Discrete Choice Modeling
William Greene
Stern School of Business, New York University

Lab Session 2 Assignment
Binary Choice Modeling

This exercise will involve estimating and analyzing binary choice models. We will analyze the panel probit, manufacturing innovation data. The data set is PanelProbit.lpj. To begin, load these data. To save you some typing, most of the commands for this exercise are contained in the file LabAssignment-2.lim.

1. **Cluster Estimator.** This is a panel data set. Do the standard errors of the probit estimator need 'correction?' This exercise computes the standard covariance matrix and the 'cluster corrected' covariance matrix and compares them. Describe your findings.

```
Sample ; All $
Namelist ; X = One,IMUM,FDIUM,SP,LogSALES $
Probit ; Lhs = IP ; Rhs = X $
Matrix ; Var0 = Varb $  (Uncorrected covariance matrix)
Probit ; Lhs = IP ; Rhs = X ; Cluster = 5 $
Matrix ; VarPanel = Varb $ (Corrected covariance matrix)
? PCTDIFF is the percentage difference between the standard errors
Matrix ; SD0 = Diag(Var0)  ; Diff = Vecd(VarPanel) - Vecd(Var0)
       ; List ; PctDiff = 100*<SD0>*Diff$
```

2. **Robust Covariance Matrix**. You can also compute a 'robust,' sandwich style asymptotic covariance matrix. This estimator would only be robust to heteroscedasticity – though we are unsure what that would mean in the probit setting.

```
Probit ; Lhs = IP ; Rhs = X $
Matrix ; Var0 = Varb $  (Uncorrected covariance matrix)
Probit  ; Lhs = IP  ; Rhs = X ;  RobustVC $
Matrix  ; VarHet = Varb $
Matrix ; SD0 = Diag(Var0)  ; Diff = Vecd(VarHet) - Vecd(Var0)
       ; List ; PctDiff = 100*<SD0>*Diff$ - Init(5,1,100) $
```

3. **Marginal Effect for a Quadratic.** Marginal effects in the binary choice models are complicated functions of the parameters and the data. They are more so when the index function contains complex functions of the data. Suppose, for example,

$$P = \Phi(\boldsymbol{\beta}'\mathbf{x} + \alpha_0 \text{logSales} + \alpha_1 \text{logSales}^2).$$

The marginal effect of logSales, which is the effect on the probability of a one percent change in sales is

$$\partial P/\partial \text{logSales} = \phi(\boldsymbol{\beta}'\mathbf{x} + \alpha_0 \text{logSales} + \alpha_1 \text{logSales}^2) \times (\alpha_0 + 2\alpha_1 \text{logSales})$$

It is possible to program this computation into the WALD command. But, it is easier to use the built in function to obtain the result.

```
Create   ; LogS2  = LogSales^2 $
Namelist ; X2     = One,IMUM,FDIUM,SP,LogSALES,LogS2 $
Probit   ; Lhs    = IP ; Rhs = X2 ; Mar $
Wald     ; Start  = b ; Var = Varb
         ; Labels = beta0,beta1,beta2,beta3,a0,a1
         ; Fn1    = n01(beta0'X2)*(a0+2*a1*Logsales) ? partial wrt logsales
         ; Fn2    = n01(beta0'X2)* beta2  ? partial wrt fdium
         ; Average (computes average partial effects) $
? Easier way
probit   ; Lhs = ip ; rhs = One,IMUM,FDIUM,SP,logsales,logsales^2$
partial  ; effects: logsales / fdium $
? Extension to look more closely at partial effects
Partial  ; effects: logsales & fdium = .05(.05)1 ; plot(ci) $
```

4. **Heteroscedasticity**. The following suggests how to incorporate heteroscedasticity in the binary logit (or probit – by changing the command) model:

Logit ; Lhs = IP ; Rhs = X ; Het ; Hfn = RAWMTL; Marginal Effects $

(1) Note the effect on the coefficients and how the marginal effects are decomposed.
(2) Repeat the computation with **;Hfn = LogSales**. Note the effect on the estimates and significance levels. The difference between the reported marginal effects and the results from the PARTIALS command is that the former is computed at the means of the data while the second is averaged over all observations. To reproduce the results at the means, we add ;Means to the PARTIALS command.

```
Namelist ; X = one,imum,fdium,sp,logsales $
Logit   ; Lhs = IP ; Rhs = X ; Het
        ; Hfn = RAWMTL; Marginal Effects $
Partial ; Function= lgp((b1+b2*IMUM+b3*FDIUM+b4*SP+b5*LogSALES)/exp(c1*rawmtl))
        ; Labels  = b1,b2,b3,b4,b5,c1
        ; parameters = b
        ; covariance = varb
        ; effects: rawmtl / logsales $
Logit   ; Lhs = IP ; Rhs = X ; Het
        ; Hfn = LogSales; Marginal Effects $
Partial ; Function=lgp((b1+b2*IMUM+b3*FDIUM+b4*SP+b5*LogSALES)
                     /exp(c1*LogSales))
        ; Labels   = b1,b2,b3,b4,b5,c1
        ; parameters = b
        ; covariance = varb
        ; effects: logsales ; means$
```

5. **Nonparametric and Semiparametric Estimation**. There are numerous alternative estimators you can use for analyzing binary choices. Interpretation of the results of these models requires some careful thought – but estimation is very straightforward. Estimation of these is generally very computer intensive, so we use only a subset of the sample – one year of the data. Compare the table of correct and incorrect predictions produced by **PROBIT** and **MSCORE**. (The other estimators do not produce enough information to generate predictions for individual observations.)

```
Namelist ; X0 = IMUM,FDIUM,SP,LogSALES $
Namelist ; X = One,X0 $
Reject ; New ; T > 1 $ (Use only first year of data)
?  Fully Parametric
PROBIT ; Lhs = IP ; Rhs = X $
?  Semiparametric:  Maximum Score
MSCORE ; Lhs = IP ; Rhs = X $
```

```
Semiparametric ; LHS = IP ; Rhs = X0 $ (Klein and Spady.)
?  Nonparametric, Kernel density regression estimator
?  Note, the nonparametric estimator can only have one RHS variable
NPREG ; LHS = IP ; Rhs = LogSales $
```

6. **Creating a Plot of Probabilities**. The following will demonstrate how to use NLOGIT to produce the plot shown in the class discussion.

```
Reject     ; New ; T > 1 $
Probit     ; Lhs = IP ; Rhs = one,IMUM,FDIUM,SP,logsales $
Calc       ; Low = .5*Min(LogSales) ; High = 1.5*Max(LogSales)
           ; inc = .05*(high-low) $$
partials ; simulate ; effects: logsales & logsales = Low(inc)high
     ;plot(ci)  ;title=Simulation of Innovation Probabilities vs. Log Sales$
```

7. **Testing for Structural Change**. It might be interesting to test whether underlying structural of the model has changed over the five year period of the data. Consider the structure

$$P_{it} = F(\boldsymbol{\beta}_t'\mathbf{x}_{it}),\ i = 1,\ldots,1270,\ t = 1,\ldots,5 \text{ (1993 to 1997)}$$

which allows for different coefficient vectors in each year. We are interested in testing the hypothesis

$$H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta}_3 = \boldsymbol{\beta}_4 = \boldsymbol{\beta}_5$$
$$H_1: \text{ not } H_0.$$

In a linear regression context, this would be a 'Chow' test and would be tested with an F test. Since this is not a linear regression model, we can't use the F test here. The easiest way to do this test is with a likelihood ratio test. The strategy is to fit the restricted model (pool the 5 years of data) and the unrestricted model (estimate the model separately for each year), and compare the log likelihoods. The log likelihood for the unrestricted model is the sum of the five years. Here is how you can automate this computation. The last part of the last CALC displays the 95% critical value from the chi-squared table.

```
Sample ; All $
Namelist ; X = One,IMUM,FDIUM,SP,LogSALES $
Probit ; Lhs = IP ; Rhs = X ; quietly $  (We suppress the model results)
Calc ; Logl0 = Logl ; Logl1 = 0 ; i = 0 $
Procedure
Include ; New ; T = i $
Probit ; Lhs = IP ; Rhs = X $
Calc ; Logl1 = Logl1 + Logl $
EndProc $
Execute ; i = 1,5 ; Silent $ (This suppresses the individual year results.)
Calc ; List ; Chisq = 2*(Logl1 - Logl0) ; Df = 4*Col(X) ; Ctb(.95,df) $
```

Carry out the test. What do you conclude? Should the null hypothesis be rejected? Repeat the test using a logit model instead of a probit model. Does the conclusion change? Try the exercise again while adding the sector dummy variables to the model. To do these, it is only necessary to change the model name from PROBIT to LOGIT, or the NAMELIST command by adding variables to it.

8. **Hypothesis Tests:** This exercise will illustrate the three methods of carrying out hypothesis tests.

```
Reject   ; New ; T < 5 $
Namelist ; X = One,IMUM,FDIUM,LogSales $
Namelist ; Sectors = RawMtl,InvGood$
Probit   ; Lhs = IP ; Rhs = X $
Calc     ; Logl0 = LogL $
Probit   ; Lhs = IP ; Rhs = X ; Het ; Hfn = Sectors
         ; Start = b,0,0 ; Maxit = 0 $
Probit   ; Lhs = IP ; Rhs = X,Sectors ; Parameters ; test: sectors = 0$
Calc     ; KX = Col(X) ; K1 = KX + 1 ; Kc = Col(Sectors); K = KX + KC$
Matrix   ; c = B(K1:K) ; vc = Varb(K1:K , K1:K) $
Matrix   ; List ; Wald = c'<vc>c  $
Calc     ; List ; Ctb(.95,2) $
Wald     ; start = b ; Var = Varb ; labels=KX_d,Kc_c
              ; fn1 = c1 - 0 ; Fn2 = c2 - 0 $
```

The model is $y_i^* = \boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i$

$$\varepsilon \sim N[0,\sigma_i^2],\ \sigma_i = \exp(\boldsymbol{\gamma}'\mathbf{z}_i).$$

$$y_i = 1(y_i^* > 0]$$

The various testing procedures shown estimate γ and test whether $\boldsymbol{\gamma} = 0$, in which case $\sigma_i^2 = 1$. Carry out the tests, and determine whether the null hypothesis, H_0:$\boldsymbol{\gamma} = 0$, should be rejected.

10. **Simulation:** Using the binary choice model simulator, examine how an increase in FDIUM of 50% would affect the probability of innovation.

```
Probit        ; Lhs = IP ; Rhs=one,logsales,imum,fdium $
BinaryChoice  ; Lhs = IP ; Rhs = one,logsales,imum,fdium
                           ; model=probit  ; start=b
                           ; scenario: fdium * = 1.5 ;plot:fdium $
```