Chapter 13

Limited Dependent Variables

Section 13.1

Limited Dependent Variables

Conventional regression methods require that the dependent variable be observed on a continuous scale. But there are instances wherein the dependent variable is limited in scope:

(1) Qualitative (discrete)

(2) Censored or truncated

(3) Integer valued (count data)

References:

Greene (2008)

Maddala (1983)

<u>Models wherein the dependent variables correspond to choices</u>

- **Probit Models**
- **Logit Models**

<u>Censored Response Models</u> (Dependent variables are discontinuous)

- **Tobit Models**
- **Heckman Sample Selection Procedure**

<u>Count Data Models</u> (Dependent Variables are integers)

Section 13.2

Probit/Logit Models

Binary Choice Models

Dependent variable takes on two values

Often the dependent variable represents the occurrence of an event, or a choice between two alternatives.

Example: The dependent variable Y corresponds to employment status. Individuals in the sample are either employed or not. The individuals differ in age (X1) educational attainment (X2), race (X3), marital status (X4), and perhaps other observable characteristics (Z).

$$Y_i = a_0 + a_1 X1_i + a_2 X2_i + a_3 X3_i + a_4 X4_i + a_5 Z_i + \varepsilon_i$$

Dichotomous Choices:

- **Linear Probability Model**
- **Probit Model**
- **Logit Model**

Modeling Binary Choices:

(1) Participate or do not participate in a government food assistance program

(2) Buy or do not buy a food or beverage product (e.g. organic milk)

(3) Vote yes or no in elections conditional on voting

(4) Report or fail to report income

(5) Employed or Unemployed

Objective: conduct a profile of individuals or households who make one choice or the alternative

Commonalities:

(1) Use of cross-sectional data

(2) seek probabilities conditional on explanatory variables

(3) determine how changes in explanatory variables affect probabilities (marginal effects)

Linear Probability Model

$$Y_i = X_i^T \beta + \varepsilon_i$$

OLS yields consistent and unbiased estimates of β

Deficiencies:

- heteroscedasticity of disturbance terms

- distribution of disturbance terms are non-normal

- allows $\hat{y}_i$ to fall outside the interval of 0 to 1,

Use of monotonic transformations to guarantee predictions lie in the unit interval

Mechanics of Monotonic Transformations

Let $Z_i = X_i^T\beta$ and let Z^* be a random variable with probability density function f.

$$y = 1 \text{ if } Z_i \geq Z_i^* \text{ or } 0 \text{ if } Z_i < Z_i^*$$

$$P(y_i = 1 | Z_i) = P(Z_i^* \leq Z_i) = F(Z_i)$$

$$P(y_i = 0 | Z_i) = P(Z_i^* > Z_i) = 1 - F(Z_i)$$

Probit Model

$$P_i(y_i = 1) = F(z_i) = \int_{-\infty}^{z_i} (2\Pi)^{-\frac{1}{2}} EXP\left(-\frac{s^2}{2}\right) ds,$$

$$-\infty < Z_i < \infty$$

$$\text{Marginal Effects : } \frac{\partial P_i}{\partial X_i} = \left(\frac{\partial F}{\partial Z_i}\right)\left(\frac{\partial Z_i}{\partial X_i}\right) = f(Z_i)\beta$$

$\mathrm{F}(\mathrm{Z_i}) \rightarrow$ Standard Normal Cumulative Distribution Function

$\mathrm{f}(\mathrm{Z_i}) \rightarrow$ Standard Normal Density Function

Long history in biometrics-Finney

$$\mathrm{f}(\mathrm{Z_i}) = \left(\frac{1}{(2\Pi)^{\frac{1}{2}}}\right) \mathrm{EXP}\left(\frac{-\mathrm{Z_i^2}}{2}\right), -\infty < \mathrm{Z_i} < \infty$$

Logit Model

Early work of Berkson

$$P_i(y_i = 1) = F(Z_i) = \frac{e^{Z}i}{(1+e^{Z}i)}, -\infty < Z_i < \infty$$

$$\text{Marginal Effects}: \frac{\partial P_i}{\partial X_i} = f(Z_i)\beta$$

$\mathrm{F}(\mathrm{Z_i}) \rightarrow$ Logistic Cumulative Distribution Function

$\mathrm{f}(\mathrm{Z_i}) \rightarrow$ Logistic Density Function

$$\mathrm{f}(\mathrm{Z_i}) = \frac{\mathrm{e^{Z}i}}{(1+\mathrm{e^{Z}i})^2}, -\infty < \mathrm{Z_i} < \infty$$

Alternatively,

$$Z_i = \log\left(\frac{P_i}{(1-P_i)}\right) = X_i^T\beta$$ The $\log\left(\frac{P_i}{(1-P_i)}\right)$ is called the logit.

Section 13.3

Computational Methods and Statistical Considerations for Empirical Analysis

Computational Methods and Statistical Considerations for Empirical Analysis

Computational Methods

Most common characteristics of qualitative and censored response models is that parameter estimation is usually carried out via some maximum-likelihood algorithm. The likelihood functions are often times the product of a series of density and distribution functions. The objective is to find the estimator $\hat{\beta}$ that maximizes the likelihood of observing the pattern of choices in the sample. An important feature of the maximum likelihood approach is the reliance on individual rather than grouped observations.

Maximum likelihood estimation procedure assures the large-sample properties of consistency and asymptotic normality of the parameter vector β so that conventional tests of significance are applicable.

Estimation of Binary Choice Models

Likelihood Function

$$L = P(y_1, \ldots, y_n | Z) = P(y_1 | Z_1) \ldots P(y_n | Z_n)$$

$$L = F(Z_1) \ldots F(Zn_1)(1 - F(Zn_1 + 1)) \ldots (1 - F(Z_n))$$

$$\log L = \sum_{i=1}^{n_1} \log P_i + \sum_{i=n_1+1}^{n} \log(1 - P_i)$$

-P_i is either the standard normal cumulative distribution function or the logistic cumulative distribution function.

- To obtain the estimator of $\beta, (\hat{\beta}_{ML})$ differentiate *log L* with respect to β, set the result to zero, and solve the system of normal equations.

- Iterative Methods: the Quasi-Newton Method and the Newton-Raphson Method

-To obtain the estimator of asymoptotic variance-covariance matrix of find the second-order derivative of log L with respect to β, find $(\hat{\beta}_{ML})$, the expectation of this expression, and evaluate the expectation at $\beta = (\hat{\beta}_{ML})$.

- The PROC QLIM procedure uses maximum likelihood methods. Initial starting values for the nonlinear optimizations typically are based on OLS estimates.

Goodness-of-Fit Measures for Binary Choice Models

McFadden (1974) suggested a likelihood ratio index that is analogous to the R^2 in the linear regression model:

$$R_M^2 = 1 - \frac{\ln L}{\ln L_0} \qquad \text{(most popular)}$$

where L is the maximum value of the likelihood function and L_0 is the value of the likelihood function when all regression coefficients except the intercept term are zero.

Estrella's (1998) measure:

$$R_{E1}^2 = 1 - \left(\frac{\ln L}{\ln L_0} \right)^{-\frac{2}{N}\ln L_0}$$

An alternative measure suggested by Estrella (1998) is:

$$R_{E2}^2 = 1 - \left[\left(\ln L - k \right) / \ln L_0 \right]^{-\frac{2}{N} \ln L_0}$$

where *N* is the number of observations used, and *k* represents the number of estimated parameters.

Other goodness-of-fit measures are summarized as follows:

$$R^2_{CU1} = 1 - \left(\frac{L_0}{L}\right)^{\frac{2}{N}} \quad \text{(Cragg} - \text{Uhler1)}$$

$$R^2_{CU2} = \frac{1 - (L_0 / L)^{\frac{2}{N}}}{1 - L_0^{\frac{2}{N}}} \quad \text{(Cragg} - \text{Uhler2)}$$

$$R^2_{A} = \frac{2(\ln L - \ln L_0)}{2(\ln L - \ln L_0) + N} \quad \text{(Aldrich} - \text{Nelson)}$$

$$R_{VZ}^2 = R_A^2 \frac{2\ln L_0 - N}{2\ln L_0} \text{ (Veall – Zimmermann)}$$

$$R_{MZ}^2 = \frac{\sum_{i=1}^{N} (\hat{y}_i - \bar{\hat{y}}_i)^2}{N + \sum_{i=1}^{N} (\hat{y}_i - \bar{\hat{y}}_i)^2} \text{ (McKelvey – Zavoina)}$$

where $\hat{y}_i = x_i^{'} \overline{\beta}$ and $\bar{\hat{y}}_i = \sum_{i=1}^{N} \hat{y}_i / N$.

Correct Classification of Decision-Makers

If the estimated probability if greater than .5 and the first alternative is selected, the decision is correctly classified; if the estimated probability is less than .5 and the second alternative is selected, the decision is correctly classified; we seek maximum proportion of correct classifications of outcomes.

But in many cases, the appropriate cutoff may not be 0.5

$$\text{appropriate cutoff} = \frac{\text{\# of observations which conform to } Y = 1}{\text{total number of observations}}$$

See Park and Capps (1997), Briggeman (2002), Alviola and Capps (2010)

Expectation-Prediction Table or Prediction-Success Table

PREDICTED	ACTUAL 0	ACTUAL 1
0	a	b
1	c	d

Number of right predictions = a + d

Percentage of right predictions $= \frac{(a+d)}{a+b+c+d} x100$

The fraction of y=1 observations that are correctly predicted is termed the sensitivity $\left(\frac{d}{b+d}\right)$

The fraction of y=0 observations that are correctly predicted is known as the specificity $\left(\frac{a}{a+c}\right)$

Section 13.4

SAMPLE PROBLEM: Use of Probit Analysis

SAMPLE PROBLEM: Use of Probit Analysis

$$P(\text{YESVM} = 1) = f(\text{PUB12, PUB34, PUB5, PRIV, YEARS, SCHOOL, LINC, PTCON})$$

Key Products:

(1) Goodness-of-fit measures

(2) Test of goodness-of-fit

(3) Parameter estimates/statistical significance (*β*)

(4) *z* values, linear combination of parameter estimates with individual observations (*xbeta*)

(5) Marginal effects f(z)*β.*

(6) Probability of alternatives conditional on right-hand side variables, either *F(z) or 1-F(z)*

(7) Inverse Mills ratio *f(z)/F(z)* or *f(z)/1-F(z)*

```
* Qualitative Choice Model Example;
* To vote yes or no dealing with the use of bonds to
 fund public schools in a local district;
* use of the Logit Model;
* Use of the Probit Model;
* 95 observations;
* Variable names FAM PUB12 PUB34 PUB5 PRIV YEARS
 SCHOOL LINC PTCON YESVM;
* Dependent variable YESVM;
* YESVM=1 if individual votes yes, 0 otherwise;
* FAM refers to the particular number of the
 individual voter, 1 to 95;
* PUB12=1 if 1 or 2 children attend public school
 within the individual voter's family, 0 otherwise;
* PUB34=1 if 3 or 4 children attend public school
 within the individual voter's family, 0 otherwise;
```


continued...

* PUB5=1 if 5 or more children attend public school within the individual voter's family, 0 otherwise;
* PRIV=1 if the family has 1 or more children attending private school, 0 otherwise;
* SCHOOL=1 if the individual voter is employed as a teacher either in private or public school, 0 otherwise;
* YEARS refeers to the number of years the individual voter has lived in the community;
* LINC refers to the natural log of annual household income, in dollars;
* LPTCON refers to the natural log of property taxes paid per year, in dollars;

```
options nodate;
* descriptive statistics;
proc means data=voting n mean median std min max;
  var pub12 pub34 pub5 priv years school inc ptcon;
* ols regression model;
proc reg data=voting;
 model yesvm = pub12 pub34 pub5 priv years school linc
  lptcon / dw dwprob vif collin;
 test pub12=0, pub34=0, pub5=0;

* binary probit model;
proc qlim data=voting;
 model yesvm = pub12 pub34 pub5 priv years school linc
  lptcon / discrete(d=normal);
 output out=probitresults marginal mills prob xbeta;
* goodness-of-fit test;
test pub12=0, pub34=0, pub5=0, priv=0, years=0,
  school=0, linc=0, lptcon=0 / all;
```

probit model

χ^2 test analogous to the F-test in the conventional single-equation econometric model

```
* test of whether children attending either private or public
school influences voting behavior;
test pub12=0, pub34=0, pub5=0, priv=0 / lr;

proc print data=probitresults; var yesvm xbeta_yesvm
prob_yesvm mills_yesvm;
run;

proc print data=probitresults; var fam meff_p2_pub12
meff_p2_pub34 meff_p2_pub5;
run;

proc print data=probitresults; var fam meff_p2_priv
meff_p2_years meff_p2_school;
run;

proc print data=probitresults; var fam meff_p2_linc
meff_p2_lptcon;
run;

proc means n mean median; var meff_p2_pub12 meff_p2_pub34
meff_p2_pub5
       meff_p2_priv meff_p2_years meff_p2_school meff_p2_linc
meff_p2_lptcon;
run;
```

```
* cutoff equal to # of times yesvm=1 relative to the sample
size (95);
data final; set probitresults;
ap00=0; ap01=0; ap10=0; ap11=0;
if prob_yesvm < .6210526 and yesvm=0 then ap00=1;
if prob_yesvm < .6210526 and yesvm=1 then ap10=1;
if prob_yesvm > .6210526 and yesvm=0 then ap01=1;
if prob_yesvm > .6210526 and yesvm=1 then ap11=1;
proc means data=final n mean median sum ; var ap00 ap10 ap01
ap11 yesvm;
run;
* cutoff equal to 0.5;
data final; set probitresults;
ap00=0; ap01=0; ap10=0; ap11=0;
if prob_yesvm < .5 and yesvm=0 then ap00=1;
if prob_yesvm < .5 and yesvm=1 then ap10=1;
if prob_yesvm > .5 and yesvm=0 then ap01=1;
if prob_yesvm > .5 and yesvm=1 then ap11=1;
proc means data=final n mean median sum ; var ap00 ap10 ap01
ap11 yesvm;
run;
```

The MEANS Procedure

Variable	N	Mean	Median	Std Dev	Minimum	Maximum
pub12	95	0.4842105	0	0.5024018	0	1.0000000
pub34	95	0.3157895	0	0.4672955	0	1.0000000
pub5	95	0.0421053	0	0.2018947	0	1.0000000
priv	95	0.1052632	0	0.3085203	0	1.0000000
years	95	8.5157895	5.0000000	9.5157911	1.0000000	49.0000000
school	95	0.1157895	0	0.3216698	0	1.0000000
inc	95	23093.30	22493.91	8871.35	3999.80	50011.09
ptcon	95	1079.97	1149.98	307.5520124	400.0141814	1799.92

OLS Estimates: Linear Probability Model

The REG Procedure
Dependent Variable: yesvm

Number of Observations Read	95
Number of Observations Used	95

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	3.82493	0.47812	2.22	0.0336
Error	86	18.53296	0.21550		
Corrected Total	94	22.35789			

Root MSE	0.46422	R-Square	0.1711
Dependent Mean	0.62105	Adj R-Sq	0.0940
Coeff Var	74.74718		

OLS Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	1	-0.38672	1.48589	-0.26	0.7953	0
pub12	1	0.10668	0.14759	0.72	0.4718	2.39835
pub34	1	0.21658	0.16171	1.34	0.1840	2.49080
pub5	1	0.10118	0.26480	0.38	0.7033	1.24675
priv	1	-0.06788	0.16806	-0.40	0.6873	1.17273
years	1	-0.00551	0.00549	-1.00	0.3182	1.18876
school	1	0.31384	0.15895	1.97	0.0515	1.14028
linc	1	0.37777	0.14080	2.68	0.0087	1.46693
lptcon	1	-0.41293	0.18429	-2.24	0.0276	1.48792

No evidence of collinearity.

Collinearity Diagnostics

Number	Eigenvalue	Condition Index	Proportion of Variation: Intercept	pub12	pub34	pub5	priv
1	4.61511	1.00000	0.00004699	0.00494	0.00429	0.00151	0.00435
2	1.19731	1.96331	6.153522E-7	0.00559	0.05000	0.15439	0.16551
3	1.02635	2.12052	2.05793E-7	0.07328	0.05448	0.26553	0.03526
4	0.82757	2.36150	1.345239E-8	0.01097	0.02146	0.34014	0.44851
5	0.73070	2.51317	0.00000729	0.00542	0.05434	0.00003362	0.15883
6	0.50144	3.03376	0.00004526	0.02247	0.01137	0.00006634	0.00889
7	0.09991	6.79656	0.00096295	0.86125	0.75958	0.22453	0.16839
8	0.00093990	70.07304	0.07078	0.00110	0.01376	0.01349	0.00034555
9	0.00066672	83.19906	0.92815	0.01498	0.03072	0.00032184	0.00991

Collinearity Diagnostics

	Proportion of Variation			
Number	years	school	linc	lptcon
1	0.01194	0.00613	0.00005235	0.00006303
2	0.00015787	0.16034	9.410293E-7	0.00000107
3	0.00540	0.02432	2.293479E-7	1.095453E-7
4	0.00211	0.03420	2.010786E-8	8.77066E-10
5	0.01477	0.68304	0.00000976	0.00001259
6	0.77601	0.00549	0.00006579	0.00008179
7	0.02306	0.00014149	0.00101	0.00119
8	0.02017	0.00748	0.36194	0.94111
9	0.14639	0.07887	$\boxed{0.63692}$	0.05755

Linc behaves similarly to the intercept

The REG Procedure

Dependent Variable: yesvm

Durbin-Watson D	2.015
Pr < DW	0.5365
Pr > DW	0.4635
Number of Observations	95
1st Order Autocorrelation	-0.020

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

No evidence of autocorrelation.

The REG Procedure

```
Pub12 = 0, Pub34 = 0, Pub5 = 0, Pnv = 0;
```

Source	DF	Mean Square	F Value	Pr > F
Numerator	3	0.14088	0.65	0.5828
Denominator	86	0.21550		

The # of children attending school has no effect on the probability of voting yes.

Probit Model
The QLIM Procedure

Discrete Response Profile of yesvm

Index	Value	Frequency	Percent	
1	0	36	37.89	# of individuals who voted no
2	1	59	62.11	# of individuals who voted yes

Model Fit Summary

Number of Endogenous Variables	1
Endogenous Variable	yesvm
Number of Observations	95
Log Likelihood	-53.14333
Maximum Absolute Gradient	1.72601E-6
Number of Iterations	23
Optimization Method	Quasi-Newton
AIC	124.28665
Schwarz Criterion	147.27155

Goodness-of-Fit Measures

Measure	Value	Formula
Likelihood Ratio (R)	19.787	2 * (LogL - LogL0)
Upper Bound of R (U)	126.07	- 2 * LogL0
Aldrich-Nelson	0.1724	R / (R+N)
Cragg-Uhler 1	0.188	1 - exp(-R/N)
Cragg-Uhler 2	0.2559	(1-exp(-R/N)) / (1-exp(-U/N))
Estrella	0.2027	1 - (1-R/U)^(U/N)
Adjusted Estrella	0.0188	1 - ((LogL-K)/LogL0)^(-2/N*LogL0)
McFadden's LRI	0.1569	R / U
Veall-Zimmermann	0.3023	(R * (U+N)) / (U * (R+N))
McKelvey-Zavoina	0.3756	

N = # of observations, K = # of regressors

The QLIM Procedure

Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-2.956379	4.502313	-0.66	0.5114
pub12	1	0.368112	0.429459	0.86	0.3914
pub34	1	0.691198	0.472186	1.46	0.1432
pub5	1	0.295463	0.759204	0.39	0.6971
priv	1	-0.211166	0.481407	-0.44	0.6609
years	1	-0.015759	0.015282	-1.03	0.3025
school	1	1.584090	0.824349	1.92	0.0547
linc	1	1.314178	0.463637	2.83	0.0046
lptcon	1	-1.464281	0.640103	-2.29	0.0222

Test Results

Type	Statistic	Pr > ChiSq	Label
Wald	13.31	0.1018	pub12 = 0 , pub34 = 0 , pub5 = 0 , priv = 0 , years = 0 , school = 0 , linc = 0 , lptcon = 0
L.R. (Likelihood Ratio)	19.79	$\boxed{0.0112}$	pub12 = 0 , pub34 = 0 , pub5 = 0 , priv = 0 , years = 0 , school = 0 , linc = 0 , lptcon = 0

Goodness-of-Fit tests

Test Results

Type	Statistic	Pr > ChiSq	Label
L.M. (Lagrange Multiplier)	16.46	0.0362	pub12 = 0 , pub34 = 0 , pub5 = 0 , priv = 0 , years = 0 , school = 0 , linc = 0 , lptcon = 0
L.R. (Likelihood Ratio)	3.10	0.5413	pub12 = 0 , pub34 = 0 , pub5 = 0 , priv = 0

Goodness-of-Fit tests

Test of subset of coefficients dealing with children in public or private schools xbeta_yesvm = z linear combination of coefficients with parameter estimates for each observation

Obs	yesvm	Xbeta_ yesvm	Prob_ yesvm	Mills_ yesvm
1	$\boxed{1}$	$\boxed{1.68132}$	0.95365	0.10178
2	$\boxed{0}$	$\boxed{0.45860}$	0.32326	0.53066
3	0	0.19855	0.42131	0.67593
4	0	0.56039	0.28761	0.47863
5	1	1.78221	0.96264	0.08467
6	0	0.76366	0.22253	0.38335
7	0	-0.16956	0.56732	0.90887
8	1	0.16646	0.56610	0.69502
9	0	0.36815	0.35638	0.57923
10	1	-0.26620	0.39504	0.97472
11	1	0.21431	0.58485	0.66664
12	0	0.41163	0.34031	0.55562

$$z_1 = -2.956379 + .368112*pub12 + .691198*pub34 + .295463*pub5 - .211166*pnv - .015759*years + 1.58409*school + 1.314178*linc - 1.464281*lptcon$$

Mills_yesvm (IMR) = f(z)/F(z) if yesvm = 1
OR f(z)/1-F(z) if yesvm = 0.

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2)$$

Obs	yesvm	Xbeta_ yesvm	Prob_ yesvm	Mills_ yesvm
1	1	1.68132	0.95365	0.10178
2	0	0.45860	0.32326	0.53066
3	0	0.19855	0.42131	0.67593
4	0	0.56039	0.28761	0.47863
5	1	1.78221	0.96264	0.08467
6	0	0.76366	0.22253	0.38335
7	0	-0.16956	0.56732	0.90887
8	1	0.16646	0.56610	0.69502
9	0	0.36815	0.35638	0.57923
10	1	-0.26620	0.39504	0.97472
11	1	0.21431	0.58485	0.66664
12	0	0.41163	0.34031	0.55562

Represents the probability of voting yes: *$P(y_i=1)=F(z_i)$*
or
1 - probability of voting no: $P(y_i=0)=1-F(z_i)$

$.03573 = .368112*f(z_1)$

$z_1 = 1.68132$

$$f(z_1) = \frac{1}{\sqrt{2\pi}} \exp\left(-z_1^{2}/2\right)$$

$.06709 = .691198*f(z_1)$

$.02868 = .295463*f(z_1)$

Marginal Effects: $\hat{\beta} f(z)$

Obs	fam	Meff_P2_ pub12	Meff_P2_ pub34	Meff_P2_ pub5
1	1	0.03573	0.06709	0.02868
2	2	0.13220	0.24822	0.10611
3	3	0.14399	0.27037	0.11557
4	4	0.12552	0.23568	0.10074
5	5	0.03000	0.05634	0.02408
6	6	0.10971	0.20600	0.08806
7	7	0.14476	0.27181	0.11619
8	8	0.14483	0.27195	0.11625
9	9	0.13723	0.25768	0.11015
10	10	0.14174	0.26615	0.11377
11	11	0.14352	0.26949	0.11520
12	12	0.13493	0.25335	0.10830

-0.015759*f(z_1)

1.584090*f(z_1)

Marginal Effects: $\hat{\beta} f(z)$

-0.211166*f(z_1)

Obs	fam	Meff_P2_priv	Meff_P2_years	Meff_P2_school
1	1	-0.020497	-.001529651	0.15376
2	2	-0.075834	-.005659276	0.56888
3	3	-0.082599	-.006164108	0.61963
4	4	-0.072001	-.005373262	0.54013
5	5	-0.017211	-.001284426	0.12911
6	6	-0.062936	-.004696726	0.47212
7	7	-0.083041	-.006197092	0.62294
8	8	-0.083084	-.006200318	0.62327
9	9	-0.078723	-.005874891	0.59055
10	10	-0.081311	-.006067972	0.60996
11	11	-0.082331	-.006144088	0.61761
12	12	-0.077400	-.005776140	0.58063

The correct marginal effects are given by

$$\frac{\partial P}{\partial x} = \frac{\partial P}{\partial z} \cdot \frac{\partial z}{\partial x}$$

$\frac{\partial P}{\partial z} = f(z)$, but because

linc and *lptcon* represent the natural log of *inc* and *ptcon* respectively, $\frac{\partial z}{\partial x}$ for these variables are given by

$$\frac{\hat{\beta} linc}{inc} \text{ and } \frac{\hat{\beta} lptcon}{ptcon}$$

Technically, then we need to scale each of these parameter estimates by dividing by *inc* and *ptcon*

Marginal Effects for Inc and Ptcon

Obs	fam	Meff_P2_ linc	Meff_P2_ lptcon
1	1	0.12756	-0.14213
2	2	0.47195	-0.52585
3	3	0.51405	-0.57276
4	4	0.44810	-0.49928
5	5	0.10711	-0.11935
6	6	0.39168	-0.43641
7	7	0.51680	-0.57583
8	8	0.51707	-0.57613
9	9	0.48993	-0.54589
10	10	0.50603	-0.56383
11	11	0.51238	-0.57090
12	12	0.48169	-0.53671

```
Marginal Effects for Inc and Ptcon
                    Meff_P2_     Meff_P2_
Obs      fam          linc        lptcon
```

Obs	fam	Meff_P2_ linc	Meff_P2_ lptcon
1	1	0.12756	-0.14213
2	2	0.47195	-0.52585
3	3	0.51405	-0.57276
4	4	0.44810	-0.49928
5	5	0.10711	-0.11935
6	6	0.39168	-0.43641
7	7	0.51680	-0.57583
8	8	0.51707	-0.57613
9	9	0.48993	-0.54589
10	10	0.50603	-0.56383
11	11	0.51238	-0.57090
12	12	0.48169	-0.53671

Need to divide .12756 by exp(9.77)

need to divide -.14213 by exp(7.0475)

in order to obtain the correct marginal effects.

The MEANS Procedure

Summary of Marginal Effects Through Means of the 95 observations

Variable	Label	N	Mean
Meff_P2_pub12	Marginal effect of pub12 on the probability of yesvm=2	95	0.1166412
Meff_P2_pub34	Marginal effect of pub34 on the probability of yesvm=2	95	0.2190152
Meff_P2_pub5	Marginal effect of pub5 on the probability of yesvm=2	95	0.0936214
Meff_P2_priv	Marginal effect of priv on the probability of yesvm=2	95	-0.0669108
Meff_P2_years	Marginal effect of years on the probability of yesvm=2	95	-0.0049934
Meff_P2_school	Marginal effect of school on the probability of yesvm=2	95	0.5019396
Meff_P2_linc	Marginal effect of linc on the probability of yesvm=2	95	0.4164144
Meff_P2_lptcon	Marginal effect of lptcon on the probability of yesvm=2	95	-0.4639766

Scale these marginal effects.

The MEANS Procedure
Summary of Marginal Effects Through Medians
of the 95 observations

Variable	Label	Median
Meff_P2_pub12	Marginal effect of pub12 on the probability of yesvm=2	0.1340547
Meff_P2_pub34	Marginal effect of pub34 on the probability of yesvm=2	0.2517121
Meff_P2_pub5	Marginal effect of pub5 on the probability of yesvm=2	0.1075983
Meff_P2_priv	Marginal effect of priv on the probability of yesvm=2	-0.0768999
Meff_P2_years	Marginal effect of years on the probability of yesvm=2	-0.0057388
Meff_P2_school	Marginal effect of school on the probability of yesvm=2	0.5768745
Meff_P2_linc	Marginal effect of linc on the probability of yesvm=2	0.4785813
Meff_P2_lptcon	Marginal effect of lptcon on the probability of yesvm=2	-0.5332441

Scale these marginal effects.

Information to Generate Prediction-Success Table

Cutoff value: 59/95 = 0.6210526

The MEANS Procedure

Variable	N	Mean	Median	Sum
ap00	95	0.2947368	0	28.0000000
ap10	95	0.2631579	0	25.0000000
ap01	95	0.0842105	0	8.0000000
ap11	95	0.3578947	0	34.0000000
yesvm	95	0.6210526	1.0000000	59.0000000

Information to Generate Prediction-Success Table

Cutoff value: 0.5

The MEANS Procedure

Variable	N	Mean	Median	Sum
ap00	95	0.1894737	0	18.0000000
ap10	95	0.0736842	0	7.0000000
ap01	95	0.1894737	0	18.0000000
ap11	95	0.5473684	1.0000000	52.0000000
yesvm	95	0.6210526	1.0000000	59.0000000

Prediction-Success Table with Cutoff value 0.6210526
Probit Model

	Actual	
Predicted	**0**	**1**
0	28	25
1	8	34
	36	59

Correct predictions

(1) Number of right predictions 28 + 34 = 62

(2) Percentage of right predictions (62/95)*100 = 65.3%

(3) Sensitivity (the fraction of y = 1 observations that are correctly predicted) 34/59 = 57.6%

(4) Specificity (the fraction of y = 0 observations that are correctly predicted) 28/36 = 77.8%

Section 13.5

Sample Problem: Use of Logit Analysis

The QLIM Procedure

LOGIT Model

Discrete Response Profile of yesvm

Index	Value	Frequency	Percent
1	0	36	37.89
2	1	59	62.11

Model Fit Summary

Number of Endogenous Variables	1
Endogenous Variable	yesvm
Number of Observations	95
Log Likelihood	-53.30459
Maximum Absolute Gradient	9.53159E-8
Number of Iterations	23
Optimization Method	Quasi-Newton
AIC	124.60918
Schwarz Criterion	147.59408

Goodness-of-Fit Measures

Measure	Value	Formula
Likelihood Ratio (R)	19.465	2 * (LogL - LogL0)
Upper Bound of R (U)	126.07	- 2 * LogL0
Aldrich-Nelson	0.17	R / (R+N)
Cragg-Uhler 1	0.1853	1 - exp(-R/N)
Cragg-Uhler 2	0.2521	(1-exp(-R/N)) / (1-exp(-U/N))
Estrella	0.1995	1 - (1-R/U)^(U/N)
Adjusted Estrella	0.0154	1 - ((LogL-K)/LogL0)^(-2/N*LogL0)
McFadden's LRI	0.1544	R / U
Veall-Zimmermann	0.2982	(R * (U+N)) / (U * (R+N))
McKelvey-Zavoina	0.6212	

N = # of observations, K = # of regressors

The QLIM Procedure

LOGIT Model

Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-5.197788	7.550855	-0.69	0.4912
pub12	1	0.583293	0.687755	0.85	0.3964
pub34	1	1.125618	0.768129	1.47	0.1428
pub5	1	0.525766	1.269313	0.41	0.6787
priv	1	-0.341526	0.782980	-0.44	0.6627
years	1	-0.026111	0.026931	-0.97	0.3323
school	1	2.627088	1.410853	1.86	0.0626
linc	1	2.187099	0.788196	2.77	0.0055
lptcon	1	-2.394847	1.081372	-2.21	0.0268

Goodness-of-Fit Tests

Test Results

Type	Statistic	Pr > ChiSq	Label
Wald	12.03	0.1500	pub12 = 0 , pub34 = 0 , pub5 = 0 , priv = 0 , years = 0 , school = 0 , linc = 0 , lptcon = 0
L.R.	19.46	0.0126	pub12 = 0 , pub34 = 0 , pub5 = 0 , priv = 0 , years = 0 , school = 0 , linc = 0 , lptcon = 0
L.M.	16.25	0.0389	pub12 = 0 , pub34 = 0 , pub5 = 0 , priv = 0 , years = 0 , school = 0 , linc = 0 , lptcon = 0
L.R.	3.03	0.5527	pub12 = 0 , pub34 = 0 , pub5 = 0 , priv = 0 ,

Test of Subset of Coefficients

LOGIT Model

Obs	yesvm	Xbeta_ yesvm	Prob_ yesvm	Mills_ yesvm
1	1	2.78409	0.94181	0.05819
2	0	0.75818	0.31904	0.31904
3	0	0.32030	0.42060	0.42060
4	0	0.90080	0.28889	0.28889
5	1	2.96094	0.95078	0.04922
6	0	1.26089	0.22082	0.22082
7	0	-0.26299	0.56537	0.56537
8	1	0.28139	0.56989	0.43011
9	0	0.60324	0.35360	0.35360
10	1	-0.44840	0.38974	0.61026
11	1	0.34641	0.58575	0.41425
12	0	0.65055	0.34287	0.34287

Marginal Effects LOGIT Model

Obs	fam	Meff_P2_ pub12	Meff_P2_ pub34	Meff_P2_ pub5
1	1	0.03197	0.06169	0.02881
2	2	0.12672	0.24454	0.11422
3	3	0.14215	0.27431	0.12813
4	4	0.11983	0.23124	0.10801
5	5	0.02730	0.05268	0.02461
6	6	0.10036	0.19367	0.09046
7	7	0.14333	0.27659	0.12919
8	8	0.14297	0.27591	0.12887
9	9	0.13332	0.25728	0.12017
10	10	0.13873	0.26772	0.12505
11	11	0.14153	0.27313	0.12758
12	12	0.13142	0.25361	0.11846

Marginal Effects LOGIT Model

Obs	fam	Meff_P2_ priv	Meff_P2_ years	Meff_P2_ school
1	1	-0.018717	-.001430973	0.14397
2	2	-0.074198	-.005672655	0.57074
3	3	-0.083229	-.006363083	0.64021
4	4	-0.070160	-.005363961	0.53969
5	5	-0.015983	-.001221962	0.12295
6	6	-0.058763	-.004492588	0.45201
7	7	-0.083922	-.006416104	0.64555
8	8	-0.083713	-.006400155	0.64394
9	9	-0.078062	-.005968068	0.60047
10	10	-0.081230	-.006210252	0.62483
11	11	-0.082870	-.006335703	0.63746
12	12	-0.076949	-.005882983	0.59191

```
Marginal Effects LOGIT Model

                    Meff_P2_    Meff_P2_
 Obs     fam            linc      lptcon

   1       1         0.11986    -0.13125
   2       2         0.47516    -0.52029
   3       3         0.53299    -0.58361
   4       4         0.44930    -0.49198
   5       5         0.10235    -0.11208
   6       6         0.37631    -0.41205
   7       7         0.53743    -0.58848
   8       8         0.53609    -0.58701
   9       9         0.49990    -0.54738
  10      10         0.52019    -0.56960
  11      11         0.53069    -0.58110
  12      12         0.49277    -0.53958
```

Again, realize we need to divide Meff_P2_linc by inc(exp(9.77)) and Meff_P2_lptcon by ptcon(exp(7.0475)) in order to obtain the corect marginal effects

Summary of Marginal Effects: LOGIT Model

Variable	Label	N	Mean
Meff_P2_pub12	Marginal effect of pub12 on the probability of yesvm=2	95	0.1123932
Meff_P2_pub34	Marginal effect of pub34 on the probability of yesvm=2	95	0.2168925
Meff_P2_pub5	Marginal effect of pub5 on the probability of yesvm=2	95	0.1013085
Meff_P2_priv	Marginal effect of priv on the probability of yesvm=2	95	-0.0658078
Meff_P2_years	Marginal effect of years on the probability of yesvm=2	95	-0.0050312
Meff_P2_school	Marginal effect of school on the probability of yesvm=2	95	0.5062070
Meff_P2_linc	Marginal effect of linc on the probability of yesvm=2	95	0.4214267
Meff_P2_lptcon	Marginal effect of lptcon on the probability of yesvm=2	95	-0.4614570

Scale these marginal effects.

Summary of Marginal Effects: LOGIT Model

Variable	Label	Median
Meff_P2_pub12	Marginal effect of pub12 on the probability of yesvm=2	0.1290633
Meff_P2_pub34	Marginal effect of pub34 on the probability of yesvm=2	0.2490620
Meff_P2_pub5	Marginal effect of pub5 on the probability of yesvm=2	0.1163346
Meff_P2_priv	Marginal effect of priv on the probability of yesvm=2	-0.0755684
Meff_P2_years	Marginal effect of years on the probability of yesvm=2	-0.0057774
Meff_P2_school	Marginal effect of school on the probability of yesvm=2	0.5812875
Meff_P2_linc	Marginal effect of linc on the probability of yesvm=2	0.4839326
Meff_P2_lptcon	Marginal effect of lptcon on the probability of yesvm=2	-0.5299002

Scale these marginal effects.

Information for Prediction-Success Table LOGIT Model
Cuttoff value: 0.6210526

Variable	N	Mean	Median	Sum
ap00	95	0.2947368	0	28.0000000
ap10	95	0.2631579	0	25.0000000
ap01	95	0.0842105	0	8.0000000
ap11	95	0.3578947	0	34.0000000
yesvm	95	0.6210526	1.0000000	59.0000000

The MEANS Procedure

Variable	N	Mean	Median	Sum
ap00	95	0.1894737	0	18.0000000
ap10	95	0.0736842	0	7.0000000
ap01	95	0.1894737	0	18.0000000
ap11	95	0.5473684	1.0000000	52.0000000
yesvm	95	0.6210526	1.0000000	59.0000000

Prediction-Success Table with Cutoff value 0.6210526 LOGIT Model

	Actual	
Predicted	**0**	**1**
0	28	25
1	8	34

Correct predictions

Same performance as probit model

Parameter Estimates and Associated Standard Errors of the Variables in Sample Problem

Variable	OLS ANALYSIS	PROBIT ANALYSIS		LOGIT ANALYSIS	
	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)	Change in[a] Probability (Marginal Effects)	Parameter Estimate (Standard Error)	Change in[b] Probability (Marginal Effects)
PUB12	.10668 (.14759)	.36811 (.42946)	(.1343)	.58329 (.68775)	(.1294)
PUB34	.21658 (.16171)	.69119 (.47219)	(.2523)	1.1256 (.76813)	(.2498)
PUB5	.10118 (.26480)	.29548 (.75920)	(.1078)	.52577 (1.2693)	(.1166)
PRIV	-.06788 (.16806)	-.21116 (.48141)	(-.0770)	-.34153 (.78297)	(-.0757)
YEARS	-.0550 (.00548)	-.01575 (.01528)	(-.0057)	-.02611 (.02693)	(-.0057)
SCHOOL	.31384* (.15895)	1.5840* (.82435)	(.5783)	2.6271* (1.4103)	(.5830)
LINC	.37777* (.14080)	1.3141* (.46364)	(.4797)	2.1871* (.78796)	(.4854)
LPTCON	-.41293* (.18429)	-1.4843* (.64010)	(-.5346)	-2.3948* (1.0813)	(-.5314)
INTERCEPT	-.38672 (1.4859)	-2.9563 (4.50231)		-5.1978 (7.5500)	

* Statistically significant at $\alpha = 0.05$.

[a] At the sample means, $z_i = x_i'\hat{\beta} = .42109$. Consequently, $f(Z_i) = .3651$)

[b] At the sample means, $z_i = x_i'\hat{\beta} = .69698$. Consequently, $f(Z_i) = .2219$)

Section 13.6

Censored Response Models

Censored Response Models

Tobit model (Tobin, 1958)

Heckman sample selection procedure
(Heckman 1976, 1979)

Censored Response Models

Overview

Dependent variables are often subject to some thresholds or constraints in economic problems such as non-negativity, price support levels, and acreage or import quotas.

Application of OLS produces inconsistent estimators

Using household budget data, one is very likely to encounter zero observations, typically corresponding to expenditures

Fair amount of work in demand analysis has been geared to the “zero-expenditure” problem

Traditional Approach

- Tobit Procedure (Tobin 1958)
- McDonald and Moffitt (1980)
- McCracken and Brandt (1987)
- Capps and Love (1983)

Cragg (1987), generalization of Tobit model to allow the decision process to have two steps

- Haines *et al (*1988)
- Blaylock and Blisard (1992)
- Blisard and Blaylock (1993)
- Yen (1993)

Heckman-Type Sample Selectivity Correction

Capps and Cheng (1988)

Jensen, Kesavan, and Johnson (1992)

Heckman Procedure with Calculation of Correct Marginal Effects—Saha, Capps, and Byrne (1997)

Section 13.7

Censored Samples: Use of the Tobit Model

Censored Samples: Use of the Tobit Model

Tobin (1958)

Tobit Model

$$y_i = X_i^T \beta + \varepsilon_i \quad \text{if RHS} > 0$$
$$0 \qquad \text{if RHS} \le 0$$

$\varepsilon_{\mathrm{i}}\ iid\ N(0, \sigma^2)$ **lower limit zero**

Suppose lower bound (threshold) not zero as in the case of price support levels, but

$$\alpha_i$$

$$\mathrm{y_i} - \alpha_{\mathrm{i}} = (\mathrm{X_i^T} \beta - \alpha_{\mathrm{i}}) + \varepsilon_{\mathrm{i}} \quad \text{if RHS} > 0$$
$$0 \qquad \text{if RHS} \le 0$$

Suppose the threshold is an upper bound as in the case of acreage and import quotas

$$
-y_i + \alpha_i = \begin{cases} -(X_i^T\beta - \alpha_i) - \varepsilon_i, & \text{if RHS} < 0 \\ 0, & \text{if RHS} \geq 0 \end{cases}
$$

Suppose both lower and upper bounds exist (ceiling and floors on wage rates)

$$
y_i = \begin{cases} \alpha_1 & \text{if RHS} < \alpha_1 \\ X_i^T\beta + \varepsilon_i & \text{if } \alpha_1 \leq \text{RHS} \leq \alpha_2 \\ \alpha_2 & \text{if RHS} > \alpha_2 \end{cases}
$$

Description of the Tobit Model
(for the case of lower limit of zero)

$$Y = X\beta + \epsilon \quad \text{if } X\beta + \epsilon > 0$$
$$Y = 0 \quad \text{if } X\beta + \epsilon \leq 0 \tag{1}$$

$$E(Y) = x\beta F(z) + \sigma f(z) \tag{2}$$

$$E(Y^*) = X\beta + \sigma f(z)/F(z) \tag{3}$$

$$\partial E(Y)/\partial X = F(z)(\partial E(Y^*)/\partial X) + E(Y^*)(\partial F(z)/\partial X) = F(z)\beta \tag{4}$$

$$\partial E(Y^*)/\partial X = \beta(1 - zf(z)/F(z) - f(z)^2/F(z)^2) \tag{5}$$

$$\partial F(z)/\partial X = f(z)\,\beta/\sigma \tag{6}$$

where

X	=	a vector of regressor variables,
β	=	a vector of unknown coeffecients (Tobit coefficients)
ϵ	=	a vector of independentdent and identically disstributed normal random variables assumed to have mean zero and constant variance, σ^2,
z	=	$X\beta/\sigma$, normalized index,
f(z)	=	the standard normal density function, and
F(z)	=	the cumulative standard normal distribution function

Source: McDonald and Moffitt (1980)

Estimation of Tobit Model

$$y_i = X_i^T\beta + \varepsilon_i \quad \text{if RHS} > 0$$
$$\qquad 0 \qquad \text{if RHS} \le 0$$

$$L = \prod_{i=1}^{n_1}\left[1 - F\left(X_i^T\beta;\sigma^2\right)\right]\prod_{i=n_1=1}^{n} f\left(y_i - X_i^T\beta, \sigma^2\right)$$

$$\text{Let } z = X_i^T\beta.\ \ F\left(X_i^T\beta;\sigma^2\right) = \int_{-\infty}^{z} f(z;\sigma^2)dz$$

$$f(z;\sigma^2) = 1/(2\pi\sigma^2)^{1/2}\,\text{EXP}(-z^2/2\sigma^2)$$

$$\text{Log } L = \sum_{i=1}^{n_1}\text{Log}\left[1 - F\left(X_i^T\beta;\sigma^2\right)\right] - (n_2/2)\text{Log}(2\pi\sigma^2)$$

$$- \sum_{i=n_1+1}^{n}(y_i - X_i^T\beta)^2/2\sigma^2$$

$$n_2 = n - n_1$$

Section 13.8

Sample Problem with the Tobit Model

```
data tobitsamproblem;
input samn y x1 x2;
datalines;
1 0 .693 .693
2 11.478 1.733 .693
3 0 .693 1.386
4 0 1.733 1.386
5 0 .693 1.792
6 0 2.340 .693
7 12.404 1.733 1.792
8 0 2.340 1.386
9 12.006 2.340 1.792
10 0 .693 .693
11 0 .693 1.386
12 12.062 1.733 .693
13 0 1.733 1.386
14 0 .693 1.792
15 11.548 2.340 .693
16 0 1.733 1.792
17 11.795 2.340 1.386
18 0 2.340 1.792
19 0 1.733 1.386
20 0 .693 .693
;
```

```
options nodate;

proc means data=tobitsamproblem n mean median std min max;
      var y x1 x2;
run;
* OLS Analysis;

proc reg data=tobitsamproblem;
 model y = x1 x2;

* TOBIT Analysis;

 proc qlim data=tobitsamproblem;
 model y = x1 x2;
 endogenous y ~ censored(lb=0);
 output out=tobitresults conditional expected marginal xbeta;
 run;

 data new; set tobitresults;
 proc print data=tobitresults;
      var xbeta_y meff_x1 meff_x2 expct_y cexpct_y;
 run;
```

```
data mddecomp; set new;
 * z is the normalized index;
 * SAS does not provide z directly;
 z=xbeta_y/11.718546;
 * capfz is the cdf standard normal;
 * SAS does not provide capfz directly;
 capfz=probnorm(z);
 * fz is the standard normal density function;
 * SAS does not provide fz directly;
 fz=exp(-z**2/2)/2.5066272;
 * expected_y is the unconditional expected value of the
dependent variable;
 * expected_y also serves as the predicted value of the
dependent variable;
 * SAS captures the unconditional expected values;
 expected_y=xbeta_y*capfz+11.718546*fz;
 * cexpected_y is the conditional expected value of the
dependent variable;
 * SAS captures the conditional expected values;
```

```
* cexpected_y is the conditional expected value of the dependent
variable;
 * SAS captures the conditional expected values;
 cexpected_y=xbeta_y+11.718546*fz/capfz;
 * unconditional marginal effects;
 * SAS captures the unconditional marginal effects;
 me_yx1=capfz*15.151408;
 me_yx2=capfz*-6.313204;
 * conditional marginal effects;
 * SAS does not capture the conditional marginal effects
directly;
 cme_yx1=15.151408*(1-z*fz/capfz-(fz/capfz)*(fz/capfz));
 cme_yx2=-6.313204*(1-z*fz/capfz-(fz/capfz)*(fz/capfz));
 * the change in probability of being above the lower limit due
to changes in x1 and x2;
 * SAS does not provide the change in probability directly;
 dcapfz_x1=fz*15.151408/11.718546;
 dcapfz_x2=fz*-6.313205/11.718546;
```

```
proc print data=mddecomp;
 var z capfz fz expected_y cexpected_y;
run;

proc print data=mddecomp;
 var me_yx1 me_yx2 cme_yx1 cme_yx2 dcapfz_x1 dcapfz_x2;
run;

proc means data=mddecomp n mean median min max;
    var z capfz fz expected_y cexpected_y;
run;

proc means data=mddecomp n mean median min max;
    var me_yx1 me_yx2 cme_yx1 cme_yx2;
run;

proc means data=mddecomp n mean median min max;
    var dcapfz_x1 dcapfz_x2;
run;

 * To obtain a reasonable R2 value, initially calculate the
correlation of y and expected_y;
 * Subsequently square this correlation to obtain a proxy for R2;
proc corr data=mddecomp; var y expected_y;
run;
```

The MEANS Procedure

Variable	N	Mean	Median	Std Dev	Minimum	Maximum
y	20	3.5646500	0	5.5893852	0	12.4040000
x1	20	1.5511000	1.7330000	0.6928257	0.6930000	2.3400000
x2	20	1.2652500	1.3860000	0.4622114	0.6930000	1.7920000

The REG Procedure

OLS Estimates

Dependent Variable: y

Number of Observations Read	20
Number of Observations Used	20

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	145.03399	72.51699	2.75	0.0924
Error	17	448.54933	26.38525		
Corrected Total	19	593.58332			

Root MSE	5.13666	R-Square	0.2443
Dependent Mean	3.56465	Adj R-Sq	0.1554
Coeff Var	144.09992		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-0.01559	4.16227	-0.00	0.9971
x1	1	3.88710	1.70739	2.28	0.0360
x2	1	-1.93562	2.55928	-0.76	0.4598

The QLIM Procedure

Summary Statistics of Continuous Responses

Variable	Mean	Standard Error	Type	Lower Bound	Upper Bound	N Obs Lower Bound	N Obs Upper Bound
y	3.56465	5.589385	Censored	0		14	6

Model Fit Summary

Number of Endogenous Variables	1
Endogenous Variable	y
Number of Observations	20
Log Likelihood	-28.66930
Maximum Absolute Gradient	6.45846E-7
Number of Iterations	18
Optimization Method	Newton-Raphson
AIC	65.33859
Schwarz Criterion	69.32152

Maximum Likelihood Parameter Estimates (Tobit Estimates)

Parameter	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-23.009570	17.487114	-1.32	0.1882
x1	1	15.151408	7.925983	1.91	0.0559
x2	1	-6.313204	7.730709	-0.82	0.4141
_Sigma	1	11.718546	3.995490	2.93	0.0034

		unconditional marginal effects		unconditional expected value	conditional expected value
Obs	Xbeta_y	Meff_x1	Meff_x2	Expct_y	Cexpct_y
1	-16.8847	1.1335	-0.47231	0.39247	5.2460
2	-1.1272	6.9952	-2.91471	4.13302	8.9521
3	-21.2597	0.5276	-0.21985	0.16140	4.6349
4	-5.5023	4.8385	-2.01608	2.42996	7.6092
5	-23.8229	0.3186	-0.13277	0.09106	4.3298
6	8.0697	11.4313	-4.76312	9.77651	12.9581
7	-8.0654	3.7218	-1.55080	1.70787	6.9526
8	3.6946	9.4503	-3.93771	6.75278	10.8265
9	1.1315	8.1584	-3.39940	5.26253	9.7733
10	-16.8847	1.1335	-0.47231	0.39247	5.2460
11	-21.2597	0.5276	-0.21985	0.16140	4.6349
12	-1.1272	6.9952	-2.91471	4.13302	8.9521
13	-5.5023	4.8385	-2.01608	2.42996	7.6092
14	-23.8229	0.3186	-0.13277	0.09106	4.3298
15	8.0697	11.4313	-4.76312	9.77651	12.9581
16	-8.0654	3.7218	-1.55080	1.70787	6.9526
20	-16.8847	1.1335	-0.47231	0.39247	5.2460

xbeta_y linear combination of Tobit estimates with data associated with the RHS variables (intercept, x1, and x2)

Probability of being above the lower limit (0) conditional on x1 and x2; capfz = F(z)

SAS fails to report z, capfz, and fz

Obs	z	capfz	fz	expected_ y	cexpected_ y
1	-1.44085	0.07481	0.14129	0.39247	5.2461
2	-0.09619	0.46168	0.39710	4.13302	8.9521
3	-1.81420	0.03482	0.07695	0.16140	4.6349
4	-0.46954	0.31934	0.35730	2.42996	7.6092
5	-2.03292	0.02103	0.05052	0.09106	4.3298
6	0.68862	0.75447	0.31473	9.77651	12.9581
7	-0.68826	0.24564	0.31481	1.70787	6.9526
8	0.31528	0.62373	0.37960	6.75278	10.8265
9	0.09655	0.53846	0.39709	5.26253	9.7733
10	-1.44085	0.07481	0.14129	0.39247	5.2461
11	-1.81420	0.03482	0.07695	0.16140	4.6349
12	-0.09619	0.46168	0.39710	4.13302	8.9521
13	-0.46954	0.31934	0.35730	2.42996	7.6092
14	-2.03292	0.02103	0.05052	0.09106	4.3298
15	0.68862	0.75447	0.31473	9.77651	12.9581
16	-0.68826	0.24564	0.31481	1.70787	6.9526
17	0.31528	0.62373	0.37960	6.75278	10.8265
18	0.09655	0.53846	0.39709	5.26253	9.7733
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	7.6092
20	-1.44085	0.07481	0.14129	0.39247	5.2461

McDonald and Moffitt calculations

Obs	me_yx1	me_yx2	cme_yx1	cme_yx2	dcapfz_ x1	dcapfz_ x2
1	1.1335	-0.47231	2.34185	-0.97579	0.18268	-0.07612
2	6.9952	-2.91471	5.19602	-2.16505	0.51343	-0.21393
3	0.5276	-0.21985	1.90935	-0.79558	0.09949	-0.04146
4	4.8385	-2.01608	4.14362	-1.72654	0.46197	-0.19249
5	0.3186	-0.13277	1.70216	-0.70925	0.06532	-0.02722
6	11.4313	-4.76312	8.16237	-3.40105	0.40693	-0.16956
7	3.7218	-1.55080	3.63094	-1.51292	0.40703	-0.16960
8	9.4503	-3.93771	6.63217	-2.76346	0.49080	-0.20450
9	8.1584	-3.39940	5.83274	-2.43035	0.51341	-0.21393
10	1.1335	-0.47231	2.34185	-0.97579	0.18268	-0.07612
11	0.5276	-0.21985	1.90935	-0.79558	0.09949	-0.04146
12	6.9952	-2.91471	5.19602	-2.16505	0.51343	-0.21393
13	4.8385	-2.01608	4.14362	-1.72654	0.46197	-0.19249
14	0.3186	-0.13277	1.70216	-0.70925	0.06532	-0.02722
15	11.4313	-4.76312	8.16237	-3.40105	0.40693	-0.16956
16	3.7218	-1.55080	3.63094	-1.51292	0.40703	-0.16960
17	9.4503	-3.93771	6.63217	-2.76346	0.49080	-0.20450
18	8.1584	-3.39940	5.83274	-2.43035	0.51341	-0.21393
19	4.8385	-2.01608	4.14362	-1.72654	0.46197	-0.19249
20	1.1335	-0.47231	2.34185	-0.97579	0.18268	-0.07612

McDonald and Moffitt Calculations for unconditional and conditional marginal effects as well as for changes in probabilities of being above the lower limit due to changes in x1 and x2.

Me_yx1 and me_yx2 are calculated by SAS, but cme_yx1, cme_yx2, dcapfz_x1, and dcapfz_x2 are NOT calculated by SAS

The MEANS Procedure

Variable	N	Mean	Median	Minimum	Maximum
z	20	-0.6396700	-0.4695361	-2.0329234	0.6886242
capfz	20	0.3271071	0.3193432	0.0210301	0.7544701
fz	20	0.2678684	0.3148084	0.0505230	0.3971010
expected_y	20	3.2118830	2.4299595	0.0910571	9.7765078
cexpected_y	20	7.7710331	7.6092403	4.3298392	12.9581114

Variable	N	Mean	Median	Minimum	Maximum
me_yx1	20	4.9561334	4.8384997	0.3186361	11.4312846
me_yx2	20	-2.0650940	-2.0160790	-4.7631238	-0.1327675
cme_yx1	20	4.2793961	4.1436238	1.7021583	8.1623726
cme_yx2	20	-1.7831148	-1.7265420	-3.4010518	-0.7092458

Variable	N	Mean	Median	Minimum	Maximum
dcapfz_x1	20	0.3463385	0.4070292	0.0653233	0.5134289
dcapfz_x2	20	-0.1443104	-0.1695987	-0.2139327	-0.0272186

The CORR Procedure

2 Variables: y expected_y

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
y	20	3.56465	5.58939	71.29300	0	12.40400
expected_y	20	3.21188	3.15951	64.23766	0.09106	9.77651

Pearson Correlation Coefficients, N = 20

Prob > |r| under H0: Rho=0

	y	expected_y
y	1.00000	0.43281 0.0566
expected_y	0.43281 0.0566	1.00000

$R^2 = (.43281)^2 = .1873$

Weaknesses of Tobit Model

1. Assumes that the decision to consume is the same as the decision about how much to consume.

   According to Haines, Guilkey, and Popkin (1988), modeling food consumption decisions is a two-step process. "Ignoring the two-step nature of the decision process may hamper understanding of true behavioral patterns, lead to erroneous conclusions, and generate incorrect policy recommendations."

2. Suppose the number of zero observations is sizable. Often the Tobit procedure breaks down -- i.e. not possible to maximize the likelihood function.

Section 13.9

Heckman Sample Selection Procedure

Heckman Sample Selection Procedure

According to Heckman (1976, 1979), sample selection bias is characterized as a specification error or omitted-variable problem. Heckman subsequently proposes a technique that amounts to estimating the omitted variable and using least squares including the estimated omitted variable as a regressor; similar to the McDonald and Moffit (1980) decomposition in accord with the Tobit model,

$$E(Y^*) = X\beta + \frac{\sigma f(z)}{F(z)}$$

the omitted $\uparrow$ variable

$$\frac{f(z)}{F(z)} \rightarrow \text{the inverse of the Mills ratio (IMR)}$$

Heckman Two-Stage Procedure

In the first stage, probit analysis is used to determine the inverse of Mills ratio (IMR_{hi}) for the hth household in the ith commodity. The probit analysis employs all available observations; the dependent variable equals one if the household makes a purchase; otherwise the dependent variable is zero. The second stage involves the use of the estimated inverse Mills Ratio $(I\hat{M}R_{hi})$ as an additional regressor in the estimation equation involving the continuous, non-zero dependent variable. The appropriate estimation technique in the second-stage is either ordinary (OLS) or generalized least squares (GLS). The OLS procedure produces consistent estimates; but the GLS procedure, when implementation is possible (Heckman, pp 480-83), improves the precision of the estimates. The GLS procedure circumvents the potential heteroscedasticity problem inherent in the Heckman procedure.

Mathematically, we can characterize the probit-based Heckman-type selectivity correction as follows. In the first stage, let Z_{hi} denote an indicator variable that takes a value of one if expenditure occurs for the i^{th} commodity by the h^{th} household and zero otherwise. Denoting the normal cumulative distribution function (CDF) by Φ we have:

$$\Pr[Z_{hi} = 1] = \Phi(W_k \gamma_i) \quad \text{and}$$
$$\Pr[Z_{hi} = 0] = 1 - \Phi(W_k \gamma_i) \quad i = 1, \ldots, n; \; h = 1, \ldots, H \qquad \textbf{(1)}$$

W_h is a vector of regressors, related to this purchase decision, is the coefficient vector. The first-stage estimation γ_i provides estimates of γ_i and the inverse of Mills ratio (IMR) defined as:

$$I\hat{M}R_{hi} = \begin{cases} \left\{ \dfrac{\varphi(W_k \hat{\gamma}_i)}{\Phi(W_k \hat{\gamma}_i)} \; for \; Z_{hi} = 1 \right\} \\ \left\{ \dfrac{\varphi(W_k \hat{\gamma}_i)}{1 - \Phi(W_k \hat{\gamma}_i)} \; for \; Z_{hi} = 0 \right\} \end{cases} \qquad \textbf{(2)}$$

In the second stage, let Y_{hi} denote the expenditure of household h on commodity *i*. Then,

$$E\left[Y_{hi} \middle| Z_{hi} = 1\right] = X_h \beta_i + \alpha_i \frac{\varphi(W_h \hat{\gamma}_i)}{\Phi(W_h \hat{\gamma}_i)}$$

$$= X_h \beta_i + \alpha_i I\hat{M}R_{hi} \qquad (3)$$

X_h is a vector of regressors related to the magnitude of the expenditure on the i^{th} commodity.

Importantly, only the non-zero observations on Y_{hi} are used in the second stage.

GLS (Weighted Least Squares) With Heckman

$Let\ \hat{S}_{hi} = W_h \hat{\gamma}_i$, a scalar we estimate from stage one. Let $\hat{\lambda}_{hi} = I\hat{M}R_{hi}$.

The estimate of $I\hat{M}R_{hi}$ is inserted in equation (3) and the coefficients in (3) are estimated using either ordinary least squares (OLS) or generalized or weighted least squares (GLS). For GLS, the weight for each observation is $\left(1+\hat{\delta}_i\left(\hat{S}_{hi}\hat{\lambda}_{hi} - \hat{\lambda}_{hi}^2\right)\right)^{-\frac{1}{2}}$, where $\hat{S}_{hi}$ (xbeta) and $\hat{\lambda}_{hi}$ (Mills) are estimated in stage one, and δ_i is estimated by regressing each squared residual, v_{hi}^2 from the OLS estimation of (3), on $S_{hi}\lambda_{hi} - \lambda_{hi}^2$. Interest lies in testing whether $\hat{\delta}_i$ is significantly different from zero. If $H_0{:}\delta_i = 0$ cannot be rejected, then OLS is the correct estimation procedure. If $H_0{:}\delta_i = 0$ is rejected, then GLS on weighted least squares is the correct estimation procedure. However, nothing with this procedure ensures that the weight for each observation can be determined. That is, since the weight involves a square root, it is necessary for the expression $\left(1+\hat{\delta}_i\left(\hat{S}_{hi}\hat{\lambda}_{hi} - \hat{\lambda}_{hi}^2\right)\right)$ to be positive for all i.

Marginal Effects

Let X_{hj} denote the j^{th} regressor that is common to both W_h and X_h, the vector of regressors on stage 1 and stage 2 equations, respectively. Using (3), the estimated marginal effect of a change in the j^{th} regressor is given by

$$\hat{M}E_{hj} = \frac{\partial E\left[Y_{hi} \middle| Z_{hi} = 1\right]}{\partial X_{hj}} = \beta_j + \alpha_i \frac{\partial}{\partial X_{hj}}\left(I\hat{M}R_{hi}\right) \qquad \textbf{(4)}$$

It is evident from (4) that the marginal effect of the j^{th} regressor is composed of two parts: (a) a change in X_j which affects the probability of choosing the i^{th} commodity; this effect is captured by the second term in the right hand side of (4); and (b) a change in X_j which affects the expenditure on the i^{th} commodity; this effect, however, is conditional upon the household choosing to select the i^{th} commodity. This effect is captured by β_{ij} in (4). In the conventional marginal effect expression, only β_{ij} is considered. The degree and direction of the attendant bias in the calculation of marginal effects depends on the magnitude and sign on the second term of the right hand side of (4).

After much simplification, the correct marginal effect expression becomes

Second stage coefficient associated with correct IMR_{hi}.

xbeta

$$\hat{ME}_{hj} = \beta_{ij} - \alpha_i \boxed{\hat{\gamma}_{ij}} \{ \boxed{W_h \hat{\gamma}_i} \hat{IMR}_{hi} + (\hat{IMR}_{hi})^2 \} \qquad \textbf{(5)}$$

Second stage coefficient

Probit coefficient

Equation (5) represents the appropriate general expression for calculating marginal effects in single equations using the Heckman-type correction.

See Saha, Capps, and Byrne (1997). If α_i is not significantly different from zero, then no sample selection bias exists and the second term in equation (5) is essentially zero. Consequently, β_{ij} represents the appropriate marginal effect when α_i is not significantly different from zero.

The standard Heckman Selection model can be defined as:

$$z_i^* = w_i'\gamma + u_i$$

$$z_i = \begin{cases} 1 & \text{if } z_i^* > 0 \\ 0 & \text{if } z_i^* \le 0 \end{cases}$$

$$y_i = x_i'\beta + \varepsilon_i \text{ if } z_i = 1$$

where u_i and ε_i are jointly normal with zero mean, standard deviations of 1 and σ, and correlation of ρ. z is the variable that the selection is based on, and y is observed when z has a value of 1. Ordinary least squares regression using the observed data of y produces inconsistent estimates of β. The maximum likelihood method is used to estimate selection models.

The log-likelihood function of the Heckman selection model is written as:

$$\ell = \sum_{i \in [z_i = 0]} \ln\left[1 - \Phi(w_i'\gamma)\right]$$

$$+ \sum_{i \in [z_i = 1]} \left\{ \ln \phi\left(\frac{yi - x_1'\beta}{\sigma} \right) - \ln \sigma + \ln \Phi\left(\frac{w_i'\gamma + \rho \dfrac{yi - x_1'\beta}{\sigma}}{\sqrt{1 - \rho^2}} \right) \right\}$$

Because cross-sectional data often are used in conjunction with censored response models, the vetting of heteroscedasticity is important.

SAS allows the use of HETERO in this regard.

The heteroscedastic regression model supported by PROC QLIM is:

$$y_i = x_i'\beta + \varepsilon_i$$

$$\varepsilon_i \sim N(0, \sigma_i^2)$$

LINK=value

The functional form can be specified using the **LINK=** option. The following option values are allowed:

EXP – specifies the exponential link function

$$\sigma_i^2 = \sigma^2(1 + \exp(z_i'\gamma))$$

LINEAR – specifies the linear link function

$$\sigma_i^2 = \sigma^2(1 + z_i'\gamma)$$

When the LINK= option is not specified, the exponential link function is specified by default.

Example of Use of Hetero with Proc Qlim

Proc qlim;

Model yesvm=pub12 pub34 pub5 private school years linc lptcon / discrete (normal)

Hetero yesvm ~ inc / link = linear; (or exp)

Section 13.10

Sample Problem with the Heckman Sample Selection Procedure

Program for Heckman Sample Selection Procedure
Example with Purchase of Bottled Water

```
options nodate;
proc means data=botwater n mean median std min max;
      var bwgallons drinkbw bwexp hincome east midwest
south west white black asian other;
run;
proc reg data=botwater;
      model bwexp = hincome east midwest south black white
/ dw dwprob;
      test black=0, white=0;
      test east=0, midwest=0, south=0;
run;
proc qlim data=botwater;
 model drinkbw = hincome east midwest south black white /
discrete(d=normal);
 output out=probitresults marginal mills prob xbeta;
 * goodness-of-fit test;
test hincome=0, east=0, midwest=0, south=0, black=0, white=0
/ all;
* test of influence of region;
test east=0, midwest=0, south=0 / lr;
* test of influence of race;
test black=0, white=0 / lr;
```

```
proc means n mean median; var meff_p2_hincome
meff_p2_east meff_p2_midwest
      meff_p2_south meff_p2_black meff_p2_white
xbeta_drinkbw mills_drinkbw;
run;
* cutoff equal to # of times drinkbw=1 relative to the
sample size (7195);
data final; set probitresults;

ap00=0; ap01=0; ap10=0; ap11=0;

if prob_drinkbw < .6807505 and drinkbw=0 then ap00=1;
if prob_drinkbw < .6807505 and drinkbw=1 then ap10=1;
if prob_drinkbw > .6807505 and drinkbw=0 then ap01=1;
if prob_drinkbw > .6807505 and drinkbw=1 then ap11=1;
proc means data=final n mean median sum ; var ap00 ap10
ap01 ap11 drinkbw;

run;
* Heckman sample selection model;
data heckman; set final;
if bwexp=0 or bwgallons=0 then delete;
```

```
data heckmanfinal; set heckman;
bwprice=bwexp/bwgallons;
lbwprice=log(bwprice);
lbwgallons=log(bwgallons);
lhincome=log(hincome);

•no adjustment for sample selection bias;

proc autoreg data=heckmanfinal;
model lbwgallons = lbwprice lhincome east midwest south
black white / dw=1 normal;
run;

•adjustment for sample selection bias;

proc autoreg data=heckmanfinal;
model lbwgallons = lbwprice lhincome east midwest south
black white mills_drinkbw / dw=1 normal;
output out=finalfinal residual=rlbwgallons;
run;
```

```
* adjustment for sample selection bias with correction for
heteroscedasticity;
* initially run the auxillary equation of residuals squared
against rhsheckman;

data theend; set finalfinal;
res2=rlbwgallons*rlbwgallons;
rhsheckman=xbeta_drinkbw*mills_drinkbw-
(mills_drinkbw*mills_drinkbw);

proc autoreg data=theend;
model res2=  rhsheckman;
run;
```

```
* if the coefficient associated with rhsheckman is not
statistically different from zero then OLS is the
appropriate estimation procedure;
* but if the coefficient associated with rhsheckman is
statistically different from zero then WLS (weighted least
squares) is the appropriate estimation procedure;
```

The MEANS Procedure

Variable	N	Mean	Median	Std Dev	Minimum	Maximum
bwgallons	7195	9.7570801	1.5234380	27.2933469	0	447.26410
drinkbw	7195	0.6807505	1	0.4662183	0	1
bwexp	7195	12.0702432	2.4400000	29.0989693	0	594.76000
hincome	7195	51740.24	47500	26254.90	5000	100000
east	7195	0.2034746	0	0.4026105	0	1
midwest	7195	0.2532314	0	0.4348926	0	1
south	7195	0.3432940	0	0.4748416	0	1
west	7195	0.2000000	0	0.4000278	0	1
white	7195	0.8354413	1	0.3708076	0	1
black	7195	0.1020153	0	0.3026895	0	1
asian	7195	0.0132036	0	0.1141538	0	1
other	7195	0.0493398	0	0.2165916	0	1

OLS estimates

The REG Procedure

Dependent Variable: bwexp

Number of Observations Read	7195
Number of Observations Used	7195

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	127800	21300	25.67	<.0001
Error	7188	5963720	829.67721		
Corrected Total	7194	6091520			

Root MSE	28.80412	R-Square	0.0210
Dependent Mean	12.07024	Adj R-Sq	0.0202
Coeff Var	238.63743		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	12.66184	1.59962	7.92	<.0001
hincome	1	0.00011934	0.00001298	9.20	<.0001
east	1	-2.20390	1.07343	-2.05	0.0401
midwest	1	-4.85773	1.02385	-4.74	<.0001
south	1	-3.02216	0.96672	-3.13	0.0018
black	1	1.69372	1.74369	0.97	0.3314
white	1	-5.05458	1.41847	-3.56	0.0004

```
The REG Procedure
Dependent Variable: bwexp

Durbin-Watson D                 1.991
Pr < DW                        0.3450
Pr > DW                        0.6550
Number of Observations           7195
1st Order Autocorrelation       0.004

NOTE: Pr<DW is the p-value for testing positive autocorrelation,
and Pr>DW is the p-value for testing negative autocorrelation.
```

NOTE absence of autocorrelation

The REG Procedure

Test 1 Results for Dependent Variable bwexp

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	18622	22.44	<.0001
Denominator	7188	829.67721		

Importance of Race

The REG Procedure

Test 2 Results for Dependent Variable bwexp

Source	DF	Mean Square	F Value	Pr > F
Numerator	3	6421.40511	7.74	<.0001
Denominator	7188	829.67721		

Importance of Region

The QLIM Procedure

Probit Model

Discrete Response Profile of drinkbw

Index	Value	Frequency	Percent
1	0	2297	31.92
2	1	4898	68.08

Model Fit Summary

Number of Endogenous Variables	1
Endogenous Variable	drinkbw
Number of Observations	7195
Log Likelihood	-4411
Maximum Absolute Gradient	1.33005
Number of Iterations	13
Optimization Method	Quasi-Newton
AIC	8835
Schwarz Criterion	8883

Goodness-of-Fit Measures

Measure	Value	Formula
Likelihood Ratio (R)	191.2	2 * (LogL - LogL0)
Upper Bound of R (U)	9012.5	- 2 * LogL0
Aldrich-Nelson	0.0259	R / (R+N)
Cragg-Uhler 1	0.0262	1 - exp(-R/N)
Cragg-Uhler 2	0.0367	(1-exp(-R/N)) / (1-exp(-U/N))
Estrella	0.0265	1 - (1-R/U)^(U/N)
Adjusted Estrella	0.0246	1 - ((LogL-K)/LogL0)^(-2/N*LogL0)
McFadden's LRI	0.0212	R / U
Veall-Zimmermann	0.0466	(R * (U+N)) / (U * (R+N))
McKelvey-Zavoina	0.0456	

N = # of observations, K = # of regressors

The QLIM Procedure

Probit Model

Parameter Estimates

Parameter	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	0.545429	0.077178	7.07	<.0001
hincome	1	0.000006245	0	.	.
east	1	-0.101377	0.049520	-2.05	0.0406
midwest	1	-0.187380	0.046788	-4.00	<.0001
south	1	-0.095416	0.044715	-2.13	0.0329
black	1	0.005261	0.085012	0.06	0.9507
white	1	-0.343437	0.069131	-4.97	<.0001

Test Results

Test	Type	Statistic	Pr > ChiSq	Label
Test0	Wald	.	<.0001	hincome = 0 , east = 0 , midwest = 0 , south = 0 , black = 0 , white = 0
Test0	L.R.	191.20	<.0001	hincome = 0 , east = 0 , midwest = 0 , south = 0 , black = 0 , white = 0
Test0	L.M.	201.39	<.0001	hincome = 0 , east = 0 , midwest = 0 , south = 0 , black = 0 , white = 0

Goodness-of-Fit Tests

Test	Type	Statistic	Pr > ChiSq	Label
Test1 Region	L.R.	16.22	0.0010	east = 0 , midwest = 0, south = 0
Test2 Race	L.R.	63.06	<.0001	black = 0 , white = 0

The MEANS Procedure

Variable	Label	N	Mean
Meff_P2_hincome	Marginal effect of hincome on the probability of drinkbw=2	7195	2.1814E-6
Meff_P2_east	Marginal effect of east on the probability of drinkbw=2	7195	-0.0354
Meff_P2_midwest	Marginal effect of midwest on the probability of drinkbw=2	7195	-0.0654
Meff_P2_south	Marginal effect of south on the probability of drinkbw=2	7195	-0.0333
Meff_P2_black	Marginal effect of black on the probability of drinkbw=2	7195	0.0018
Meff_P2_white	Marginal effect of white on the probability of drinkbw=2	7195	-0.1199
Xbeta_drinkbw	X * Beta of drinkbw	7195	0.4813
Mills_drinkbw	Inverse Mills ratio of drinkbw	7195	0.5252

Variable	Label	Median
Meff_P2_hincome	Marginal effect of hincome on the probability of drinkbw=2	2.2514E-6
Meff_P2_east	Marginal effect of east on the probability of drinkbw=2	-0.0365
Meff_P2_midwest	Marginal effect of midwest on the probability of drinkbw=2	-0.0675
Meff_P2_south	Marginal effect of south on the probability of drinkbw=2	-0.0343
Meff_P2_black	Marginal effect of black on the probability of drinkbw=2	0.0018
Meff_P2_white	Marginal effect of white on the probability of drinkbw=2	-0.1238
Xbeta_drinkbw	X * Beta of drinkbw	0.4500
Mills_drinkbw	Inverse Mills ratio of drinkbw	0.5351

The MEANS Procedure

Variable	N	Mean	Median	Sum
ap00	7195	0.3192495	0	2297.00
ap10	7195	0.3364837	0	2421.00
ap01	7195	0	0	0
ap11	7195	0.3442669	0	2477.00
drinkbw	7195	0.6807505	1.0000000	4898.00

Prediction-Success Table

	Actual	
Predicted	**0**	**1**
0	2297	2421
1	0	2477
	2297	4898

Correct predictions

- Percentage of correct predictions (2297 + 2477) / 7195 = 66.4%
- Percentage of correct predictions for those households who *did not* purchase bottled water 2297 / 2297 = 100%
- Percentage of correct predictions for those households who *did* purchase bottled water 2297 / 4898 = 50.6%

The AUTOREG Procedure

Second Stage OLS Estimates

No Adjustment for Sample Selection Bias

Dependent Variable bwgallons

Ordinary Least Squares Estimates

SSE	4752397.32	DFE	4890
MSE	971.86039	Root MSE	31.17468
SBC	47654.2692	AIC	47602.2965
MAE	15.9218627	AICC	47602.326
MAPE	751.246323	Regress R-Square	0.0566
Durbin-Watson	2.0173	Total R-Square	0.0566

Miscellaneous Statistics

Statistic	Value	Prob	Label
Normal Test	477996.585	<.0001	Pr > ChiSq

Durbin-Watson Statistics

Order	DW
1	2.0173

OLS Estimates

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	26.5644	2.0709	12.83	<.0001
bwprice	1	-4.7982	0.2988	-16.06	<.0001
hincome	1	0.0000454	0.0000171	2.66	0.0079
east	1	0.0363	1.3925	0.03	0.9792
midwest	1	-3.1048	1.3497	-2.30	0.0215
south	1	-2.6573	1.2486	-2.13	0.0334
black	1	0.1895	2.1416	0.09	0.9295
white	1	-4.3204	1.7485	-2.47	0.0135

Marginal effects with no adjustment for sample selection bias

The AUTOREG Procedure
Second Stage OLS Estimates
with Adjustment for Sample Selection Bias

Dependent Variable   bwgallons

Ordinary Least Squares Estimates

SSE	4748336.46	DFE	4889
MSE	971.22857	Root MSE	31.16454
SBC	47658.5787	AIC	47600.1094
MAE	15.9199129	AICC	47600.1463
MAPE	750.083616	Regress R-Square	0.0575
Durbin-Watson	2.0169	Total R-Square	0.0575

Miscellaneous Statistics

Statistic	Value	Prob	Label
Normal Test	474157.186	<.0001	Pr > ChiSq

Durbin-Watson Statistics

Order	DW
1	2.0169

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	-32.7893	29.1006	-1.13	0.2599	
bwprice	1	-4.7942	0.2987	-16.05	<.0001	
hincome	1	0.000419	0.000183	2.28	0.0224	
east	1	-5.9751	3.2528	-1.84	0.0663	
midwest	1	-14.5349	5.7504	-2.53	0.0115	
south	1	-8.2807	3.0201	-2.74	0.0061	
black	1	0.9798	2.1755	0.45	0.6525	
white	1	-23.6028	9.5907	-2.46	0.0139	
Mills_drinkbw	1	118.3052	57.8569	2.04	0.0409	Inverse Mills ratio of drinkbw

Existence of sample selection bias
No interpretation of this coefficient

The AUTOREG Procedure

Dependent Variable res2

Ordinary Least Squares Estimates

SSE	2.24817E11	DFE	4896
MSE	45918517	Root MSE	6776
SBC	100327.287	AIC	100314.294
MAE	1524.77038	AICC	100314.296
MAPE	103691237	Regress R-Square	0.0002
Durbin-Watson	2.0172	Total R-Square	0.0002

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	989.7232	99.6189	9.94	<.0001
rhsheckman	1	507.5965	586.4630	0.87	0.3868

No evidence of heteroscedasticity
OLS is the appropriate estimation technique

	$\hat{\gamma}$ (Probit Coefficient)	$\hat{\beta}$ (Second-Stage Coefficient)	Appropriate Marginal Effect[a] (with adjustment for sample selection bias)	Appropriate Marginal Effect (no adjustment for sample selection bias)
Income	0.000006245	0.000419	0.0000285	0.0000454
East	-.101377	-5.9751	0.3648	0.0363
Midwest	-.187380	-14.5349	-2.8166	-3.1048
South	-.095416	-8.2807	-2.3136	-2.6573
Black	0.005261	0.9798	0.6508	0.1895
White	-.343437	-23.6028	-2.1250	-4.3204
Price		-4.7942	-4.7942	-4.7982
α	-----	118.3052		
Xbeta (Probit)	0.4813	from probit analysis		
IMR	0.5252	from probit analysis		

[a] $\hat{\beta} - \alpha\hat{\gamma}\left\{XBeta(\text{Probit})\ IMR + IMR^2\right\}$