Statistical and Econometric Methods for Transportation Data Analysis

Chapter 14 – Ordered Probability Models

Example 14.1
Ordered Discrete Data – Ordered Probit

A survey of 322 commuters was in the Seattle metropolitan area. 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 "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 and the observed percentage frequency of response in these 5 categories was 32.74, 21.71, 8.54, 12.10, and 24.91 respectively. To understand the factors determining commuter opinions, an ordered probit model of this survey question is appropriate.

Your task is to estimate a model of the ordered response of whether commuters 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):

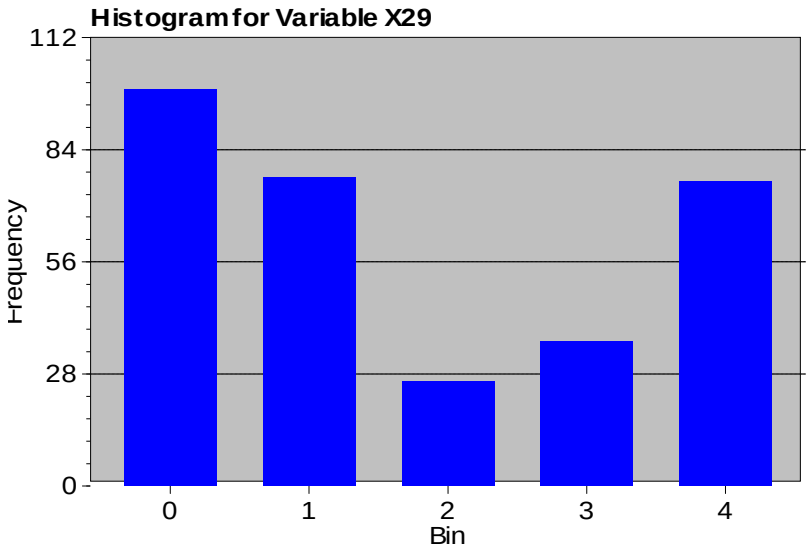
Variable Number	Explanation
x1	Usual mode of travel: 0 if drive alone, 1 if two person carpool, 2 if three or more person carpool, 3 if vanpool, 4 if bus, 5 if bicycle or walk, 6 if motorcycle, 7 if other
x2	Have used HOV lanes: 1 if yes, 0 if no
x3	If used HOV lanes, what mode is most often used: 0 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

```
--> read;nvar=42;nobs=322;file=D:Ex14-1.txt$
--> create;if(x1=0)dalone=1$
--> create;if(x33>3&x32=1)oldmen=1$
--> create;if(x35>2)college=1$
--> histogram;rhs=x29$

Histogram for X29          NOBS=   314, Too low:     0, Too high:     0
Bin  Lower limit   Upper limit       Frequency     Cumulative Frequency
================================================================================
  0         .000         1.000         99 ( .3153)          99( .3153)
  1        1.000         2.000         77 ( .2452)         176( .5605)
  2        2.000         3.000         26 ( .0828)         202( .6433)
  3        3.000         4.000         36 ( .1146)         238( .7580)
  4        4.000         5.000         76 ( .2420)         314(1.0000)
```




```
--> skip$
--> ordered;lhs=x29;rhs=one,dalone,x8,oldmen,college,x36;marginal effects$
```

```
+---------------------------------------------------------------------+
| Dependent variable is binary, y=0 or y not equal 0                  |
| Ordinary    least squares regression    Weighting variable = none   |
| Dep. var. = Y=0/Not0 Mean=   .6738351254    , S.D.=   .4696508592   |
| Model size: Observations =     279, Parameters =   6, Deg.Fr.=    273 |
| Residuals:  Sum of squares= .9972582530D+03, Std.Dev.=        1.91127 |
| Fit:        R-squared=*********, Adjusted R-squared =       -15.56131 |
| Diagnostic: Log-L =   -573.5787, Restricted(b=0) Log-L =    -184.5243 |
|             LogAmemiyaPrCrt.=    1.317, Akaike Info. Crt.=      4.155 |
+---------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  .1296715074      .44220253         .293    .7693
 DALONE    .4388033502      .27449960        1.599    .1099   .77060932
 X8        .4781738604E-01  .23267286         .206    .8372   .48028674
 OLDMEN    .6020054050E-01  .35308446         .170    .8646   .12903226
 COLLEGE   .9724386887E-01  .28975461         .336    .7372   .79211470
 X36       .3343083528E-01  .97510448E-01     .343    .7317  2.9390681

Normal exit from iterations. Exit status=0.
```

```
+---------------------------------------------+
| Ordered Probit Model                        |
| Maximum Likelihood Estimates                |
| Dependent variable                  X29     |
| Weighting variable                  ONE     |
| Number of observations              279     |
| Iterations completed                 14     |
| Log likelihood function       -397.2770     |
| Restricted log likelihood     -421.3950     |
| Chi-squared                    48.23599     |
| Degrees of freedom                    5     |
| Significance level             .0000000     |
|    Cell frequencies for outcomes            |
|  Y Count Freq  Y Count Freq  Y Count Freq    |
|  0    91 .326  1    60 .215  2    24 .086   |
|  3    34 .121  4    70 .250                 |
+---------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
	Index function for probability				
Constant	-.5807798304	.27813128	-2.088	.0368	
DALONE	1.136565726	.16430272	6.918	.0000	.77060932
X8	.2301353655	.13561406	1.697	.0897	.48028674
OLDMEN	.1968407034	.20635114	.954	.3401	.12903226
COLLEGE	.1996976747E-01	.15582658	.128	.8980	.79211470
X36	.1178065062E-01	.60437908E-01	.195	.8455	2.9390681
	Threshold parameters for index				
Mu(1)	.6231650207	.73062591E-01	8.529	.0000	
Mu(2)	.8657954320	.83104656E-01	10.418	.0000	
Mu(3)	1.240495241	.95160650E-01	13.036	.0000	

Marginal Effects for OrdProbt

Variable	X29=0	X29=1	X29=2	X29=3
ONE	.2063	.0230	-.0142	-.0415
DALONE	-.4038	-.0451	.0278	.0812
X8	-.0818	-.0091	.0056	.0164
OLDMEN	-.0699	-.0078	.0048	.0141
COLLEGE	-.0071	-.0008	.0005	.0014
X36	-.0042	-.0005	.0003	.0008

Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.

	Predicted					
Actual	0	1	2	3	4	Total
0	59	0	0	0	32	91
1	23	0	0	0	37	60
2	9	0	0	0	15	24
3	9	0	0	0	25	34
4	25	0	0	0	45	70
Total	125	0	0	0	154	279