Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

Question 1

A. Summarize on A16 variables

```
global A16 approve white hrat obrat loanprc unem male married dep
sch cosign chist pubrec mortlat1 mortlat2 vr
describe $A16
summarize $A16
```

```
. summarize $A16
```

Variable	Obs	Mean	Std. dev.	Min	Max
approve	1,790	.8748603	.3309697	0	1
white	1,790	.847486	.3596188	0	1
hrat	1,790	24.80235	7.186854	1	72
obrat	1,790	32.33759	8.299516	0	95
loanprc	1,790	.7739717	.1900165	.0210526	2.571429
unem	1,790	3.88	2.144775	1.8	10.6
male	1,777	.8103545	.3921308	0	1
married	1,787	.6552882	.4754072	0	1
dep	1,787	.761052	1.096408	0	8
sch	1,790	.772067	.4196164	0	1
cosign	1,790	.0290503	.1679944	0	1
chist	1,790	.8407821	.3659814	0	1
pubrec	1,790	.0692737	.25399	0	1
mortlat1	1,790	.0189944	.1365432	0	1
mortlat2	1,790	.0106145	.1025072	0	1
vr	1,790	.4083799	.4916714	0	1

B. Missing values in A16

Variables male, married and dep have the number of missing values 13,3 and 3 respectively, summary stat for nonmissing

```
. misstable summarize $A16
```

				Obs<.		
Variable	Obs=.	Obs>.	Obs<.	Unique values	Min	Max
male	13		1,777	2	0	1
married	3		1,787	2	0	1
dep	3		1,787	9	0	8

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

```
. summarize $A16 if male != . & dep!=. & married!= .
```

Variable	Obs	Mean	Std. dev.	Min	Max
approve	1,774	.8737317	.3322452	0	1
white	1,774	.848929	.3582191	0	1
hrat	1,774	24.81508	7.196018	1	72
obrat	1,774	32.33934	8.310612	0	95
loanprc	1,774	.7737213	.1902714	.0210526	2.571429
unem	1,774	3.886133	2.15212	1.8	10.6
male	1,774	.8105975	.3919385	0	1
married	1,774	.6555806	.4753126	0	1
dep	1,774	.7609921	1.097261	0	8
sch	1,774	.770575	.4205816	0	1
cosign	1,774	.0293123	.168728	0	1
chist	1,774	.8399098	.3667931	0	1
pubrec	1,774	.0687711	.2531359	0	1
mortlat1	1,774	.0191657	.1371459	0	1
mortlat2	1,774	.0107103	.1029637	0	1
vr	1,774	.4064262	.4913044	0	1

C. Regress approve on white, restricting to nonmissing

designating that a white individual is anticipated to have a higher possibility of obtaining a loan sanctioned.
The model in usual form is given by:

$$approve = 0.6977612 + 0.2072853\ white + u$$
$$(.0197873) \qquad (0.0214759)$$

$n = 1774$, $R2 = 0.0499$

The coefficient of *white* is 0.2072853 indicating that a white individual has a 20.72853% higher probability of loan approval. Which is practically large.
The coefficient of *white* has the p-value of 0.0000 which is less than the critical p-value of 0.05 at a 5% level of significance, indicating that the variable *white* is statistically significant at a 5% level of significance in regression model.

```
regress approve i.white if male != . & dep!=. & married!= .
outreg2 using table1, title(Table 1. Effect of Race on mortgage
approval) word label replace
```

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

```
. /* 1.C Regress approve on white , restricting to nonmissing */
. regress approve i.white if male != . & dep!=. & married!= .
```

Source	SS	df	MS	Number of obs	=	1,774
				F(1, 1772)	=	93.16
Model	9.77559282	1	9.77559282	Prob > F	=	0.0000
Residual	185.940303	1,772	.104932451	R-squared	=	0.0499
				Adj R-squared	=	0.0494
Total	195.715896	1,773	.110386856	Root MSE	=	.32393

approve	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
1.white	.2072853	.0214759	9.65	0.000	.1651645	.2494061
_cons	.6977612	.0197873	35.26	0.000	.6589522	.7365702

```
.
. outreg2 using table1, title(Table 1. Effect of Race on mortgage approval) word label replace
table1.rtf
dir : seeout
```

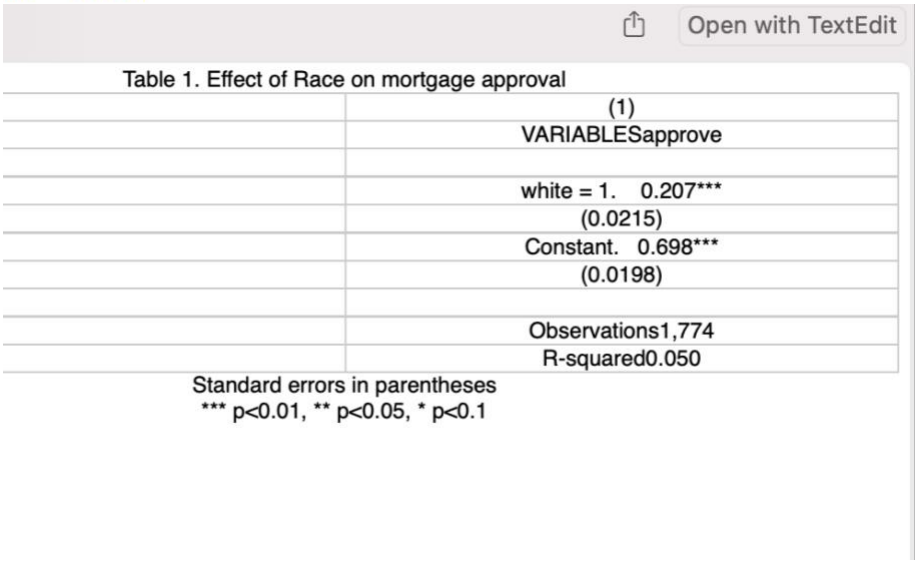

Table 1. Effect of Race on mortgage approval

	(1)
	VARIABLESapprove
	white = 1. 0.207***
	(0.0215)
	Constant. 0.698***
	(0.0198)
	Observations1,774
	R-squared0.050

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

D. Add the other variables,

The coefficient of *white* is 0.1278722 with a p-value of 0.0000 which is less than the critical p-value of 0.05 at 5% level of significance indicating that the variable white is still statistically significant at 5% level of significance. This also indicates that there is still evidence of discrimination against nonwhites.

```
. /* 1.D Add other variables regress on approve */
. regress approve i.white c.hrat c.obrat c.loanprc c.unem i.male i.married c.dep i.sch i.cos
> at2 i.vr
```

Source	SS	df	MS		
Model	33.2438816	15	2.21625877	Number of obs	= 1,774
Residual	162.472015	1,758	.092418666	F(15, 1758)	= 23.98
				Prob > F	= 0.0000
Total	195.715896	1,773	.110386856	R-squared	= 0.1699
				Adj R-squared	= 0.1628
				Root MSE	= .304

approve	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
1.white	.1278722	.0211249	6.05	0.000	.0864396	.1693048
hrat	.0017248	.0013512	1.28	0.202	-.0009254	.0043749
obrat	-.0057256	.001186	-4.83	0.000	-.0080517	-.0033995
loanprc	-.134513	.0395027	-3.41	0.001	-.2119903	-.0570356
unem	-.00726	.0034295	-2.12	0.034	-.0139862	-.0005337
1.male	-.0095075	.0199184	-0.48	0.633	-.0485738	.0295587
1.married	.0436661	.0172906	2.53	0.012	.0097539	.0775783
dep	-.0026524	.0071586	-0.37	0.711	-.0166927	.0113879
1.sch	.0068934	.0177295	0.39	0.697	-.0278798	.0416665
1.cosign	.0108263	.0435083	0.25	0.804	-.074507	.0961597
1.chist	.1428376	.0207083	6.90	0.000	.1022221	.1834532
1.pubrec	-.2449393	.0299098	-8.19	0.000	-.3036019	-.1862767
1.mortlat1	-.0579501	.0531709	-1.09	0.276	-.162235	.0463348
1.mortlat2	-.1224583	.0710272	-1.72	0.085	-.261765	.0168485
1.vr	-.0332678	.0149054	-2.23	0.026	-.0625021	-.0040336
_cons	.9281153	.0556856	16.67	0.000	.8188982	1.037332

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

Table 2. Determinants of mortgage approval

	(1)
VARIABLES	approve
white = 1	0.128***
	(0.0211)
hrat	0.00172
	(0.00135)
obrat	-0.00573***
	(0.00119)
loanprc	-0.135***
	(0.0395)
unem	-0.00726**
	(0.00343)
male = 1	-0.00951
	(0.0199)
married = 1	0.0437**
	(0.0173)
dep	-0.00265
	(0.00716)
sch = 1	0.00689
	(0.0177)
cosign = 1	0.0108
	(0.0435)
chist = 1	0.143***
	(0.0207)
pubrec = 1	-0.245***
	(0.0299)
mortlat1 = 1	-0.0580
	(0.0532)
mortlat2 = 1	-0.122*
	(0.0710)
vr = 1	-0.0333**
	(0.0149)
Constant	0.928***
	(0.0557)
Observations	1,774
R-squared	0.170

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

E. Interaction term of obrat and white

The coefficient of interaction variable *white*Orbat* is a .0087337 with a p-value of 0.00 which is less than the critical p-value of 0.05 at 5% level of significance indicating that the interaction variable white*orbat is statistically significant at 5% level of significance.

```
regress approve i.white##c.obrat c.hrat c.loanprc c.unem i.male
i.married c.dep i.sch i.cosign i.chist i.pubrec i.mortlat1 i.mortlat2
i.vr
```

Source	SS	df	MS	Number of obs	=	1,774
				F(16, 1757)	=	23.45
Model	34.4426589	16	2.15266618	Prob > F	=	0.0000
Residual	161.273237	1,757	.09178898	R-squared	=	0.1760
				Adj R-squared	=	0.1685
Total	195.715896	1,773	.110386856	Root MSE	=	.30297

approve	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
1.white	-.1693864	.0849061	-1.99	0.046	-.335914	-.0028587
obrat	-.0130207	.0023392	-5.57	0.000	-.0176085	-.0084328
white#c.obrat						
1	.0087337	.0024167	3.61	0.000	.0039938	.0134736
hrat	.0016278	.0013469	1.21	0.227	-.0010138	.0042694
loanprc	-.1415938	.0394167	-3.59	0.000	-.2189023	-.0642853
unem	-.0075504	.0034187	-2.21	0.027	-.0142555	-.0008452
1.male	-.011654	.0198593	-0.59	0.557	-.0506044	.0272964
1.married	.0431459	.0172321	2.50	0.012	.0093482	.0769436
dep	-.0034003	.0071372	-0.48	0.634	-.0173986	.010598
1.sch	.006153	.0176702	0.35	0.728	-.0285039	.0408098
1.cosign	.0196569	.0434286	0.45	0.651	-.0655202	.1048341
1.chist	.1389488	.0206657	6.72	0.000	.0984169	.1794807
1.pubrec	-.243906	.0298091	-8.18	0.000	-.3023712	-.1854409
1.mortlat1	-.0618183	.0530003	-1.17	0.244	-.1657685	.0421319
1.mortlat2	-.1391984	.0709363	-1.96	0.050	-.2783268	-.0000701
1.vr	-.0321308	.0148579	-2.16	0.031	-.0612719	-.0029898
_cons	1.194431	.0922515	12.95	0.000	1.013497	1.375365

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

Table 3. Determinants of mortgage approval with interaction term

VARIABLES	(1) approve
white = 1	-0.169**
	(0.0849)
obrat	-0.0130***
	(0.00234)
0b.white#co.obrat	0
	(0)
1.white#c.obrat	0.00873***
	(0.00242)
hrat	0.00163
	(0.00135)
loanprc	-0.142***
	(0.0394)
unem	-0.00755**
	(0.00342)
male = 1	-0.0117
	(0.0199)
married = 1	0.0431**
	(0.0172)
dep	-0.00340
	(0.00714)
sch = 1	0.00615
	(0.0177)
cosign = 1	0.0197
	(0.0434)
chist = 1	0.139***
	(0.0207)
pubrec = 1	-0.244***
	(0.0298)
mortlat1 = 1	-0.0618
	(0.0530)
mortlat2 = 1	-0.139**
	(0.0709)
vr = 1	-0.0321**
	(0.0149)
Constant	1.194***
	(0.0923)
Observations	1,774
R-squared	0.176

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

G. When Obrat= 32.339

When Obrat is equal to sample average, The effect of white on approval will be 0.1130526 with p-value of 0.000 which indicate it is statistically significant at 5% level of significance. and confidence interval (0.0709853 , 0.1551198)

```
gen obrat2=obrat-32.339
regress approve i.white##c.obrat2 c.hrat c.loanprc c.unem i.male
i.married c.dep i.sch i.cosign i.chist i.pubrec i.mortlat1
i.mortlat2 i.vr
```

Source	SS	df	MS	Number of obs	=	1,774
				F(16, 1757)	=	23.45
Model	34.4426588	16	2.15266618	Prob > F	=	0.0000
Residual	161.273237	1,757	.09178898	R-squared	=	0.1760
				Adj R-squared	=	0.1685
Total	195.715896	1,773	.110386856	Root MSE	=	.30297

approve	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
1.white	.1130526	.0214485	5.27	0.000	.0709853	.1551198
obrat2	-.0130207	.0023392	-5.57	0.000	-.0176085	-.0084328
white#c.obrat2						
1	.0087337	.0024167	3.61	0.000	.0039938	.0134736
hrat	.0016278	.0013469	1.21	0.227	-.0010138	.0042694
loanprc	-.1415938	.0394167	-3.59	0.000	-.2189023	-.0642853
unem	-.0075504	.0034187	-2.21	0.027	-.0142555	-.0008452
1.male	-.011654	.0198593	-0.59	0.557	-.0506044	.0272964
1.married	.0431459	.0172321	2.50	0.012	.0093482	.0769436
dep	-.0034003	.0071372	-0.48	0.634	-.0173986	.010598
1.sch	.006153	.0176702	0.35	0.728	-.0285039	.0408098
1.cosign	.0196569	.0434286	0.45	0.651	-.0655202	.1048341
1.chist	.1389488	.0206657	6.72	0.000	.0984169	.1794807
1.pubrec	-.243906	.0298091	-8.18	0.000	-.3023712	-.1854409
1.mortlat1	-.0618183	.0530003	-1.17	0.244	-.1657685	.0421319
1.mortlat2	-.1391984	.0709363	-1.96	0.050	-.2783268	-.0000701
1.vr	-.0321308	.0148579	-2.16	0.031	-.0612719	-.0029898
_cons	.7733562	.0594164	13.02	0.000	.656822	.8898904

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

Table 3. Determinants of mortgage approval with obrat=32.339

	(1)
VARIABLES	approve
white = 1	0.113***
	(0.0214)
obrat2	-0.0130***
	(0.00234)
0b.white#co.obrat20	
	(0)
1.white#c.obrat2	0.00873***
	(0.00242)
hrat	0.00163
	(0.00135)
loanprc	-0.142***
	(0.0394)
unem	-0.00755**
	(0.00342)
male = 1	-0.0117
	(0.0199)
married = 1	0.0431**
	(0.0172)
dep	-0.00340
	(0.00714)
sch = 1	0.00615
	(0.0177)
cosign = 1	0.0197
	(0.0434)
chist = 1	0.139***
	(0.0207)
pubrec = 1	-0.244***
	(0.0298)
mortlat1 = 1	-0.0618
	(0.0530)
mortlat2 = 1	-0.139**
	(0.0709)
vr = 1	-0.0321**
	(0.0149)
Constant	0.773***
	(0.0594)
Observations	1,774
R-squared	0.176

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

H. Probabilities and summary statistics

We can see that there are observations with probabilities of higher than 1 as the max value is 1.13. The most basic probability law states that the probability of an event occurring must be contained within the interval [0,1]. But the nature of an LPM is such that it doesn't ensure this fundamental law of probability is satisfied. Although most of the predicted probabilities from an LPM have sensible values (between 0 and 1), some predicted probabilities may have nonsensical values that are less than 0 or greater than 1.

```
predict mypred
summarize mypred
```

```
. summarize mypred

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
      mypred |      1,774    .8737317    .1393779   .0074495   1.132905

.
```

Question 2.

Load the file

- `cd "/Users/marckangah/Downloads/vote1"`
- `infile state district democA voteA expendA expendB prtystrA lexpendA lexpendB shareA using vote1.raw`

A. Estimate a model with voteA as the dependent variable and prtystrA, democA, log(expendA), and log(expendB) as explanatory variables. Obtain the OLS residuals, ui, and the summary statistics for this residual variable.

$$voted = \beta0 + \beta1*prtystrA + \beta2*democA + \beta3*\log(expendA) + \beta4*\log(expendB) + \varepsilon$$

The OLS residuals are:

$$\varepsilon = voted - (\beta0 + \beta1*prtystrA + \beta2*democA + \beta3*\log(expendA) + \beta4*\log(expendB))$$

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

```
. regress voteA prtystrA democA lexpendA lexpendB
```

Source	SS	df	MS	Number of obs	=	173
				F(4, 168)	=	169.23
Model	38822.1768	4	9705.5442	Prob > F	=	0.0000
Residual	9635.07174	168	57.3516175	R-squared	=	0.8012
				Adj R-squared	=	0.7964
Total	48457.2486	172	281.728189	Root MSE	=	7.5731

voteA	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
prtystrA	.2519175	.0712925	3.53	0.001	.1111729	.3926622
democA	3.792944	1.40652	2.70	0.008	1.016213	6.569674
lexpendA	5.779294	.3918197	14.75	0.000	5.00577	6.552819
lexpendB	-6.237836	.3974596	-15.69	0.000	-7.022495	-5.453178
_cons	37.66142	4.736036	7.95	0.000	28.3116	47.01123

The R^2 model (0.79640 is extremely higher and suggests that the model explains 80% variance in vote share percent of 'A'.

```
predict pred_voteA
predict residual_vote, residuals
summarize residual_vote
```

```
. do "/var/folders/gk/vzl_r4g114z52p957wyhtf0h0000gn/T//SD16939.000000"

. summarize residual_vote
```

Variable	Obs	Mean	Std. dev.	Min	Max
residual_v~e	173	-1.05e-08	7.484508	-18.57652	24.56568

B. Obtain both the usual OLS standard errors and the heteroskedasticity-robust versions. Are there any important differences?

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

```
. regress voteA prtystrA democA lexpendA lexpendB
```

Source	SS	df	MS	Number of obs	=	173
				F(4, 168)	=	169.23
Model	38822.1768	4	9705.5442	Prob > F	=	0.0000
Residual	9635.07174	168	57.3516175	R-squared	=	0.8012
				Adj R-squared	=	0.7964
Total	48457.2486	172	281.728189	Root MSE	=	7.5731

voteA	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
prtystrA	.2519175	.0712925	3.53	0.001	.1111729	.3926622
democA	3.792944	1.40652	2.70	0.008	1.016213	6.569674
lexpendA	5.779294	.3918197	14.75	0.000	5.00577	6.552819
lexpendB	-6.237836	.3974596	-15.69	0.000	-7.022495	-5.453178
_cons	37.66142	4.736036	7.95	0.000	28.3116	47.01123

```
. regress voteA prtystrA democA lexpendA lexpendB, vce(robust)
```

Linear regression

Number of obs	=	173
F(4, 168)	=	164.20
Prob > F	=	0.0000
R-squared	=	0.8012
Root MSE	=	7.5731

voteA	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
prtystrA	.2519175	.0660631	3.81	0.000	.1214968	.3823383
democA	3.792944	1.452168	2.61	0.010	.926096	6.659791
lexpendA	5.779294	.533142	10.84	0.000	4.726773	6.831816
lexpendB	-6.237836	.3561583	-17.51	0.000	-6.940959	-5.534714
_cons	37.66142	4.418941	8.52	0.000	28.93761	46.38522

When considering the heteroskedasticity robust standard error, the standard error of the coefficients does not change significantly, and the coefficients that were statistically significant at the 5% level of significance when taking into account the usual OLS standard error remained so after taking account for the heteroskedasticity robust standard error.

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

C. Compute the Breusch-Pagan test for heteroskedasticity. Use the F statistic and nR^2 versions and report their standard p-values. What do you conclude?

```
. hettest, rhs fstat

Breusch-Pagan/Cook-Weisberg test for heteroskedasticity
Assumption: i.i.d. error terms
Variables: All independent variables

H0: Constant variance

F(4, 168) =    2.33
 Prob > F =  0.0581
```

$H_0 \neq$ error variance is equal = no heteroskedasticity
Ha = error variance is not equal = heteroskedasticity
The F- statistics and p-value of the above test are:
F = 2.33
p = 0.0581

The F-statistic is 2.33 with a p-value of 0.058. This is greater than the critical p-value of 0.05. It indicates that although some heteroskedasticity can be observed but it is NOT statistically significant at 5% level of significance.
Now we regress $\hat{\mu}$ 2 on voteA and (voteAi)2 where the voteAi are the OLS fitted values from part 1..
The F test, with 2 and 170df, is about 2.79 with p-value= .065. This is slightly less evidence of heteroskedasticity than provided by the B-P test, but the conclusion is very similar.

```
gen ressq=(residual_vote)^2
gen sqpred_voteA=pred_voteA^2
regress ressq pred_voteA sqpred_voteA
```

```
. regress ressq pred_voteA sqvote

      Source |       SS           df       MS      Number of obs   =       173
-------------+----------------------------------   F(2, 170)       =      2.79
       Model |  37152.5749         2  18576.2875   Prob > F        =    0.0645
    Residual |  1133582.99       170  6668.13521   R-squared       =    0.0317
-------------+----------------------------------   Adj R-squared   =    0.0203
       Total |  1170735.56       172   6806.6021   Root MSE        =    81.659

------------------------------------------------------------------------------
       ressq | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
  pred_voteA |  -4.263682   2.166534    -1.97   0.051    -8.540455    .0130913
      sqvote |   .0357354   .0212419     1.68   0.094    -.0061964    .0776672
       _cons |   171.8584   53.14213     3.23   0.001     66.95499    276.7619
------------------------------------------------------------------------------
```

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

Question 3

a) Expressions of variance of men and women

Given that the variance of u in the model is given by $var(\mu/male) = exp(\alpha_0 + \alpha_1 * male)$, the expression for the variance of males and females will be given by:

$$var(male) = exp(\alpha_0 + \alpha_1 * male)$$
$$var(female) = exp(\alpha_0)$$

The null hypothesis states that there is no difference in variance of μ is between males and females while alternative hypothesis states that there is a significant difference in variance of μ is between males and females.
In other words:

H0: $Var(u|x_1,..x_n) = \sigma^2$ data is homoscedastic

H1: There is heteroskedasticity present

b) Model building

The model is
$sleep = 3840.85 - .16 * totwrk - 11.71 * educ - 8.70 * age + .13 * agesq - .02 * yngkid + 87.75 * male$.
The model shows that there is a significant difference in average sleeping minutes between males and females with males having 87 more minutes sleeping in a week than their female counterparts at 5% level of significance when other factors are held constant. There is a notable difference in variance of mu between males and females with variance in sleeping among males being higher than that of females.

```
regress sleep totwrk educ age agesq yngkid male
predict res, r
```

Source	SS	df	MS	Number of obs	=	706
				F(6, 699)	=	16.30
Model	17092058.6	6	2848676.43	Prob > F	=	0.0000
Residual	122147777	699	174746.462	R-squared	=	0.1228
				Adj R-squared	=	0.1152
Total	139239836	705	197503.313	Root MSE	=	418.03

sleep	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
totwrk	-0.1634	0.0182	-8.997	0.0000	-0.1991	-0.1278
educ	-11.7133	5.8720	-1.995	0.0465	-23.2420	-0.1845
age	-8.6974	11.3291	-0.768	0.4429	-30.9405	13.5457
agesq	0.1284	0.1347	0.954	0.3405	-0.1360	0.3928
yngkid	-0.0228	50.2764	-0.000	0.9996	-98.7337	98.6881
male	87.7546	34.6679	2.531	0.0116	19.6888	155.8203
_cons	3840.8521	239.4139	16.043	0.0000	3370.7955	4310.9087

Reihaneh Moghisi
Jean Marc Kangah
Yagaize Anyanwu

	sleep	pred_s~p	resid_s~p
1.	3113	3079.405	33.59556
2.	2920	2798.047	121.9526
3.	2670	3135.421	-465.4208
4.	3083	2936.247	146.7534
5.	3448	3312.451	135.5492

c) Test for difference in variance

The coefficient of male is negative and a very high value (-28849.63) indicating that the variance of error is higher for female than for males. However, this difference is NOT statistically significant with p-value of 0.291 at 5% not even 20% percent of significance level.

Hence p-value >0.05 The result shows that there is no statistical difference in variance between males and females.

```
regress sleep totwrk educ age agesq yngkid male
predict res, r
gen ressq = res*res
regress ressq male
```

```
5 . regress ressq male
```

Source	SS	df	MS	Number of obs	=	706
				F(1, 704)	=	1.12
Model	1.4430e+11	1	1.4430e+11	Prob > F	=	0.2909
Residual	9.0942e+13	704	1.2918e+11	R-squared	=	0.0016
				Adj R-squared	=	0.0002
Total	9.1086e+13	705	1.2920e+11	Root MSE	=	3.6e+05

ressq	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
male	-28849.63	27296.51	-1.06	0.291	-82441.94	24742.69
_cons	189359.2	20546.36	9.22	0.000	149019.8	229698.7