

Answer Sheet Cover Page Midterm Examination Semester 2/2020

(Readable handwriting and printed version are acceptable)

Student Name Nattaphong Tianipar

Student ID 614041486

Course ID EE 426 Course Title Economics 2

Lecturer Assoc. Prof. Dr. Tatra Jantarakulchai

Exam date 29/7/21 Time 8.00-10.00

Total pages 22

***Students are responsible for checking the number of pages before submitting the answer sheets.
Incomplete submissions caused by carelessness will not be accepted as an excuse for resubmission**

Student Signature [Signature]

Date 29/7/21

1

Statistics/Data Analysis (R)

Statistics/Data Analysis (R) 14.0
MP - Parallel Edition

Copyright 1985-2015 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC <http://www.stata.com>
979-696-4600 stata@stata.com
979-696-4601 (fax)

Single-user 8-core Stata perpetual license:
Serial number: 10699393
Licensed to: T

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

- 1 . use "C:\Users\com\Downloads\04\Midterm_q1_4.dta", clear
(1980 Census housing data)
- 2 . log using "C:\Users\com\Downloads\04\ee426_q1.smcl"

name: <unnamed>
log: C:\Users\com\Downloads\04\ee426_q1.smcl
log type: smcl
opened on: 21 Mar 2021, 09:23:14

With OLS the estimators will be biased,
Inconsistent, and inefficient

a.

- 3 . reg y1 y2 x3

Source	SS	df	MS	Number of obs	=	50
Model	53701.8594	2	26850.9297	F(2, 47)	=	108.06
Residual	11678.5606	47	248.480013	Prob > F	=	0.0000
				R-squared	=	0.8214
				Adj R-squared	=	0.8138
Total	65380.42	49	1334.29429	Root MSE	=	15.763

y1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y2	.0009745	.0001932	5.04	0.000	.0005858 .0013632
x3	.0073336	.0010819	6.78	0.000	.005157 .0095101
_cons	44.21682	16.45559	2.69	0.010	11.11245 77.32119

- 4 . reg y2 y1 x1 x2

Source	SS	df	MS	Number of obs	=	50
Model	8.2479e+09	3	2.7493e+09	F(3, 46)	=	31.72
Residual	3.9872e+09	46	86678794	Prob > F	=	0.0000
				R-squared	=	0.6741
				Adj R-squared	=	0.6529
Total	1.2235e+10	49	249695445	Root MSE	=	9310.1

y2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y1	302.0238	45.88109	6.58	0.000	209.67 394.3776
x1	239.351	130.6043	1.83	0.073	-23.54177 502.2437
x2	-.0004897	.0003274	-1.50	0.142	-.0011488 .0001694
_cons	-36028.9	8783.98	-4.10	0.000	-53710.14 -18347.67

5 . reg3 (y1 y2 x3) (y2 y1 x1 x2),2sls inst(x1 x2 x3)

Two-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	F-Stat	P
y1	50	2	16.1045	0.8136	94.38	0.0000
y2	50	3	9316.633	0.6737	25.29	0.0000

6.)

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
y1						
y2	.0012516	.000508	2.46	0.016	.0002429	.0022604
x3	.0062856	.0020867	3.01	0.003	.0021419	.0104294
_cons	51.21041	20.54627	2.49	0.014	10.40958	92.01123
y2						
y1	290.4038	59.13173	4.91	0.000	172.9799	407.8277
x1	259.0293	145.1328	1.78	0.078	-29.17563	547.2343
x2	-.0005039	.0003308	-1.52	0.131	-.0011608	.000153
_cons	-34552.04	9984.793	-3.46	0.001	-54379.86	-14724.22

Endogenous variables: y1 y2
Exogenous variables: x1 x2 x3

6 Reduced form: $y_{1t} : \tilde{\pi}_{10} + \tilde{\pi}_{11} X_{1t} + \tilde{\pi}_{12} X_{2t} + \tilde{\pi}_{13} X_{3t} + u_{1t}$
 $y_{2t} : \tilde{\pi}_{20} + \tilde{\pi}_{21} X_{1t} + \tilde{\pi}_{22} X_{2t} + \tilde{\pi}_{23} X_{3t} + u_{2t}$

7 . ivregress 2sls y1 x3 (y2= x1 x2 x3)

Instrumental variables (2SLS) regression

Number of obs = 50
Wald chi2(2) = 200.81
Prob > chi2 = 0.0000
R-squared = 0.8136
Root MSE = 15.614

y1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
y2	.0012516	.0004925	2.54	0.011	.0002863	.0022169
x3	.0062856	.0020231	3.11	0.002	.0023204	.0102509
_cons	51.21041	19.92035	2.57	0.010	12.16724	90.25357

Instrumented: y2
Instruments: x3 x1 x2

8 . ivregress 2sls y2 x1 x2 (y1= x1 x2 x3)

Instrumental variables (2SLS) regression

Number of obs = 50
Wald chi2(3) = 82.47
Prob > chi2 = 0.0000
R-squared = 0.6737
Root MSE = 8936.2

y2	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
y1	290.4038	56.71716	5.12	0.000	179.2402	401.5674
x1	259.0293	139.2065	1.86	0.063	-13.81037	531.8691
x2	-.0005039	.0003173	-1.59	0.112	-.0011258	.000118
_cons	-34552.04	9577.077	-3.61	0.000	-53322.76	-15781.31

Instrumented: y1
Instruments: x1 x2 x3

9 . reg y1 x1 x2 x3

Source	SS	df	MS	Number of obs	=	50
Model	49028.5831	3	16342.861	F(3, 46)	=	45.97
Residual	16351.8369	46	355.474716	Prob > F	=	0.0000
				R-squared	=	0.7499
				Adj R-squared	=	0.7336
Total	65380.42	49	1334.29429	Root MSE	=	18.854

y1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x1	.5447944	.2531595	2.15	0.037	.0352108	1.054378
x2	-6.27e-07	6.57e-07	-0.95	0.345	-1.95e-06	6.96e-07
x3	.0096631	.0011563	8.36	0.000	.0073355	.0119906
_cons	12.59677	19.09739	0.66	0.513	-25.84429	51.03783

10 . predict y1hat, xb

11 . reg y2 x1 x2 x3

Source	SS	df	MS	Number of obs	=	50
Model	6.5854e+09	3	2.1951e+09	F(3, 46)	=	17.87
Residual	5.6497e+09	46	122819623	Prob > F	=	0.0000
				R-squared	=	0.5382
				Adj R-squared	=	0.5081
Total	1.2235e+10	49	249695445	Root MSE	=	11082

y2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x1	417.2397	148.8071	2.80	0.007	117.7065	716.773
x2	-.0006859	.0003863	-1.78	0.082	-.0014635	.0000916
x3	2.806191	.6796891	4.13	0.000	1.438048	4.174334
_cons	-30893.89	11225.45	-2.75	0.008	-53489.54	-8298.235

12 . predict y2hat, xb

13 . reg y1 y2hat x3

Source	SS	df	MS	Number of obs	=	50
Model	48956.6738	2	24478.3369	F(2, 47)	=	70.05
Residual	16423.7462	47	349.441408	Prob > F	=	0.0000
				R-squared	=	0.7488
				Adj R-squared	=	0.7381
Total	65380.42	49	1334.29429	Root MSE	=	18.693

y1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
y2hat	.0012516	.0005896	2.12	0.039	.0000654	.0024379
x3	.0062856	.0024221	2.60	0.013	.0014129	.0111583
_cons	51.21041	23.84915	2.15	0.037	3.232101	99.18871

14 . reg y2 y1hat x1 x2

Source	SS	df	MS	Number of obs	=	50
Model	6.5854e+09	3	2.1951e+09	F(3, 46)	=	17.87
Residual	5.6497e+09	46	122819634	Prob > F	=	0.0000
				R-squared	=	0.5382
				Adj R-squared	=	0.5081
Total	1.2235e+10	49	249695445	Root MSE	=	11082

y2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
y1hat	290.4038	70.33888	4.13	0.000	148.8189	431.9886
x1	259.0294	172.6396	1.50	0.140	-88.47608	606.535
x2	-.0005039	.0003935	-1.28	0.207	-.001296	.0002882
_cons	-34552.03	11877.2	-2.91	0.006	-58459.58	-10644.47

C. 15 . reg3 (y1 y2 x3) (y2 y1 x1 x2), 3sls inst(x1 x2 x3)

Three-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
y1	50	2	15.61389	0.8136	200.81	0.0000
y2	50	3	8954.393	0.6723	82.71	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
y1						
y2	.0012516	.0004925	2.54	0.011	.0002863	.0022169
x3	.0062856	.0020231	3.11	0.002	.0023204	.0102509
_cons	51.21041	19.92035	2.57	0.010	12.16724	90.25357
y2						
y1	285.0294	55.84722	5.10	0.000	175.5709	394.488
x1	270.6433	137.5542	1.97	0.049	1.042027	540.2445
x2	-.0004181	.0002751	-1.52	0.129	-.0009573	.0001212
_cons	-34463.42	9575.687	-3.60	0.000	-53231.42	-15695.42

Endogenous variables: y1 y2
Exogenous variables: x1 x2 x3

16 . *Single equation*

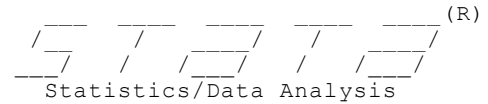
2SLS advantage: more consistent and asymptotically efficient compared to OLS
disadvantage: misspecified model can lead to unuseable estimation

System equations

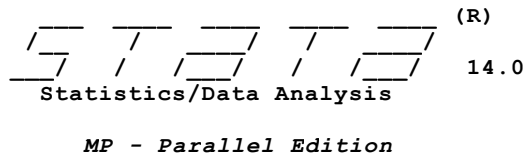
3SLS advantage: more consistent and asymptotically efficient compared to 2SLS
disadvantage: misspecified model can lead to unuseable estimation

OLS $\rightarrow$ OLS
2SLS $\rightarrow$ OLS + IV
3SLS $\rightarrow$ OLS + IV + FGLS

} with endogeneity exists $\Rightarrow$ All estimators will be biased but with 2SLS and 3SLS Asymptotic property both will be consistent and Asymptotic efficient



2



Copyright 1985-2015 StataCorp LP
 StataCorp
 4905 Lakeway Drive
 College Station, Texas 77845 USA
 800-STATA-PC <http://www.stata.com>
 979-696-4600 stata@stata.com
 979-696-4601 (fax)

Single-user 8-core Stata perpetual license:
 Serial number: 10699393
 Licensed to: T

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

```
1 . use "C:\Users\com\Downloads\04\Midterm_q2_4.dta", clear
2 . log using "C:\Users\com\Downloads\04\ee426_q2.smcl"
```

```
name: <unnamed>
log: C:\Users\com\Downloads\04\ee426_q2.smcl
log type: smcl
opened on: 21 Mar 2021, 09:42:23
```

```
3 .
```

```
4 .
5 .
```

```
6
```

a

```
7 . nl (lnC = {lnramda}-({beta}/{alpha})*ln({theta}*R^(-{alpha})+(1-{theta})*W^(-{alpha}))), init
> (lnramda 1 theta 0.5 beta 0.5 alpha -0.5)
(obs = 250)
```

```
Iteration 0: residual SS = 315.7594
Iteration 1: residual SS = 298.0254
Iteration 2: residual SS = 297.5135
Iteration 3: residual SS = 297.5127
Iteration 4: residual SS = 297.5127
Iteration 5: residual SS = 297.5127
Iteration 6: residual SS = 297.5127
Iteration 7: residual SS = 297.5127
```

Source	SS	df	MS
Model	44.833222	3	14.9444072
Residual	297.51272	246	1.20940131
Total	342.34594	249	1.37488331

```
Number of obs = 250
R-squared = 0.1310
Adj R-squared = 0.1204
Root MSE = 1.099728
Res. dev. = 752.9683
```

lnC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
/lnramda	2.150055	.7760475	2.77	0.006	.6215102	3.678601
/beta	.8120362	.1479719	5.49	0.000	.5205828	1.10349
/alpha	-1.070749	.7145536	-1.50	0.135	-2.478172	.3366745
/theta	.225971	.2422622	0.93	0.352	-.2512017	.7031437

Parameter lnramda taken as constant term in model & ANOVA table

```

8 . est store lognls1
9 . sca sigma=1/(1-(_b[/theta]))
10 . sca list sigma
    sigma = 1.2919413
11 . test (_b[/theta]=0) (_b[/alpha]=0) (_b[/beta]=0)

( 1) [theta]_cons = 0
( 2) [alpha]_cons = 0
( 3) [beta]_cons = 0

```

F(3, 246) = 37.78
 Prob > F = 0.0000 < 0.05 H_0 rejected $\theta \neq 0$ $\alpha \neq 0$ $\beta \neq 0$

6.) 12 . nl (lnC = {lnramda} - ({beta}/{alpha}) * ln({theta} * R^(-{alpha}) + (1-{theta}) * W^(-{alpha}))), ini
 > t(lnramda 0.5 theta 0.1 beta 0.1 alpha -0.1)
 (obs = 250)

```

Iteration 0: residual SS = 4522.524
Iteration 1: residual SS = 2782.494
Iteration 2: residual SS = 499.7826
Iteration 3: residual SS = 492.9609
Iteration 4: residual SS = 333.3437
Iteration 5: residual SS = 299.3888
Iteration 6: residual SS = 297.9689
Iteration 7: residual SS = 297.5269
Iteration 8: residual SS = 297.5129
Iteration 9: residual SS = 297.5127
Iteration 10: residual SS = 297.5127
Iteration 11: residual SS = 297.5127
Iteration 12: residual SS = 297.5127
Iteration 13: residual SS = 297.5127

```

Source	SS	df	MS		
Model	44.833222	3	14.9444072	Number of obs =	250
Residual	297.51272	246	1.20940131	R-squared =	0.1310
				Adj R-squared =	0.1204
				Root MSE =	1.099728
Total	342.34594	249	1.37488331	Res. dev. =	752.9683

lnC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
/lnramda	2.150055	.7760475	2.77	0.006	.6215102	3.678601
/beta	.8120362	.1479719	5.49	0.000	.5205828	1.10349
/alpha	-1.070749	.7145523	-1.50	0.135	-2.47817	.3366723
/theta	.2259711	.2422623	0.93	0.352	-.2512019	.7031441

Parameter lnramda taken as constant term in model & ANOVA table

Different sets at initial values can give different estimates results

```
13 . nl (lnC = {lnramda}-({beta}/{alpha})*ln({theta}*R^(-{alpha})+(1-{theta})*W^(-{alpha}))), ini
> t(lnramda 1 theta 0.5 beta 0.5 alpha -0.5) eps(1e-1)
(obs = 250)
```

```
Iteration 0: residual SS = 315.7594
Iteration 1: residual SS = 298.0254
Iteration 2: residual SS = 297.5135
```

Source	SS	df	MS	Number of obs =	250
Model	44.832469	3	14.9441564	R-squared =	0.1310
Residual	297.51347	246	1.20940437	Adj R-squared =	0.1204
Total	342.34594	249	1.37488331	Root MSE =	1.099729
				Res. dev. =	752.9689

lnC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
/lnramda	2.150245	.7760538	2.77	0.006	.6216872 3.678802
/beta	.8119886	.1478579	5.49	0.000	.5207596 1.103218
/alpha	-1.066218	.679886	-1.57	0.118	-2.405358 .2729223
/theta	.2263338	.2278233	0.99	0.321	-.2223993 .675067

Parameter beta taken as constant term in model & ANOVA table

```
14 . nl (lnC = {lnramda}-({beta}/{alpha})*ln({theta}*R^(-{alpha})+(1-{theta})*W^(-{alpha}))), ini
> t(lnramda 1 theta 0.5 beta 0.5 alpha -0.5) eps(1e-15) iter(40)
(obs = 250)
```

```
Iteration 0: residual SS = 315.7594
Iteration 1: residual SS = 298.0254
Iteration 2: residual SS = 297.5135
Iteration 3: residual SS = 297.5127
Iteration 4: residual SS = 297.5127
Iteration 5: residual SS = 297.5127
Iteration 6: residual SS = 297.5127
Iteration 7: residual SS = 297.5127
Iteration 8: residual SS = 297.5127
Iteration 9: residual SS = 297.5127
Iteration 10: residual SS = 297.5127
Iteration 11: residual SS = 297.5127
Iteration 12: residual SS = 297.5127
Iteration 13: residual SS = 297.5127
Iteration 14: residual SS = 297.5127
Iteration 15: residual SS = 297.5127
Iteration 16: residual SS = 297.5127
Iteration 17: residual SS = 297.5127
```

Source	SS	df	MS	Number of obs =	250
Model	11566.902	4	2891.72555	R-squared =	0.9749
Residual	297.51272	246	1.20940131	Adj R-squared =	0.9745
Total	11864.415	250	47.4576596	Root MSE =	1.099728
				Res. dev. =	752.9683

lnC	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
/lnramda	2.150055	.7760475	2.77	0.006	.6215102 3.678601
/beta	.8120362	.1479719	5.49	0.000	.5205828 1.10349
/alpha	-1.070749	.7145544	-1.50	0.135	-2.478174 .3366759
/theta	.2259709	.2422621	0.93	0.352	-.2512016 .7031435

With different convergence value can give the different estimated results

15 .

d.) With lag term the estimated result will be easier to compute and more robust

3

(R)
Statistics/Data Analysis

(R)
Statistics/Data Analysis 14.0
MP - Parallel Edition

Copyright 1985-2015 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC <http://www.stata.com>
979-696-4600 stata@stata.com
979-696-4601 (fax)

Single-user 8-core Stata perpetual license:
Serial number: 10699393
Licensed to: T

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

```
1 . use "C:\Users\com\Downloads\04\Midterm_q3_4.dta", clear
2 . log using "C:\Users\com\Downloads\04\ee426_q3.smcl"
```

```
name: <unnamed>
log: C:\Users\com\Downloads\04\ee426_q3.smcl
log type: smcl
opened on: 21 Mar 2021, 10:59:04
```

```
3 . do "C:\Users\com\Downloads\04\Midterm_q3 - ml_logit.do"
```

a.

```
4 . program ml_logit
   1.   args lnf theta
   2.   quietly replace `lnf'=ln(1/(1+exp(-`theta')))) if $ML_y1==1
   3.   quietly replace `lnf'=ln(1-(1/(1+exp(-`theta')))) if $ML_y1==0
   4. end

5 .
end of do-file

6 . ml model lf ml_logit (y=x1 x2 x3)

7 . ml maximize
```

```
initial:      log likelihood = -277.25887
alternative:  log likelihood = -266.63079
rescale:      log likelihood = -266.63079
Iteration 0:  log likelihood = -266.63079
Iteration 1:  log likelihood = -218.78011
Iteration 2:  log likelihood = -217.46841
Iteration 3:  log likelihood = -217.46691
Iteration 4:  log likelihood = -217.46691
```

Log likelihood = -217.46691

```
Number of obs      =      400
Wald chi2(3)       =      68.29
Prob > chi2        =      0.0000
```

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.2808583	.1167717	2.41	0.016	.0519899	.5097267
x2	-.5823592	.0721582	-8.07	0.000	-.7237868	-.4409317
x3	-.2448167	.1282455	-1.91	0.056	-.4961733	.0065399
_cons	.2458087	.1949767	1.26	0.207	-.1363385	.627956

```
8 . est store ur
9 . ml model lf ml_logit (y=x*), tech(bhhh)
10 . ml maximize
```

```
initial:      log likelihood = -277.25887
alternative:  log likelihood = -266.63079
rescale:      log likelihood = -266.63079
Iteration 0:  log likelihood = -266.63079
Iteration 1:  log likelihood = -218.31365
Iteration 2:  log likelihood = -217.47308
Iteration 3:  log likelihood = -217.46697
Iteration 4:  log likelihood = -217.46691
Iteration 5:  log likelihood = -217.46691
```

```

                                     Number of obs      =          400
                                     Wald chi2(3)         =          66.57
                                     Prob > chi2          =          0.0000

Log likelihood = -217.46691
```

y	OPG		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
x1	.2808443	.1189786	2.36	0.018	.0476505	.5140381
x2	-.5823674	.0740101	-7.87	0.000	-.7274246	-.4373102
x3	-.2448195	.1221053	-2.00	0.045	-.4841415	-.0054974
_cons	.2458228	.1962393	1.25	0.210	-.1387992	.6304448

```
11 . est store bhhh
12 . ml model lf ml_logit (y=x), tech(bfgs)
    x ambiguous abbreviation
    r(111);
13 . ml model lf ml_logit (y=x*), tech(bfgs)
14 . ml maximize
```

```
initial:      log likelihood = -277.25887
alternative:  log likelihood = -266.63079
rescale:      log likelihood = -266.63079
Iteration 0:  log likelihood = -266.63079
Iteration 1:  log likelihood = -258.52576 (backed up)
Iteration 2:  log likelihood = -254.22702 (backed up)
Iteration 3:  log likelihood = -250.56611
Iteration 4:  log likelihood = -249.60074
Iteration 5:  log likelihood = -228.08779
Iteration 6:  log likelihood = -218.74128
Iteration 7:  log likelihood = -217.47501
Iteration 8:  log likelihood = -217.46735
Iteration 9:  log likelihood = -217.46693
Iteration 10: log likelihood = -217.46691
```

```

                                     Number of obs      =          400
                                     Wald chi2(3)         =          68.30
                                     Prob > chi2          =          0.0000

Log likelihood = -217.46691
```

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.2809273	.1167753	2.41	0.016	.0520519	.5098026
x2	-.5824227	.0721628	-8.07	0.000	-.7238592	-.4409862
x3	-.2448659	.1282494	-1.91	0.056	-.4962301	.0064983
_cons	.2458029	.1949809	1.26	0.207	-.1363527	.6279584

15 . est store bfgs

16 . est table ur bhhh bfgs, star(.1 .05 .01) stat(N ll chi2 p)

Variable	ur	bhhh	bfgs
x1	.28085833**	.28084432**	.28092727**
x2	-.58235923***	-.5823674***	-.58242267***
x3	-.2448167*	-.24481947**	-.24486589*
_cons	.2458087	.24582284	.24580287
N	400	400	400
ll	-217.46691	-217.46691	-217.46691
chi2	68.290298	66.569347	68.297136
p	9.915e-15	2.315e-14	9.882e-15

legend: * p<.1; ** p<.05; *** p<.01

17 . do "C:\Users\com\Downloads\04\Midterm_q3 - ml_probit.do"

```
18 . program ml_probit
    1.   args lnf theta
    2.   tempvar z
    3.   quietly g double `z'=`theta'
    4.   quietly replace `lnf'=ln(normal(`z')) if $ML_y1==1
    5.   quietly replace `lnf'=ln(1-normal(`z')) if $ML_y1==0
    6.   end
```

19 .
end of do-file

20 . ml model lf ml_probit (y=x1 x2)

21 . ml maximize

```
initial:      log likelihood = -277.25887
alternative:  log likelihood = -271.85123
rescale:      log likelihood = -266.80559
Iteration 0:  log likelihood = -266.80559
Iteration 1:  log likelihood = -219.83398
Iteration 2:  log likelihood = -219.35724
Iteration 3:  log likelihood = -219.35696
Iteration 4:  log likelihood = -219.35696
```

```
Log likelihood = -219.35696

Number of obs      =      400
Wald chi2(2)       =      76.45
Prob > chi2        =      0.0000
```

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.1585088	.069309	2.29	0.022	.0226656	.294352
x2	-.3350693	.039372	-8.51	0.000	-.412237	-.2579016
_cons	.1398805	.1162269	1.20	0.229	-.08792	.3676809

Using different algorithm
will estimates a different
results because of different
computation function.

b.) LR test is more
preferable because this method
used 2 optimization which give
more accurate results.

c.) Because, MLE distribution is χ^2
Ans F have different distribution

```

22 . do "C:\Users\com\Downloads\04\Midterm_q3 - ml_probit_het.do"

23 . program ml_probit_het
    1.   args lnf theta sigma
    2.   tempvar z s
    3.   quietly g double `s'=exp(`sigma')
    4.   quietly g double `z'=`theta'/`s'
    5.   quietly replace `lnf'=ln(normal(`z')) if $ML_y1==1
    6.   quietly replace `lnf'=ln(1-normal(`z')) if $ML_y1==0
    7. end

24 .
    end of do-file

25 .

```

```

26 . ml model lf ml_probit_het (y=x1 x2) (x3, noconstant)

27 . ml maximize

```

```

initial:      log likelihood = -277.25887
alternative:  log likelihood = -271.834
rescale:      log likelihood = -266.8011
rescale eq:   log likelihood = -266.61459
Iteration 0:  log likelihood = -266.61459
Iteration 1:  log likelihood = -241.98721
Iteration 2:  log likelihood = -233.69206
Iteration 3:  log likelihood = -219.04695
Iteration 4:  log likelihood = -217.26474
Iteration 5:  log likelihood = -217.02239
Iteration 6:  log likelihood = -217.02113
Iteration 7:  log likelihood = -217.02113

```

```

Log likelihood = -217.02113
Number of obs   = 400
Wald chi2(2)    = 74.28
Prob > chi2     = 0.0000

```

	y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
eq1							
	x1	.1744698	.0664826	2.62	0.009	.0441662	.3047733
	x2	-.3334951	.0399312	-8.35	0.000	-.4117589	-.2552314
	_cons	.0952407	.1122231	0.85	0.396	-.1247127	.315194
eq2							
	x3	-.2116171	.0998641	-2.12	0.034	-.407347	-.0158872

```

28 .

```

```

29 . est store hetprob
30 . do "C:\Users\com\Downloads\04\Midterm_q3 - ml_probit.do"
31 . program ml_probit
    program ml_probit already defined
    r(110);

    end of do-file

    r(110);
32 . ml model lf ml_probit (y=x1 x2)
33 . ml maximize

```

```

initial:      log likelihood = -277.25887
alternative:  log likelihood = -271.85123
rescale:      log likelihood = -266.80559
Iteration 0:  log likelihood = -266.80559
Iteration 1:  log likelihood = -219.83398
Iteration 2:  log likelihood = -219.35724
Iteration 3:  log likelihood = -219.35696
Iteration 4:  log likelihood = -219.35696

```

```

Log likelihood = -219.35696
Number of obs   = 400
Wald chi2(2)    = 76.45
Prob > chi2     = 0.0000

```

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.1585088	.069309	2.29	0.022	.0226656	.294352
x2	-.3350693	.039372	-8.51	0.000	-.412237	-.2579016
_cons	.1398805	.1162269	1.20	0.229	-.08792	.3676809

```

34 . est store probit
35 . lr test hetprob probit
command lr is unrecognized
r(199);

```

d.)

```

36 . lrtest hetprob probit

Likelihood-ratio test
(Assumption: probit nested in hetprob)

```

```

LR chi2(1) = 4.67
Prob > chi2 = 0.0307

```

ca. as H₀ rejected

```

37 .

```

There exists heteroskedasticity in this case.

e.) Z-test is used because at the asymptotic distribution that give a normal distribution so Z can be applied. We don't have $\hat{\rho}^2$ because at non-linear property. E-likelihoods should be employed.

①

(R)

Statistics/Data Analysis

(R)

Statistics/Data Analysis 14.0

MP - Parallel Edition

Copyright 1985-2015 StataCorp LP
 StataCorp
 4905 Lakeway Drive
 College Station, Texas 77845 USA
 800-STATA-PC <http://www.stata.com>
 979-696-4600 stata@stata.com
 979-696-4601 (fax)

Single-user 8-core Stata perpetual license:
 Serial number: 10699393
 Licensed to: T

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

- 1 . use "C:\Users\com\Downloads\04\Midterm_q4-1_4.dta", clear
- 2 . log using "C:\Users\com\Downloads\04\Untitled.smcl", replace

name: <unnamed>
 log: C:\Users\com\Downloads\04\Untitled.smcl
 log type: smcl
 opened on: 21 Mar 2021, 14:36:45

- 3 . tsset time
 time variable: time, 1 to 1326
 delta: 1 unit
- 4 . gen dr=f.r-r
 (1 missing value generated)
- 5 . gmm (dr-{alpha}-{beta}*r) ((dr-{alpha}-{beta}*r)*r) ((dr-{alpha}-{beta}*r)^2-{sigma2}*r^(2*{gamma}
 > a))) (((dr-{alpha}-{beta}*r)^2-{sigma2}*r^(2*{gamma}))*r) winitial(identity)
 note: 1 missing value returned for equation 1 at initial values
 note: 1 missing value returned for equation 2 at initial values
 note: 1 missing value returned for equation 3 at initial values
 note: 1 missing value returned for equation 4 at initial values

Step 1

numerical derivatives are approximate
 flat or discontinuous region encountered
 Iteration 0: GMM criterion Q(b) = .00001285
 Iteration 1: GMM criterion Q(b) = 9.555e-06 (backed up)
 Iteration 2: GMM criterion Q(b) = 6.114e-06 (not concave)
 Iteration 3: GMM criterion Q(b) = 5.940e-06 (backed up)
 Iteration 4: GMM criterion Q(b) = 5.935e-06 (backed up)

Step 2

Iteration 0: GMM criterion Q(b) = .00875867
 could not calculate numerical derivatives -- flat or discontinuous region encountered
 could not calculate numerical derivatives -- flat or discontinuous region encountered
 r(430);

6 .

⑨ c.) with endogeneity Q29 will be biased inconsistent

4

Statistics/Data Analysis (R)

Statistics/Data Analysis (R) 14.0
MP - Parallel Edition

Copyright 1985-2015 StataCorp LP
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC <http://www.stata.com>
979-696-4600 stata@stata.com
979-696-4601 (fax)

Single-user 8-core Stata perpetual license:
Serial number: 10699393
Licensed to: T

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

```
1 . use "C:\Users\com\Downloads\04\Midterm_q5_4.dta", clear
2 . log using "C:\Users\com\Downloads\04\ee426_q5.smcl"
```

```
name: <unnamed>
log: C:\Users\com\Downloads\04\ee426_q5.smcl
log type: smcl
opened on: 21 Mar 2021, 11:31:29
```

```
3 . findit fitstat
4 . logit y x1 x2 x3
```

```
Iteration 0: log likelihood = -115.39801
Iteration 1: log likelihood = -108.82885
Iteration 2: log likelihood = -108.51434
Iteration 3: log likelihood = -108.51329
Iteration 4: log likelihood = -108.51329
```

Logistic regression

```
Number of obs = 215
LR chi2(3) = 13.77
Prob > chi2 = 0.0032
Pseudo R2 = 0.0597
```

Log likelihood = -108.51329

Overall test is significant

used for comparison

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.006117	.0026021	2.35	0.019	.001017	.0112169
x2	.0195677	.0154868	1.26	0.206	-.0107858	.0499212
x3	.1721685	.0801424	2.15	0.032	.0150924	.3292447
_cons	.7341525	.20264	3.62	0.000	.3369854	1.13132

```
5 . fitstat
```

Measures of Fit for **logit** of **y**

Individual test

Log-Lik Intercept Only:	-115.398	Log-Lik Full Model:	-108.513
D(211):	217.027	LR(3):	13.769
		Prob > LR:	0.003
McFadden's R2:	0.060	McFadden's Adj R2:	0.025
Maximum Likelihood R2:	0.062	Cragg & Uhler's R2:	0.094
McKelvey and Zavoina's R2:	0.137	Efron's R2:	0.056
Variance of y*:	3.814	Variance of error:	3.290
Count R2:	0.772	Adj Count R2:	0.000
AIC:	1.047	AIC*n:	225.027
BIC:	-916.178	BIC':	2.342

6 . probit y x1 x2 x3

Iteration 0: log likelihood = -115.39801
 Iteration 1: log likelihood = -108.55308
 Iteration 2: log likelihood = -108.46762
 Iteration 3: log likelihood = -108.46758
 Iteration 4: log likelihood = -108.46758

Probit regression	Number of obs	=	215
	LR chi2(3)	=	13.86
	Prob > chi2	=	0.0031
Log likelihood = -108.46758	Pseudo R2	=	0.0601

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
x1	.0035049	.0014387	2.44	0.015	.0006852 .0063246
x2	.0102078	.0082103	1.24	0.214	-.005884 .0262996
x3	.0948062	.0427046	2.22	0.026	.0111068 .1785057
_cons	.4573707	.1225489	3.73	0.000	.2171792 .6975622

7 . fitstat

Measures of Fit for **probit** of y

Log-Lik Intercept Only:	-115.398	Log-Lik Full Model:	-108.468
D(211):	216.935	LR(3):	13.861
		Prob > LR:	0.003
McFadden's R2:	0.060	McFadden's Adj R2:	0.025
Maximum Likelihood R2:	0.062	Cragg & Uhler's R2:	0.095
McKelvey and Zavoina's R2:	0.141	Efron's R2:	0.056
Variance of y*:	1.165	Variance of error:	1.000
Count R2:	0.772	Adj Count R2:	0.000
AIC:	1.046	AIC*n:	224.935
BIC:	-916.269	BIC':	2.251

We cannot compare the estimated coefficients because of the difference in distribution

8 . mfx

Marginal effects after probit
 y = Pr(y) (predict)
 = .79072583

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]	X
x1	.0010081	.00041	2.49	0.013	.000213 .001803	60.8632
x2	.0029359	.00234	1.25	0.210	-.001654 .007526	-1.60773
x3	.0272677	.01206	2.26	0.024	.003632 .050904	1.63138

Change in X_{11} by 1 unit (From 60.8632 to 61.8632) $\hat{P}(y=1)$ will increase 0.001 or 0.1% holding X_2, X_3 constant.

9 . mfx, at(median)

Marginal effects after probit
y = Pr(y) (predict)
= .68501037

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]	X
x1	.001245	.00055	2.27	0.023	.00017 .00232	.753216
x2	.0036261	.00293	1.24	0.216	-.002121 .009374	.54305
x3	.0336783	.01576	2.14	0.033	.002788 .064568	.170897

10 . qui logit y x1 x2 x3

11 . test x1 x2 x3

(1) [y]x1 = 0
(2) [y]x2 = 0
(3) [y]x3 = 0

chi2(3) = 10.78
Prob > chi2 = 0.0130

12 . qui probit y x1 x2 x3

13 . test x1 x2 x3

(1) [y]x1 = 0
(2) [y]x2 = 0
(3) [y]x3 = 0

chi2(3) = 12.04
Prob > chi2 = 0.0073

14 . test (x1=x2)

(1) [y]x1 - [y]x2 = 0

chi2(1) = 0.70
Prob > chi2 = 0.4044 > 0.05

15 . qui logit y x1 x2 x3

16 . test (x1=x2)

(1) [y]x1 - [y]x2 = 0

chi2(1) = 0.81
Prob > chi2 = 0.3679 > 0.05

Fails to reject H_0 that is, $\beta_1 = \beta_2$

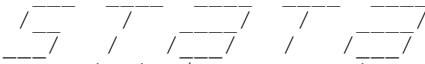
17 . log close

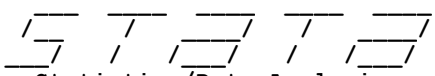
name: <unnamed>
log: C:\Users\com\Downloads\04\ee426_q5.smcl
log type: smcl
closed on: 21 Mar 2021, 11:56:24

18 .

2.) Because the change in threshold can lead to different interpretation whether it is $y_i \geq 0.5$ or < 0.5 . And it will change the value of computed R^2 with different computation.

6.

(R)

 Statistics/Data Analysis

(R)

Statistics/Data Analysis
MP - Parallel Edition

14.0 Copyright 1985-2015 StataCorp LP
 StataCorp
 4905 Lakeway Drive
 College Station, Texas 77845 USA
 800-STATA-PC <http://www.stata.com>
 979-696-4600 stata@stata.com
 979-696-4601 (fax)

Single-user 8-core Stata perpetual license:
 Serial number: 10699393
 Licensed to: T

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

```
1 . use "C:\Users\com\Downloads\04\Midterm_q6_4.dta", clear
2 . log using "C:\Users\com\Downloads\04\ee426_q6.smcl"
```

```
name: <unnamed>
log: C:\Users\com\Downloads\04\ee426_q6.smcl
log type: smcl
opened on: 21 Mar 2021, 11:57:47
```

```
3 .
4 . xtset id
    panel variable: id (balanced)
5 . xtset id t
    panel variable: id (strongly balanced)
    time variable: t, 1 to 12
    delta: 1 unit
6 . xtglm y x1 x2 x3
```

Cross-sectional time-series FGLS regression

Coefficients: **generalized least squares**
 Panels: **homoskedastic**
 Correlation: **no autocorrelation**

Estimated covariances	=	1	Number of obs	=	1,200
Estimated autocorrelations	=	0	Number of groups	=	100
Estimated coefficients	=	4	Time periods	=	12
			Wald chi2(3)	=	29882.41
Log likelihood	=	-6211.29	Prob > chi2	=	0.0000

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.3183087	.0109146	29.16	0.000	.2969164	.3397011
x2	1.320714	.0149495	88.34	0.000	1.291413	1.350014
x3	-.4642183	.011884	-39.06	0.000	-.4875104	-.4409262
_cons	-136.2145	6.645908	-20.50	0.000	-149.2402	-123.1887

```

7 . xtglm y x1 x2 x3, igls panels(heteroskedastic)
Iteration 1: tolerance = .00590371
Iteration 2: tolerance = .00177471
Iteration 3: tolerance = .00055434
Iteration 4: tolerance = .000177
Iteration 5: tolerance = .00005721
Iteration 6: tolerance = .00001862
Iteration 7: tolerance = 6.085e-06
Iteration 8: tolerance = 1.993e-06
Iteration 9: tolerance = 6.534e-07
Iteration 10: tolerance = 2.144e-07
Iteration 11: tolerance = 7.040e-08

```

Cross-sectional time-series FGLS regression

Coefficients: **generalized least squares**
Panels: **heteroskedastic**
Correlation: **no autocorrelation**

Estimated covariances	=	100	Number of obs	=	1,200
Estimated autocorrelations	=	0	Number of groups	=	100
Estimated coefficients	=	4	Time periods	=	12
			Wald chi2(3)	=	33298.40
Log likelihood	=	-6165.009	Prob > chi2	=	0.0000

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.3168738	.0100633	31.49	0.000	.2971501	.3365975
x2	1.311977	.0139878	93.79	0.000	1.284562	1.339393
x3	-.4630102	.011065	-41.84	0.000	-.4846972	-.4413232
_cons	-132.2058	6.140582	-21.53	0.000	-144.2411	-120.1705

```

8 . est store panelhet

```

```

9 . xtglm y x1 x2 x3

```

Cross-sectional time-series FGLS regression

Coefficients: **generalized least squares**
Panels: **homoskedastic**
Correlation: **no autocorrelation**

Estimated covariances	=	1	Number of obs	=	1,200
Estimated autocorrelations	=	0	Number of groups	=	100
Estimated coefficients	=	4	Time periods	=	12
			Wald chi2(3)	=	29882.41
Log likelihood	=	-6211.29	Prob > chi2	=	0.0000

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.3183087	.0109146	29.16	0.000	.2969164	.3397011
x2	1.320714	.0149495	88.34	0.000	1.291413	1.350014
x3	-.4642183	.011884	-39.06	0.000	-.4875104	-.4409262
_cons	-136.2145	6.645908	-20.50	0.000	-149.2402	-123.1887

```
10 . eststore panel
    command eststore is unrecognized
    r(199);

11 . est store panel

12 . local df=e(N_g)-1

13 . lrtest panelhet, df(`df')
```

Heteroskedasticity exists,
and this causes the model
to be inefficient

Likelihood-ratio test
(Assumption: panel nested in panelhet)

LR chi2(99) = 92.56
Prob > chi2 = 0.6629 > 0.05

```
14 . xtreg y x1 x2 x3, fe
```

Fixed-effects (within) regression
Group variable: **id**

Number of obs = 1,200
Number of groups = 100

R-sq:

within = 0.9502
between = 0.9848
overall = 0.8846

Obs per group:
min = 12
avg = 12.0
max = 12

corr(u_i, Xb) = 0.8057

F(3,1097) = 6980.72
Prob > F = 0.0000

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x1	.1014426	.0044866	22.61	0.000	.0926393	.1102459
x2	.6951028	.0085422	81.37	0.000	.678342	.7118637
x3	-.4953168	.0041895	-118.23	0.000	-.5035372	-.4870965
_cons	208.659	4.373144	47.71	0.000	200.0784	217.2397
sigma_u	123.46108					
sigma_e	14.376737					
rho	.98662139	(fraction of variance due to u_i)				

F test that all u_i=0: F(99, 1097) = 96.47 Prob > F = 0.0000

```
15 . est store fe
```

```
16 . xtreg y x1 x2 x3, re
```

Random-effects GLS regression
Group variable: **id**

Number of obs = 1,200
Number of groups = 100

R-sq:

within = 0.8956
between = 0.9953
overall = 0.9543

Obs per group:
min = 12
avg = 12.0
max = 12

corr(u_i, X) = 0 (assumed)

Wald chi2(3) = 8707.50
Prob > chi2 = 0.0000

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	.2162513	.0090869	23.80	0.000	.1984414	.2340613
x2	1.028607	.0152844	67.30	0.000	.9986503	1.058564
x3	-.4779339	.0091412	-52.28	0.000	-.4958503	-.4600176
_cons	25.24251	8.000206	3.16	0.002	9.562392	40.92262
sigma_u	12.351151					
sigma_e	14.376737					
rho	.42464732	(fraction of variance due to u_i)				

```
17 . est store re
18 . hausman fe re
```

	Coefficients			
	(b) fe	(B) re	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
x1	.1014426	.2162513	-.1148087	.
x2	.6951028	1.028607	-.3335042	.
x3	-.4953168	-.4779339	-.0173829	.

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= -1451.21 chi2<0 ==> model fitted on these
data fails to meet the asymptotic
assumptions of the Hausman test;
see suest for a generalized test

```
19 . log close
    name: <unnamed>
    log: C:\Users\com\Downloads\04\ee426_q6.smcl
    log type: smcl
    closed on: 21 Mar 2021, 12:12:55
```

20 . $H_0 : \beta_{FE} = \beta_{RE}$

$H_a : \beta_{FE} \neq \beta_{RE}$

H_0 is rejected so, fixed effects is applied, so more appropriate model would be FE model

c.) The difference between fixed effects and first difference is that fixed effects will use mean to be fixed and compare with any period but for first period it uses only t and t-1 periods to compare

Fixed effects and Random effects differ with the fact that for fixed effects unknown α correlates with x_i but Random effects is not

d.)