PROCEDURE FOR ORDERED PROBIT & LOGIT MODELS: INSTRUCTIONS


Instructions

Copy the execution file, ORDER.SHA, the data file, ORDER.DAT, and the procedure (read only) file, ORDER.PRO, into your working directory. Then, open the execution file, ORDER.SHA, using Shazam.

Provide information about your model and data by changing names and numbers in ORDER.SHA as explained below:

1. PAR xxxx: This command is used to increase the memory size that can be used by Shazam. Usually, you will not need this command, but will need to increase this value (from the default value) when prompted by an error message. For example, a data set with 30,000 observations on four variables would need a PAR value at least 10000. The maximum value that you can increase PAR to will depend on your computer’s memory capacity. (If you have to change the PAR value, remove the asterisk (*) in front of the command after changing the value.)

2. Specify your working directory, where the procedure and data files are located, after “WD:” as illustrated below.

3. Provide the sample range (1 to 921 for the current example) after “SAM:” as shown below.

4. Provide the data file name (within the parentheses) and the names of the variables in the order of the columns in your data set at the READ command. The “SKIPLINES=” option is used to skip lines with labels on the top of the data file. In the current data file, only first line is used for labels. So, “SKIPLINES=1” is used.

5. State whether you want to estimate ordered probit or ordered logit.

6. It is optional, but you can obtain descriptive statistics of variables by providing their names after the STAT command.

7. Provide the name of the discrete choice variable (usually denoted Y). Here, you read that variable as “rating” in the READ command. So, type “rating” after “GENR Y=” command. It is assumed that the choice values are integers starting from 0 or 1 and increasing by one. In the current data set, rating=1,2,3,4,5,6, or 7.

8. The value you provide at “JLOW=” is the value of Y for the lowest choice (either 0 or 1) in your data file.

9. The number of choices is indicated at “JNUM=”.  

10. The number of the explanatory variables in the index equation excluding intercept (K-1) should be indicated at “K_1=”. In the present example, there are five explanatory variables used in the index equation. The maximum number is 20 for the attached procedure file.

10. Provide the names of the explanatory variables at “X1=”, “X2=”, and “X3=” etc. Add more lines like “GENR X6=name” if there are more explanatory variables.

11. Do not change any content in the part between the two lines “DO NOT CHANGE ANY PART BELOW THIS LINE” and “DO NOT CHANGE ANY PART ABOVE THIS LINE”.

12. If you want to examine the predicted probabilities, the predicted choice variable (Y), and/or predicted index value (IHAT) for some observations, add the PRINT command at the end of the execution file (just above the STOP command). The following command will show the predicted probabilities for each of the seven choices, predicted choice, and predicted index for the first 10 observations. (To see only a subset of the sample, change “[SAM]” at the SAMPLE command with the range of your choice (as shown below).

    SAMPLE 1 10
    PRINT P1HAT P2HAT P3HAT P4HAT YHAT
The Shazam execution file (ORDER.SHA)

*************
ORDERED CHOICE MODELS
*************

* This file and the accompanying procedure file, ORDER.PRO, have been created by
* Daehoon Nahm
* Department of Economics
* Macquarie University
* Australia
*
**
* Model
**

* Ystar = b0+b1*X1+b2*X2... + u
* Y = 1 if Ystar <= 0
* Y = 2 if 0 < Ystar <= r1
* Y = 3 if r1 < Ystar <= r2
* Y = 4 if r2 < Ystar etc.

* If u ~ N(0,1) --> ordered probit model
* If u ~ logistic distribution --> ordered logit model

?=SET NOECHO
*
* Provide necessary information by changing the words and numbers in the lines enclosed by
* "///.........\\" and "\\.........///".
*?=PAR 10000
*=?SET NODOECHO
*=?SET NOOUTPUT
*=?SET NOWARN
*
* Provide below the working Directory where the PROCEDURE file, ORDERP.PRO, is saved:
* (Example: "C:\works", without the quotation marks)
*
**
FILE PATH [WD]
FILE PROC [WD]ORDER.PRO
*
* Read the Data
**

* Provide the sample range.
**
SAMPLE [SAM]
*
* Provide name of the data file & a list of the variables in the data file in order of appearance.
**
READ(order.dat) booklev ebit invgrade logsales marklev rating reta wka /skiplines=1
*
* Indicate whether "probit" or "logit".
**
MODEL:logit
*
* Provide a list of the variables for descriptive statistics.
**
STAT rating invgrade booklev wka reta ebit logsales
*
* The name of the discrete ordered choice variable (1,2,...,M)
**
GENR Y=rating
*
* The lowest value of the ordered choice variable (usually 0 or 1)
**
GEN1 JLOW=1
*
* The number of choices (minimum=3, maximum=7)
ORDERED CHOICE MODELS

The Number of Explanatory Variables (excluding the intercept, maximum=20)

Provide the names of the explanatory variables - one in each row after "=".

Set echo

If needed, predicted probabilities (P1HAT, P2HAT, etc.) and predicted Y (YHAT)
can be printed by adding the print command, for example,

PRINT P1HAT P2HAT P3HAT P4HAT YHAT

at the bottom of this file.

If needed, add print command below, just above the STOP command.
The predicted variables that can be printed below are:
- Predicted probabilities: Y1HAT, Y2HAT, Y3HAT, ...
- Predicted choice: YHAT (1, 2, 3, ...)
- Predicted index value: IHAT

Sample

Sample

Print P1HAT P2HAT P3HAT P4HAT P5HAT P6HAT P7HAT
Print YHAT IHAT

STOP
OUTPUT (part)

<table>
<thead>
<tr>
<th>INDEX FUNCTION COEFFICIENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>COEFF</td>
</tr>
<tr>
<td>B0</td>
</tr>
<tr>
<td>B1</td>
</tr>
<tr>
<td>B2</td>
</tr>
<tr>
<td>B3</td>
</tr>
<tr>
<td>B4</td>
</tr>
<tr>
<td>B5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>THRESHOLD PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAMMA1</td>
</tr>
<tr>
<td>GAMMA2</td>
</tr>
<tr>
<td>GAMMA3</td>
</tr>
<tr>
<td>GAMMA4</td>
</tr>
<tr>
<td>GAMMA5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LOG-LIKELIHOOD FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOGL</td>
</tr>
<tr>
<td>LOGL1</td>
</tr>
<tr>
<td>LOGL0</td>
</tr>
<tr>
<td>MCFAD_R2</td>
</tr>
</tbody>
</table>

READ USES FORMAT:(A15)
READ USES FORMAT:(5X,A10)

<table>
<thead>
<tr>
<th>PREDICTION SUCCESS TABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>Predict</td>
</tr>
<tr>
<td>Y2</td>
</tr>
<tr>
<td>Y3</td>
</tr>
<tr>
<td>Y4</td>
</tr>
<tr>
<td>Y5</td>
</tr>
<tr>
<td>Y6</td>
</tr>
<tr>
<td>Y7</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

NUMBER OF RIGHT PREDICTIONS = 478.0000

PERCENTAGE OF RIGHT PREDICTIONS = 0.5190011