Paper 279-25

Qualitative and Limited Dependent Variable Models Using the New QLIM Procedure

Minbo Kim, SAS Institute Inc., Cary, NC

Abstract

There are many applications of limited dependent variable and discrete choice models in a wide variety of areas, including economics, finance, marketing, political science, and sociology, to name a few. This paper introduces multinomial discrete choice and simultaneous equations models and presents how to estimate these models using PROC QLIM. Three data sets are used for illustration.

Introduction

It is not uncommon to encounter econometric models where dependent variables are limited or qualitative. These discrete choice and limited dependent variable models need to be analyzed using more complicated methods than usual continuous models. The SAS/ETS® QLIM procedure is developed to analyze mainly cross-sectional data, though you can use the QLIM procedure for panel or time-series data. PROC QLIM can analyze the following models:

- binary probit and logit
- ordinal probit
- binary probit and logit with heteroskedasticity
- multinomial logit
- nested logit
- mixed multinomial logit
- univariate and bivariate poisson regression
- negative binomial regression
- tobit (censored and truncated regression)
- simultaneous equations model
- endogenous switching regression

Introductory Example

For multinomial data analysis, it is important to organize data series in an appropriate way. Here is a simple binary data set that illustrates how you can estimate the multinomial logit model using PROC QLIM. The first five observations of a simulated data set (Ben-Akiva and Lerman 1985, p. 88) are shown as follows:

Example of Binary Choice Data

```
id    auto    transit    ttdif    cchoice    alt

 1    52.9       4.4      48.5    Transit     0
 2     4.1      28.5     -24.4    Transit     0
 3     4.1      86.9     -82.8    Auto        1
 4    56.2      31.6      24.6    Transit     0
 5    51.8      20.2      31.6    Transit     0
```

The two travel alternatives in the data set are auto and transit. The attribute of travel time is recorded as AUTO for automobile travel and TRANSIT for transit travel. The binary choice data can be estimated using the conditional logit model. For conditional logit estimation, we created the new data set and the first ten observations are shown in the following table:

Rearranged Binary Data

```
id    autodum    ttime    cchoice    mode    choice

 1       1        52.9    Transit      1        0
 1       0         4.4    Transit      2        1
 2       1         4.1    Transit      1        0
 2       0        28.5    Transit      2        1
 3       1         4.1    Auto         1        1
 3       0        86.9    Auto         2        0
 4       1        56.2    Transit      1        0
 4       0        31.6    Transit      2        1
 5       1        51.8    Transit      1        0
 5       0        20.2    Transit      2        1
```

The new variable (TTIME) takes the value of automobile travel time (AUTO) for the first record of each individual while it contains transit travel time (TRANSIT) for the second record. We use an alternative-specific constant (AUTODUM) for conditional logit estimation.

The probability of choosing the auto mode (MODE=1) is

$$P(y_i = 1) = \frac{\exp(\beta_1 \mathrm{AUTODUM} + \beta_2 \mathrm{TTIME})}{1 + \exp(\beta_1 \mathrm{AUTODUM} + \beta_2 \mathrm{TTIME})}$$

PROC QLIM estimates the conditional logit model with the TYPE=CLOGIT option. Maximum likelihood estimation uses the Newton-Raphson optimization technique (OPTMETHOD=NR), and the standard errors are calculated from the Hessian metrix (COVEST=HESS).

```
proc qlim data=travel2;
   model mode = autodum ttime / noint type=clogit id=id
       choice=choice covest=hess optmethod=nr;
   endogenous discrete=(mode 1 2);
run;
```

Binary Choice Modeling Using Multinomial Logit

Variable Name	Parameter Estimate	Standard Error	t statistic
AUTODUM	-0.2376	0.7505	-0.32
TTIME	-0.0531	0.0206	-2.57
Summary Table			
Log L	-6.166		
# Observations	21		
# Records	42		
AIC	16.3321		

As explained in the next section, the binary choice logit can estimate the binomial choice model using the following specification:

$$\log \left[\frac{P(y_i = 1)}{1 - P(y_i = 1)} \right] = (\mathbf{x}_{iA} - \mathbf{x}_{iT})\boldsymbol{\beta} = \mathbf{x}_{iD}\boldsymbol{\beta}$$

where $\mathbf{x}_{iA}$ = (1, AUTO), $\mathbf{x}_{iT}$ = (0, TRANSIT), and $\mathbf{x}_{iD}$ = (1, TTDIF). The explanatory variables, $\mathbf{x}_{iA}$ and $\mathbf{x}_{iT}$, are stored in the separate line with the variable TTIME when multinomial modeling is used. For binomial logit estimation, one observation per person is needed since the probability of transit choice, $P(y_i = 0)$, can be calculated from the probability of automobile choice, $P(y_i = 1)$.

The binary logit and probit models are estimated using the maximum likelihood method. Parameter estimates are summarized in the following table. The predicted probabilities are generated by the OUTPUT statement using the formula

$$P(y_i = 1) = \frac{\exp(\hat{\beta}_1 + \hat{\beta}_2 \mathrm{TTDIF})}{1 + \exp(\hat{\beta}_1 + \hat{\beta}_2 \mathrm{TTDIF})} \quad \text{(logit)}$$

$$P(y_i = 1) = \Phi(\hat{\beta}_1 + \hat{\beta}_2 \mathrm{TTDIF}) \quad \text{(probit)}$$

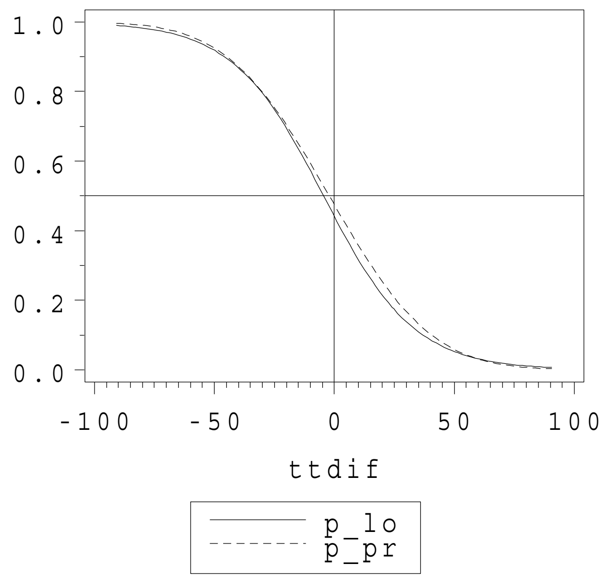
where $\Phi(\cdot)$ is the distribution function of standard normal variables and $\hat{\beta}_1$ and $\hat{\beta}_2$ are maximum likelihood estimates.

```
proc qlim data=travel1;
   model alt = ttdif /
         type=blogit covest=hess optmethod=nr;
   endogenous discrete=(alt 0 1);
   output out=blg p=p_lo;
   model alt = ttdif /
         type=bprobit covest=hess optmethod=nr;
   endogenous discrete=(alt 0 1);
   output out=bpr p=p_pr;
run;
```

Binary Choice Modeling Using Logit & Probit

Variable Name	Parameter Estimate	Standard Error	t statistic
	Logit Estimates		
INTERCEPT	-0.2376	0.7505	-0.32
TTDIF	-0.0531	0.0206	-2.57
Summary Table			
Log L	-6.166		
# Observations	21		
# Records	21		
AIC	16.3321		
	Probit Estimates		
INTERCEPT	-0.0644	0.3992	-0.16
TTDIF	-0.0300	0.0103	-2.92
Summary Table			
Log L	-6.1652		
# Observations	21		
# Records	21		
AIC	16.3303		

The binomial logit estimates are exactly the same as the preceding multinomial logit estimates, as expected. The probit estimates are different, since the error variance of probit is normalized as 1. The prediction probabilities are displayed. The plot indicates that auto mode would be chosen if the travel by automobile takes less time than the travel by transit (TTDIF < 0).



Discrete Choice Modeling

Binary choice modeling is used to analyze binary response variable models. The regression model with binary responses can be written as

$$y_i^* = \mathbf{x}_i'\beta + \epsilon_i$$

where only the sign of the dependent variable is observed as follows:

$$\begin{array}{rcl} y_i & = & 1 \quad \text{if} \quad y_i^* > 0 \\ & = & 0 \quad \text{otherwise} \\ \epsilon_i & \sim & \mathrm{standardnormalwithCDF:} \\ & & \Phi(x) = \displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)dt \quad \text{(probit)} \\ \mathrm{or} & & \\ & & \mathrm{logisticwithCDF:} \\ & & \Lambda(x) = \dfrac{\exp(x)}{1+\exp(x)} \quad \mathrm{(logit)} \end{array}$$

For the binary logit model, the probability of one choice is denoted

$$P(y_i = 1) = \frac{\exp(\mathbf{x}_i'\beta)}{1+\exp(\mathbf{x}_i'\beta)}$$

The probability of an alternative choice can be derived easily as $P(y_i = 0) = 1 - P(y_i = 1)$.

When the dependent variable takes multiple discrete values, the multinomial logit model can be used to analyze unordered categorical response variables. Standard multinomial (or conditional) logit models can be used to calculate a probability of choice j among $J + 1$ alternatives.

$$P(y_i = j) = \frac{\exp(\mathbf{x}_{ij}'\boldsymbol{\beta})}{1 + \sum_{k=1}^{J} \exp(\mathbf{x}_{ik}'\boldsymbol{\beta})}$$

The conditional logit model can also be derived from utility maximization. Let the random utility function be defined as

$$U_{ij} = V_{ij} + \epsilon_{ij}$$

where V_{ij} is a non-stochastic utility function. If we assume linear utility function, then $V_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta}$. The error disturbances are assumed to have iid Gumbel (log Weibull or type I extreme value) distribution with distribution function, $\exp(-\exp(-\epsilon_{ij}))$. Then the event $\{y_i = j\}$ can be expressed using a random utility function as follows:

$$U_{ij} \geq \mathbf{max}_{k \in C_i, k \neq j} U_{ik}$$

Using properties of the Gumbel distribution, the probability of choosing an alternative j from n_i choices of individual i can be derived from utility maximization:

$$\begin{array}{rcl} P_i(j) & = & P[\mathbf{x}_{ij}'\boldsymbol{\beta} + \epsilon_{ij} \geq \mathbf{max}_{k \in C_i}(\mathbf{x}_{ik}'\boldsymbol{\beta} + \epsilon_{ik})] \\ & = & \dfrac{\exp(\mathbf{x}_{ij}'\boldsymbol{\beta})}{\sum_{k \in C_i} \exp(\mathbf{x}_{ik}'\boldsymbol{\beta})} \end{array}$$

When explanatory variables contain only individual characteristics, the multinomial logit model is sometimes defined by

$$\begin{array}{rcl} P(y_i = j) & = & \dfrac{\exp(\mathbf{x}_i'\boldsymbol{\beta}_j)}{1 + \sum_{k=1}^{J} \exp(\mathbf{x}_i'\boldsymbol{\beta}_k)}, \qquad j \geq 2 \\ P(y_i = 1) & = & \dfrac{1}{1 + \sum_{k=1}^{J} \exp(\mathbf{x}_i'\boldsymbol{\beta}_k)} \end{array}$$

This type of multinomial choice modeling has a couple of weaknesses: it has too many parameters and it is difficult to interpret.

The odds of success or failure is an important concept in binary and multinomial modeling. The log-odds ratio is defined

$$\log \left[\frac{P(y_i = j)}{P(y_i = k)} \right] = (\mathbf{x}_{ij} - \mathbf{x}_{ik})'\boldsymbol{\beta}$$

This expression also shows that the multinomial logit has the property of the independence of irrelevant alternatives (IIA). In other words, the log-odds ratio is only affected by choice j and k.

Multidimensional Discrete Choice Modeling

Suppose a decision maker has a multidimensional choice C_i, where $C_i = C_1 \times C_2 \times \cdots C_L - C_i^*$ and C_i^* is defined as the set of infeasible elements in $C_1 \times C_2 \times \cdots C_L$ for the decision maker i. When elements in C_i share unobserved and observed attributes, utilities of these elements are not independent. In this case, nested logit modeling is appropriate for the multinomial choice data analysis. Multinomial probit modeling can be another approach. It is convenient to introduce several notations to explain nested logit modeling. Assume that there are L levels, with 1 representing the lowest branch of the tree and L denoting the stem of the tree. The index of a node at level h is represented by $(j_h, \cdots, j_L)$. Let π_h denote a node at level $h+1$, where $\pi_h = (j_{h+1}, \cdots, j_L)$. The choice set C_{π_h} contains choices that belong to branches below the node π_h. The notation C_{π_h} can also be used to express a set of indices below π_h. Note that C_{π_0} is a set with a single element, while C_{π_L} represents a choice set containing all possible alternatives. The probability of choice at level h is written

$$P_i(j_h|\pi_h) =$$

$$\frac{\exp[(\mathbf{x}^h_{i,j_h\pi_h})'\boldsymbol{\beta}^h + \sum_{k\in C_{j_h\pi_h}} \Im_{k,j_h\pi_h}]}{\sum_{j\in C_{\pi_h}} \exp[(\mathbf{x}^h_{i,j\pi_h})'\boldsymbol{\beta}^h + \sum_{k\in C_{j_h\pi_h}} \Im_{k,j\pi_h}]}$$

where $\mathbf{x}^h_{i,\pi_{h-1}}$ is the vector of variables for observation i related with the node π_{h-1} and

$$\Im_{k,j_h\pi_h} = I_{k,j_h\pi_h}\theta_{k,j_h\pi_h}$$

The inclusive value of the level $h+1$ is written

$$\begin{aligned} I_{\pi_h} &= \ln \sum_{j\in C_{\pi_h}} \exp[(\mathbf{x}^h_{i,j\pi_h})'\boldsymbol{\beta}^h + \\ &\qquad \sum_{k\in C_{j_h\pi_h}} I_{k,j\pi_h}\theta_{k,j\pi_h}] \end{aligned}$$

$$0 \le \theta_{k,\pi_1} \le \cdots \le \theta_{k,\pi_{L-1}} \le 1$$

When the decision level is at 1, there are no inclusive values. Therefore, the conditional probability is defined as

$$P_i(j_1|\pi_1) = \frac{\exp[(\mathbf{x}^h_{i,j_1\pi_1})'\boldsymbol{\beta}^1]}{\sum_{j\in C_{\pi_1}} \exp[(\mathbf{x}^1_{i,j\pi_1})'\boldsymbol{\beta}^1]}$$

The utility function of a decision maker can be specified as

$$U_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \xi_{ij} + \epsilon_{ij}$$

where the error component, ξ_{ij}, is correlated among alternatives and heteroskedastic and another error component, ϵ_{ij}, independently and identically distributed. When we assume that $\xi_{ij} = \mathbf{z}'_{ij}\boldsymbol{\gamma}$, random coefficients, $\boldsymbol{\gamma}$, have the mixing distribution with the probability density function $f(\boldsymbol{\gamma}|\boldsymbol{\theta})$. The choice probability of an alternative j is

$$P_i(j) = \int Q_i(j|\boldsymbol{\gamma})f(\boldsymbol{\gamma}|\boldsymbol{\theta})d\boldsymbol{\gamma}$$

where

$$Q_i(j|\boldsymbol{\gamma}) = \frac{\exp(\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\boldsymbol{\gamma})}{\sum_{k\in C_i}\exp(\mathbf{x}'_{ik}\boldsymbol{\beta} + \mathbf{z}'_{ik}\boldsymbol{\gamma})}$$

Mixed multinomial logit models are estimated by simulation-based methods (simulated maximum likelihood or method of simulated moments). The simulated probability is written

$$\tilde{P}_i(j) = \frac{1}{S}\sum_{s=1}^{S} Q_i(j|\boldsymbol{\gamma}^s)$$

where S is the number of simulation replications drawn from density $f(\cdot)$ and $\boldsymbol{\gamma}^s$ is a simulated value. The simulated log-likelihood function is calculated as $\log\tilde{L} = \sum_{i=1}^{N}\log(\tilde{P}_i(j))$. The simulated probability is an unbiased estimate of $P_i(j)$, and its variance decreases as replications (S) increase. Therefore, it is critical to find out the most efficient method of simulation. Train (1999) argues that quasi-random numbers (Halton sequences) can provide much better simulation-based estimation method than pseudo random numbers. His finding is that 100 Halton draws have smaller simulation variance than 1000 random draws. However, the properties of Halton sequences have not been fully investigated.

Multinomial Logit Example

Hensher and Greene analyze travel mode choice for travel between Sydney and Melbourne. They used data on four travel modes: airplane, train, bus, and car. The data set contains 210 observations. Refer to Greene (2000) for more detailed data descriptions. In this example, we assume that the choice of transportation mode is a function of terminal waiting time (TTIME), in-vehicle time (INVT), in-vehicle cost

(INVC), generalized cost measure (GC), and income (INCOME). The variable INCOME contains zero value if the choice is not airplane. Original data is rescaled dividing by 100. The multinomial choice models are estimated using the conditional logit, nested logit, and mixed multinomial logit log-likelihood functions. The log-likelihood function of the mixed multinomial logit is calculated using simulation with 125 Halton sequence draws. All coefficients are assumed to be random with normal distribution when the mixed logit is estimated. In the following table, the mean and standard deviation of random coefficients are shown while the conditional and nested logit models estimate fixed parameters. It is interesting to observe that parameter estimates of conditional logit and mixed logit models are almost identical when we assume that the random parameters have the normal distribution. Asymptotic standard errors are reported in parentheses.

Multinomial Logit Modeling

Parameter	Conditional Logit	Nested Logit	Mixed Logit
TTIME	-3.5606	-2.4540	-3.5606
	(0.4725)	(0.4253)	(0.4725)
TTIME$_\sigma$			0.0213
			(0.6408)
INVC	-2.6820	-0.4933	-2.6820
	(1.4694)	(3.7622)	(1.4695)
INVC$_\sigma$			0.0093
			(1.0168)
INVT	-0.6202	-1.6618	-0.6202
	(0.1857)	(0.5769)	(0.1857)
INVT$_\sigma$			-0.0001
			(0.0506)
GC	3.4269	2.4142	3.4269
	(1.3884)	(3.6993)	(1.3885)
GC$_\sigma$			-0.0046
			(0.7071)
INCOME	1.2462	3.7627	1.2464
	(0.6637)	(0.9876)	(0.6637)
INCOME$_\sigma$			0.0201
			(0.8404)
INCLUDE_1		2.1817	
		(0.4304)	
INCLUDE_2		0.1273	
		(0.0946)	
$\log L$	-242.394	-193.666	-242.393

The QLIM statement is given as follows:

```
proc qlim data=travel id=id;
   /*-- conditional logit --*/
   model mode = ttime invc invt gc income /
         noint type=clogit choice=decision;
   endogenous discrete=(mode 1 2 3 4);
   /*-- nested logit --*/
   model upmode mode = ttime invc invt gc income /
         type=nestedlogit choice=decision;
   utility u(1, 2 3 4 @ 2) = ttime invc invt gc,
           u(2, 1 2) = income;
   endogenous discrete=(mode 1 2 3 4, upmode 1 2);
   /*-- mixed logit --*/
   model mode = ttime invc invt gc income /
         type=mixedlogit choice=decision
         mixed=(normalvar=ttime invc invt gc income,
                randnum=halton);
   endogenous discrete=(mode 1 2 3 4);
run;
```

Heckman's Two-Equation Model Example

Heckman's two-equation simultaneous system is defined

$$\begin{aligned} y_{1i} &= y_{2i}^*\beta_1 + \mathbf{x}_{1i}'\gamma_1 + \delta_1 d_i + u_{1i} \\ y_{2i}^* &= y_{1i}\beta_2 + \mathbf{x}_{2i}'\gamma_2 + \delta_2 d_i + u_{2i} \\ d_i &= \begin{cases} 1 & \text{if } y_{2i}^* > 0 \\ 0 & \text{otherwise} \end{cases} \\ \delta_2 &= -\beta_2\delta_1 \end{aligned}$$

where $\mathbf{x}_{1i}$ and $\mathbf{x}_{2i}$ are $K_1 \times 1$ and $K_2 \times 1$ exogenous variable vectors, and (u_{1i}, u_{2i}) are *iid* bivariate normal variables. The reduced form of this system of two equations can be written

$$\begin{aligned} \mathbf{y}_1 &= \delta_1\mathbf{d} + X\boldsymbol{\pi}_1 + \mathbf{e}_1 \\ \mathbf{y}_2^* &= X\boldsymbol{\pi}_2 + \mathbf{e}_2 \end{aligned}$$

where X consists of all distinct column vectors of X_1 and X_2 and $(e_{1i}, e_{2i}) \sim N\left(0, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & 1 \end{bmatrix}\right)$. Amemiya (1978) suggested the two-stage estimation method using the following equations:

$$\boldsymbol{\pi}_1 = J_1\boldsymbol{\gamma}_1 + \boldsymbol{\pi}_2\beta_1$$

$$\boldsymbol{\pi}_2 = J_2\boldsymbol{\gamma}_2 + \boldsymbol{\pi}_1\beta_2$$

where J_1 and J_2 are selection matrices for X_1 and X_2, respectively.

Copley et al. (1994) studied the relation of quality of audit services (QUALINDX) with audit fees (LNFEE) using Heckman's two-equation simultaneous equations model with $\delta_1 = 0$. They find that the audit fee is positively related with quality of service in the audit supply equation, while there is a negative relationship between the demand for audit quality and audit service fees. However, single equation modeling does not reveal this relationship because of simultaneous equations bias. Their specification of simultaneous equations system is the following:

$$\begin{aligned} \text{QUALINDX} = f_1(&\text{LNSIZE}, \text{FINOFFCL}, \\ &\text{GOVT}, \text{LNFEE}) \end{aligned}$$

$$\text{LNFEE} = f_2(\text{LNSIZE}, \text{LNBIDS}, \text{COSTWGHT}, \text{TENURE}, \text{QUALINDX})$$

Refer to Copley et al. for more details on data and variable definition. We randomly selected 70% of the data set (sample size = 118) that was used by Copley et al. and reported two-stage OLS and GLS estimates. The estimates in the demand equation show sizable difference when we compare with those of Copley et al. The OLS coefficient estimate of LNFEE is not significant though the sign is the same as the GLS. Deis and Hill (1998) carried out bootstrapping. Their finding is that some asymptotic t-statistics of the Amemiya's GLS estimates are inflated by 55% when they use Copley et al.'s data set with sample size 118.

```
proc qlim data=account system;
   model qualindx = lnsize finoffcl govt lnfee;
   model lnfee = lnsize lnbids costwght tenure qualindx;
   endogenous lnfee discrete=(qualindx 0 1);
   instruments lnsize lnbids finoffcl
               govt costwght tenure;
run;
```

Simultaneous Equations Estimates of Demand Function for Audit Quality

Variable Name	Parameter Estimate	Standard Error	t statistic
GLS Estimates (sample size = 77)			
INTERCEPT	1.9326	2.4245	0.80
LNSIZE	0.7934	0.4924	1.61
FINOFFCL	0.5279	0.5911	0.89
GOVT	2.2846	1.0190	2.24
LNFEE	-1.5916	1.0341	-1.54
OLS Estimates (sample size = 77)			
INTERCEPT	0.8625	2.0647	0.42
LNSIZE	0.4248	0.5008	0.85
FINOFFCL	0.4851	0.3270	1.48
GOVT	1.7400	0.8290	2.10
LNFEE	-0.8607	1.0715	-0.80

Simultaneous Equations Estimates of Supply Function for Audit Quality

Variable Name	Parameter Estimate	Standard Error	t statistic
GLS Estimates (sample size = 77)			
INTERCEPT	2.5884	1.0489	2.47
LNSIZE	0.4208	0.0730	5.76
LNBIDS	-0.1123	0.2235	-0.50
COSTWGHT	-0.4687	0.3584	-1.31
TENURE	0.0887	0.0428	2.07
QUALINDX	0.4944	0.3188	1.55
OLS Estimates (sample size = 77)			
INTERCEPT	2.5543	1.0514	2.43
LNSIZE	0.4292	0.0752	5.70
LNBIDS	-0.1209	0.2243	-0.54
COSTWGHT	-0.4626	0.3586	-1.29
TENURE	0.0892	0.0428	2.09
QUALINDX	0.5021	0.3192	1.57

Conclusion

In this paper, three examples are given to introduce how the QLIM procedure can be used to solve real problems. However, there are many other interesting features. For example, you can use PROC QLIM for count data and limited dependent variable modeling. Predicted values and marginal effects are calculated with the OUTPUT statement in the QLIM procedure. You can also use PROC QLIM to fit switching regression models.

References

Amemiya, T. (1978), "The Estimation of a Simultaneous Equation Generalized Probit Model," *Econometrica*, 46, 1193–1205.

Amemiya, T. (1985), *Advanced Econometrics*, Cambridge: Harvard University Press.

Ben-Akiva, M. and Lerman, S.R. (1985), *Discrete Choice Analysis*, Cambridge: MIT Press.

Brownstone, D. and Train, K. (1999), "Forecasting New Product Penetration with Flexible Substitution Patterns," *Journal of Econometrics*, 89, 109–129.

Copley, P.A., Doucet, M.S., and Gaver, K.M. (1994), "A Simultaneous Equations Analysis of Quality Control Review Outcomes and Engagement Fees for Audits of Recipients of Federal Financial Assistance," *The Accounting Review*, 69, 244–256.

Deis, D.R. and Hill, R.C. (1998), "An Application of the Bootstrap Method to the Simultaneous Equations

Model of the Demand and Supply of Audit Services," *Contemporary Accounting Research*, 15, 83–99.

Greene, W.H. (2000), *Econometric Analysis*, 4th ed., Upper Saddle River, N.J.: Prentice Hall.

Heckman, J.J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46, 931–959.

Lee, L.-F. (1981), "Simultaneous Equations Models with Discrete and Censored Dependent Variables," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. C.F. Manski and D. McFadden, Cambridge: MIT Press.

Train, K. (1999), "Halton Sequences for Mixed Logit," working paper, University of California, Berkeley.

Contact Information

Minbo Kim, SAS Institute Inc., SAS Campus Drive, Cary, NC 27513. Email Minbo.Kim@sas.com