Homework 2

João Pedro Bastos

October 27, 2022

Question 1

a) Estimate the linear probability (regression) model explaining $DELINQUENT$ as a function of the remaining variables. Use White robust standard errors. Are the signs of the estimated coefficients reasonable?

```
. reg DELINQUENT LVR REF INSUR RATE AMOUNT CREDIT TERM ARM, vce(robust)

Linear regression                               Number of obs     =      1,000
                                                F(8, 991)         =      43.39
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3363
                                                Root MSE          =     .32673
```

DELINQUENT	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
LVR	.0016239	.0006752	2.40	0.016	.0002988	.0029489
REF	-.0593237	.0240256	-2.47	0.014	-.1064706	-.0121768
INSUR	-.4815849	.0303694	-15.86	0.000	-.5411807	-.4219891
RATE	.0343761	.0098194	3.50	0.000	.0151068	.0536454
AMOUNT	.023768	.0144509	1.64	0.100	-.0045898	.0521259
CREDIT	-.0004419	.0002073	-2.13	0.033	-.0008487	-.0000351
TERM	-.0126195	.003556	-3.55	0.000	-.0195976	-.0056414
ARM	.1283239	.0276932	4.63	0.000	.0739798	.1826681
_cons	.6884913	.2285064	3.01	0.003	.2400792	1.136903

The coefficients seem reasonable: having a loan insurance (less risk), a higher credit score (history of good payments), a longer term loan (lower individual payments, all else equal), all reduce the probability of delinquency. A refinance loan also reduces the probability, presumably because refinancing usually happens with more favorable terms.

On the other hand, having a higher a loan amount either in absolute terms ($AMOUNT$) or relative to the property value (LVR), having an adjustable rate (more uncertainty), and having a higher rate (higher cost) all contribute to a greater delinquency probability.

b) Use probit to estimate the model in (a). Are the signs and significance of the estimated coefficients the same as for the linear probability model?

```
. probit DELINQUENT LVR REF INSUR RATE AMOUNT CREDIT TERM ARM

Iteration 0:   log likelihood =   -499.013
Iteration 1:   log likelihood = -338.38904
Iteration 2:   log likelihood = -332.81547
Iteration 3:   log likelihood = -332.79661
Iteration 4:   log likelihood = -332.79661

Probit regression                                   Number of obs =  1,000
                                                    LR chi2(8)    = 332.43
                                                    Prob > chi2   = 0.0000
Log likelihood = -332.79661                         Pseudo R2     = 0.3331

------------------------------------------------------------------------------
  DELINQUENT | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         LVR |   .0076007   .0045911     1.66   0.098    -.0013977    .0165991
         REF |  -.2884561   .1259446    -2.29   0.022    -.5353029   -.0416092
       INSUR |  -1.772714   .1158088   -15.31   0.000    -1.999695   -1.545733
        RATE |   .1711988   .0438147     3.91   0.000     .0853236    .2570741
      AMOUNT |    .121236   .0615491     1.97   0.049      .000602    .2418701
      CREDIT |  -.0019131   .0010638    -1.80   0.072    -.0039981    .0001718
        TERM |  -.0775769   .0198396    -3.91   0.000    -.1164618    -.038692
         ARM |   .8091109   .2077119     3.90   0.000      .402003    1.216219
       _cons |    .964646   1.088121     0.89   0.375    -1.168033    3.097325
------------------------------------------------------------------------------
```

All the coefficients have the same sign and are significant (with the exception of the constant, which is not significant in the probit model. $AMOUNT$ becomes significant at the 5% level in the probit (10% in OLS) and LVR experiences the opposite: significant at the 10% level in probit, 5% in the OLS. The other coefficients suffer very minor changes in significance.

(c) Compute the predicted value of DELINQUENT for the 500th and 1,000th observations using both the linear probability model and the probit model. Interpret the values.

Table 1: Predicted Results

Obs.	$DELINQ$	Pred. OLS	Pred. Probit
500	0	.1827828	.1404525
1000.	0	.5785297	.6167872

Obs.	$DELINQ$	LVR	REF	$INSUR$	$RATE$	$AMOUNT$	$CREDIT$	$TERM$	ARM
500.	0	70	1	1	10.95	.854	509	30	1
1000.	0	88.2	1	0	7.65	2.91	624	30	1

Given their respective values presented in the above table, the OLS model predicts a probability of delinquency equal to 18.2% for loan 500 and 57.8% for loan 1000. Whereas the probit model predicts a probability of delinquency equal to 14% to loan 500 and 61.6% to loan 1000, given their respective values presented in the above table.

(d) Construct a histogram of $CREDIT$. Using both linear probability and probit models, calculate the probability of delinquency for $CREDIT$ = 500,600, and 700 for a loan of \$250,000 ($AMOUNT$ = 2.5). For the other variables, loan to value ratio (LVR) is 80%, initial interest rate is 8%, indicator variables take the value one, and $TERM$ = 30. Discuss similarities and differences among the predicted probabilities from the two models.

Figure 1: Credit Histogram

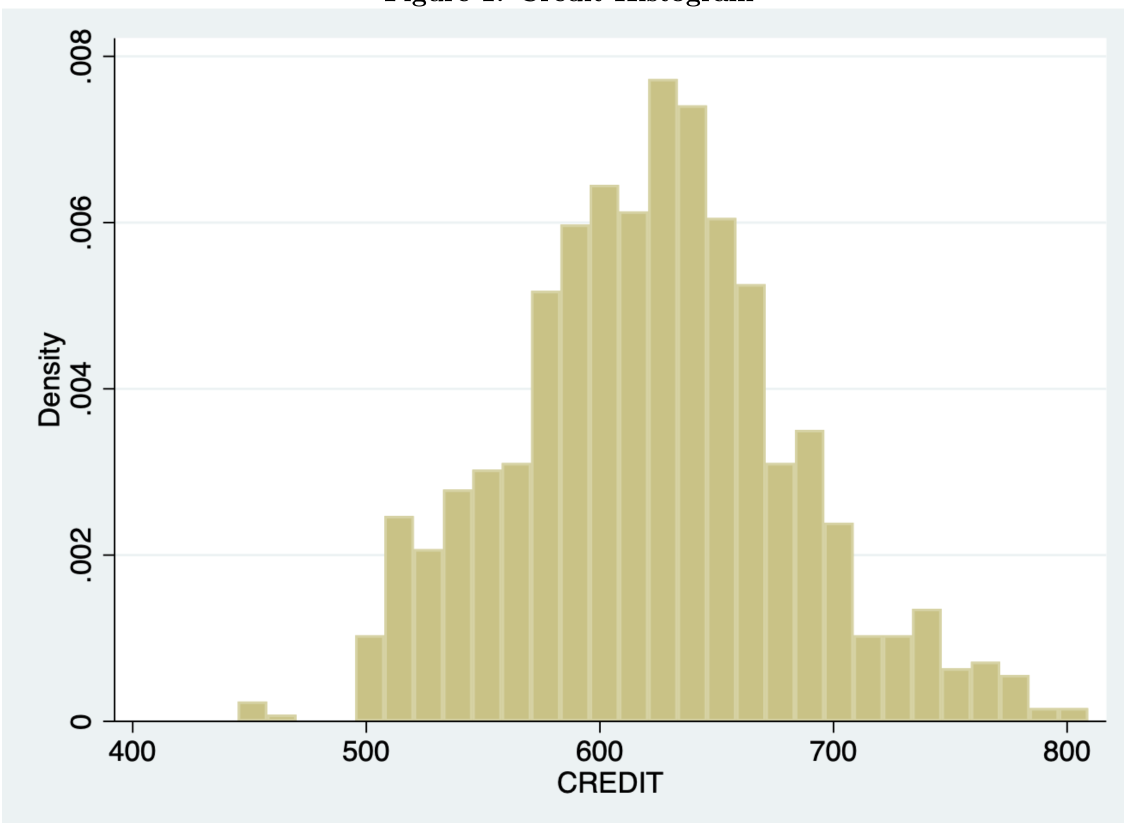

Table 2: Predicted Results (in %) for different levels of $CREDIT$

Model	$CREDIT = 500$	$CREDIT = 600$	$CREDIT = 700$
LPM	27.5 %	23.1%	18.7%
Probit	26.2%	22.1%	18.5%

Given the values of the other variables described above, the probability of delinquency predicted by the linear probability model is equal to 27.5% when credit score is 500, 23.1% when credit score is 600, and 18.7% when credit score is 700. The probability of delinquency predicted by the probit model is equal to 26.2% when credit score is 500, 22.1% when credit score is 600, and 18.5% when credit score is 700.

Both models predict very similar results, with the probability of delinquency diminishing as credit score increases. However, the effect of credit score is linear for linear probability model: each 100 point increase in credit score will reduce the probability of delinquency by 4.4%. For the probit model, this effect in non-linear. An increase in credit score from 500 to 600 will reduce the probability if delinquency by 4.1%, while an increase in credit score from 600 to 700 is associated with a smaller reduction of 3.6%.

(e) **Compute the marginal effect of CREDIT on the probability of delinquency for CREDIT = 500,600, and 700, given that the other explanatory variables take the values in (d). Discuss the interpretation of the marginal effect.**

```
. margins, dydx(CREDIT) at (CREDIT =(500(100)700) AMOUNT=2.5 LVR=80 RATE=8 TERM=30 REF=1 ARM=1)

Average marginal effects                                   Number of obs = 1,000
Model VCE: OIM

Expression: Pr(DELINQUENT), predict()
dy/dx wrt:  CREDIT
1._at: LVR    =  80
       REF    =   1
       RATE   =   8
       AMOUNT = 2.5
       CREDIT = 500
       TERM   =  30
       ARM    =   1
2._at: LVR    =  80
       REF    =   1
       RATE   =   8
       AMOUNT = 2.5
       CREDIT = 600
       TERM   =  30
       ARM    =   1
3._at: LVR    =  80
       REF    =   1
       RATE   =   8
       AMOUNT = 2.5
       CREDIT = 700
       TERM   =  30
       ARM    =   1
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
CREDIT						
_at						
1	-.0004292	.0002588	-1.66	0.097	-.0009365	.0000781
2	-.0003881	.0002132	-1.82	0.069	-.0008061	.0000298
3	-.0003482	.0001694	-2.05	0.040	-.0006803	-.0000161

Given the values of the other variables described in the output, a 1 point increase in credit score will reduce the probability of delinquency by 0.0042% when credit score is 500, 0.0038% when credit score is 600, and 0.0034% when credit score is 700. As discussed in part (d), this shows that the marginal effect is non-linear, as is depends on the initial value of the credit score. In this specific cases, the higher credit scores decrease the probability of delinquency at a decreasing rate.

(f) Construct a histogram of LVR. Using both linear probability and probit models, calculate the probability of delinquency for $LVR = 20$ and $LVR = 80$, with $CREDIT = 600$ and other variables set as they are in (d). Compare and contrast the results.

Figure 2: Loan Amount to Value of Property Ratio (LVR) Histogram

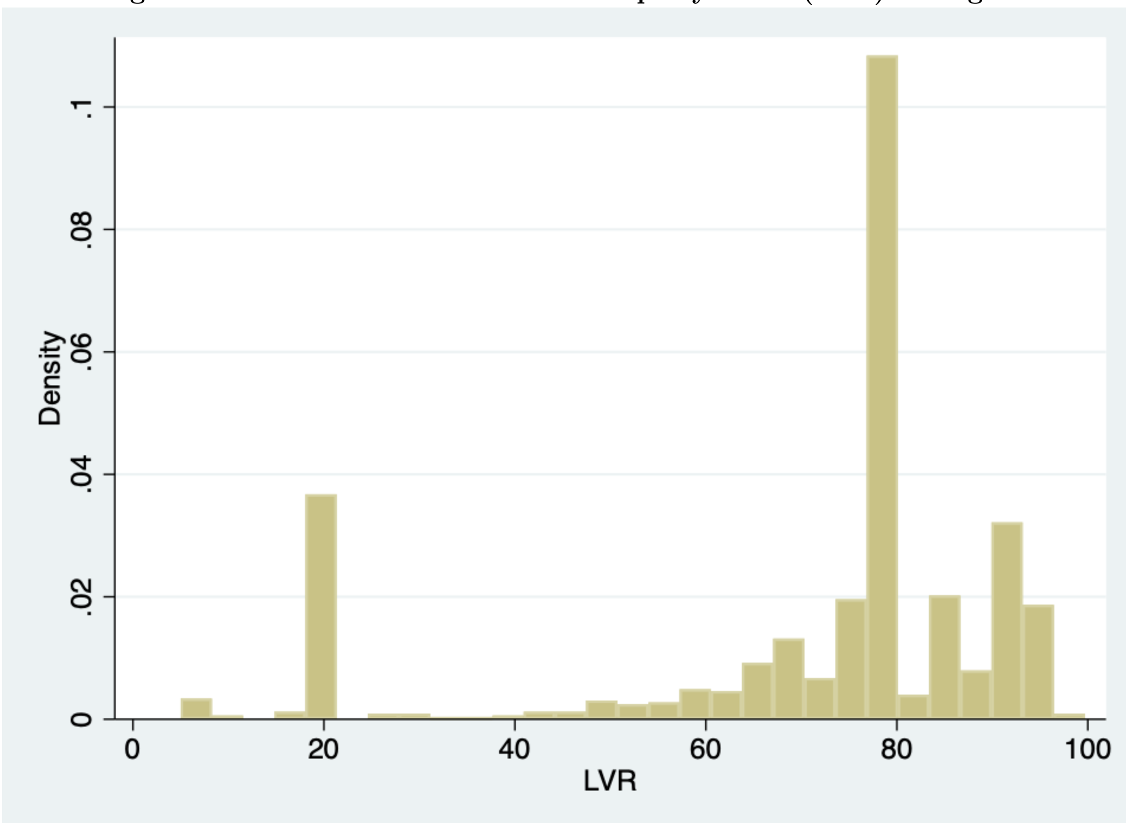

Table 3: Predicted Results (in %) for different levels of LVR

Model	$LVR = 20$	$LVR = 80$
LPM	13.39%	23.13%
Probit	14.05%	22.18%

Again, the results are very similar for both models, where a higher loan to property value ratio increases the probability of delinquency, but there are some differences as well.

The LPM model indicates a slightly lower probability at $LVR = 20$ and a slightly higher at $LVR = 80$. In the LPM model, a 60 percentage point increase in LVR increases the probability of delinquency by 9.74%.

The probit model on the other hand indicates a slightly higher probability at $LVR = 20$ and slightly lower probability at $LVR = 80$. That implies that the marginal effect within that range (from 20 to 80) is relatively smaller, increasing the probability of delinquency by 8.13%.

(g) Compare the percentage of correct predictions from the linear probability model and the probit model, using a predicted probability of 0.5 as the threshold.

We consider the predictions as correct if the predicted probability is $< .5$ and we observe no delinquency, or if the predicted probability is $> .5$ and there was a delinquency. The table below reports the results.

As it can be seen, both models have very similar ratios of success in predicting the correct outcome. The Linear Probability Model predicted the correct outcome 858 times and the Probit model 855 times, so only a difference of 3 correct predictions.

Table 4: Percentage of Correct and Incorrect Predictions

Model	Incorrect	Correct
LPM	14.2%	85.8%
Probit	14.5%	85.5%

(h) As a loan officer, you wish to provide loans to customers who repay on schedule and are not delinquent. Suppose you have available to you the first 500 observations in the data on which to base your loan decision on the second 500 applications (501–1,000). Is using the probit model, with a threshold of 0.5 for the predicted probability the best decision rule for deciding on loan applications? If not, what is a better rule.

Using the first 500 observations for our sample, we estimate the two models again, and the results are presented below.

```
. reg DELINQUENT LVR REF INSUR RATE AMOUNT CREDIT TERM ARM if id<501, vce(robust)

Linear regression                               Number of obs     =        500
                                                F(8, 491)         =      32.39
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3785
                                                Root MSE          =     .35068

------------------------------------------------------------------------------
             |               Robust
  DELINQUENT | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         LVR |   .0025827    .001191     2.17   0.031     .0002425    .0049228
         REF |  -.0783842   .0379617    -2.06   0.039    -.1529717   -.0037967
       INSUR |  -.4880244   .0387692   -12.59   0.000    -.5641985   -.4118504
        RATE |    .057459   .0153479     3.74   0.000     .0273034    .0876146
      AMOUNT |   .0232445   .0193415     1.20   0.230    -.0147579    .0612468
      CREDIT |   -.000138   .0003353    -0.41   0.681    -.0007968    .0005209
        TERM |  -.0180095   .0064574    -2.79   0.005    -.0306971    -.005322
         ARM |   .1829719   .0385865     4.74   0.000     .1071569    .2587869
       _cons |   .4065923   .3863153     1.05   0.293    -.3524427    1.165627
------------------------------------------------------------------------------
```

```
. probit DELINQUENT LVR REF INSUR RATE AMOUNT CREDIT TERM ARM

Iteration 0:   log likelihood =   -499.013
Iteration 1:   log likelihood = -338.38904
Iteration 2:   log likelihood = -332.81547
Iteration 3:   log likelihood = -332.79661
Iteration 4:   log likelihood = -332.79661

Probit regression                                       Number of obs =  1,000
                                                        LR chi2(8)    = 332.43
                                                        Prob > chi2   = 0.0000
Log likelihood = -332.79661                             Pseudo R2     = 0.3331

------------------------------------------------------------------------------
  DELINQUENT | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         LVR |   .0076007   .0045911     1.66   0.098    -.0013977    .0165991
         REF |  -.2884561   .1259446    -2.29   0.022    -.5353029   -.0416092
       INSUR |  -1.772714   .1158088   -15.31   0.000    -1.999695   -1.545733
        RATE |   .1711988   .0438147     3.91   0.000     .0853236    .2570741
      AMOUNT |    .121236   .0615491     1.97   0.049      .000602    .2418701
      CREDIT |  -.0019131   .0010638    -1.80   0.072    -.0039981    .0001718
        TERM |  -.0775769   .0198396    -3.91   0.000    -.1164618    -.038692
         ARM |   .8091109   .2077119     3.90   0.000      .402003    1.216219
       _cons |    .964646   1.088121     0.89   0.375    -1.168033    3.097325
------------------------------------------------------------------------------
```

As in (g), we calculate the number of correct prediction if the predicted probability is $< .5$ and we observe no delinquency, or if the predicted probability is $> .5$ and there was a delinquency. The table below reports the results. The difference is that we are doing an out-of-sample prediction, using the last 500 observations that were not used to estimate the model.

The table below shows that the percentage of successful predictions diminishes substantially, which was expected, given that we are not only reducing the size of our sample for estimating the model but also requiring it to predict out of sample.

Again, the LPM model does slightly better at predicting successfully, with 3 more correct predictions than the the Probit model. If our only task is to predict delinquency or not using the .5% threshold, the LPM model would be a very slightly better rule.

Table 5: Percentage of Correct and Incorrect Out-of-Sample Predictions

Model	Incorrect	Correct
LPM	42.4%	57.6%
Probit	42.7%	57.3%

Question 2

(a) Estimate a multinomial logit model explaining *PSECHOICE*. **Use the group who did not attend college as the base group. Use as explanatory variables** *GRADES*, *FAMINC*, *FEMALE*, **and** *BLACK*. **Are the estimated coefficients statistically significant?**

```
. mlogit PSECHOICE GRADES FAMINC FEMALE BLACK, baseoutcome(1)

Iteration 0:   log likelihood = -1018.6575
Iteration 1:   log likelihood = -858.70649
Iteration 2:   log likelihood = -846.95257
Iteration 3:   log likelihood = -846.75617
Iteration 4:   log likelihood = -846.75601
Iteration 5:   log likelihood = -846.75601

Multinomial logistic regression                         Number of obs =  1,000
                                                        LR chi2(8)    = 343.80
                                                        Prob > chi2   = 0.0000
Log likelihood = -846.75601                             Pseudo R2     = 0.1688
```

PSECHOICE	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
1	(base outcome)					
2						
GRADES	-.3089701	.0552152	-5.60	0.000	-.4171899	-.2007503
FAMINC	.0118943	.003928	3.03	0.002	.0041956	.0195931
FEMALE	.1169483	.1949887	0.60	0.549	-.2652224	.4991191
BLACK	.5679813	.4295461	1.32	0.186	-.2739136	1.409876
_cons	1.937035	.491135	3.94	0.000	.9744285	2.899642
3						
GRADES	-.7272638	.0566698	-12.83	0.000	-.8383346	-.6161929
FAMINC	.0204678	.0038319	5.34	0.000	.0129574	.0279781
FEMALE	-.1337162	.1932327	-0.69	0.489	-.5124453	.2450129
BLACK	1.607127	.4079379	3.94	0.000	.807583	2.40667
_cons	4.962637	.4744651	10.46	0.000	4.032702	5.892572

Most coefficients are significant with the exception of the female dummy when explaining 4-year college outcome. For the 2-year college outcome, the female and black dummies are not significant.

(b) Compute the estimated probability that a white male student with median values of $GRADES$ and $FAMINC$ will attend a four-year college.

```
. margins, predict(outcome(3)) at((medians) GRADES FAMINC FEMALE=0 BLACK=0)

Adjusted predictions                                    Number of obs = 1,000
Model VCE: OIM

Expression: Pr(PSECHOICE==3), predict(outcome(3))
At: GRADES = 6.64 (median)
    FAMINC = 42.5 (median)
    FEMALE =    0
    BLACK  =    0
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf. interval]	
_cons	.5239466	.027215	19.25	0.000	.4706062	.577287

(c) Compute the probability ratio that a white male student with median values of $GRADES$ and $FAMINC$ will attend a four-year college rather than not attend any college.

We first need to calculate the probability of no attending any college:

```
. margins, predict(outcome(1)) at((medians) GRADES FAMINC FEMALE=0 BLACK=0)

Adjusted predictions                                    Number of obs = 1,000
Model VCE: OIM

Expression: Pr(PSECHOICE==1), predict(outcome(1))
At: GRADES = 6.64 (median)
    FAMINC = 42.5 (median)
    FEMALE =    0
    BLACK  =    0
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf. interval]	
_cons	.1920783	.0213003	9.02	0.000	.1503304	.2338263

The probability ratio is this given by:

$$\frac{Pr(y = 4YCollege)}{Pr(y = NoCollege)} = \frac{0.5239466}{0.1920783} = 4.2896 \tag{1}$$

(d) Compute the change in probability of attending a four-year college for a white male student with median $FAMIN$C whose $GRADES$ change from 6.64 (the median value) to 4.905 (top 25th percentile).

```
. margins, at((medians) FAMINC GRADES=6.53039 FEMALE=0 BLACK=0)

Adjusted predictions                                    Number of obs = 1,000
Model VCE: OIM

1._predict: Pr(PSECHOICE==1), predict(pr outcome(1))
2._predict: Pr(PSECHOICE==2), predict(pr outcome(2))
3._predict: Pr(PSECHOICE==3), predict(pr outcome(3))

At: GRADES = 6.53039
    FAMINC =    42.5 (median)
    FEMALE =       0
    BLACK  =       0

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    _predict |
          1  |   .1823658   .0208929     8.73   0.000     .1414164    .2233152
          2  |    .278903   .0233568    11.94   0.000     .2331245    .3246816
          3  |   .5387311   .0272494    19.77   0.000     .4853233     .592139
------------------------------------------------------------------------------

. margins, at((medians) FAMINC GRADES=4.905 FEMALE=0 BLACK=0)

Adjusted predictions                                    Number of obs = 1,000
Model VCE: OIM

1._predict: Pr(PSECHOICE==1), predict(pr outcome(1))
2._predict: Pr(PSECHOICE==2), predict(pr outcome(2))
3._predict: Pr(PSECHOICE==3), predict(pr outcome(3))

At: GRADES = 4.905
    FAMINC =  42.5 (median)
    FEMALE =     0
    BLACK  =     0

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    _predict |
          1  |   .0759824   .0139192     5.46   0.000     .0487013    .1032635
          2  |   .1920101   .0227243     8.45   0.000     .1474714    .2365489
          3  |   .7320075   .0266109    27.51   0.000     .6798511    .7841639
------------------------------------------------------------------------------
```

The difference in probability is given by 0.7320075 – 0.5387311 = 0.1932764

(e) From the full data set create a subsample, omitting the group who attended a two-year college. Estimate a logit model explaining student's choice between attending a four-year college and not attending college, using the same explanatory variables in (a). Compute the probability ratio that a white male student with median values of GRADES and FAMINC will attend a four-year college rather than not attend any college. Compare the result to that in (c).

```
. mlogit PSECHOICE GRADES FAMINC FEMALE BLACK if PSECHOICE!=2, baseoutcome(1)

Iteration 0:   log likelihood = -455.22643
Iteration 1:   log likelihood = -324.38318
Iteration 2:   log likelihood = -309.98175
Iteration 3:   log likelihood = -309.55773
Iteration 4:   log likelihood =  -309.5561
Iteration 5:   log likelihood =  -309.5561

Multinomial logistic regression                         Number of obs =    749
                                                        LR chi2(4)    = 291.34
                                                        Prob > chi2   = 0.0000
Log likelihood = -309.5561                              Pseudo R2     = 0.3200
```

PSECHOICE	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
1	(base outcome)					
3						
GRADES	-.7272205	.0616125	-11.80	0.000	-.8479788	-.6064622
FAMINC	.0182128	.0038987	4.67	0.000	.0105716	.0258541
FEMALE	-.1313463	.2036464	-0.64	0.519	-.5304858	.2677932
BLACK	1.37962	.422851	3.26	0.001	.5508473	2.208393
_cons	5.069643	.504986	10.04	0.000	4.079889	6.059397

```
. margins, predict(outcome(3)) at((medians) GRADES FAMINC FEMALE=0 BLACK=0)

Adjusted predictions                                     Number of obs = 749
Model VCE: OIM

Expression: Pr(PSECHOICE==3), predict(outcome(3))
At: GRADES = 6.42 (median)
    FAMINC = 42.5 (median)
    FEMALE =    0
    BLACK  =    0
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
_cons	.7640356	.0278927	27.39	0.000	.7093669	.8187043

The probability ratio of attending a 4-year college vs. not attending college at all is given by:

$$\frac{Pr(y = 4YCollege)}{Pr(y = NoCollege)} = \frac{Pr(y = 4YCollege)}{1 - Pr(y = 4YCollege)} = \frac{0.7640356}{1 - 0.7640356} = 3.237927 \tag{2}$$

Question 3

(a) In addition to $PRICE$, the data file contains variables indicating whether the product was "featured" at the time ($FEATURE$) or whether there was a store display ($DISPLAY$). Estimate a conditional logit model explaining choice of soda using $PRICE$, $DISPLAY$, and $FEATURE$ as explanatory variables. Discuss the signs of the estimated coefficients and their significance.

```
. asclogit CHOICE PRICE FEATURE DISPLAY, case(ID) alternatives(ALT) noconstant
note: variable PRICE has 315 cases that are not alternative-specific; there is no within-case variability.
note: variable FEATURE has 449 cases that are not alternative-specific; there is no within-case variability.
note: variable DISPLAY has 873 cases that are not alternative-specific; there is no within-case variability.

Iteration 0:   log likelihood = -1842.1342
Iteration 1:   log likelihood = -1822.2841
Iteration 2:   log likelihood = -1822.2267
Iteration 3:   log likelihood = -1822.2267

Alternative-specific conditional logit         Number of obs      =      5,466
Case ID variable: ID                           Number of cases    =       1822

Alternatives variable: ALT                     Alts per case: min =          3
                                                              avg =        3.0
                                                              max =          3

                                                  Wald chi2(3)    =     308.35
Log likelihood = -1822.2267                       Prob > chi2     =     0.0000
```

CHOICE	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
ALT						
PRICE	-1.744454	.1799323	-9.70	0.000	-2.097115	-1.391793
FEATURE	-.0106038	.0799373	-0.13	0.894	-.1672781	.1460705
DISPLAY	.4624477	.0930481	4.97	0.000	.2800768	.6448185

The coefficient associated with $PRICE$ is negative and significant, meaning that a higher price diminishes the probability of being chosen. The coefficient on $DISPLAY$ is positive and significant, presumably because it draws greater consumer attention to a specific brand.

Finally, the $FEATURE$ variable has a negative coefficient. The fact we observe a smaller probability of being chosen for a featured good is counter-intuitve. However, this effect is not significant, so we cannot reject the null hypothesis that being featured does not affect the at all the probability of being chosen.

(b) Compute the probability ratio of choosing *Coke* relative to *Pepsi* and *7Up* if the price of each is \$1.25 and no display or feature is present.

```
. estat mfx, at (PRICE=1.25 DISPLAY=0 FEATURE=0)
```

Pr(choice = Pepsi|1 selected) = .33333333

variable	dp/dx	Std. err.	z	P>|z|	[95% C.I.	]	X
PRICE							
Pepsi	-.387656	.039985	-9.70	0.000	-.466026	-.309287	1.25
7Up	.193828	.019992	9.70	0.000	.154644	.233013	1.25
Coke	.193828	.019992	9.70	0.000	.154644	.233013	1.25
FEATURE							
Pepsi	-.002356	.017764	-0.13	0.894	-.037173	.03246	0
7Up	.001178	.008882	0.13	0.894	-.01623	.018586	0
Coke	.001178	.008882	0.13	0.894	-.01623	.018586	0
DISPLAY							
Pepsi	.102766	.020677	4.97	0.000	.062239	.143293	0
7Up	-.051383	.010339	-4.97	0.000	-.071647	-.03112	0
Coke	-.051383	.010339	-4.97	0.000	-.071647	-.03112	0

Pr(choice = 7Up|1 selected) = .33333333

variable	dp/dx	Std. err.	z	P>|z|	[95% C.I.	]	X
PRICE							
Pepsi	.193828	.019992	9.70	0.000	.154644	.233013	1.25
7Up	-.387656	.039985	-9.70	0.000	-.466026	-.309287	1.25
Coke	.193828	.019992	9.70	0.000	.154644	.233013	1.25
FEATURE							
Pepsi	.001178	.008882	0.13	0.894	-.01623	.018586	0
7Up	-.002356	.017764	-0.13	0.894	-.037173	.03246	0
Coke	.001178	.008882	0.13	0.894	-.01623	.018586	0
DISPLAY							
Pepsi	-.051383	.010339	-4.97	0.000	-.071647	-.03112	0
7Up	.102766	.020677	4.97	0.000	.062239	.143293	0
Coke	-.051383	.010339	-4.97	0.000	-.071647	-.03112	0

Pr(choice = Coke|1 selected) = .33333333

variable	dp/dx	Std. err.	z	P>|z|	[95% C.I.	]	X
PRICE							
Pepsi	.193828	.019992	9.70	0.000	.154644	.233013	1.25
7Up	.193828	.019992	9.70	0.000	.154644	.233013	1.25
Coke	-.387656	.039985	-9.70	0.000	-.466026	-.309287	1.25
FEATURE							
Pepsi	.001178	.008882	0.13	0.894	-.01623	.018586	0
7Up	.001178	.008882	0.13	0.894	-.01623	.018586	0
Coke	-.002356	.017764	-0.13	0.894	-.037173	.03246	0
DISPLAY							
Pepsi	-.051383	.010339	-4.97	0.000	-.071647	-.03112	0
7Up	-.051383	.010339	-4.97	0.000	-.071647	-.03112	0
Coke	.102766	.020677	4.97	0.000	.062239	.143293	0

The probability ratio is given by:

$$\frac{Pr(y = Coke)}{Pr(y = Pepsi)} = \frac{Pr(y = Coke)}{Pr(y = 7Up)} \frac{0.333}{0.333} = 0.5 \tag{3}$$

(c) Compute the probability ratio of choosing *Coke* relative to *Pepsi* and *7Up* if the

price of each is $1.25, a display is present for *Coke* but not for the others, and none of the items is featured.

```
. estat mfx, at (PRICE=1.25 Coke:DISPLAY=1 Pepsi:DISPLAY=0 7Up:DISPLAY=0 FEATURE=0)

Pr(choice = Pepsi|1 selected) = .27871022

variable  |   dp/dx   Std. err.     z     P>|z|  [    95% C.I.    ]        X
----------+-------------------------------------------------------------------
PRICE     |
    Pepsi | -.350689   .040498   -8.66   0.000  -.430064  -.271314      1.25
      7Up |  .135508   .021051    6.44   0.000   .094249   .176767      1.25
     Coke |  .215181   .021391   10.06   0.000   .173255   .257107      1.25
----------+-------------------------------------------------------------------
FEATURE   |
    Pepsi | -.002132   .016057   -0.13   0.894  -.033602   .029339         0
      7Up |  .000824   .006193    0.13   0.894  -.011314   .012962         0
     Coke |  .001308   .009864    0.13   0.895  -.018025   .020641         0
----------+-------------------------------------------------------------------
DISPLAY   |
    Pepsi |  .092966   .016356    5.68   0.000   .060908   .125024         0
      7Up | -.035923   .004269   -8.41   0.000   -.04429  -.027555         0
     Coke | -.057044   .012087   -4.72   0.000  -.080734  -.033353         1
------------------------------------------------------------------------------

Pr(choice = 7Up|1 selected) = .27871022

variable  |   dp/dx   Std. err.     z     P>|z|  [    95% C.I.    ]        X
----------+-------------------------------------------------------------------
PRICE     |
    Pepsi |  .135508   .021051    6.44   0.000   .094249   .176767      1.25
      7Up | -.350689   .040498   -8.66   0.000  -.430064  -.271314      1.25
     Coke |  .215181   .021391   10.06   0.000   .173255   .257107      1.25
----------+-------------------------------------------------------------------
FEATURE   |
    Pepsi |  .000824   .006193    0.13   0.894  -.011314   .012962         0
      7Up | -.002132   .016057   -0.13   0.894  -.033602   .029339         0
     Coke |  .001308   .009864    0.13   0.895  -.018025   .020641         0
----------+-------------------------------------------------------------------
DISPLAY   |
    Pepsi | -.035923   .004269   -8.41   0.000   -.04429  -.027555         0
      7Up |  .092966   .016356    5.68   0.000   .060908   .125024         0
     Coke | -.057044   .012087   -4.72   0.000  -.080734  -.033353         1
------------------------------------------------------------------------------

Pr(choice = Coke|1 selected) = .44257956

variable  |   dp/dx   Std. err.     z     P>|z|  [    95% C.I.    ]        X
----------+-------------------------------------------------------------------
PRICE     |
    Pepsi |  .215181   .021391   10.06   0.000   .173255   .257107      1.25
      7Up |  .215181   .021391   10.06   0.000   .173255   .257107      1.25
     Coke | -.430362   .042783  -10.06   0.000  -.514215  -.346509      1.25
----------+-------------------------------------------------------------------
FEATURE   |
    Pepsi |  .001308   .009864    0.13   0.895  -.018025   .020641         0
      7Up |  .001308   .009864    0.13   0.895  -.018025   .020641         0
     Coke | -.002616   .019728   -0.13   0.895  -.041282    .03605         0
----------+-------------------------------------------------------------------
DISPLAY   |
    Pepsi | -.057044   .012087   -4.72   0.000  -.080734  -.033353         0
      7Up | -.057044   .012087   -4.72   0.000  -.080734  -.033353         0
     Coke |  .114087   .024174    4.72   0.000   .066706   .161468         1
------------------------------------------------------------------------------
```

The probability ratio is given by:

$$\frac{Pr(y = Coke|DISPLAY = 1)}{Pr(y = Pepsi|DISPLAY = 0)} = \frac{Pr(y = Coke|DISPLAY = 1)}{Pr(y = 7Up|DISPLAY = 0)} = \frac{0.44257956}{0.27871022} = 1.58795 \qquad (4)$$

(d) Compute the change in the probability of purchase of each type of soda if the price of *Coke* changes from \$1.25 to \$1.30, with the prices of the *Pepsi* and *7Up* remaining at \$1.25. Assume that a display is present for *Coke*, but not for the others, and none of the items is featured.

We create the same table as in (c) using:

```
estat mfx, at (Coke:PRICE=1.30 Pepsi:PRICE=1.25 7Up:PRICE=1.25 Coke:DISPLAY=1 Pepsi:DISPLAY=0
7Up:DISPLAY=0 FEATURE=0).
```

The full table is omitted for space, but the probabilities are represented in the table below:

Table 6: Probability of Choosing Different Sodas: Conditional Logit Model

Price of Coke	$Pr(y = Coke)$[a]	$Pr(y = Pepsi)$	$Pr(y = 7Up)$
\$ 1.25	0.44257956	0.27871022	0.27871022
\$ 1.30	0.4211822	0.2894089	0.2894089
Change	-0.02139736	+0.01069868	+0.01069868

[a]Note: Coke has a display, Pepsi and 7Up do not.

(e) Add the alternative specific "intercept" terms for *Pepsi* and *7Up* to the model in (a). Estimate the conditional logit model. Compute the probability ratio in (c) based upon these new estimates.

```
Alternative-specific conditional logit         Number of obs      =      5,466
Case ID variable: ID                           Number of cases    =       1822

Alternatives variable: ALT                     Alts per case: min =          3
                                                              avg =        3.0
                                                              max =          3

                                                  Wald chi2(3)    =     302.93
Log likelihood = -1811.3543                       Prob > chi2     =     0.0000

------------------------------------------------------------------------------
      CHOICE | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
ALT          |
       PRICE |  -1.849186   .1886595    -9.80   0.000    -2.218952    -1.47942
     FEATURE |  -.0408576   .0830752    -0.49   0.623    -.2036821    .1219669
     DISPLAY |   .4726786   .0935445     5.05   0.000     .2893346    .6560225
-------------+----------------------------------------------------------------
Pepsi        |
       _cons |   .2840865   .0625595     4.54   0.000     .1614722    .4067008
-------------+----------------------------------------------------------------
7Up          |
       _cons |   .0906629   .0639666     1.42   0.156    -.0347094    .2160352
-------------+----------------------------------------------------------------
Coke         |  (base alternative)
------------------------------------------------------------------------------
```

```
. estat mfx, at (PRICE=1.25 Coke:DISPLAY=1 Pepsi:DISPLAY=0 7Up:DISPLAY=0 FEATURE=0)
```

Pr(choice = Pepsi|1 selected) = .32985

variable	dp/dx	Std. err.	z	P>|z|	[95% C.I.]		X
PRICE							
Pepsi	-.408761	.046806	-8.73	0.000	-.500498	-.317023	1.25
7Up	.16581	.024974	6.64	0.000	.116862	.214758	1.25
Coke	.242951	.02675	9.08	0.000	.190522	.295379	1.25
FEATURE							
Pepsi	-.009032	.018341	-0.49	0.622	-.044979	.026916	0
7Up	.003664	.007365	0.50	0.619	-.010772	.018099	0
Coke	.005368	.010982	0.49	0.625	-.016157	.026893	0
DISPLAY							
Pepsi	.104485	.018887	5.53	0.000	.067467	.141503	0
7Up	-.042383	.005483	-7.73	0.000	-.05313	-.031637	0
Coke	-.062102	.013761	-4.51	0.000	-.089072	-.035131	1

Pr(choice = 7Up|1 selected) = .2718402

variable	dp/dx	Std. err.	z	P>|z|	[95% C.I.]		X
PRICE							
Pepsi	.16581	.024974	6.64	0.000	.116862	.214758	1.25
7Up	-.366034	.040076	-9.13	0.000	-.444582	-.287485	1.25
Coke	.200224	.018788	10.66	0.000	.163399	.237048	1.25
FEATURE							
Pepsi	.003664	.007365	0.50	0.619	-.010772	.018099	0
7Up	-.008087	.016348	-0.49	0.621	-.040129	.023954	0
Coke	.004424	.008989	0.49	0.623	-.013194	.022042	0
DISPLAY							
Pepsi	-.042383	.005483	-7.73	0.000	-.05313	-.031637	0
7Up	.093563	.016297	5.74	0.000	.061622	.125505	0
Coke	-.05118	.011129	-4.60	0.000	-.072992	-.029368	1

Pr(choice = Coke|1 selected) = .39830979

variable	dp/dx	Std. err.	z	P>|z|	[95% C.I.]		X
PRICE							
Pepsi	.242951	.02675	9.08	0.000	.190522	.295379	1.25
7Up	.200224	.018788	10.66	0.000	.163399	.237048	1.25
Coke	-.443174	.043583	-10.17	0.000	-.528595	-.357754	1.25
FEATURE							
Pepsi	.005368	.010982	0.49	0.625	-.016157	.026893	0
7Up	.004424	.008989	0.49	0.623	-.013194	.022042	0
Coke	-.009792	.019969	-0.49	0.624	-.048931	.029347	0
DISPLAY							
Pepsi	-.062102	.013761	-4.51	0.000	-.089072	-.035131	0
7Up	-.05118	.011129	-4.60	0.000	-.072992	-.029368	0
Coke	.113282	.024644	4.60	0.000	.06498	.161583	1

The probability ratios are given by:

$$\frac{Pr(y = Coke|DISPLAY = 1)}{Pr(y = Pepsi|DISPLAY = 0)} = \frac{0.39830979}{0.32985} = 1.2075482 \tag{5}$$

$$\frac{Pr(y = Coke|DISPLAY = 1)}{Pr(y = 7Up|DISPLAY = 0)} = \frac{.39830979}{0.2718402} = 1.465235 \tag{6}$$

(f) Based on the estimates in (e), calculate the effects of the price change in (d) on the choice probability for each brand.

As in (d), the full table is omitted for space, but the probabilities are represented in the table below:

Table 7: Probability of Choosing Different Sodas: Conditional Logit Model

Price of Coke	$Pr(y = Coke)$[a]	$Pr(y = Pepsi)$	$Pr(y = 7Up)$
\$ 1.25	0.39830979	0.32985	0.2718402
\$ 1.30	0.37637297	0.3418759	0.28175114
Change	-0.021936	+0.012025	+0.009910

[a]Note: Coke has a display, Pepsi and 7Up do not.

(g) Estimate a nested logit model with cola and non-cola nest, and repeat (d).

```
RUM-consistent nested logit regression        Number of obs      =      5,466
Case variable: ID                             Number of cases    =       1822

Alternative variable: ALT                     Alts per case: min =          3
                                                             avg =        3.0
                                                             max =          3

                                                 Wald chi2(7)    =     179.89
Log likelihood = -1801.6476                      Prob > chi2     =     0.0000

------------------------------------------------------------------------------
      CHOICE | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
ALT          |
     FEATURE |  -.0370077   .0897056    -0.41   0.680    -.2128276    .1388121
-------------+----------------------------------------------------------------
ALT equations
-------------+----------------------------------------------------------------
Pepsi        |
       PRICE |  -2.306575   .3183694    -7.24   0.000    -2.930567   -1.682582
     DISPLAY |   .5644912   .1513013     3.73   0.000     .2679461    .8610363
       _cons |          0  (base)
-------------+----------------------------------------------------------------
7Up          |
       PRICE |  -1.967287   .2342082    -8.40   0.000    -2.426327   -1.508247
     DISPLAY |   .3273523    .130295     2.51   0.012     .0719788    .5827259
       _cons |  -.4797802   .3959623    -1.21   0.226    -1.255852    .2962916
-------------+----------------------------------------------------------------
Coke         |
       PRICE |  -1.229497   .2913356    -4.22   0.000    -1.800504   -.6584893
     DISPLAY |   .6303467   .1552597     4.06   0.000     .3260432    .9346503
       _cons |  -1.587008   .4672671    -3.40   0.001    -2.502835   -.6711816
-------------+----------------------------------------------------------------
dissimilarity parameters
-------------+----------------------------------------------------------------
/type        |
    Cola_tau |   1.040659   .1422981                         .76176    1.319558
   Other_tau |          1   75542.11                      -148058.8    148060.8
------------------------------------------------------------------------------
LR test for IIA (tau=1): chi2(2) = 0.08                   Prob > chi2 = 0.9586
```

Table 8: Probability of Choosing Different Sodas: Nested Logit Model

Price of Coke	$Pr(y = Coke)$[a]	$Pr(y = Pepsi)$	$Pr(y = 7Up)$
\$ 1.25	0.4320086	0.297075	0.2709164
\$ 1.30	0.4174125	0.3045048	0.2780827
Change	-0.0145961	+0.0074298	+0.0071663

[a]Note: Coke has a display, Pepsi and 7Up do not.

Question4

(a) Use an ordered probit to explain the probability of PSECHOICE as a function of GRADES. Calculate the probability that a student will choose no college, a two-year college, and a four-year college if the student's grades are the median value, GRADES = 6.64. Recompute these probabilities assuming that GRADES = 4.905. Discuss the probability changes. Are they what you anticipated? Explain.

```
. oprobit PSECHOICE GRADES

Iteration 0:   log likelihood = -1018.6575
Iteration 1:   log likelihood = -876.21962
Iteration 2:   log likelihood = -875.82172
Iteration 3:   log likelihood = -875.82172

Ordered probit regression                               Number of obs =  1,000
                                                        LR chi2(1)    = 285.67
                                                        Prob > chi2   = 0.0000
Log likelihood = -875.82172                             Pseudo R2     = 0.1402
```

PSECHOICE	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
GRADES	-.3066252	.0191735	-15.99	0.000	-.3442045	-.2690459
/cut1	-2.9456	.1468283			-3.233378	-2.657822
/cut2	-2.089993	.1357681			-2.356094	-1.823893

The predicted probabilities (presented below) are in line with the expected once we consider that the grades are in an inverted scale; that is, better grades have smaller values. With that in mind, it makes sense that better grades (though smaller in value) have a positive effect on the probability that the student has a more valuable degree. A student in the 25th percentile has a 71.6% probability of going to a 4-year college, whereas one with the median grades has only a 52.1% probability.

```
. margins, at(GRADES=6.64)

Adjusted predictions                                    Number of obs = 1,000
Model VCE: OIM

1._predict: Pr(PSECHOICE==1), predict(pr outcome(1))
2._predict: Pr(PSECHOICE==2), predict(pr outcome(2))
3._predict: Pr(PSECHOICE==3), predict(pr outcome(3))

At: GRADES = 6.64

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    _predict |
          1  |   .1815145   .0130109    13.95   0.000     .1560136    .2070154
          2  |   .2969523    .015928    18.64   0.000     .2657339    .3281706
          3  |   .5215332   .0170868    30.52   0.000     .4880437    .5550227
------------------------------------------------------------------------------

. margins, at(GRADES=4.95)

Adjusted predictions                                    Number of obs = 1,000
Model VCE: OIM

1._predict: Pr(PSECHOICE==1), predict(pr outcome(1))
2._predict: Pr(PSECHOICE==2), predict(pr outcome(2))
3._predict: Pr(PSECHOICE==3), predict(pr outcome(3))

At: GRADES = 4.95

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    _predict |
          1  |    .076674   .0093575     8.19   0.000     .0583336    .0950144
          2  |   .2069197   .0138072    14.99   0.000     .1798581    .2339813
          3  |   .7164063   .0185133    38.70   0.000     .6801209    .7526916
------------------------------------------------------------------------------
```

(b) Expand the ordered probit model to include family income ($FAMINC$), family size ($FAMSIZ$), and the indicator variables $BLACK$ and $PARCOLL$. Discuss the estimates and their signs and significance.

```
. oprobit PSECHOICE GRADES FAMINC FAMSIZ BLACK PARCOLL

Iteration 0:   log likelihood = -1018.6575
Iteration 1:   log likelihood = -841.36397
Iteration 2:   log likelihood = -839.86676
Iteration 3:   log likelihood = -839.86473
Iteration 4:   log likelihood = -839.86473

Ordered probit regression                               Number of obs =  1,000
                                                        LR chi2(5)    = 357.59
                                                        Prob > chi2   = 0.0000
Log likelihood = -839.86473                             Pseudo R2     = 0.1755
```

PSECHOICE	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
GRADES	-.2952923	.0202251	-14.60	0.000	-.3349328	-.2556518
FAMINC	.0052525	.001322	3.97	0.000	.0026615	.0078435
FAMSIZ	-.0241215	.0301846	-0.80	0.424	-.0832822	.0350391
BLACK	.7131312	.1767871	4.03	0.000	.3666348	1.059628
PARCOLL	.4236226	.1016424	4.17	0.000	.2244071	.6228381
/cut1	-2.595845	.2045863			-2.996827	-2.194864
/cut2	-1.694591	.1971365			-2.080971	-1.30821

(c) Test the joint significance of the variables added in (b) using a likelihood ratio test.

As the likelihood ratio test presented below shows, the added variables are jointly significant at the 99.9% level.

```
. lrtest rest unrest

Likelihood-ratio test
Assumption: rest nested within unrest

 LR chi2(4) =  71.91
Prob > chi2 = 0.0000
```

(d) Compute the probability that a black student from a household of four members, including a parent who went to college, and household income of \$52,000, will attend a four-year college if (i) $GRADES = \mathbf{6.64}$ and (ii) $GRADES = \mathbf{4.905}$.

```
. margins, at (BLACK=1 FAMINC=52 FAMSIZ=4 PARCOLL=1 GRADES=(6.64,4.95))

Adjusted predictions                                    Number of obs = 1,000
Model VCE: OIM

1._predict: Pr(PSECHOICE==1), predict(pr outcome(1))
2._predict: Pr(PSECHOICE==2), predict(pr outcome(2))
3._predict: Pr(PSECHOICE==3), predict(pr outcome(3))

1._at: GRADES  = 6.64
       FAMINC  =   52
       FAMSIZ  =    4
       BLACK   =    1
       PARCOLL =    1
2._at: GRADES  = 4.95
       FAMINC  =   52
       FAMSIZ  =    4
       BLACK   =    1
       PARCOLL =    1
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
_predict#_at						
1 1	.0256775	.0115091	2.23	0.026	.0031199	.048235
1 2	.0071916	.0040241	1.79	0.074	-.0006955	.0150788
2 1	.1218153	.0322475	3.78	0.000	.0586113	.1850193
2 2	.0538255	.0196685	2.74	0.006	.0152759	.0923751
3 1	.8525072	.0432158	19.73	0.000	.7678058	.9372087
3 2	.9389829	.0235394	39.89	0.000	.8928465	.9851193

The outcomes of interest are the last two lines (3,1) and (3,2). A black student (with the other variables at the values described above) has a 85.25% probability of going to a 4-year college if the has median grades and a 93.89% probability of the same outcome if the is the top 25th percentile of grades.

(e) Repeat (d) for a "non-black" student and discuss the differences in your findings.

```
. margins, at (BLACK=0 FAMINC=52 FAMSIZ=4 PARCOLL=1 GRADES=(6.64,4.95))

Adjusted predictions                                    Number of obs = 1,000
Model VCE: OIM

1._predict: Pr(PSECHOICE==1), predict(pr outcome(1))
2._predict: Pr(PSECHOICE==2), predict(pr outcome(2))
3._predict: Pr(PSECHOICE==3), predict(pr outcome(3))

1._at: GRADES  = 6.64
       FAMINC  =   52
       FAMSIZ  =    4
       BLACK   =    0
       PARCOLL =    1
2._at: GRADES  = 4.95
       FAMINC  =   52
       FAMSIZ  =    4
       BLACK   =    0
       PARCOLL =    1
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
_predict#_at						
1 1	.1083463	.0169399	6.40	0.000	.0751447	.1415478
1 2	.0414223	.0086218	4.80	0.000	.0245238	.0583207
2 1	.2607997	.0201004	12.97	0.000	.2214037	.3001958
2 2	.160955	.0179199	8.98	0.000	.1258327	.1960773
3 1	.630854	.0320606	19.68	0.000	.5680164	.6936916
3 2	.7976228	.0247454	32.23	0.000	.7491227	.8461229

For a "non-black" student with the all other characteristics being the same, these probabilities are lower. A "non-black" student has with median grades has a 63.08% probability of attending a 4-year college, or a 79.76% probability if their grades are on the top 25th percentile.

Question 5

1. Estimate the following model

$$logspend = \beta_1 + \beta_2 \ln income + \beta_3 Age + \beta_4 Adepcnt + \beta_5 ownrent + \epsilon \tag{7}$$

(a) Using OLS, what is the effect of 10% increase in income on credit card expenditure?

```
. reg logspend logincome age adepcnt ownrent
```

Source	SS	df	MS
Model	2166.35661	4	541.589151
Residual	18551.6408	10,494	1.76783313
Total	20717.9974	10,498	1.97351852

Number of obs	=	10,499
F(4, 10494)	=	306.36
Prob > F	=	0.0000
R-squared	=	0.1046
Adj R-squared	=	0.1042
Root MSE	=	1.3296

logspend	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
logincome	1.121208	.0325205	34.48	0.000	1.057462	1.184954
age	-.0145581	.0014062	-10.35	0.000	-.0173146	-.0118017
adepcnt	-.0272734	.0111879	-2.44	0.015	-.0492038	-.0053429
ownrent	-.2033712	.0297326	-6.84	0.000	-.2616527	-.1450897
_cons	-3.363318	.2433664	-13.82	0.000	-3.840362	-2.886273

A 10% increase in income is associated with a 11.2% increase in credit card expenditure, all else equal.

(b) Using Censored regression, what is the effect of 10% increase in income on credit card expenditure?

```
Tobit regression                                Number of obs     =  10,499
                                                     Uncensored =  10,251
Limits: Lower =    1                               Left-censored =     248
        Upper = +inf                              Right-censored =       0

                                                LR chi2(4)        = 1186.53
                                                Prob > chi2       =  0.0000
Log likelihood = -17455.665                     Pseudo R2         =  0.0329
```

logspend	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
logincome	1.094635	.0314454	34.81	0.000	1.032996	1.156274
age	-.0144051	.0013544	-10.64	0.000	-.01706	-.0117501
adepcnt	-.0256828	.010766	-2.39	0.017	-.0467862	-.0045793
ownrent	-.1913027	.0286148	-6.69	0.000	-.2473931	-.1352123
_cons	-3.153538	.2352271	-13.41	0.000	-3.614628	-2.692448
var(e.logspend)	1.633056	.0230068			1.588575	1.678782

A 10% increase in income is associated with a 10.94% increase in credit card expenditure, all else equal.

(c) Using Heckman Two-Step Estimator, what the is effect of 10% increase in income on credit card expenditure?

```
Heckman selection model                         Number of obs     =     13,444
(regression model with sample selection)              Selected    =     10,499
                                                      Nonselected =      2,945

                                                Wald chi2(4)      =     457.11
Log likelihood = -24165.62                      Prob > chi2       =     0.0000
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
logspend						
logincome	.6813308	.0372162	18.31	0.000	.6083884	.7542732
age	-.0116583	.0016161	-7.21	0.000	-.0148257	-.0084908
adepcnt	.0580111	.0127608	4.55	0.000	.0330005	.0830217
ownrent	-.3394852	.0340503	-9.97	0.000	-.4062226	-.2727478
_cons	.4844651	.2788994	1.74	0.082	-.0621676	1.031098
cardhldr						
logincome	.3844089	.0276048	13.93	0.000	.3303045	.4385133
age	.0019958	.00122	1.64	0.102	-.0003954	.004387
adepcnt	-.0902368	.0092953	-9.71	0.000	-.1084553	-.0720183
ownrent	.1976502	.0254879	7.75	0.000	.1476948	.2476055
_cons	-2.284821	.2056409	-11.11	0.000	-2.68787	-1.881772
/athrho	-1.720274	.0320709	-53.64	0.000	-1.783131	-1.657416
/lnsigma	.4965339	.0082761	60.00	0.000	.4803131	.5127548
rho	-.937896	.0038598			-.9450309	-.9298683
sigma	1.643017	.0135978			1.61658	1.669885
lambda	-1.540979	.0168915			-1.574085	-1.507872

LR test of indep. eqns. (rho = 0): chi2(1) = 1021.31   Prob > chi2 = 0.0000

A 10% increase in income is associated with a 6.81% increase in credit card expenditure, all else equal.

2. Create a subsample where only credit cardholders appear and do the following:
(a) Estimate the above model using OLS. What is the difference in credit card spending between home owner and renter?

```
. reg logspend logincome age adepcnt ownrent if cardhldr==1

      Source |       SS           df       MS      Number of obs   =    10,499
-------------+----------------------------------   F(4, 10494)     =    306.36
       Model |  2166.35661         4  541.589151   Prob > F        =    0.0000
    Residual |  18551.6408    10,494  1.76783313   R-squared       =    0.1046
-------------+----------------------------------   Adj R-squared   =    0.1042
       Total |  20717.9974    10,498  1.97351852   Root MSE        =    1.3296

------------------------------------------------------------------------------
    logspend | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   logincome |   1.121208   .0325205    34.48   0.000     1.057462    1.184954
         age |  -.0145581   .0014062   -10.35   0.000    -.0173146   -.0118017
     adepcnt |  -.0272734   .0111879    -2.44   0.015    -.0492038   -.0053429
     ownrent |  -.2033712   .0297326    -6.84   0.000    -.2616527   -.1450897
       _cons |  -3.363318   .2433664   -13.82   0.000    -3.840362   -2.886273
------------------------------------------------------------------------------
```

A renter has a credit card expenditure 22.55%[1] lower than a homeowner, on average.

(b) Estimate the above model using truncated regression. What is the difference in credit card spending between home owner and renter?

```
. truncreg logspend logincome age adepcnt ownrent if cardhldr==1, ll(1)
(248 obs truncated)

Fitting full model:

Iteration 0:   log likelihood = -15858.701
Iteration 1:   log likelihood =  -15858.46
Iteration 2:   log likelihood =  -15858.46

Truncated regression
Limit: Lower =    1                             Number of obs     =     10,251
       Upper = +inf                             Wald chi2(4)      =    1179.92
Log likelihood =  -15858.46                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
    logspend | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
   logincome |   .9829195   .0292028    33.66   0.000     .9256831    1.040156
         age |  -.0138854   .0012399   -11.20   0.000    -.0163155   -.0114553
     adepcnt |  -.0157937   .0097809    -1.61   0.106     -.034964    .0033765
     ownrent |  -.1493865   .0259945    -5.75   0.000    -.2003349   -.0984382
       _cons |  -2.242699   .2181561   -10.28   0.000    -2.670277   -1.815121
-------------+----------------------------------------------------------------
      /sigma |   1.141809   .0081287   140.47   0.000     1.125877    1.157741
------------------------------------------------------------------------------
```

A renter has a credit card expenditure 16.11% lower than a homeowner, on average.

[1] Using: Effect $= (e^{\beta_k} - 1) \times 100$

3. Now we are interested in explaining the number of major derogatory reports as function of log income, age, the number of dependents, home ownership status and ratio of monthly credit card expenditure to yearly income.

(a) Estimate this model using Poisson regression for credit cardholders only. What is the effect of 10% increase in income on the expected value (mean) of the number of major derogatory reports? Is Poisson regression a good specification for the data at hand?

```
. poisson minordrg logincome age adepcnt ownrent exp_inc if cardhldr==1

Iteration 0:   log likelihood = -6371.1793
Iteration 1:   log likelihood =  -6371.179

Poisson regression                                      Number of obs = 10,499
                                                        LR chi2(5)    = 289.28
                                                        Prob > chi2   = 0.0000
Log likelihood = -6371.179                              Pseudo R2     = 0.0222
```

minordrg	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
logincome	.2511779	.0503345	4.99	0.000	.1525241	.3498318
age	.0152978	.0020865	7.33	0.000	.0112084	.0193873
adepcnt	.0789871	.0162789	4.85	0.000	.047081	.1108932
ownrent	.2370228	.0486273	4.87	0.000	.141715	.3323306
exp_inc	.7503179	.1620077	4.63	0.000	.4327887	1.067847
_cons	-4.297388	.3806971	-11.29	0.000	-5.043541	-3.551236

```
. mfx

Marginal effects after poisson
      y  = Predicted number of events (predict)
         =  .20766452
```

variable	dy/dx	Std. err.	z	P>|z|	[95%	C.I.]	X
loginc~e	.0521607	.01041	5.01	0.000	.031756	.072565	7.76554
age	.0031768	.00043	7.40	0.000	.002335	.004018	33.6749
adepcnt	.0164028	.00337	4.87	0.000	.009801	.023005	.99038
ownrent*	.0495816	.01021	4.85	0.000	.029565	.069598	.479093
exp_inc	.1558144	.03354	4.64	0.000	.090068	.221561	.090744

(*) dy/dx is for discrete change of dummy variable from 0 to 1

A 10% increase in income is associated with 0.0104% increase in derogatory reports.

```
. summarize minordrg, detail

                          MINORDRG
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0              0
 5%            0              0
10%            0              0       Obs              13,444
25%            0              0       Sum of wgt.      13,444

50%            0                      Mean           .2905385
                        Largest       Std. dev.      .7676199
75%            0              7
90%            1              8       Variance       .5892402
95%            2              9       Skewness       3.733436
99%            4             11       Kurtosis       21.84461
```

The poisson model requires the assumption that the mean of the dependent variable is equal to its variance. As it can be seen above, it is not suitable as the variance is more than twice the value of the mean. We should use the negative binomial model instead.

(b) Estimate this model using negative binomial regression for credit cardholders only. What is the effect of 10% increase in income on the expected value (mean) of the number of major derogatory reports?

```
Negative binomial regression                        Number of obs =  13,444
                                                    LR chi2(5)    =  329.25
Dispersion: mean                                    Prob > chi2   =  0.0000
Log likelihood = -9223.5203                         Pseudo R2     =  0.0175

------------------------------------------------------------------------------
    minordrg | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
   logincome |   .0559386   .0655147     0.85   0.393    -.0724678     .184345
         age |   .0148228   .0029555     5.02   0.000     .0090301    .0206154
     adepcnt |   .1178443   .0215002     5.48   0.000     .0757047    .1599839
     ownrent |   .1592426   .0591841     2.69   0.007     .0432439    .2752413
     exp_inc |  -3.910639   .2963569   -13.20   0.000    -4.491487    -3.32979
       _cons |  -2.810009   .4904226    -5.73   0.000    -3.771219   -1.848798
    cardhldr |          1  (offset)
-------------+----------------------------------------------------------------
    /lnalpha |   1.625203   .0381283                      1.550473    1.699933
-------------+----------------------------------------------------------------
       alpha |   5.079448   .1936708                      4.713697    5.473579
------------------------------------------------------------------------------
LR test of alpha=0: chibar2(01) = 3308.53                Prob >= chibar2 = 0.000
```

```
. mfx

Marginal effects after nbreg
      y  = Predicted number of events (predict)
         =  .20679726
```

variable	dy/dx	Std. err.	z	P>|z|	[95% C.I.	]	X
loginc~e	.0621476	.01491	4.17	0.000	.032925	.09137	7.76554
age	.0036212	.00064	5.68	0.000	.002372	.004871	33.6749
adepcnt	.0156676	.00479	3.27	0.001	.006283	.025052	.99038
ownrent*	.0458399	.01357	3.38	0.001	.019234	.072445	.479093
exp_inc	.1714965	.05613	3.06	0.002	.06149	.281503	.090744

```
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

A 10% increase in income is associated with 0.0621% increase in derogatory reports.

(c) Estimate the two models taking into account the truncation.What is the effect of 10% increase in income on the expected value (mean) of the number of major derogatory reports?

(i) Poisson

```
. poisson minordrg logincome age adepcnt ownrent exp_inc, offset(cardhldr)

Iteration 0:   log likelihood = -10877.824
Iteration 1:   log likelihood = -10877.787
Iteration 2:   log likelihood = -10877.787

Poisson regression                                      Number of obs = 13,444
                                                        LR chi2(5)    = 560.18
                                                        Prob > chi2   = 0.0000
Log likelihood = -10877.787                             Pseudo R2     = 0.0251
```

minordrg	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
logincome	.0216212	.0378257	0.57	0.568	-.0525158	.0957583
age	.0114671	.0016232	7.06	0.000	.0082858	.0146485
adepcnt	.105451	.0123339	8.55	0.000	.081277	.1296249
ownrent	.1673314	.0366663	4.56	0.000	.0954668	.2391959
exp_inc	-3.265179	.2400541	-13.60	0.000	-3.735676	-2.794681
_cons	-2.638729	.2836133	-9.30	0.000	-3.1946	-2.082857
cardhldr	1	(offset)				

```
. mfx

Marginal effects after poisson
      y  = Predicted number of events (predict)
         =  .25791819
```

variable	dy/dx	Std. err.	z	P>|z|	[95% C.I.	]	X
loginc~e	.0055765	.00975	0.57	0.567	-.013535	.024688	7.72481
age	.0029576	.00042	7.10	0.000	.002141	.003774	33.4718
adepcnt	.0271977	.00317	8.57	0.000	.020981	.033414	1.01726
ownrent*	.0435277	.00959	4.54	0.000	.024737	.062318	.455965
exp_inc	-.842149	.06114	-13.77	0.000	-.961988	-.72231	.070974
cardhldr	(offset)						.780943

(*) dy/dx is for discrete change of dummy variable from 0 to 1

A 10% increase in income is associated with 0.0055% increase in derogatory reports.

(ii) Negative Binomial

```
Negative binomial regression                    Number of obs = 13,444
                                                LR chi2(5)    = 329.25
Dispersion: mean                                Prob > chi2   = 0.0000
Log likelihood = -9223.5203                     Pseudo R2     = 0.0175
```

minordrg	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
logincome	.0559386	.0655147	0.85	0.393	-.0724678	.184345
age	.0148228	.0029555	5.02	0.000	.0090301	.0206154
adepcnt	.1178443	.0215002	5.48	0.000	.0757047	.1599839
ownrent	.1592426	.0591841	2.69	0.007	.0432439	.2752413
exp_inc	-3.910639	.2963569	-13.20	0.000	-4.491487	-3.32979
_cons	-2.810009	.4904226	-5.73	0.000	-3.771219	-1.848798
cardhldr	1	(offset)				
/lnalpha	1.625203	.0381283			1.550473	1.699933
alpha	5.079448	.1936708			4.713697	5.473579

LR test of alpha=0: chibar2(01) = 3308.53   Prob >= chibar2 = 0.000

```
Marginal effects after nbreg
      y  = Predicted number of events (predict)
         =  .30547998
```

variable	dy/dx	Std. err.	z	P>|z|	[95%	C.I.]	X
loginc~e	.0170881	.02	0.85	0.393	-.022117	.056294	7.72481
age	.0045281	.00091	5.00	0.000	.002754	.006302	33.4718
adepcnt	.0359991	.0066	5.45	0.000	.023057	.048941	1.01726
ownrent*	.0490395	.01839	2.67	0.008	.012996	.085083	.455965
exp_inc	-1.194622	.09283	-12.87	0.000	-1.37656	-1.01268	.070974
cardhldr	(offset)						.780943

(*) dy/dx is for discrete change of dummy variable from 0 to 1

A 10% increase in income is associated with 0.017% increase in derogatory reports.