

Illustration 1

Aim:

- Explaining annual health care expenditures (*med*) and the annual number of medical consultations (*mdu*) as a function of health insurance characteristics (*lcoins*), socio-economic factors (*female, age, lfam, child*) and health status variables (*ndisease*)

Details:

- Cameron and Trivedi (2005), ch. 16.6

Illustration 1 – Question 1

```
. summarize med
```

Variable	Obs	Mean	Std. Dev.	Min	Max
med	20186	171.5892	698.2689	0	39182.02

```
. summarize med if med>0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
med	15733	220.1551	784.1543	.4276146	39182.02

- Non-negative outcome
- High proportion of zeros: 4453 observations (22.1% of the sample units)

Illustration 1 – Question 2.1

```
. drop if year!=1  
(14548 observations deleted)
```

```
. poisson med lcoins ndisease female age lfam child, robust  
(...)
```

		Robust				[95% Conf. Interval]	
med	Coef.	Std. Err.	z	P> z			
lcoins	-.0478877	.0245805	-1.95	0.051	-.0960646	.0002892	
ndisease	.0281757	.0060746	4.64	0.000	.0162696	.0400817	
female	.2263111	.1192456	1.90	0.058	-.007406	.4600282	
age	.0097846	.003701	2.64	0.008	.0025307	.0170384	
lfam	-.0928784	.0882497	-1.05	0.293	-.2658446	.0800879	
child	-.4112285	.2444742	-1.68	0.093	-.8903892	.0679321	
_cons	4.631177	.2039135	22.71	0.000	4.231514	5.03084	

```
. estimates store poisson
```

Illustration 1 – Question 2.2

```
. poisson med lcoins ndisease female age lfam child if med>0, robust
note: you are responsible for interpretation of noncount dep. variable
(...)
```

```
Poisson regression                               Number of obs   =       4451
                                                Wald chi2(6)    =       66.48
                                                Prob > chi2     =       0.0000
Log pseudolikelihood = -1087511.2                Pseudo R2      =       0.0657
```

		Robust				[95% Conf. Interval]	
	med	Coef.	Std. Err.	z	P> z		
	lcoins	-.0200287	.0246889	-0.81	0.417	-.068418	.0283606
	ndisease	.0210867	.0062785	3.36	0.001	.0087812	.0333923
	female	.1474704	.1177851	1.25	0.211	-.0833842	.3783249
	age	.0087803	.0036967	2.38	0.018	.0015349	.0160258
	lfam	-.0823594	.0902764	-0.91	0.362	-.2592978	.0945791
	child	-.3678983	.2481063	-1.48	0.138	-.8541777	.1183811
	_cons	4.91825	.2098069	23.44	0.000	4.507036	5.329464

```
. estimates store poisson0
```

Illustration 1 – Question 2.3

```
. gen lmed=log(med)
(1187 missing values generated)
```

```
. regress lmed lcoins ndisease female age lfam child if med>0, robust
(...)
```

```
-----+-----
```

	lmed	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
	lcoins	-.054901	.0099633	-5.51	0.000	-.0744341	-.0353679
	ndisease	.030717	.0030927	9.93	0.000	.0246537	.0367803
	female	.1199091	.0417646	2.87	0.004	.0380298	.2017885
	age	.0072817	.0020791	3.50	0.000	.0032055	.0113578
	lfam	-.1357759	.0442937	-3.07	0.002	-.2226136	-.0489382
	child	-.4972463	.0737896	-6.74	0.000	-.6419107	-.3525819
	_cons	3.954487	.1043663	37.89	0.000	3.749877	4.159096

```
-----+-----
```

```
. estimates store log0
```

Illustration 1 – Question 2.4

```
. gen lmed1=log(med+1)
```

```
. regress lmed1 lcoins ndisease female age lfam child, robust
```

```
(...)
```

```
-----
```

		Robust				[95% Conf. Interval]	
	lmed1	Coef.	Std. Err.	t	P> t		
	lcoins	-.1485372	.0126915	-11.70	0.000	-.1734174	-.1236569
	ndisease	.0584591	.0039132	14.94	0.000	.0507878	.0661304
	female	.3064252	.0535977	5.72	0.000	.201353	.4114974
	age	.0089512	.0027704	3.23	0.001	.0035202	.0143821
	lfam	-.1806985	.0556686	-3.25	0.001	-.2898305	-.0715666
	child	-.4759264	.0946324	-5.03	0.000	-.6614425	-.2904103
	_cons	3.000465	.1346348	22.29	0.000	2.736529	3.264401

```
-----
```

```
. estimates store log1
```

Illustration 1 – Question 2 (summary)

```
. estimates table poisson poisson0, b star(0.1 0.05 0.01)
```

Variable	poisson	poisson0
lcoins	-.04788772*	-.0200287
ndisease	.02817566***	.02108674***
female	.22631112*	.14747037
age	.00978456***	.00878032**
lfam	-.09287836	-.08235938
child	-.41122852*	-.3678983
_cons	4.6311773***	4.9182498***

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

```
. estimates table log0 log1, b star(0.1 0.05 0.01)
```

Variable	log0	log1
lcoins	-.05490096***	-.14853716***
ndisease	.03071698***	.0584591***
female	.11990912***	.3064252***
age	.00728165***	.00895117***
lfam	-.13577591***	-.18069853***
child	-.4972463***	-.47592639***
_cons	3.9544865***	3.0004653***

Illustration 1 – Question 3.1

```
. summarize mdu
```

Variable	Obs	Mean	Std. Dev.	Min	Max
mdu	5638	2.877971	4.332918	0	69

```
. tabulate mdu
```

number			
face-to-fac			
e md visits	Freq.	Percent	Cum.
0	1,729	30.67	30.67
1	1,047	18.57	49.24
2	814	14.44	63.68
3	511	9.06	72.74
4	385	6.83	79.57
5	271	4.81	84.37
6	188	3.33	87.71
7	156	2.77	90.48
8	134	2.38	92.85
9	80	1.42	94.27
10	56	0.99	95.26
11	56	0.99	96.26
12	32	0.57	96.83

```
(...)
```


Illustration 1 – Question 3.2.1

```
. poisson mdu lcoins ndisease female age lfam child
```

```
Poisson regression                               Number of obs   =           5638
                                                LR chi2(6)      =           2296.50
                                                Prob > chi2     =             0.0000
Log likelihood = -17173.058                    Pseudo R2       =             0.0627
```

```
-----+-----
```

mdu	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
lcoins	-.0682563	.0038214	-17.86	0.000	-.0757462	-.0607665
ndisease	.03538	.0010286	34.40	0.000	.0333639	.037396
female	.1276798	.0164078	7.78	0.000	.0955211	.1598386
age	.0041167	.0007841	5.25	0.000	.0025798	.0056536
lfam	-.141855	.0154716	-9.17	0.000	-.1721789	-.1115312
child	.056673	.0278998	2.03	0.042	.0019903	.1113557
_cons	.7463066	.0385653	19.35	0.000	.67072	.8218932

```
-----+-----
```

```
. estimates store cpoissonml
```

Illustration 1 – Question 3.2.2

```
. poisson mdu lcoins ndisease female age lfam child, robust
```

```
Poisson regression                               Number of obs   =       5638
                                                Wald chi2(6)    =       388.20
                                                Prob > chi2     =       0.0000
Log pseudolikelihood = -17173.058                Pseudo R2      =       0.0627
```

```
-----+-----
```

		Robust				
mdu	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
lcoins	-.0682563	.0094538	-7.22	0.000	-.0867854 -.0497273	
ndisease	.03538	.002696	13.12	0.000	.0300959 .0406641	
female	.1276798	.04043	3.16	0.002	.0484385 .2069211	
age	.0041167	.0019155	2.15	0.032	.0003624 .007871	
lfam	-.141855	.0336317	-4.22	0.000	-.207772 -.0759381	
child	.056673	.0625248	0.91	0.365	-.0658734 .1792194	
_cons	.7463066	.0861806	8.66	0.000	.5773956 .9152175	

```
-----+-----
```

```
. estimates store cpoissonqml
```

Illustration 1 – Question 3.2.3

```
. nbreg mdu lcoins ndisease female age lfam child, dispersion(constant)
```

```
Negative binomial regression          Number of obs   =       5638
                                      LR chi2(6)       =       529.02
Dispersion      = constant            Prob > chi2     =       0.0000
Log likelihood = -12128.041           Pseudo R2      =       0.0213
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
mdu						
lcoins	-.076432	.0069326	-11.02	0.000	-.0900197	-.0628442
ndisease	.0309127	.0019493	15.86	0.000	.0270922	.0347333
female	.1447002	.029679	4.88	0.000	.0865305	.2028699
age	.0027029	.0014494	1.86	0.062	-.0001378	.0055436
lfam	-.1235042	.0276393	-4.47	0.000	-.1776761	-.0693323
child	-.0049686	.0499784	-0.10	0.921	-.1029245	.0929873
_cons	.8537349	.0697031	12.25	0.000	.7171195	.9903504
/lndelta	1.278115	.0319841			1.215427	1.340802
delta	3.589865	.1148187			3.371733	3.822109

```
Likelihood-ratio test of delta=0:  chibar2(01) = 1.0e+04 Prob>=chibar2 = 0.000
```

Overdispersion test:
the hypothesis of
correct specification
of the Poisson mode
is rejected

```
. estimates store cnegbin1
```

Illustration 1 – Question 3.2.4

```
. nbreg mdu lcoins ndisease female age lfam child, dispersion(mean)
```

```
Negative binomial regression          Number of obs   =       5638
                                      LR chi2(6)       =       459.10
Dispersion      = mean                Prob > chi2      =       0.0000
Log likelihood = -12163.001           Pseudo R2       =       0.0185
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
mdu						
lcoins	-.0719827	.0083402	-8.63	0.000	-.0883292	-.0556362
ndisease	.0391488	.002618	14.95	0.000	.0340175	.0442801
female	.1229883	.0350213	3.51	0.000	.0543479	.1916288
age	.0035899	.0017692	2.03	0.042	.0001223	.0070574
lfam	-.1834739	.036481	-5.03	0.000	-.2549754	-.1119723
child	.0842864	.0621692	1.36	0.175	-.0375631	.2061359
_cons	.7638038	.087809	8.70	0.000	.5917012	.9359063
/lnalpha	.2193579	.0274111			.1656333	.2730826
alpha	1.245277	.0341344			1.18014	1.314009

```
Likelihood-ratio test of alpha=0:  chibar2(01) = 1.0e+04 Prob>=chibar2 = 0.000
```

Overdispersion test:
the hypothesis of
correct specification
of the Poisson mode
is rejected

```
. estimates store cnegbin2
```

Illustration 1 – Question 3.2 (summary)

```
. estimates table cpoissonml cpoissonqml cnegbin1 cnegbin2, b star(0.1 0.05 0.01)
```

Variable	cpoissonml	cpoissonqml	cnegbin1	cnegbin2
-----+-----				
mdu				
lcoins	-.06825635***	-.06825635***	-.07643198***	-.0719827***
ndisease	.03537998***	.03537998***	.03091273***	.03914881***
female	.12767984***	.12767984***	.1447002***	.12298834***
age	.0041167***	.0041167**	.0027029*	.00358987**
lfam	-.14185505***	-.14185505***	-.1235042***	-.18347387***
child	.05667301**	.05667301	-.00496864	.0842864
_cons	.74630658***	.74630658***	.85373493***	.76380377***
-----+-----				
lndelta				
_cons			1.2781147***	
-----+-----				
lnalpha				
_cons				.21935795***
-----+-----				

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

Illustration 1 – Question 3.3

```
. estimates restore cpoissonml  
(results cpoissonml are active now)
```

Formulas:

$$Pr(Y_i = y|x_i) = \frac{e^{-\lambda_i} \lambda_i^y}{y!}, \quad \lambda_i = E(Y|X) = \exp(x_i' \beta)$$

- $Pr(Y_i = 0|x_i) = e^{-\lambda_i}$
- $Pr(Y_i = 1|x_i) = e^{-\lambda_i} \lambda_i$
- $Pr(Y_i \geq 2|x_i) = 1 - Pr(Y_i = 0|x_i) - Pr(Y_i = 1|x_i)$

```
. scalar lambda0 = exp(_b[_cons] + _b[lcoins]*log(1) + _b[age]*50 +  
_b[lfam]*log(3))
```

```
. scalar lambda1 = exp(_b[_cons] + _b[lcoins]*log(51) + _b[age]*50 +  
_b[lfam]*log(3))
```

```
. scalar lambda2 = exp(_b[_cons] + _b[lcoins]*log(101) + _b[age]*50 +  
_b[lfam]*log(3))
```

Note that lcoins is defined in the dataset as $\ln(\text{coins}+1)$ to avoid discarding 0's

Illustration 1 – Question 3.3 (cont.)

```
. display lambda0  
2.2173168
```

```
. display lambda1  
1.695412
```

```
. display lambda2  
1.6181549
```

```
. display exp(-lambda0)  
.10890092
```

```
. display exp(-lambda1)  
.18352361
```

```
. display exp(-lambda2)  
.19826418
```

```
. display exp(-lambda0)*lambda0  
.24146784
```

```
. display exp(-lambda1)*lambda1  
.31114812
```

```
. display exp(-lambda2)*lambda2  
.32082215
```

```
. display 1-0.10890092-0.24146784  
.64963124
```

```
. display 1-0.18352361-0.31114812  
.50532827
```

```
. display 1-0.19826418-0.32082215  
.48091367
```

Illustration 1 – Question 3.3 (cont.)

<i>coins:</i>	0	50	100
$E(mdu \dots)$	2.2	1.7	1.6
$Pr(mdu = 0 \dots)$	10.9	18.4	19.8
$Pr(mdu = 1 \dots)$	24.1	31.1	32.1
$Pr(mdu \geq 2 \dots)$	65.0	50.5	48.1

Illustration 1 – Question 4.1

```
. xtset id year
      panel variable:  id (unbalanced)
      time variable:  year, 1 to 5, but with gaps
                  delta:  1 unit
```

```
. xtdescribe
      id:  125024, 125025, ..., 632167          n =          5908
      year:  1, 2, ..., 5                      T =           5
      Delta(year) = 1 unit
      Span(year)  = 5 periods
      (id*year uniquely identifies each observation)
```

```
Distribution of T_i:  min      5%      25%      50%      75%      95%      max
                    1        2        3        3        5        5        5
      Freq.  Percent  Cum.  |  Pattern
-----+-----
      3710    62.80   62.80 |  111..
      1584    26.81   89.61 |  11111
       156     2.64   92.25 |  1....
       147     2.49   94.74 |  11...
        79     1.34   96.07 |  ..1..
        66     1.12   97.19 |  .11..
        33     0.56   97.75 |  ..111
        33     0.56   98.31 |  .1111
        29     0.49   98.80 |  ...11
        71     1.20  100.00 |  (other patterns)
-----+-----
      5908   100.00           |  XXXXX
```

Illustration 1 – Question 4.2.1

```
. poisson mdu lcoins ndisease female age lfam child, vce(cluster id)
```

```
Poisson regression                               Number of obs   =       20186
                                                Wald chi2(6)    =       476.93
                                                Prob > chi2     =       0.0000
Log pseudolikelihood = -62579.401                Pseudo R2      =       0.0609
                                                (Std. Err. adjusted for 5908 clusters in id)
```

```
-----
```

		Robust				[95% Conf. Interval]	
mdu	Coef.	Std. Err.	z	P> z			
lcoins	-.0808023	.0080013	-10.10	0.000	-.0964846	-.0651199	
ndisease	.0339334	.0026024	13.04	0.000	.0288328	.039034	
female	.1717862	.0342551	5.01	0.000	.1046473	.2389251	
age	.0040585	.0016891	2.40	0.016	.000748	.0073691	
lfam	-.1481981	.0323434	-4.58	0.000	-.21159	-.0848062	
child	.1030453	.0506901	2.03	0.042	.0036944	.2023961	
_cons	.748789	.0785738	9.53	0.000	.5947872	.9027907	

```
-----
```

```
. estimates store pooled
```

Note that s.e. are almost 4 times larger than their standard version, that overlooks overdispersion

Illustration 1 – Question 4.2.2

```
. xtpoisson mdu lcoins ndisease female age lfam child, re  
(...)
```

mdu	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
lcoins	-.0878258	.0068682	-12.79	0.000	-.1012873	-.0743642
ndisease	.0387629	.0022046	17.58	0.000	.034442	.0430839
female	.1667192	.0286298	5.82	0.000	.1106058	.2228325
age	.0019159	.0011134	1.72	0.085	-.0002663	.0040982
lfam	-.1351786	.0260022	-5.20	0.000	-.186142	-.0842152
child	.1082678	.0341477	3.17	0.002	.0413396	.1751961
_cons	.7574177	.0618346	12.25	0.000	.6362241	.8786112
/lnalpha	.0251256	.0209586			-.0159526	.0662038
alpha	1.025444	.0214919			.984174	1.068444

```
Likelihood-ratio test of alpha=0: chibar2(01) = 3.9e+04 Prob>=chibar2 = 0.000
```

```
. estimates store RE
```

Illustration 1 – Question 4.2.3

```
. xtpoisson mdu lcoins ndisease female age lfam child, fe
note: 265 groups (265 obs) dropped because of only one obs per group
note: 666 groups (2130 obs) dropped because of all zero outcomes
note: lcoins dropped because it is constant within group
note: ndisease dropped because it is constant within group
note: female dropped because it is constant within group
(...)
```

```
Log likelihood = -24173.211          Wald chi2(3)          =          19.24
                                   Prob > chi2            =          0.0002
```

```
-----+-----
```

mdu	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	-.0112009	.0039024	-2.87	0.004	-.0188494	-.0035523
lfam	.0877134	.0554606	1.58	0.114	-.0209874	.1964141
child	.1059867	.0437744	2.42	0.015	.0201905	.1917829

```
-----+-----
```

```
. estimates store FE
```

Illustration 1 – Question 4.2 (summary)

```
. estimates table pooled RE FE, b star(0.1 0.05 0.01)
```

Variable	pooled	RE	FE
-----+-----			
mdu			
lcoins	-.08080226***	-.08782576***	
ndisease	.0339334***	.03876295***	
female	.17178621***	.16671918***	
age	.00405853**	.00191594*	-.01120087***
lfam	-.1481981***	-.13517858***	.08771336
child	.10304526**	.10826785***	.10598669**
_cons	.74878896***	.7574177***	
-----+-----			
lnalpha			
_cons		.02512562	

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

Note that relative to pooled and RE results, FE:

- Coefficients differ substantially (allow correlation between individual heterogeneity and time-varying covariates)
- Tend to present larger s.e. for identified coefficients (less significance)

Illustration 1 – Question 4.3

```
. hausman FE RE
```

```
----- Coefficients -----
      |      (b)      (B)      (b-B)      sqrt(diag(V_b-V_B))
      |      EF      EA      Difference      S.E.
-----+-----
      |
age   |   -.0112009   .0019159   -.0131168   .0037402
lfam  |   .0877134  -.1351786   .2228919   .0489874
child |   .1059867   .1082678   -.0022812   .0273886
-----+-----
```

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

```
Test:  Ho:  difference in coefficients not systematic
```

```
      chi2(3) = (b-B)' [(V_b-V_B)^(-1)] (b-B)
              = 29.73
Prob>chi2 = 0.0000
```

→ The hypothesis of random effects is rejected

Illustration 2 – Question 1

Analisis based on panel data, but it is implemented using the cross-section framework in questions 1-4. Focus on SME's (implies dropping LE)

```
. Describe ...
```

```
-----  
variable name   type      format    label      variable label  
-----  
id              int       %8.0g     Firm id  
YEAR            int       %8.0g     Year  
LEV_ST1         float     %9.0g     Short-term debt / (STD+LTD+Equity)  
LEV_LT1         float     %9.0g     Long-term debt / (STD+LTD+Equity)  
LEV1            float     %9.0g     Total debt / (STD+LTD+Equity)  
COLLAT1         float     %9.0g     Tangible assets / Total assets  
COLLAT2         float     %9.0g     (Tangible assets + Stock) / Total assets  
SIZE1           float     %9.0g     Log(Total assets)  
SIZE2           float     %9.0g     Log(Sales)  
PROF1           float     %9.0g     EBIT / Total assets  
PROF2           float     %9.0g     EBITDA / Total assets  
GROWTH1         float     %9.0g     Sales growth rate  
GROWTH2         float     %9.0g     Total assets growth rate  
AGE             int       %8.0g     YEAR - Foundation year  
LE              byte      %8.0g     =1 se large firm  
MicE            byte      %8.0g     =1 se micro firm  
SE              byte      %8.0g     =1 se small firm  
MedE            byte      %8.0g     =1 se medium firm
```

Illustration 2 – Question 1

```
. drop if LE==1  
(1922 observations deleted)
```

```
. summarize LEV_LT1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
LEV_LT1	30304	.0734047	.1668595	0	.9982489

```
. count if LEV_LT==0  
23027
```

```
. display 23027/30304  
.75986668
```


Illustration 2 – Question 2

```
. glm LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE, family(binomial) link(logit)
vce(cluster id)
(...)
```

```
Log pseudolikelihood = -6448.71996      AIC      = .4259979
                                      BIC      = -304709.7
```

(Std. Err. adjusted for 4995 clusters in id)

LEV_LT1	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
SIZE2	.3160809	.0150444	21.01	0.000	.2865946	.3455673
COLLAT2	.453255	.1394787	3.25	0.001	.1798818	.7266282
PROF1	-3.629699	.2667238	-13.61	0.000	-4.152469	-3.10693
GROWTH2	.0004846	.0002476	1.96	0.050	-6.89e-07	.0009699
AGE	-.0045993	.0017012	-2.70	0.007	-.0079335	-.0012651
_cons	-6.823869	.2264392	-30.14	0.000	-7.267681	-6.380056

```
. estimates store LOGIT
```

Illustration 2 – Question 3.1, 3.2 & 3.3

3.1 Logit model:

$$E(LEV_LT_i|x_i) = \frac{e^{x_i'\beta}}{1 + e^{x_i'\beta}}$$

```
. estimates restore LOGIT  
(results LOGIT are active now)
```

```
. scalar xb = _b[_cons] + _b[SIZE2]*13.54 + _b[COLLAT2]*0.41 + _b[PROF1]*0.07 +  
_b[GROWTH2]*15.03 + _b[AGE]*19
```

```
. display exp(xb)/(1+exp(xb))  
.06341858
```

3.2 ...

3.3 ...

Illustration 2 – Question 4

```
. quietly glm LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE, family(binomial)  
link(logit) vce(cluster id)
```

```
. predict xbL, xb
```

```
. gen xbL2=xbL^2
```

```
. gen xbL3=xbL^3
```

```
. quietly glm LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE xbL2 xbL3 ,  
family(binomial) link(logit) vce(cluster id)
```

```
. test xbL2 xbL3
```

```
( 1)  [LEV_LT1]xbL2 = 0
```

```
( 2)  [LEV_LT1]xbL3 = 0
```

```
             chi2( 2) = 131.99  
Prob > chi2 = 0.0000
```

Illustration 2 – Question 4

```
. quietly glm LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE, family(binomial)  
link(probit) vce(cluster id)
```

```
. predict xbP, xb
```

```
. gen xbP2=xbP^2
```

```
. gen xbP3=xbP^3
```

```
. quietly glm LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE xbP2 xbP3 ,  
family(binomial) link(probit) vce(cluster id)
```

```
. test xbP2 xbP3
```

```
( 1)  [LEV_LT1]xbP2 = 0
```

```
( 2)  [LEV_LT1]xbP3 = 0
```

```
      chi2( 2) = 120.35  
Prob > chi2 = 0.0000
```

Illustration 2 – Question 4

```
. quietly glm LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE, family(binomial)
link(logit) vce(cluster id)
. margins, dydx (SIZE2 COLLAT2 PROF1 GROWTH2 AGE) atmean
Conditional marginal effects          Number of obs      =      30,304
Model VCE      : Robust
Expression    : Predicted mean LEV_LT1, predict()
dy/dx w.r.t.  : SIZE2 COLLAT2 PROF1 GROWTH2 AGE
at            : SIZE2          =      13.53998 (mean)
              : COLLAT2       =       .4105925 (mean)
              : PROF1         =       .0678864 (mean)
              : GROWTH2      =       15.034 (mean)
              : AGE          =       19.38932 (mean)
```

	Delta-method					
	dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]	
SIZE2	.0188751	.0008384	22.51	0.000	.0172319	.0205184
COLLAT2	.0270666	.0080893	3.35	0.001	.011212	.0429213
PROF1	-.2167517	.0160444	-13.51	0.000	-.2481981	-.1853052
GROWTH2	.0000289	.0000148	1.95	0.051	-9.42e-08	.000058
AGE	-.0002747	.0001021	-2.69	0.007	-.0004747	-.0000746

Illustration 2 – Question 4

```
. margins, dydx (SIZE2 COLLAT2 PROF1 GROWTH2 AGE)
```

```
Average marginal effects          Number of obs      =      30,304
```

```
Model VCE      : Robust
```

```
Expression     : Predicted mean LEV_LT1, predict()
```

```
dy/dx w.r.t.  : SIZE2 COLLAT2 PROF1 GROWTH2 AGE
```

```
-----
```

		Delta-method				
	dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]	
SIZE2	.0210336	.0010403	20.22	0.000	.0189947 .0230724	
COLLAT2	.0301618	.0091286	3.30	0.001	.01227 .0480536	
PROF1	-.2415379	.0185096	-13.05	0.000	-.277816 -.2052598	
GROWTH2	.0000322	.0000165	1.95	0.051	-8.76e-08 .0000646	
AGE	-.0003061	.0001139	-2.69	0.007	-.0005292 -.0000829	

```
-----
```

Illustration 2 – Question 4

```
. quietly glm LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE, family(binomial)
link(probit) vce(cluster id)
. margins, dydx (SIZE2 COLLAT2 PROF1 GROWTH2 AGE) atmean
Conditional marginal effects          Number of obs      =      30,304
Model VCE      : Robust
Expression     : Predicted mean LEV_LT1, predict()
dy/dx w.r.t.  : SIZE2 COLLAT2 PROF1 GROWTH2 AGE
at             : SIZE2          =      13.53998 (mean)
                COLLAT2       =       .4105925 (mean)
                PROF1         =       .0678864 (mean)
                GROWTH2      =       15.034 (mean)
                AGE           =       19.38932 (mean)
```

	Delta-method					
	dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]	
SIZE2	.0201057	.0009364	21.47	0.000	.0182704	.0219409
COLLAT2	.0245778	.008263	2.97	0.003	.0083827	.040773
PROF1	-.2265281	.017486	-12.95	0.000	-.2608001	-.1922561
GROWTH2	.0000398	.0000184	2.17	0.030	3.79e-06	.0000758
AGE	-.0002937	.0001078	-2.72	0.006	-.0005049	-.0000824

Illustration 2 – Question 4

```
. margins, dydx (SIZE2 COLLAT2 PROF1 GROWTH2 AGE)
```

```
Average marginal effects          Number of obs      =      30,304
```

```
Model VCE      : Robust
```

```
Expression     : Predicted mean LEV_LT1, predict()
```

```
dy/dx w.r.t.  : SIZE2 COLLAT2 PROF1 GROWTH2 AGE
```

```
-----
```

		Delta-method			[95% Conf. Interval]	
	dy/dx	Std. Err.	z	P> z		
SIZE2	.0213241	.0010521	20.27	0.000	.019262	.0233862
COLLAT2	.0260673	.0088125	2.96	0.003	.008795	.0433395
PROF1	-.2402559	.0189853	-12.65	0.000	-.2774664	-.2030453
GROWTH2	.0000422	.0000195	2.17	0.030	4.04e-06	.0000804
AGE	-.0003115	.0001144	-2.72	0.006	-.0005357	-.0000873

```
-----
```

```
.
```


Illustration 2 – Question 5

```
. xtset id YEAR
      panel variable:  id (unbalanced)
      time variable:  YEAR, 1995 to 2001, but with gaps
                  delta:  1 unit
```

```
. . xtdescribe
      id:  1, 2, ..., 5514                n =      4995
      YEAR: 1995, 1996, ..., 2001        T =         7
      Delta(YEAR) = 1 unit
      Span(YEAR) = 7 periods
      (id*YEAR uniquely identifies each observation)
```

```
Distribution of T_i:  min      5%      25%      50%      75%      95%      max
                    1         2         6         7         7         7         7

      Freq.  Percent  Cum.  |  Pattern
-----+-----
      3208    64.22   64.22 | 1111111
      176     3.52   67.75 | 111111.
      167     3.34   71.09 | .111111
      80      1.60   72.69 | ..11111
      80      1.60   74.29 | 11111..
      76      1.52   75.82 | ...1111
      71      1.42   77.24 | 1.11111
      62      1.24   78.48 | 11.1111
```

Illustration 2 – Question 5

```
. xtset id YEAR
      panel variable:  id (unbalanced)
      time variable:  YEAR, 1995 to 2001, but with gaps
                  delta: 1 unit
. gen LEVt= LEV_LT1/(1- LEV_LT1)

. xtpoisson LEVt SIZE2 COLLAT2 PROF1 GROWTH2 AGE, fe vce(bootstrap)
(...)
Conditional fixed-effects Poisson regression      Number of obs      =      13245
Group variable: id                               Number of groups   =      2048
(...)
                                         (Replications based on 2048 clusters in id)
-----
```

	Observed	Bootstrap			Normal-based	
LEVt	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
SIZE2	-.2563746	.1857317	-1.38	0.167	-.620402	.1076529
COLLAT2	.7169194	.4518515	1.59	0.113	-.1686932	1.602532
PROF1	-1.205703	2.287238	-0.53	0.598	-5.688608	3.277201
GROWTH2	.0024675	.0016065	1.54	0.125	-.0006813	.0056162
AGE	.0402399	.0242907	1.66	0.098	-.007369	.0878487

```
-----
```

Illustration 2 – Question 5

```
. gen LEVtlink=ln(LEV_LT1/(1- LEV_LT1))  
(23027 missing values generated)
```

```
. xtreg LEVtlink SIZE2 COLLAT2 PROF1 GROWTH2 AGE, fe vce(bootstrap)  
(running xtreg on estimation sample)  
(...)
```

```
Fixed-effects (within) regression          Number of obs      =      7277  
Group variable: id                       Number of groups   =      2064  
(...)
```

(Replications based on 2064 clusters in id)

	Observed	Bootstrap			Normal-based	
LEVtlink	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
SIZE2	-.0741736	.0541096	-1.37	0.170	-.1802265	.0318793
COLLAT2	-.1839106	.1936327	-0.95	0.342	-.5634237	.1956025
PROF1	-1.928813	.4180013	-4.61	0.000	-2.74808	-1.109545
GROWTH2	.0017424	.0005736	3.04	0.002	.0006181	.0028667
AGE	-.0132349	.0073683	-1.80	0.072	-.0276765	.0012067
_cons	.4652777	.8023379	0.58	0.562	-1.107276	2.037831

(...)

Illustration 3 – Question 1.1

```
. gen DIVIDA=LEV_LT1>0
```

```
. probit DIVIDA SIZE2 COLLAT2 PROF1 GROWTH2 AGE, vce(cluster id)
```

```
(...)
```

```
Probit regression                               Number of obs   =       30304
                                                Wald chi2(5)    =       1125.30
                                                Prob > chi2     =         0.0000
Log pseudolikelihood = -14363.976              Pseudo R2      =         0.1401
```

(Std. Err. adjusted for 4995 clusters in id)

```
-----
```

		Robust				[95% Conf. Interval]	
DIVIDA	Coef.	Std. Err.	z	P> z			
SIZE2	.3309445	.011001	30.08	0.000	.309383	.3525061	
COLLAT2	.754961	.07028	10.74	0.000	.6172147	.8927074	
PROF1	-2.114841	.176704	-11.97	0.000	-2.461175	-1.768508	
GROWTH2	.000226	.0001309	1.73	0.084	-.0000307	.0004826	
AGE	.0011396	.0011065	1.03	0.303	-.001029	.0033082	
_cons	-5.510938	.1548431	-35.59	0.000	-5.814425	-5.207451	

```
-----
```

```
. estimates store DP1
```

Illustration 3 – Question 1.1 (cont.)

```
. glm LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE if LEV_LT>0, family(binomial)
link(logit) vce(cluster id)
(...)
```

```
Log pseudolikelihood = -3187.347033
```

AIC	=	.8776548
BIC	=	-63243.49

(Std. Err. adjusted for 2064 clusters in id)

LEV_LT1	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
SIZE2	-.185588	.0140524	-13.21	0.000	-.2131301	-.1580459
COLLAT2	-.7261906	.0998168	-7.28	0.000	-.9218279	-.5305533
PROF1	-1.523694	.2894609	-5.26	0.000	-2.091027	-.9563615
GROWTH2	.0030029	.0004916	6.11	0.000	.0020395	.0039663
AGE	-.0071666	.0013872	-5.17	0.000	-.0098856	-.0044477
_cons	2.389845	.2057352	11.62	0.000	1.986611	2.793078

```
. estimates store DP2
```

Illustration 3 – Question 1.2

```
. tobit LEV_LT1 SIZE2 COLLAT2 PROF1 GROWTH2 AGE, ll(0) vce(cluster id)
(...)
```

(Std. Err. adjusted for 4995 clusters in id)

	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
LEV_LT1						
SIZE2	.1228862	.0040375	30.44	0.000	.1149725	.1308
COLLAT2	.2341081	.0318929	7.34	0.000	.1715967	.2966195
PROF1	-.9789402	.0758608	-12.90	0.000	-1.127631	-.8302498
GROWTH2	.0001581	.0000812	1.95	0.052	-1.13e-06	.0003173
AGE	-.0003276	.0004359	-0.75	0.452	-.001182	.0005267
_cons	-2.064098	.0565249	-36.52	0.000	-2.174889	-1.953306
/sigma	.4669837	.0080894			.4511282	.4828392

```
Obs. summary:      23027 left-censored observations at LEV_LT1<=0
                  7277  uncensored observations
                  0 right-censored observations
```

```
. estimates store TOBIT
```

Illustration 3 – Question 2.1

Two-part model:

$$E(LEV_LT_i|x_i) = \Phi(x_i'\theta) \frac{e^{x_i'\beta}}{1 + e^{x_i'\beta}}$$

```
. estimates restore DP1  
(results DP1 are active now)
```

```
. scalar xb1 = _b[_cons] + _b[SIZE2]*13.54 + _b[COLLAT2]*0.41 + _b[PROF1]*0.07  
+ _b[GROWTH2]*15.03 + _b[AGE]*19
```

```
. estimates restore DP2  
(results DP2 are active now)
```

```
. scalar xb2 = _b[_cons] + _b[SIZE2]*13.54 + _b[COLLAT2]*0.41 + _b[PROF1]*0.07  
+ _b[GROWTH2]*15.03 + _b[AGE]*19
```

```
. display normal(xb1)*exp(xb2)/(1+exp(xb2))  
.06985232
```

Illustration 3 – Question 2.1 (cont.)

Tobit model:

$$E(LEV_LT_i|x_i) = \Phi\left(\frac{x_i'\beta}{\sigma}\right)x_i'\beta + \sigma\phi\left(\frac{x_i'\beta}{\sigma}\right)$$

```
. estimates restore TOBIT
(results TOBIT are active now)

. scalar xbt = _b[_cons] + _b[SIZE2]*13.54 + _b[COLLAT2]*0.41 + _b[PROF1]*0.07
+ _b[GROWTH2]*15.03 + _b[AGE]*19

. display normal(xbt/_b[/sigma])*xbt+_b[/sigma]*normalden(xbt/_b[/sigma])
.05549759
```


Illustration 3 – Question 2.2

Two-part model:

$$Pr(LEV_LT_i > 0|x_i) = \Phi(x_i'\theta)$$

```
. display normal(xb1)  
.19950111
```

Tobit model:

$$Pr(LEV_LT_i > 0|x_i) = \Phi\left(\frac{x_i'\beta}{\sigma}\right)$$

```
. estimates restore TOBIT  
(results TOBIT are active now)  
  
. display normal(xbt/_b[/sigma])  
.20998609
```

Illustration 3 – Question 2.3

Two-part model:

$$E(LEV_LT_i | x_i, Y_i > 0) = \frac{e^{x_i' \beta}}{1 + e^{x_i' \beta}}$$

```
. display exp(xb2) / (1+exp(xb2))  
.35013498
```

Tobit model:

$$E(LEV_LT_i | x_i, Y_i > 0) = x_i' \beta + \sigma \lambda \left(\frac{x_i' \beta}{\sigma} \right)$$

```
. estimates restore TOBIT  
(results TOBIT are active now)
```

```
. display xbt+_b[/sigma]*normalden(xbt/_b[/sigma])/normal(xbt/_b[/sigma])  
.26429175
```

Illustration 4

Description of share2:

```
. summarize share2
  Variable |      Obs      Mean   Std. Dev.   Min       Max
-----+-----
  share2  |     2724   .0122429   .024919     0   .1927609

. count if share2==0
1688

. display 1688/2724
.61967695
```

Illustration 4 – Question 1

```
. tobit share2 age nadults nkids nkids2 lnx agelnx nadlnx, ll(0)
```

```
Tobit regression                               Number of obs   =       2724
                                                LR chi2(7)      =       170.18
                                                Prob > chi2     =       0.0000
Log likelihood = 758.70053                    Pseudo R2      =      -0.1263
```

share2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	-.1258528	.0241782	-5.21	0.000	-.1732624	-.0784432
nadults	.01537	.0380475	0.40	0.686	-.0592349	.089975
nkids	.0042697	.0013247	3.22	0.001	.0016723	.0068671
nkids2	-.0099719	.0054713	-1.82	0.068	-.0207002	.0007565
lnx	-.0444314	.0068893	-6.45	0.000	-.0579402	-.0309225
agelnx	.0088221	.0017832	4.95	0.000	.0053256	.0123187
nadlnx	-.0006007	.0027501	-0.22	0.827	-.0059933	.0047918
_cons	.5899797	.0934268	6.31	0.000	.4067849	.7731745
/sigma	.0479951	.0011832			.0456751	.0503151

```
Obs. summary:      1688 left-censored observations at share2<=0
                   1036 uncensored observations
                   0 right-censored observations
```

Illustration 4 – Question 2

```
. probit d2 age nadults nkids nkids2 lnx agelnx nadlnx
```

```
(...)
```

```
Probit regression
```

```
Number of obs = 2724
```

```
LR chi2(7) = 101.68
```

```
Prob > chi2 = 0.0000
```

```
Log likelihood = -1758.5011
```

```
Pseudo R2 = 0.0281
```

d2	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	-2.479476	.5573451	-4.45	0.000	-3.571853	-1.3871
nadults	.6532982	.8664414	0.75	0.451	-1.044896	2.351492
nkids	.0878242	.0305731	2.87	0.004	.027902	.1477463
nkids2	-.2184835	.1224591	-1.78	0.074	-.4584988	.0215319
lnx	-.6162177	.1615998	-3.81	0.000	-.9329476	-.2994879
agelnx	.1732215	.0410448	4.22	0.000	.0927751	.2536678
nadlnx	-.0365886	.0625613	-0.58	0.559	-.1592064	.0860293
_cons	8.077725	2.197003	3.68	0.000	3.771679	12.38377

Illustration 4 – Question 2

```
. reg share2 age nadults nkids nkids2 lnx agelnx nadlnx if share2>0
```

Source	SS	df	MS			
Model	.15794176	7	.022563109	Number of obs = 1036		
Residual	.867670386	1028	.000844037	F(7, 1028) = 26.73		
				Prob > F = 0.0000		
				R-squared = 0.1540		
				Adj R-squared = 0.1482		
				Root MSE = .02905		
share2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	-.0314662	.0205632	-1.53	0.126	-.0718168	.0088844
nadults	-.0130266	.0324149	-0.40	0.688	-.0766334	.0505803
nkids	.0012847	.0010541	1.22	0.223	-.0007837	.0033531
nkids2	-.0034369	.004556	-0.75	0.451	-.0123771	.0055033
lnx	-.0335767	.0054672	-6.14	0.000	-.0443049	-.0228484
agelnx	.0022097	.001516	1.46	0.145	-.000765	.0051844
nadlnx	.0011125	.002345	0.47	0.635	-.003489	.0057141
_cons	.4896596	.0740595	6.61	0.000	.3443345	.6349847

Illustration 4 – Question 3

```
. probit d2 age nadults nkids nkids2 lnx agelnx nadlnx bluecol whitecol  
(...)
```

```
Probit regression                               Number of obs   =       2724  
                                                LR chi2(9)      =       108.91  
                                                Prob > chi2     =        0.0000  
Log likelihood = -1754.886                    Pseudo R2      =        0.0301
```

d2	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	-2.482997	.5596009	-4.44	0.000	-3.579794	-1.386199
nadults	.4852072	.8717378	0.56	0.578	-1.223367	2.193782
nkids	.0812827	.0308297	2.64	0.008	.0208575	.1417078
nkids2	-.2116622	.1230496	-1.72	0.085	-.452835	.0295106
lnx	-.6320797	.1631954	-3.87	0.000	-.9519368	-.3122225
agelnx	.1747353	.0413054	4.23	0.000	.0937781	.2556924
nadlnx	-.0253409	.0629228	-0.40	0.687	-.1486674	.0979856
bluecol	.2064198	.0834317	2.47	0.013	.0428967	.3699428
whitecol	.0215348	.069428	0.31	0.756	-.1145416	.1576111
_cons	8.244464	2.211077	3.73	0.000	3.910833	12.57809

Illustration 4 – Question 3

```

. predict xb, xb
. gen Mills=normalden(xb)/normal(xb)
. reg share2 age nadults nkids nkids2 lnx agelnx nadlnx Mills if share2>0

```

Source	SS	df	MS	Number of obs = 1036		
Model	.158145344	8	.019768168	F(8, 1027)	=	23.40
Residual	.867466802	1027	.000844661	Prob > F	=	0.0000
-----+-----				R-squared	=	0.1542
-----+-----				Adj R-squared	=	0.1476
Total	1.02561215	1035	.00099093	Root MSE	=	.02906

share2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	-.0172991	.0354385	-0.49	0.626	-.0868392	.052241
nadults	-.0174379	.0336488	-0.52	0.604	-.0834661	.0485903
nkids	.0007643	.0014951	0.51	0.609	-.0021695	.0036982
nkids2	-.0020755	.005335	-0.39	0.697	-.0125443	.0083932
lnx	-.0301094	.0089325	-3.37	0.001	-.0476375	-.0125814
agelnx	.0012243	.0025157	0.49	0.627	-.0037122	.0061608
nadlnx	.001365	.0024016	0.57	0.570	-.0033476	.0060775
Mills	-.0090179	.0183686	-0.49	0.624	-.0450622	.0270263
_cons	.4515814	.1072597	4.21	0.000	.2411081	.6620547

Illustration 5 – Question 1

```
. gen y=log( hrrate)>log(3.6)
. gen socd1234=socd1+socd2+socd3+socd4
. gen empmont= empmon/10
. logit y empmont part socd1234 hqd6 married size25 lppltw
(...)
```

```
Logistic regression                                Number of obs   =          4095
                                                    LR chi2(7)      =          421.39
                                                    Prob > chi2     =           0.0000
Log likelihood = -1223.9981                        Pseudo R2       =           0.1469
```

	y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
	+						
	empmont	.0266874	.0079159	3.37	0.001	.0111725	.0422024
	part	-1.178631	.1154354	-10.21	0.000	-1.404881	-.9523821
	socd1234	1.142527	.1865971	6.12	0.000	.7768034	1.508251
	hqd6	-.6375414	.1134435	-5.62	0.000	-.8598866	-.4151962
	married	.3557298	.1093082	3.25	0.001	.1414897	.5699699
	size25	1.023301	.1089946	9.39	0.000	.8096759	1.236927
	lppltw	-1.127114	.4766336	-2.36	0.018	-2.061298	-.1929288
	_cons	1.847127	.1300033	14.21	0.000	1.592325	2.101929

Illustration 5 – Question 1

```
. predict xbLogit,xb
. gen xbLogit2= xbLogit^2
. gen xbLogit3= xbLogit^3

. quietly logit y empmont part socd1234 hqd6 married size25 lppltw xbLogit2
xbLogit3

. test xbLogit2 xbLogit3

( 1) [y]xbLogit2 = 0
( 2) [y]xbLogit3 = 0

           chi2( 2) =      5.72
Prob > chi2 =      0.0574
```

Illustration 5 – Question 1

```
. probit y empmont part socd1234 hqd6 married size25 lppltw  
(...)
```

```
Probit regression                               Number of obs   =           4095  
                                                LR chi2(7)      =           413.14  
                                                Prob > chi2     =           0.0000  
Log likelihood = -1228.1222                    Pseudo R2      =           0.1440
```

```
-----+-----  
          y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]  
-----+-----  
    empmont |   .0130406   .0038505     3.39   0.001   .0054937   .0205876  
      part  |  -.5965585   .0590365    -10.10  0.000  -.7122678  -.4808491  
socd1234   |   .5665749   .088828     6.38   0.000   .3924751   .7406747  
      hqd6  |  -.3519214   .061924     -5.68   0.000  -.4732903  -.2305525  
    married |   .1857742   .0584009     3.18   0.001   .0713105   .3002379  
    size25  |   .532082    .057272     9.29   0.000   .4198309   .6443332  
    lppltw  |  -.6578885   .2947191     -2.23   0.026  -1.235527  -.0802496  
      _cons |   1.066556   .0682845    15.62   0.000   .932721    1.200391  
-----+-----
```

Illustration 5 – Question 1 & 2.1

```
. predict xbProbit,xb

. gen xbProbit2= xbProbit^2

. gen xbProbit3= xbProbit^3

. quietly probit y empmont part socd1234 hqd6 married size25 lppltw xbProbit2
xbProbit3

. test xbProbit2 xbProbit3

( 1) [y]xbProbit2 = 0
( 2) [y]xbProbit3 = 0

           chi2( 2) =    13.60
Prob > chi2 =    0.0011
```

2.1

- Logit model: robust to CB sampling – slope parameter consistent
- Probit model: CB sampling renders standard ML estimators inconsistent
- It is likely that the results of the RESET test reflect these issues

Illustration 5 – Question 2.2

```
. gen w=.055/.1118*(1-y)+(1-.055)/(1-.1118)*y
```

```
. probit y empmont part socd1234 hqd6 married size25 lppltw [pweight=w]
```

```
Probit regression                               Number of obs       =        4,095
                                                Wald chi2(7)        =        294.40
                                                Prob > chi2         =         0.0000
Log pseudolikelihood = -761.75842              Pseudo R2           =         0.1268
```

		Robust				[95% Conf. Interval]	
	y	Coef.	Std. Err.	z	P> z		
empmont		.011869	.0041034	2.89	0.004	.0038265	.0199115
part		-.5463279	.0528645	-10.33	0.000	-.6499404	-.4427155
socd1234		.5181078	.0827642	6.26	0.000	.3558929	.6803227
hqd6		-.3183259	.0567485	-5.61	0.000	-.4295509	-.2071008
married		.1730612	.0533792	3.24	0.001	.0684399	.2776825
size25		.4859684	.0504289	9.64	0.000	.3871296	.5848072
lppltw		-.6414934	.2531809	-2.53	0.011	-1.137719	-.145268
_cons		1.458513	.0584089	24.97	0.000	1.344034	1.572992

Illustration 5 – Question 2.2

Alternatively: using a just identified GMM estimator (multiplies the weight by the RS moment condition)

```
. gmm ( (.055/.1118*(1-y) + (1-.055) / (1-  
.1118) *y) *normalden({beta1}*empmont+{beta2}*part+{beta3}*socd1234+{beta4}  
*hqd6+{beta5}*married+{beta6}*lppltw+{beta7}*size25+{beta0}) * (y-  
normprob({beta1}*empmont+{beta2}*part+{beta3}*socd1234+{beta4}  
*hqd6+{beta5}*married+{beta6}*lppltw+{beta7}*size25+{beta0})  
) / (normprob({beta1}*empmont+{beta2}*part+{beta3}*socd1234+{beta4}  
*hqd6+{beta5}*married+{beta6}*lppltw+{beta7}*size25+{beta0}) * (1-  
normprob({beta1}*empmont+{beta2}*part+{beta3}*socd1234+{beta4}  
*hqd6+{beta5}*married+{beta6}*lppltw+{beta7}*size25+{beta0})  
)) , instruments(empmont part socd1234 hqd6 married lppltw  

```

Illustration 5 – Question 2.2

GMM estimation

Number of parameters = 8

Number of moments = 8

Initial weight matrix: Unadjusted

Number of obs = 4095

GMM weight matrix: Robust

```
-----
```

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
/beta1	.011869	.0041029	2.89	0.004	.0038274	.0199106
/beta2	-.5463279	.052858	-10.34	0.000	-.6499277	-.4427281
/beta3	.5181078	.0827541	6.26	0.000	.3559127	.6803029
/beta4	-.3183259	.0567416	-5.61	0.000	-.4295373	-.2071144
/beta5	.1730612	.0533727	3.24	0.001	.0684527	.2776697
/beta6	-.6414934	.25315	-2.53	0.011	-1.137658	-.1453286
/beta7	.4859684	.0504227	9.64	0.000	.3871417	.5847952
/beta0	1.458513	.0584018	24.97	0.000	1.344048	1.572978

```
-----
```

Instruments for equation 1: empmont part socd1234 hqd6 married lppltw size25
_cons