Conclusion:

compare the above result with the previous one in question a. (i.e. when n = 100, and the correlation of x1 and x2 is 0.9), we can see that the standard error decreased sharply, indicating improved precision. Therefore, more data can mitigate the multicolinearity problem.

(2) omitted variable

Using mcovb.r, where x2 is omitted in regression
Set the correlation of x1 and x2 to 0.9

Results: (When n = 1000):
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs

(Intercept)	X[, c(1, 3)]1	X[, c(1, 3)]2
0.9998388	4.7022631	3.9976531

True standard errors across 1000 simulation runs

(Intercept)	X[, c(1, 3)]1	X[, c(1, 3)]2
0.06306916	0.06096479	0.06125538

Average estimated standard errors across 1000 simulation runs

(Intercept)	X[, c(1, 3)]1	X[, c(1, 3)]2
0.06101046	0.06107748	0.06110853

True t-stat across 1000 simulation runs

(Intercept)	X[, c(1, 3)]1	X[, c(1, 3)]2
15.85561	32.80582	65.30039

Average estimated t-stat across 1000 simulation runs

(Intercept)	X[, c(1, 3)]1	X[, c(1, 3)]2
16.38799	76.98850	65.41890

Conclusion:

compare the above result with the previous one in question d. (i.e. when n = 100, x2 is omitted in regression, and the correlation of x1 and x2 is 0.9), we can see that the bias of coefficient estimate remains, though larger sample makes the estimates more precise (i.e. smaller standard error).
Therefore, more data cannot mitigate the omitted variable problem.

(3) selection on dependent variable

Using mcselect.r, where all observations in which y is greater than its sample mean are deleted prior to regression

Results: (When n = 1000):

```
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs
(Intercept)    selectX1    selectX2    selectX3
  0.2921318   1.7976649   2.6946022   3.5922726

True standard errors across 1000 simulation runs
(Intercept)    selectX1    selectX2    selectX3
 0.09419149  0.06749164  0.07515566  0.08378698
Average estimated standard errors across 1000 simulation runs
(Intercept)    selectX1    selectX2    selectX3
 0.09462103  0.06615658  0.07266369  0.08109542

True t-stat across 1000 simulation runs
(Intercept)    selectX1    selectX2    selectX3
   10.61667    29.63330    39.91715    47.74012
Average estimated t-stat across 1000 simulation runs
(Intercept)    selectX1    selectX2    selectX3
   3.087388   27.172880   37.083199   44.296859
```

Conclusion:

compare the above result with the previous one in question g. (i.e. when n = 100, and selection on y is performed), we can see that the bias in estimate remains, though though larger sample makes the estimates more precise (i.e. smaller standard error).
Therefore, more data cannot mitigate the problem of selection bias.

(4) heteroskedasticity

Using mchet.r, where Gamma1 is set to 1, making the model heteroskedastic.
g <- c(log(2),1)

Results: (When n = 1000):

```
True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs
(Intercept)         X1          X2          X3
   1.001697    2.004948    2.998419    4.004096
```

True standard errors across 1000 simulation runs
(Intercept) X1 X2 X3
0.05769831 0.08154747 0.05774037 0.05748110
Average estimated standard errors across 1000 simulation runs
(Intercept) X1 X2 X3
0.05731711 0.05732163 0.05738789 0.05739419

True t-stat across 1000 simulation runs
(Intercept) X1 X2 X3
17.33153 24.52559 51.95671 69.58809
Average estimated t-stat across 1000 simulation runs
(Intercept) X1 X2 X3
17.47640 34.97716 52.24828 69.76484

Conclusion:

compare the above result with the previous one in question h. (i.e. when n = 100, and heteroskedasticity exists), we can see that for variable x1, the discrepancy between the "true" and estimated value of coefficient still remains.
Therefore, more data cannot mitigate the heteroskedasticity problem.

(5) serial correlation (autocorrelation)

Using mcautocor.r, where autocorrelation exists

Set Rho = 0.9 and Rho*Xk = 0.9 for all k
rho <- 0.9
rhoX <- c(0.9, 0.9, 0.9)

Results: (When n = 1000):

True parameters: 1,2,3,4
Average LS estimate across 1000 simulation runs
(Intercept) X1 X2 X3
0.9976276 1.9977917 3.0002845 4.0026392

True standard errors across 1000 simulation runs
(Intercept) X1 X2 X3
0.08713448 0.05290644 0.05434943 0.05484860
Average estimated standard errors across 1000 simulation runs
(Intercept) X1 X2 X3
0.06025713 0.04479123 0.04481994 0.04486801

True t-stat across 1000 simulation runs
(Intercept) X1 X2 X3
11.47651 37.80258 55.19837 72.92802
Average estimated t-stat across 1000 simulation runs
(Intercept) X1 X2 X3
16.55618 44.60230 66.94084 89.20919

Conclusion:

compare the above result with the previous one in question i. (i.e. when n = 100, and autocorrelation exists in both y and X), we can see that the deiscrepancy between "true" and estimated standard error still remains. Therefore, more data cannot mitigate autocorrelation problem.