Transport Data Analysis and Modeling Methodologies

Lab Session #15
(Ordered Discrete Data – Bivariate Ordered Probit)
Based on Example 14.1

A survey of 250 commuters was in the Seattle metropolitan area (this sample is reduced from the 322 given in the book due to the elimination of some missing data). The survey's intent was to gather information on commuters' opinions of high-occupancy vehicle (HOV) lanes (lanes that are restricted for use by vehicles with 2 or more occupants). The variables available from this survey are given on the attached table.

Among the questions asked, commuters were asked whether they agreed with the statement "Existing HOV lanes are being adequately used." (variable number x28 in the table) and "HOV lanes should be open to all vehicles, regardless of vehicle occupancy level." (variable number x29 in the table). The question provided ordered responses of; strongly disagree, disagree, neutral, agree, agree strongly. These two questions are obviously interrelated. To understand the factors determining these two commuter opinions, a bivariate ordered probit model of these survey questions are appropriate.

Your task is to estimate a bivariate ordered response model of whether commuters believe existing HOV lanes are being adequately used and whether they believe HOV lanes should be open to all vehicles, regardless of vehicle occupancy level. Your solution to this problem should include:

1. The results of your best model specification.

2. A discussion of the logical process that led you to the selection of your final specification. (e.g. Discuss the theory behind the inclusion of your selected variables). Include t-statistics and justify the sign of your variables.

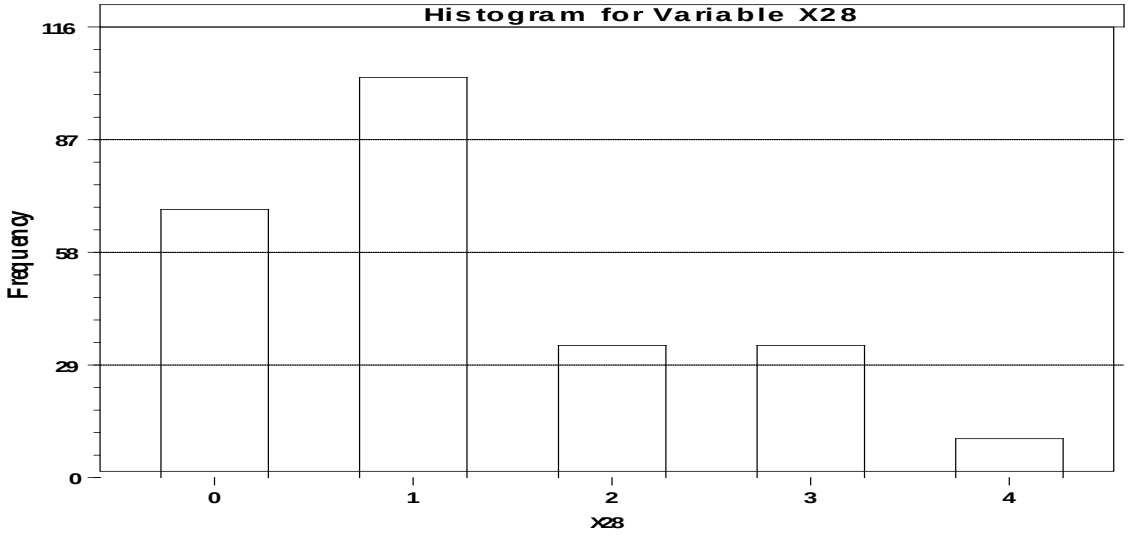
Variables available for your specification are (in file Ex14-1.txt):

Variable Number	Explanation
x1	Usual mode of travel: 0 if drive alone, 1 if two person carpool, 2 if three or more person carpool, 3 if vanpool, 4 if bus, 5 if bicycle or walk, 6 if motorcycle, 7 if other
x2	Have used HOV lanes: 1 if yes, 0 if no
x3	If used HOV lanes, what mode is most often used: 0 in a bus, 1 in two person carpool, 2 in three or more person carpool, 3 in vanpool, 4 alone in vehicle, 5 on motorcycle
x4	Sometimes eligible for HOV lane use but do not use: 1 if yes, 0 if no
x5	Reason for not using HOV lanes when eligible: 0 if slower than regular lanes, 1 if too much trouble to change lanes, 2 if HOV lanes are not safe, 3 if traffic moves fast enough, 4 if forget to use HOV lanes, 5 if other
x6	Usual mode of travel one year ago: 0 if drive alone, 1 if two person carpool, 2 if three or more person carpool, 3 if vanpool, 4 if bus, 5 if bicycle or walk, 6 if motorcycle, 7 if other
x7	Commuted to work in Seattle a year ago: 1 if yes, 0 if no
x8	Have flexible work start times: 1 if yes, 0 if no
x9	Changed departure times to work in the last year: 1 if yes, 0 if no
x10	On average, number of minutes leaving earlier for work relative to last year
x11	On average, number of minutes leaving later for work relative to last year
x12	If changed departure times to work in the last year, reason why: 0 if change in travel mode, 1 if increasing traffic congestion, 2 if change in work start time, 3 if presence of HOV lanes, 4 if change in residence, 5 if change in lifestyle, 6 if other
x13	Changed route to work in the last year: 1 if yes, 0 if no
x14	If changed route to work in the last year, reason why: 0 if change in travel mode, 1 if increasing traffic congestion, 2 if change in work start time, 3 if presence of HOV lanes, 4 if change in residence, 5 if change in lifestyle, 6 if other
x15	Usually commute to or from work on Interstate 90: 1 if yes, 0 if no

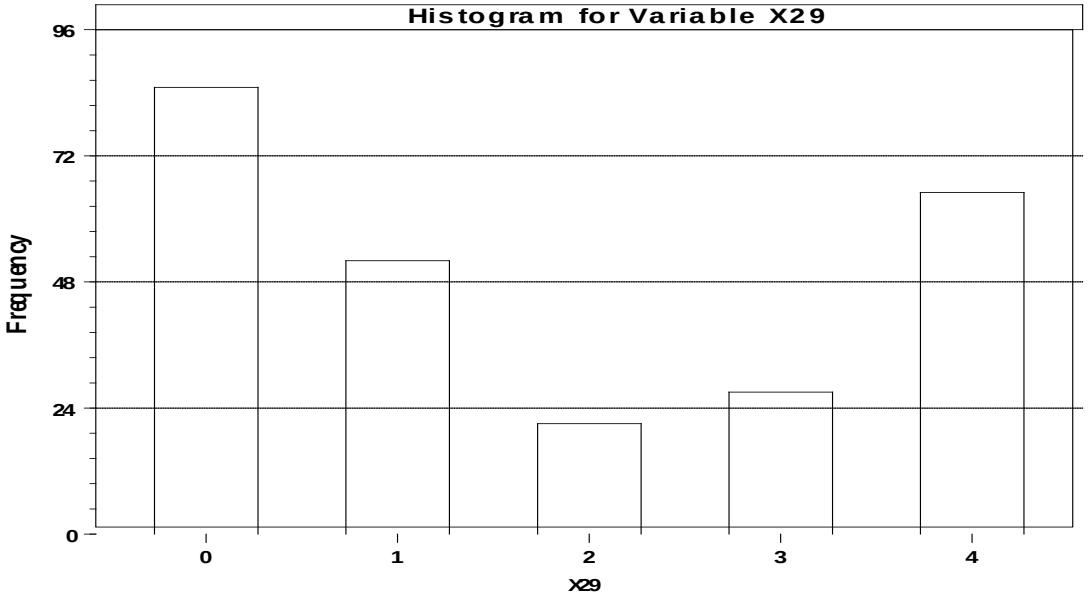
x16	Usually commuted to or from work on Interstate 90 last year: 1 if yes, 0 if no
x17	On your past five commutes to work, how often have you used HOV lanes
x18	On your past five commutes to work, how often did you drive alone
x19	On your past five commutes to work, how often did you carpool with one other person
x20	On your past five commutes to work, how often did you carpool with two or more people
x21	On your past five commutes to work, how often did you take a vanpool
x22	On your past five commutes to work, how often did you take a bus
x23	On your past five commutes to work, how often did you bicycle or walk
x24	On your past five commutes to work, how often did you take a motorcycle
x25	On your past five commutes to work, how often did you take a mode other than those listed in variables 18 through 24
x26	On your past five commutes to work, how often have you changed route or departure time
x27	HOV lanes save all commuters time: 0 if strongly disagree, 1 if disagree, 2 if neutral, 3 if agree, 4 if agree strongly
x28	Existing HOV lanes are being adequately used: 0 if strongly disagree, 1 if disagree, 2 if neutral, 3 if agree, 4 if agree strongly
x29	HOV lanes should be open to all traffic: 0 if strongly disagree, 1 if disagree, 2 if neutral, 3 if agree, 4 if agree strongly
x30	Converting some regular lanes to HOV lanes is a good idea: 0 if strongly disagree, 1 if disagree, 2 if neutral, 3 if agree, 4 if agree strongly
x31	Converting some regular lanes to HOV lanes is a good idea only if it is done before traffic congestion becomes serious: 0 if strongly disagree, 1 if disagree, 2 if neutral, 3 if agree, 4 if agree strongly
x32	Gender: 1 if male, 0 if female
x33	Age in years: 0 if under 21, 1 if 22 to 30, 2 if 31 to 40, 3 if 41 to 50, 4 if 51 to 64, 5 if 65 or greater

x34	Annual household income (US dollars per year): 0 if no income, 1 if 1 to 9,999, 2 if 10,000 to 19,999, 3 if 20,000 to 29,999, 4 if 30,000 to 39,999, 5 if 40,000 to 49,999, 6 if 50,000 to 74,999, 7 if 75,000 to 100,000, 8 if over 100,000
x35	Highest level of education: 0 if did not finish high school, 1 if high school, 2 if community college or trade school, 3 if college/university, 4 if post college graduate degree
x36	Number of household members
x37	Number of adults in household (aged 16 or more)
x38	Number of household members working outside the home
x39	Number of licensed motor vehicles in the household
x40	Postal zip code of work place
x41	Postal zip code of home
x42	Type of survey comment left by respondent regarding opinions on HOV lanes: 0 if no comment on HOV lanes, 1 if comment not in favor of HOV lanes, 2 comment positive toward HOV lanes but critical of HOV lane policies, 3 comment positive toward HOV lanes, 4 neutral HOV lane comment

```
--> RESET
Initializing NLOGIT Version 4.0.1 (January 1, 2007).
--> read;nvar=42;nobs=250;file=D:\old_drive_d\new_laptop\CE697N-disk\SURVEYS-...
--> create;if(x1=0)dalone=1$
--> create;if(x33>3&x32=1)oldmen=1$
--> create;if(x35>2)college=1$
--> histogram;rhs=x28$
```



--> histogram;rhs=x29$



--> skip$

--> oprobit;lhs=x28;rhs=one,dalone,x8,oldmen,college,x37$

```
 ***************************************************************************
 * NOTE: Deleted      3 observations with missing data. N is now    247 *
 ***************************************************************************

Normal exit from iterations. Exit status=0.
+--------------------------------------------+
| Ordered Probability Model                  |
| Maximum Likelihood Estimates               |
| Model estimated: Dec 01, 2014 at 02:12:54PM.|
| Dependent variable                  X28    |
| Weighting variable                 None    |
| Number of observations              247    |
| Iterations completed                 13    |
| Log likelihood function       -331.1284    |
| Number of parameters                   9    |
| Restricted log likelihood     -344.5389    |
| McFadden Pseudo R-squared      .0389230    |
```

```
| Chi squared                    26.82096     |
| Degrees of freedom                    5     |
| Prob[ChiSqd > value] =         .6180822E-04 |
| Underlying probabilities based on Normal    |
+---------------------------------------------+

+---------------------------------------------+
| Ordered Probability Model                   |
| Cell frequencies for outcomes               |
|  Y Count Freq  Y Count Freq  Y Count Freq   |
|  0    67 .271  1   102 .412  2    34 .137   |
|  3    34 .137  4    10 .040                 |
+---------------------------------------------+
+--------+-------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+----------------+--------+--------+----------+
---------+Index function for probability
 Constant|    1.28072277       .26305890     4.869   .0000
 DALONE  |    -.76240151       .16231855    -4.697   .0000    .76923077
 X8      |    -.21665090       .13931260    -1.555   .1199    .48178138
 OLDMEN  |    -.36231352       .20630507    -1.756   .0791    .13765182
 COLLEGE |     .11175834       .17192354      .650   .5157    .78542510
 X37     |     .01042439       .07924773      .132   .8953   2.17004049
---------+Threshold parameters for index
 Mu(1)   |    1.15563280       .08133363    14.209   .0000
 Mu(2)   |    1.62668407       .09294369    17.502   .0000
 Mu(3)   |    2.52064132       .15641923    16.115   .0000

--> matrix;b1=b;mu1=mu$
--> oprobit;lhs=x29;rhs=one,dalone,x8,oldmen,college,x37$

 ************************************************************************
 * NOTE: Deleted      3 observations with missing data. N is now    247 *
 ************************************************************************

Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Ordered Probability Model                   |
| Maximum Likelihood Estimates                |
| Model estimated: Dec 01, 2014 at 02:12:55PM.|
| Dependent variable                    X29   |
| Weighting variable                   None   |
| Number of observations                247   |
| Iterations completed                   14   |
| Log likelihood function       -344.8131     |
| Number of parameters                   9     |
| Restricted log likelihood     -368.7304     |
| McFadden Pseudo R-squared      .0648639     |
| Chi squared                   47.83456     |
| Degrees of freedom                     5     |
| Prob[ChiSqd > value] =         .0000000     |
| Underlying probabilities based on Normal    |
+---------------------------------------------+

+---------------------------------------------+
| Ordered Probability Model                   |
| Cell frequencies for outcomes               |
|  Y Count Freq  Y Count Freq  Y Count Freq   |
|  0    85 .344  1    51 .206  2    21 .085   |
|  3    27 .109  4    63 .255                 |
+---------------------------------------------+
```

```
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
---------+Index function for probability
 Constant|    -.37364383       .28228760    -1.324   .1856
 DALONE  |    1.16029608       .18285356     6.345   .0000    .76923077
 X8      |     .29455571       .14508997     2.030   .0423    .48178138
 OLDMEN  |     .25666162       .20645987     1.243   .2138    .13765182
 COLLEGE |     .02815170       .18150798      .155   .8767    .78542510
 X37     |    -.12921433       .08135103    -1.588   .1122   2.17004049
---------+Threshold parameters for index
 Mu(1)   |     .60337156       .06927361     8.710   .0000
 Mu(2)   |     .85165310       .07679550    11.090   .0000
 Mu(3)   |    1.19603541       .08992381    13.301   .0000
```

```
--> matrix;b2=b;mu2=mu$
--> oprobit;lhs=x28,x29
    ;rh1=one,dalone,x8,oldmen,college,x37
    ;rh2=one,dalone,x8,oldmen,college,x37
    ;start=b1,mu1,b2,mu2,0$

 ************************************************************************
 * NOTE: Deleted      3 observations with missing data. N is now    247 *
 ************************************************************************


Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Bivariate Ordered Probit Model              |
| Maximum Likelihood Estimates                |
| Model estimated: Dec 01, 2014 at 02:13:39PM.|
| Dependent variable             BivOrdPr     |
| Weighting variable                 None     |
| Number of observations              247     |
| Iterations completed                 25     |
| Log likelihood function       -629.4278     |
| Number of parameters                 19     |
| Restricted log likelihood     -671.7713     |
| McFadden Pseudo R-squared      .0630326     |
| Chi squared                    84.68703     |
| Degrees of freedom                   19     |
| Prob[ChiSqd > value] =         .0000000     |
+---------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
---------+   Index function for Probability Model for X28
 Constant|    1.27326613       .29175468     4.364   .0000
 DALONE  |    -.79058351       .17950706    -4.404   .0000    .76923077
 X8      |    -.19733572       .14245869    -1.385   .1660    .48178138
 OLDMEN  |    -.36715232       .23789879    -1.543   .1228    .13765182
 COLLEGE |     .10702937       .17714690      .604   .5457    .78542510
 X37     |     .01752451       .08375187      .209   .8343   2.17004049
---------+   Index function for Probability Model for X29
 Constant|    -.33771361       .28204837    -1.197   .2312
 DALONE  |    1.15031327       .18591310     6.187   .0000    .76923077
 X8      |     .25726540       .14620114     1.760   .0785    .48178138
 OLDMEN  |     .27214882       .20851183     1.305   .1918    .13765182
 COLLEGE |     .01503028       .17334284      .087   .9309    .78542510
 X37     |    -.12881210       .07899864    -1.631   .1030   2.17004049
---------+Threshold Parameters for Probability Model for X28
 MU(01)  |    1.15206973       .09693405    11.885   .0000
 MU(02)  |    1.63580197       .11324054    14.445   .0000
```

```
 MU(03)  |    2.58743785       .21769260    11.886   .0000
---------+Threshold Parameters for Probability Model for X29
 LMDA(01)|     .57780387       .07502241     7.702   .0000
 LMDA(02)|     .82822812       .09014013     9.188   .0000
 LMDA(03)|    1.18462838       .10134477    11.689   .0000
---------+Disturbance Correlation = RHO(1,2)
 RHO(1,2)|    -.31281826       .02590515   -12.076   .0000
```

```
+----------------------------------------------------------------------+
|Cross Tabulation                                                      |
|Row variable is X28      (Out of range 0-49:      0)                  |
|Number of Rows =  5      (X28      =  0 to  4)                        |
|Col variable is X29      (Out of range 0-49:      0)                  |
|Number of Cols =  5      (X29      =  0 to  4)                        |
|Chi-squared independence tests:                                       |
|Chi-squared[   0] =     .00000   Prob value =  .00000                 |
|G-squared  [   0] =     .00000   Prob value =  .00000                 |
+----------------------------------------------------------------------+
|Joint Frequencies for Row Variable X28       Column Variable X29      |
+--------+--------+-----------------------------------------------------+
|X28     | Total  |     0      1      2      3      4                   |
+--------+--------+-----------------------------------------------------+
|       0|    67  |     7      9      4      7     40                   |
|       1|   102  |    27     24     10     18     23                   |
|       2|    34  |    18      8      7      1      0                   |
|       3|    34  |    24      9      0      1      0                   |
|       4|    10  |     9      1      0      0      0                   |
+--------+--------+-----------------------------------------------------+
|   Total|   247  |    85     51     21     27     63                   |
+--------+--------+-----------------------------------------------------+
```