May, 2006

A Review of LIMDEP 9.0 and NLOGIT 4.0

Joseph Hilbe, *Arizona State University*

A Review of LIMDEP 9.0 and NLOGIT 4.0

Joseph M. Hilbe

1. OVERVIEW

LIMDEP, an acronym for "limited dependent variable models," was initially developed some 25 years ago by William Greene of New York University's Stern School of Business. His initial goal was to model Tobit regression and related models. Throughout the subsequent years, Greene has continued to maintain responsibility for the programming of all statistical procedures and functions. Marketed under the corporate name, Econometric Software, Inc., LIMDEP has steadily grown to become perhaps the world's premiere econometric software package.

NLOGIT is an extension of LIMDEP's nested logit model, which itself is an extension of LIMDEP's multinomial regression command. In 1996 Greene decided to develop NLOGIT as stand-alone package having extensive discrete choice capabilities. Included in the package are commands for nested logit models, multinomial logit and probit models, heteroscedastic extreme value models, random parameters logit models, covariance heterogeneity models, and latent class models. Version 4.0 has added generalized nested logit, kernel logit, random parameters mixed logit, and a generalized maximum entropy estimator. Additional enhancements have been added as well.

A unique feature inherent in the NLOGIT package is its simulation facility. Analysts may use a built-in simulator to fit a model, use the model to predict a set of choices for a sample, and then test how those particular choices would change if various attributes of the choices are altered. A number of other simulation options are provided to the user. This is a very nice feature of the package, and is one that many would find reason enough to purchase the program.

2. LIMDEP: THE PACKAGE

LIMDEP Version 9.0 and NLOGIT 4.0 are planned to be released in early summer 2006. The software reviewed here has been finalized; the delay in the release relates to some additions to, and subsequent printing of, the reference manuals.

2.1 Costs

The respective single-user costs (U.S. dollars) are, as of this writing, estimated to be:

Academic

- LIMDEP (only): $595 (update from 8.0: $395)
- NLOGIT (includes LIMDEP): $795 (update from 8.0/3.0: $595, $495

Government/nonprofit

- LIMDEP (only): $795 (update from 8.0: $495)
- NLOGIT (includes LIMDEP): $1095 (update from 8.0/3.0: $795, $595)

Corporate

- LIMDEP (only): $895 (update from 8.0: $595)
- NLOGIT (includes LIMDEP): $1195 (update from 8.0/3.0: $895, $695)

Extra Reference Manuals: LIMDEP $ 99; NLOGIT (includes LIMDEP): $125

2.2 Reference Manuals

- LIMDEP 9.0 Reference Guide
- LIMDEP 9.0 Econometric Modeling Guide, Volume 1
- LIMDEP 9.0 Econometric Modeling Guide, Volume 2
- NLOGIT 4.0 Reference Guide

Documentation for LIMDEP/NLOGIT consists of some 2,000 pages of reference and modeling guide information. Included are the typical instructions of setup, data management, and program syntax guidelines. In addition, the manuals provide LIMDEP users with summary background in econometrics as well as numerous examples of application. Complete technical details are provided for each procedure, including historical background, mathematical theory, and detailed explanation and examples of the various options. Many of the procedures in LIMDEP/NLOGIT cannot be found in other commercial software. More than a few are cutting-edge models.

I should mention a recently published text that can well be considered a reference text for NLOGIT. *Applied Choice Analysis: A Primer* (2005) is authored by David Hensher (University of Sidney), John Rose (University of Sidney), and LIMDEP/NLOGIT author William Greene (New York University). A hefty 717 pages in length, it is published by Cambridge University Press as a paperback. The text is readable by anyone having taken (and presumably passed) an Introductory to Statistics course or having an equivalent background. In a step-by-step manner, the text proceeds from defining mean and variance to delving into complex nested and mixed logit regression models. Case studies are provided throughout the text with the aim of making the concepts accessible to the reader. It fulfills this goal. But for the purpose of this review, the text can well be considered an ancillary Econometric Guide for NLOGIT.

2.3 System Requirements

Both LIMDEP and NLOGIT are MS Windows programs and Windows 98 or higher is required. Installation requires approximately six megabytes of disk space. Sixteen megabytes of RAM is a minimum needed to run programs effectively. As with other statistical packages, the more memory available, the faster the calculations will run.

Joseph M. Hilbe is Software Reviews Editor for *The American Statistician* and has the dual position of Professor Emeritus, University of Hawaii, and Adjunct Professor of Sociology and Statistics, Arizona State University, Tempe, AZ (E-mail: *jhilbe@aol.com*). I wish to thank Professor William Greene, author of LIMDEP and NLOGIT, for his kind assistance in providing me with a crash course in using the two packages. I needed this initial help when attempting to demonstrate some of the more complex capabilities of the software. I also wish to thank the referees who offered valuable feedback, especially with respect to the type of examples used to demonstrate the capabilities of the two packages.

 DOI: 10.1198/000313006X110492

LIMDEP/NLOGIT allows a maximum of 900 variables and as many observations as is permitted by memory. There is no set limitation. Commands, when entered on the command line, may contain up to 2,500 characters. With respect to models, the number of allowable parameters is 150. This should be sufficient for nearly every realistic model. Additionally, SURE and three-stage least-square models (3SLS) allow up to 30 equations. WALD (a command providing for the calculation of nonlinear functions of parameters and their standard errors, such as marginal effects for binary choice models), generalized least squares (NSURE), and generalized mixed models (GMM) each have a 20-equation limit.

The majority of LIMDEP/NLOGIT models either are, or provide for, panel data modeling. The packages have no set limits on the number of groups or clusters in fixed and random effects models. Limitation is due only to computer memory. Moreover, fixed and random effects models may have up to 150 regressors in both and 50,000 and unlimited numbers of groups, respectively. The time series/cross-section version of the seemingly unrelated regressions model may have up to 100 groups. Ten intervals are allowed for time varying covariates in proportional hazard models.

Finally, LIMDEP provides the user with the possibility of having 100 active matrices in memory at a time as well as 50 defined scalars. Matrices allow up to 50,000 cells.

NLOGIT, as a discrete choice modeling facility, permit 100 alternatives, with 5 trunks, 10 limbs, and 25 branches per tree. Other limitations and system capabilities can be found in either the reference manual or on the Web site: *www.nlogit.com*.

3. ECONOMETRIC MODELING GUIDES

I think it appropriate to detail the contents of the two Econometric Modeling Guides that come with the software. As previously mentioned, they provide more history and theoretical background for the various models described than the majority of other software package manuals in existence. Summarizing the guides will help provide the reader with a sense of the scope of LIMDEP. NLOGIT comes only with a reference guide, but it has been written in such a manner that it may also be considered as an econometric guide. As previously mentioned, the new text, *Applied Choice Analysis*, may also be considered as a NLOGIT econometric guide.

The LIMDEP Guide has a total of 38 chapters. The commands or procedures are arranged by "modeling framework," not by the alphabetical order of command names. As a result, the first chapter, referred to as E1, provides an overview of econometric modeling. E2, the second chapter, provides a rather detailed look at LIMDEP's descriptive statistical capabilities—in particular, descriptive statistics as it relates to panel data. Subsequent chapters provide the user with the information required to model defined data as desired.

I should also note that neither the Reference Manual nor the Econometric Modeling Guides have numbered pages. This may be rather disconcerting at first. But once the user gets accustomed to this reference strategy, they will find that they do not miss page numbers. All referencing in LIMDEP/NLOGIT is done by referring to R* or E* locations. Modeling categories are given E-numbers, and models nested under the respective category are specified by a decimal. There is a two-decimal limit for all referenced models. For example, negative binomial modeling is found in E20.3.1. There is no page number associated with that reference. E20 discusses "Models for Count Data"; 20.3 specifies "Models with Over- and Underdispersion: The Negative binomial and Gamma models." E20.3.1. deals specifically with the basic negative binomial model. E20.3.2 discusses "A heterogeneous Negative Binomial Model", and so forth. The logic of the guide makes good sense after using it a short time.

To provide a sense of the types of models addressed by LIMDEP, I'll give the chapter title names included in the Econometric Guides. Understand, though, that each chapter is differentiated into numerous subdivisions.

- E1. Estimation of Econometric Models: Generalities
- E2. Descriptive Statistics for Cross Section and Panel Data
- E3. Descriptive Statistics for Time Series Data
- E4. Scatter Diagrams and Plotting Tools
- E5. The Linear Regression Model
- E6. Non- and Semiparametric Linear Models
- E7. Hierarchical, Random Parameters and Latent Class Linear Models
- E8. Heteroscedasticity and ARCH/GARCH Models
- E9. Autocorrelation in the Linear Model
- E10. Time Series/Cross Section and Covariance Structure Linear Models
- E11. Linear Models for Panel Data
- E12. ARIMA, ARMAX and Distributed Lag Models
- E13. The Box-Cox Regression Model
- E14. Nonlinear Least Squares
- E15. 2SLS, Nonlinear 2SLS, Instrumental Variables and GMM Estimation
- E16. Linear and Nonlinear Systems of Regression Equations
- E17. Nonlinear Panel Data Models: Generalities
- E18. Models for Binary Chioce
- E19. Non- and Semiparametric Models for Binary Choice
- E20. Panel Data Models for Binary Choice
- E21. Bivariate and Multivariate Probit and Partial Observability Models
- E22. Ordered Choice Models
- E23. Multinomial Logit Models
- E24. Models for Count Data
- E25. Heterogeneity, Zero Inflation and Sample Selection in Count Models
- E26. Panel Data Models for Counts
- E27. Censored Data and Truncated Distributions
- E28. Panel Data Models for Censoring and Truncation
- E29. Loglinear Models
- E30. Sample Selection Models
- E31. Sample Selection in Nonlinear Models
- E32. Propensity Score Matching and Switching Regressions
- E33. Frontier Models and Efficiency Analysis
- E34. Nonparametric Analysis of Duration Data
- E35. Proportional Hazard Models
- E36. Parametric Duration Models
- E37. Nonlinear Optimization
- E38. Analysis of Nonlinear Functions

Each model is documented with all of the relevant formulas and mathematics that constitute the model and its fit. Diagnostic analysis is stressed for all models.

A complete suite of random number generators and numerous probability functions are also provided with the software. Together with its matrix modeling and programming features, LIMDEP can be used to expand its already remarkable modeling capabilities.

4. VERSION ENHANCEMENTS

Version 8 was released in 2002, nearly four years ago. In that time Professor Greene has added many new enhancements to the basic LIMDEP and NLOGIT packages. I'll summarize only the major enhancements—far too many housekeeping and minor fixes and additions have been made to mention here.

4.1 LIMDEP 9.0

Significant new regression models are:

- Binary: dynamic probit.
- Count: generalized Poisson, Polya-Aeppli, negative binomial with sample selection.
- Loglinear: binomial, power.
- Extensions to OLS: QREG for Quantile, nested random effects, random effects with exponential heteroscedasticity, 2SLS for panel data.
- Ordered choice: bivariate ordered probit, polychloric correlation, hierarchical ordered probit (HOPIT), zero inflated ordered probit (ZIOP, ZIHOP).
- Duration: Cox with time varying covariates and other options, parametric with parameter heterogeneity, parameter with sample selection. Binomial and multinomial: random effects and common (true) random effects, dynamic multinomial logit, generalized maximum entropy estimator.
- Stochastic frontier: Battesse and Coelli time varying inefficiency, truncation and heteroscedasticity, exponential and gamma with heterogeneity, sample selected stochastic frontier, Alvarez et al. scale, Alvarez et al. management.
- Multilevel and multiple effects random parameter: Blundell/Griffith/Windmeijer GMM estimators for count models with panel data.

4.2 NLOGIT 4.0

Significant extensions and new features are:

- New models: generalized nested logit, kernel logit, CLOGIT generalized maximum entropy estimator, propensity score matching.
- Modeling choice strategies: the program detects ignored attributes and adjusts the model appropriately without incorrectly assuming a value of zero.
- Choice models simulator: simulated probabilities, arc elasticities.
- Enhancements:
    a) random parameters (mixed) logit: enhancements to the method of simulation and random parameter specification including the calculation of willingness-to-pay estimates.
    b) HEV model: variance heterogeneity
    c) Nested logit : 1-line data set-up
    d) Latent class: proportions
    e) Output: elasticities (means and standard deviations computed for elasticities), robust standard errors (added to all appropriate NLOGIT models).

The new kernel logit model, not found in any other commercial package, is a variety of random effects model in which kernels are defined to be specific random effects that are distributed across alternatives according to a specific tree structure. A well-described example is provided in the NLOGIT reference manual.

5. EXAMPLES

I will present two examples using LIMDEP, and one using NLOGIT. Each example is aimed to provide the reader with a solid sense of how the package actually works. Instead of using the selection buttons and model-specific entry areas, I'll manipulate data and functions, as well as models, using the command line to enter instructions. LIMDEP uses "commands" that are typed into an editor in the form of a document, and submitted to the processor one line at a time, or in a kind of batch. Although nearly every model and option may be selected with the mouse, advanced users may find that they have more direct control over the modeling tasks when typing in, rather then selecting, commands. Additionally, each example is presented as a research task requiring the use of a specific LIMDEP or NLOGIT procedure. Presenting examples in this manner may help the reader obtain a feel for the software, and obtain a good sense of what the software can do.

5.1 Example 1: Binary Choice Logit Analysis

This application will illustrate estimation and analysis of a binary choice logit model.

5.1.1 Data Setup

The data used for this exercise are from the study by Riphahn, Wambach, and Million (2003). The raw data are published on the *Journal of Applied Econometrics* data archive Web site *http://qed.econ.queensu.ca/jae/*.

The URL for the data file is

http://qed.econ.queensu.ca/jae/2003-v18.4/riphahn-wambach-million/

which provides links to a text file which describes the data,

http://qed.econ.queensu.ca/jae/2003-v18.4/riphahn-wambach-million/readme.rwm.txt

and the raw data, themselves, which are in text form, zipped in the file

http://qed.econ.queensu.ca/jae/2003-v18.4/riphahn-wambach-million/rwm-data.zip

The .zip file contains the single data file `rwm.data`, which I have renamed `rwm.txt` so that it may be read into the software as a Windows text file.

The data file (Code 1) contains raw data on variables (original names).

```
id       person - identification number
female   female = 1; male = 0
year     calendar year of the observation
age      age in years
hsat     health satisfaction, coded 0 (low)
         - 10 (high)
handdum  handicapped = 1; otherwise = 0
handper  degree of handicap in
         percent (0 - 100)
hhninc   household nominal monthly net income
         in German marks / 1000
hhkids   children under age 16 in the
         household = 1; otherwise = 0
```

```
educ      years of schooling
married   married = 1; otherwise = 0
haupts    highest schooling degree is Hauptschul
          degree = 1; otherwise = 0
reals     highest schooling degree is Realschul
          degree = 1; otherwise = 0
fachhs    highest schooling degree is
          Polytechnical degree = 1; otherwise = 0
abitur    highest schooling degree is
          Abitur = 1; otherwise = 0
univ      highest schooling degree is
          university degree = 1; otherwise = 0
working   employed = 1; otherwise = 0
bluec     blue collar employee = 1;
          otherwise = 0
whitec    white collar employee = 1;
          otherwise = 0
self      self employed = 1; otherwise = 0
beamt     civil servant = 1; otherwise = 0
docvis    number of doctor visits in
          last three months
hospvis   number of hospital visits in
          last calendar year
public    insured in public health insurance
          = 1; otherwise = 0
addon     insured by add-on insurance = 1;
          otherswise = 0
```

Code 1. Data file.

The data file contains 27,326 observations. They are an unbalanced panel, with group sizes ranging from 1 to 7 with frequencies

$$T_i: \begin{array}{l} 1 = 1525,\ \ 2 = 2158,\ \ 3 = 825,\ \ 4 = 926, \\ 5 = 1051,\ \ 6 = 1000,\ \ 7 = 987. \end{array}$$

I additionally transformed the data as follows:

1. Year dummy variables were created.
2. Created variable $\texttt{sex} = 1$ for male, 2 for female equals $\texttt{female} + 1$.
3. Income $= \texttt{hhninc}$ was divided by 1,000 to improve numerical calculations.
4. Age squared $= \texttt{age}^2/1{,}000$ as recommended by the authors.
5. For purpose of illustrating a binary choice model, the count of doctor visits, `docvis`, was used to create doctor $= 1(\texttt{docvis} > 0)$. Figure 1 is a histogram representing the count of visits.
6. Create a group count variable which repeats, for each observation in a group, the number of observations in a group. LIMDEP relies on this variable in its panel data treatments.

Code 2 shows the complete set of instructions I submitted to LIMDEP in order to prepare the dataset for the analysis.

```
Read
;nobs=27326
;nvar=25
;names=id,female,year,age,hsat,handdum,handper,
   hhninc,hhkids,educ,married,haupts,
   reals,fachhs,abitur,univ,working,
   bluec,whitec,self,beamt,docvis,hospvis,
   public,addon
;file="E:\LIMDEP WIP\rwm.txt"
? Note that since the data in the file are space
  delimited, this format
? statement is not necessary. But, it speeds
  up the input substantially.
;format=(f5.0,f2.0,f5.0,f3.0,f10.6,f11.7,f11.6,
   f11.4,f2.0,
   f9.5,11f2.0,f4.0,f3.0,2f2.0)$
Create ; sex = 1 + female $
Create ; y84 = (year=1984) ; y85 = (year=1985)
       ; y86 = (year=1986) ; y87 = (year=1987)
       ; y88 = (year=1988) ; y91 = (year=1991)
       ; y94 = (year=1994) $
Create ; hhninc=hhninc/1000$
Create ; agesq=age*age/1000$
Matrix ; groupt = gsiz(id) $
Create ; ti = groupt(id) $
Histogram ; rhs=docvis $      <SEE GRAPHIC BELOW>
Create ; doctor=docvis>0$
```

Code 2. Instructions submitted to LIMDEP to prepare the dataset.

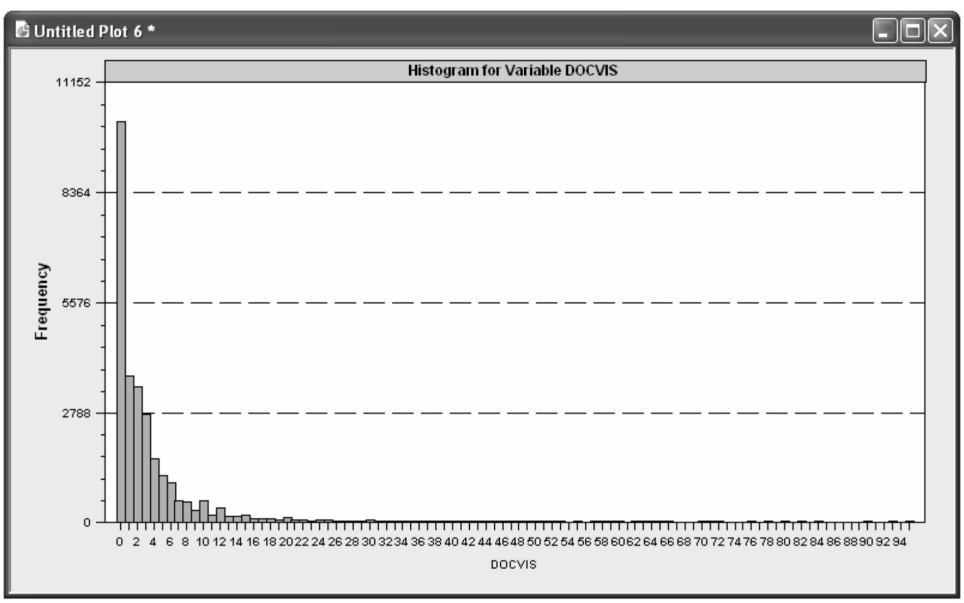

Figure 1. Count of doctor visits.

The descriptive statistics reported by Riphahn et al (2003, p. 393) were matched exactly with the command, or instruction

```
Dstat ; Rhs = *
      ; Str = sex $
```

5.1.2 Model Estimation

The analysis was primarily concerned with to the count variables `docvis` and `hospvis`. The independent variables, or predictors, in the model are listed as follows: `constant`, `age`, `agesq`, `hsat`, `handdum`, `handper`, `haupts`, `married`, `educ`, `hhninc`, `hhkids`, `self`, `beamt`, `bluec`, `working`, `public`, `addon`.

For convenience, we equate the list to a single name, X, with the instruction

```
Namelist;
x  = one,age,agesq,hsat,
handdum,handper,
haupts,married,educ,
hhninc,hhkids,self,
beamt,bluec,working,
public,addon$
```

A logit model for whether the individual visited the doctor is fit for men, women, and the entire sample: Because the output is voluminous, only the third set of results is shown in Code 3. In the third set, we requested the marginal effects for the independent variables and a plot of two interesting curves that summarize the model fit. These are discussed below.

```
Logit ; For [Sex=1] ;
Lhs=doctor ; Rhs = x $
Logit ; For [Sex=2] ;
Lhs=doctor ; Rhs = x $
Logit ; Lhs=doctor ;
Rhs = x ;
ROC ; Marginal
Effects $
```

```
+-------------------------------------------+
| Multinomial Logit Model                   |
| Maximum Likelihood Estimates              |
| Model estimated: Jan 26, 2006 at 09:37:38PM.|
| Dependent variable               DOCTOR   |
| Weighting variable                 None   |
| Number of observations            14243   |
| Iterations completed                  6   |
| Log likelihood function       -8799.681   |
| Number of parameters                 17   |
| Info. Criterion: AIC =          1.23804   |
|   Finite Sample: AIC =          1.23804   |
| Info. Criterion: BIC =          1.24707   |
| Info. Criterion:HQIC =          1.24104   |
| Restricted log likelihood     -9771.398   |
| Chi squared                    1943.435   |
| Degrees of freedom                   16   |
| Prob[ChiSqd > value] =         .0000000   |
| Hosmer-Lemeshow chi-squared =   5.68077   |
| P-value=  .68294 with deg.fr. =       8   |
+-------------------------------------------+
+---------+-------------+---------------+--------+---------+----------+
|Variable | Coefficient   | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+---------------+--------+---------+----------+
          Characteristics in numerator of Prob[Y = 1]
 Constant      4.21826769      .35149616    12.001   .0000
 AGE           -.10556051      .01600273    -6.596   .0000    42.6528119
 AGESQ         1.32642651      .18664810     7.107   .0000    1.94627522
 HSAT          -.28484261      .00960521   -29.655   .0000    6.92436176
 HANDDUM       -.16012654      .05723012    -2.798   .0051     .22729482
 HANDPER        .01605031      .00162781     9.860   .0000    8.13370920
 HAUPTS        -.07519730      .05298680    -1.419   .1558     .60113740
 MARRIED        .21954098      .05131566     4.278   .0000     .76514779
 EDUC          -.02507532      .01049038    -2.390   .0168    11.7286996
 HHNINC         .01532292      .01147802     1.335   .1819    3.59054065
 HHKIDS        -.10428907      .04458408    -2.339   .0193     .41297479
 SELF          -.36118089      .07019411    -5.145   .0000     .08565611
 BEAMT         -.05901049      .07648141     -.772   .4404     .11781226
 BLUEC         -.00954889      .04838517     -.197   .8436     .34023731
 WORKING        .07502718      .06790234     1.105   .2692     .85031243
 PUBLIC         .03190543      .07104566      .449   .6534     .86105455
 ADDON          .39060497      .14069013     2.776   .0055     .01755248
```

Code 3. Program output for the third command.

```
+----------------------------------------------------------------+
| Information Statistics for Discrete Choice Model.              |
|                            M=Model MC=Constants Only   M0=No Model |
| Criterion F (log L)    -8799.68076        -9771.39843   -9872.49529 |
| LR Statistic vs. MC     1943.43535             .00000        .00000 |
| Degrees of Freedom        16.00000             .00000        .00000 |
| Prob. Value for LR          .00000             .00000        .00000 |
| Entropy for probs.      8799.68076         9771.39843    9872.49529 |
| Normalized Entropy          .89133             .98976       1.00000 |
| Entropy Ratio Stat.     2145.62907          202.19372        .00000 |
| Bayes Info Criterion   17752.38585        19695.82120   19898.01492 |
| BIC - BIC(no model)     2145.62907          202.19372        .00000 |
| Pseudo R-squared            .09945             .00000        .00000 |
| Pct. Correct Prec.        64.81781             .00000      50.00000 |
| Means:        y=0    y=1    y=2    y=3    y=4    y=5    y=6   y>=7 |
| Outcome     .4405  .5595  .0000  .0000  .0000  .0000  .0000  .0000 |
| Pred.Pr     .4405  .5595  .0000  .0000  .0000  .0000  .0000  .0000 |
| Notes: Entropy computed as Sum(i)Sum(j)Pfit(i,j)*logPfit(i,j).  |
|        Normalized entropy is computed against M0.               |
|        Entropy ratio statistic is computed against M0.          |
|        BIC = 2*criterion - log(N)*degrees of freedom.           |
|        If the model has only constants or if it has no constants, |
|        the statistics reported here are not useable.            |
+----------------------------------------------------------------+
```

Code 4. Second set of results that summarize the model fit, related to the entropy of the outcomes and the estimated distribution.

The program output for the third command is shown in Code 3. The first set of results is the maximum likelihood estimates of the model parameters. The usual results are shown. The leading set of diagnostics shows a collection of standard "fit measures" (see below for more on this), as well as the overall likelihood ratio test statistic for the hypothesis that the coefficients are all zero (except for the constant term). Do not let the model title, "Multinomial Logit Model", mislead you. LIMDEP uses the more general multinomial regression algorithm to calculate maximum likelihood estimates of any discrete choice model, be it two levels (binary) or more (multinomial).

The second set of results, shown in Code 4, summarizes the model fit on a different level, related to the entropy of the outcomes and the estimated distribution. Code 4 also displays the "pseudo R-squared." Though this is not a fit measure for the model, it is, nonetheless commonly reported in empirical results. Notice that the BIC statistic, but not the AIC, is provided in the output. The BIC depends on a calculated deviance statistic, not on the log-likelihood.

The third set of results, shown in Code 5(a)–(b), display the partial effects for the variables in the model. For continuous variables such as education and income, these are the partial derivatives of the conditional mean (estimated probability). For the dummy variables, marked in the output, these effects are computed by evaluating the probability of a response with all predictors held at their means, then with the dummy predictor equal to zero, then equal to one, and evaluating the difference. In all cases, the delta method is used to estimate the asymptotic standard errors.

The final tables show various true fit measures for the model. These are a variety of analyses of how well the predicted outcomes of the model match the actual ones when the prediction is computed by using fitted value = 1, when fitted probability is greater than $P*$. The default value of $P*$ is 0.5, but the user has the ability to change the default.

The additional graphical output has two parts. Figure 2, labeled "Plot 8", compares the ability of the model to predict ones and zeros correctly versus incorrectly for the range of values of $P*$ from zero to one. Figure 3, labeled "Plot 7" is the receiver operating curve, or ROC, for the model. This curve also describes the ability of the model to correctly predict the outcomes. It is a plot of the proportion of ones correctly predicted for different values of the threshold $P*$.

```
+-----------------------------------------+
| Partial derivatives of probabilities with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used are All Obs.           |
+-----------------------------------------+
+---------+-------------+---------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |Elasticity|
+---------+-------------+---------------+--------+---------+----------+
         Characteristics in numerator of Prob[Y = 1]
 Constant      1.02854085      .08547206    12.034   .0000
 AGE           -.02573884      .00390077    -6.598   .0000   -1.89756597
 AGESQ          .32342278      .04549468     7.109   .0000    1.08801572
 HSAT          -.06945322      .00232560   -29.865   .0000    -.83125148
 HANDDUM       -.03904368      .01394954    -2.799   .0051    -.01533913
 HANDPER        .00391355      .00039524     9.902   .0000     .05501991
         Marginal effect for dummy variable is P|1 - P|0.
 HAUPTS        -.01831123      .01288416    -1.421   .1553    -.01902618
         Marginal effect for dummy variable is P|1 - P|0.
 MARRIED        .05392628      .01267463     4.255   .0000     .07131914
 EDUC          -.00611412      .00255796    -2.390   .0168    -.12394930
 HHNINC         .00373619      .00279871     1.335   .1819     .02318724
         Marginal effect for dummy variable is P|1 - P|0.
 HHKIDS        -.02545922      .01089466    -2.337   .0194    -.01817310
         Marginal effect for dummy variable is P|1 - P|0.
 SELF          -.08942949      .01750357    -5.109   .0000    -.01324035
         Marginal effect for dummy variable is P|1 - P|0.
 BEAMT         -.01443683      .01877018     -.769   .4418    -.00293983
         Marginal effect for dummy variable is P|1 - P|0.
 BLUEC         -.00232886      .01180335     -.197   .8436    -.00136957
         Marginal effect for dummy variable is P|1 - P|0.
 WORKING        .01836446      .01668056     1.101   .2709     .02699088
         Marginal effect for dummy variable is P|1 - P|0.
 PUBLIC         .00779319      .01738320      .448   .6539     .01159862
         Marginal effect for dummy variable is P|1 - P|0.
 ADDON          .09143750      .03124200     2.927   .0034     .00277411
+---------------------+
| Marginal Effects for|
+----------+----------+
| Variable | All Obs. |
+----------+----------+
| ONE      |  1.02854 |
| AGE      |  -.02574 |
| AGESQ    |   .32342 |
| HSAT     |  -.06945 |
| HANDDUM  |  -.03904 |
| HANDPER  |   .00391 |
| HAUPTS   |  -.01831 |
| MARRIED  |   .05393 |
| EDUC     |  -.00611 |
| HHNINC   |   .00374 |
| HHKIDS   |  -.02546 |
| SELF     |  -.08943 |
| BEAMT    |  -.01444 |
| BLUEC    |  -.00233 |
| WORKING  |   .01836 |
| PUBLIC   |   .00779 |
| ADDON    |   .09144 |
+----------+----------+

+----------------------------------------+
| Fit Measures for Binomial Choice Model |
| Logit model for variable DOCTOR        |
+----------------------------------------+
| Proportions P0= .440497   P1= .559503  |
| N =   14243 N0=    6274   N1=    7969  |
| LogL = -8799.68076 LogL0 = -9771.3984  |
| Estrella = 1-(L/L0)^(-2L0/n) = .13387  |
+----------------------------------------+
|     Efron |  McFadden  |  Ben./Lerman  |
|    .12635 |    .09945  |       .56909  |
|    Cramer | Veall/Zim. |      Rsqrd_ML |
|    .12581 |    .20757  |       .12755  |
+----------------------------------------+
| Information  Akaike I.C. Schwarz I.C.  |
| Criteria    17599.36390   17599.37293  |
+----------------------------------------+
```

Code 5a. Third set of results with partial effects for the variables.

```
+-----------------------------------------------------------+
|Predictions for Binary Choice Model.  Predicted value is |
|1 when probability is greater than  .500000, 0 otherwise.|
|Note, column or row total percentages may not sum to     |
|100% because of rounding. Percentages are of full sample.|
+------+---------------------------------+----------------+
|Actual|         Predicted Value         |                |
|Value |       0                1        | Total Actual   |
+------+----------------+----------------+----------------+
|  0   |   3752 ( 26.3%)|   2522 ( 17.7%)|   6274 ( 44.0%)|
|  1   |   2489 ( 17.5%)|   5480 ( 38.5%)|   7969 ( 56.0%)|
+------+----------------+----------------+----------------+
|Total |   6241 ( 43.8%)|   8002 ( 56.2%)|  14243 (100.0%)|
+------+----------------+----------------+----------------+

=======================================================================
Analysis of Binary Choice Model Predictions Based on Threshold =  .5000
-----------------------------------------------------------------------
Prediction Success
-----------------------------------------------------------------------
Sensitivity = actual 1s correctly predicted                     68.766%
Specificity = actual 0s correctly predicted                     59.802%
Positive predictive value = predicted 1s that were actual 1s    68.483%
Negative predictive value = predicted 0s that were actual 0s    60.119%
Correct prediction = actual 1s and 0s correctly predicted       64.818%
-----------------------------------------------------------------------
Prediction Failure
-----------------------------------------------------------------------
False pos. for true neg. = actual 0s predicted as 1s            40.198%
False neg. for true pos. = actual 1s predicted as 0s            31.234%
False pos. for predicted pos. = predicted 1s actual 0s          31.517%
False neg. for predicted neg. = predicted 0s actual 1s          39.881%
False predictions = actual 1s and 0s incorrectly predicted      35.182%
=======================================================================
```

Code 5b. Third set of results with partial effects for the variables.

The area under the darkened curve gives a measure of the model fit. It is the proportion of the unit area of the full box that is under the ROC curve.

Figure 4 is generated by a tool that can be used after a binary choice model is estimated. The command

```
Binary choice  ; Lhs=doctor ; Rhs=x
  ; Model=logit ; start=b;plot:hsat(0,10)$
```

instructs LIMDEP to compute the fitted probabilities from the model while holding all variables at their sample means, and varying HSAT over the range zero to 10. This gives a more detailed picture than the marginal effect of how the predicted probability is affected by the variation of the indicated variable.

5.2 Example 2: Using Programming Features

To provide a comparatively brief illustration of LIMDEP's programming capabilities, I have programmed, from scratch, an estimator of a loglinear model that is already contained in the package. This shows the creation of a procedure, or routine, and also illustrates the matrix programming language. The application is based, once again, on the German health care data.

Looking back to Figure 1, we can observe that the histogram resembles what would be observed if the dependent variable were generated by a discrete geometric distribution, defined as

$$\text{Prob}[Y = y_i] = \lambda_i^{yi}(1+\lambda_i)^{-(1+yi)}, \qquad Y = 0, 1, \ldots \quad \text{and} \quad \lambda_i > 0.$$

I model $\lambda_i = \exp(\boldsymbol{\gamma}'\mathbf{x}_i)$ where $\boldsymbol{\gamma}$ is the parameter vector to be estimated and $\mathbf{x}_i$ is the $K \times 1$ vector of independent variables, which includes a constant term. In this model, $E[y_i|\mathbf{x}_i] = \lambda_i$. I will use Newton's method to estimate $\boldsymbol{\gamma}$ by maximum likelihood. The log-likelihood and its derivatives are

$$F \ = \ \log L = \Sigma_i y_i(\boldsymbol{\gamma}'\mathbf{x}_i) - (1+y_i)\log(1+\lambda_i)$$

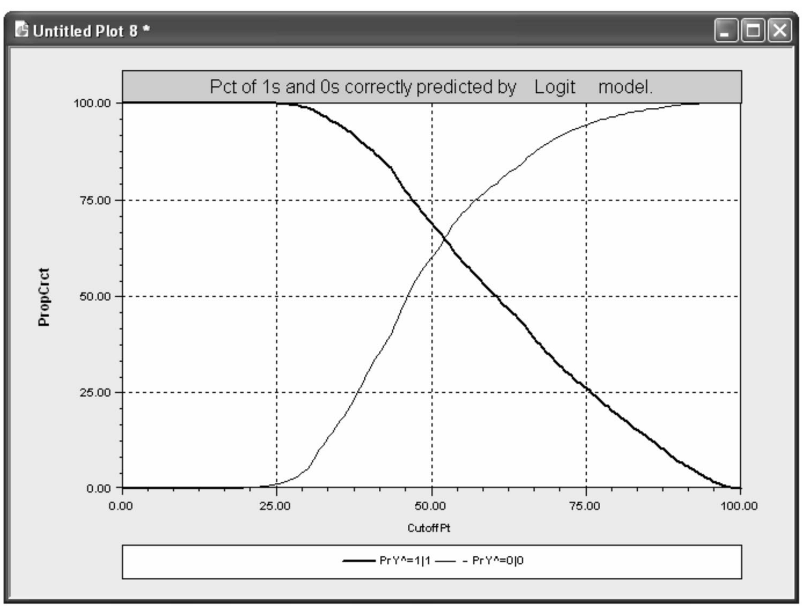

Figure 2. Prediction success for binary choice model.

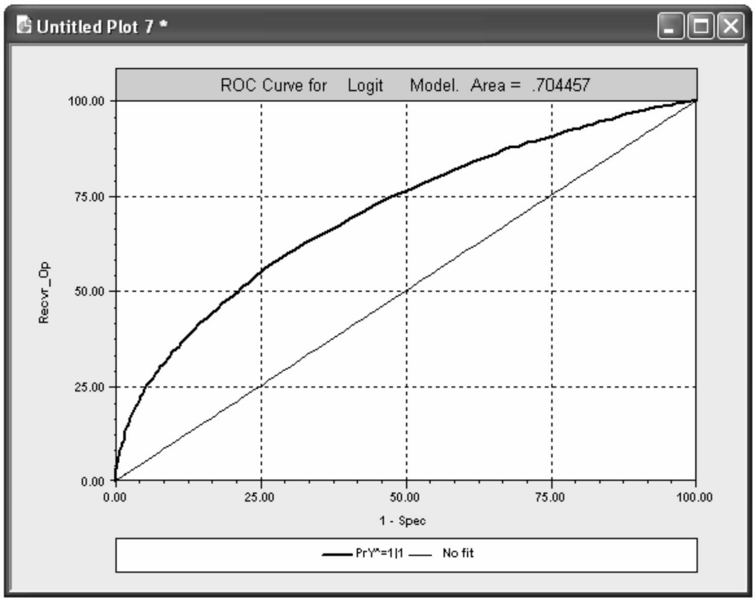

Figure 3. ROC curve for logit model.

$$\mathbf{g} \; = \; \partial \log L/\partial\boldsymbol{\gamma} = \Sigma_i \mathbf{g}_i \mathbf{x}_i$$

where $\mathbf{g}_i = y_i - (1 + y_i)\lambda_i/(1 + \lambda_i)$

$$\mathbf{H} \; = \; \partial^2 \log L/\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}' = \Sigma h_i \mathbf{x}_i \mathbf{x}_i'$$

where $h_i = -(1 + y_i)\lambda_i/(1 + \lambda_i)^2$.

The iteration is

$$\mathbf{c}^{(m)} = \mathbf{c}^{(m-1)} - [\mathbf{H}_{(m-1)}]^{-1}\mathbf{g}_{(m-1)},$$

where $\mathbf{c}$ is the estimate of $\boldsymbol{\gamma}$, "(m)" indicates the coefficient vector at iteration m, and "$(m - 1)$" indicates a computation of the gradient or Hessian based on $\mathbf{c}^{(m-1)}$. To obtain starting values for the iterations, I note that if all coefficients except for the constant term were zero, the maximum likelihood estimate of the constant term would be $a^{(0)} = \log \overline{y}$. Zeros are used for the remaining starting values. Finally, convergence of the iterations is assessed using the scale invariant measure $t = \mathbf{g}'\mathbf{H}^{-1}\mathbf{g}$.

Code 6 shows the code that sets up the procedure and does the estimation (top) and the results of executing the procedure (bottom).

There are other ways to accomplish this kind of optimization. LIMDEP has a built in MAXIMIZE command that allows the user to specify their own log-likelihood function. It automates the iterative procedure programmed above. The command to obtain the same results would be as follows:

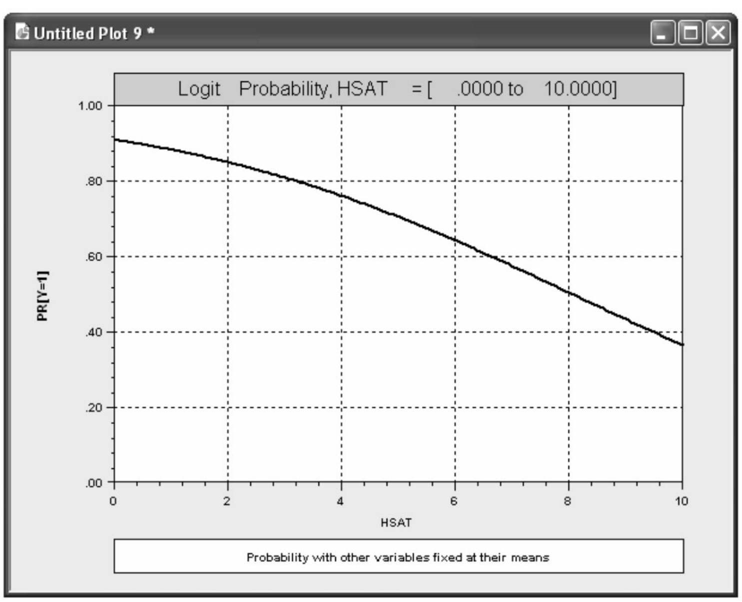

Figure 4. Marginal effects in logit model.

```
? Geometric regression for a discrete dependent variable
? ===================================================================
? This defines the RHS of the equation.
   Namelist ;xi=one,female,age,agesq,hsat,handdum,handper,married,educ,
                hhninc,hhkids,self,beamt,bluec,working,public,addon$
? This defines the LHS of the model
   Create    ; yi = docvis $
? -------------------------------------------------------------------
? The rest of the routine is generic - independent of the number of
? variables or observations.
? -------------------------------------------------------------------
? For starting values, use the MLE of the constant term if there were
? no regressors, and 0 for the remaining coefficients.
   Calc       ; a0 = log(xbr(yi)) ; k1 = Col(xi)-1 $
   Matrix     ; c0 = [a0,k1_0] ; c0 = c0'$ (Column vector. Used later.)
? Now set up procedure to fit the geometric model using Newton's method.
   Calc       ; test = 1 $  (Convergence criterion for iterations)
   Matrix     ; c = c0 $    starting vector for coefficients
? Define iterations
? --------------------------------------------------------------------
Procedure     ? First compute observation specific terms
Create        ; xc = xi'c ; li = exp(xc)         ? index and lambda(i)
              ; fi = yi * xc - (1+yi)*log(1+li)  ? function value
              ; gi = yi - (1+yi)*li/(1+li)       ? first derivative
              ; hi = -(1+yi)*li/(1+li)^2 $       ? second derivative
?               Function and derivatives are sums and matrix products
Calc          ; f = sum(fi) $     ? Could also use Matrix ; f = fi'1 $
Matrix        ; g = Xi'gi         ? first derivative vector
              ; H = -Xi'[hi]Xi    ? (negative) of second derivatives
              ; d = <H>*g         ? change vector
              ; c = c + d     $ ? iteration. new = old + inv(H) * g
Calc          ; list ; test = g'd $  Test for convergence using g'<H>g
EndProcedure
? --------------------------------------------------------------------
?  Execution to estimate model
?
Execute       ; while test > .00000001 $
?
?  Display the results
?
Matrix  ; Stat (c,<H>,Xi) $  (Estimate, covariance matrix, names)

    TEST    =  .82328999436903230D+04
    TEST    =  .33401344750368540D+03
    TEST    =  .25815773494424160D+01
    TEST    =  .33694117461187850D-03
    TEST    =  .80223569879432650D-11


TEST>.00000001

+---------------------------------------------------+
|Number of observations in current sample =   27326 |
|Number of parameters computed here       =      17 |
|Number of degrees of freedom             =   27309 |
+---------------------------------------------------+
+---------+-------------+---------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+-------------+---------------+--------+---------+
 Constant     2.91068243      .13793951    21.101   .0000
 FEMALE        .34548207      .01637507    21.098   .0000
 AGE         -.03492189      .00592814    -5.891   .0000
 AGESQ         .45208921      .06761539     6.686   .0000
 HSAT        -.22252983      .00328141   -67.815   .0000
 HANDDUM       .04123176      .02147213     1.920   .0548
 HANDPER       .00581661      .00047008    12.374   .0000
 MARRIED       .04971995      .01892647     2.627   .0086
 EDUC        -.01061506      .00380751    -2.788   .0053
 HHNINC      -.01399059      .00454970    -3.075   .0021
 HHKIDS      -.09511910      .01773246    -5.364   .0000
 SELF        -.23924209      .03374704    -7.089   .0000
 BEAMT         .04204389      .03692631     1.139   .2549
 BLUEC       -.00520479      .02056494     -.253   .8002
 WORKING       .02331360      .01979648     1.178   .2389
 PUBLIC        .11120969      .03001372     3.705   .0002
 ADDON         .03626687      .05349417      .678   .4978
```

Code 6. Code for loglinear model estimator (above). Output (below).

```
Calc
 ; a0 = log(xbr(yi))
 ; k = Col(x)
 ; k1 = k - 1 $
Matrix

  ; c0 = [a0,k1_0]
  ; c0 = c0'$
  (Column vector)
Maximize
  ; start = c0
  ; labels = k_c
  ; fcn = xtc = c1'x |
   ? Recursive defini-
tion of function
  lmi = exp(xtc) |
? using subfunctions
  yi*xtc -
(1+yi)*log(1+lmi) $
```

Results for these instructions are given in Code 7. However, this approach will be slower than the first routine because: (1) Maximize uses numeric rather than analytic derivatives, requiring many more function evaluations; (2) it uses the BFGS algorithm rather than Newton-Raphson (for this globally concave log-likelihood, Newton's method is more efficient); and (3) within the iterations, it does a rather elaborate line search, which increases computation time.

Finally, LIMDEP has a built in estimator for this particular model. One can use

```
Loglinear  ; Lhs = Docvis
 ; Rhs = Xi
 ; Model = Geometric $
```

Results are shown in Code 8.

5.3 Example 3: Discrete Choice Modeling in NLOGIT

The data for this example consist of a simulated dataset on brand choice. The choice situation modeled relates to individual choices among shoe brands. The universal choice set contains 20 "brands,", characterized by attributes:

```
PRICE:   Coded .25,.3,
        .35,.4,.45,.5
QUALITY: Low (0) or
         High (1)
STYLE:   Traditional (0)
         or Modern (1).
```

```
Line search does not improve fn. Exit iterations. Status=3
Check derivatives (with ;OUTPUT=3). This may be a solution
if several iterations have been computed, not if only one.
+-------------------------------------------+
| User Defined Optimization                 |
| Maximum Likelihood Estimates              |
| Model estimated: Jan 27, 2006 at 09:27:43AM.|
| Dependent variable             Function   |
| Weighting variable                 None   |
| Number of observations            27326   |
| Iterations completed                 26   |
| Log likelihood function        58428.85   |
| Number of parameters                  0   |
| Info. Criterion: AIC =         -4.27643   |
|   Finite Sample: AIC =         -4.27643   |
| Info. Criterion: BIC =         -4.27643   |
| Info. Criterion:HQIC =         -4.27643   |
| Restricted log likelihood      .0000000   |
| Chi squared                    116857.7   |
| Degrees of freedom                   17   |
| Prob[ChiSqd > value] =         .0000000   |
+-------------------------------------------+
+---------+--------------+----------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+--------------+----------------+--------+---------+
 C1            2.91068347      .11010260    26.436   .0000
 C2             .34548205      .01232148    28.039   .0000
 C3           -.03492192      .00459087    -7.607   .0000
 C4             .45208948      .05255838     8.602   .0000
 C5           -.22252984      .00252267   -88.212   .0000
 C6             .04123175      .01617050     2.550   .0108
 C7             .00581661      .00036849    15.785   .0000
 C8             .04971995      .01402846     3.544   .0004
 C9           -.01061508      .00298238    -3.559   .0004
 C10          -.01399059      .00358670    -3.901   .0001
 C11          -.09511911      .01305774    -7.285   .0000
 C12          -.23924211      .02329241   -10.271   .0000
 C13            .04204382      .02683883     1.567   .1172
 C14          -.00520482      .01568111     -.332   .7400
 C15            .02331362      .01545701     1.508   .1315
 C16            .11120954      .02333417     4.766   .0000
 C17            .03626695      .05253751      .690   .4900
```

Code 7. Results of using Maximize command.

Individuals are also simulated, with characteristics

```
SEX    Male $=$ 1,
       Female $=$ 0
AGE    Younger (0) or
       Older (1)
INCOME   Continuous, ranges
         from 0.2 to 1.0.
```

The data are simulated so that PRICE and QUALITY are correlated and AGE and INCOME are correlated. STYLE varies randomly among the brands and SEX varies randomly across individuals. Random effects specific to the individuals have also been built into the simulated data. (The data were simulated within LIMDEP using a program that is available by writing to *info@limdep.com*. The simulated data were then written to an ASCII file, then the internal simulated data were cleared and the ASCII data in the file were read into the program for the purpose of this example. This way, only the externally visible digits in the dataset are used in generating the empirical results shown below.)

Each choice situation consisted of an offer of 3 of the 20 brands, or a fourth choice, NONE, for a total of $J = 4$. We shall illustrate estimation of a discrete choice model with these data. Each individual makes a choice in $T = 5$ choice situations.

The basic model is a multinomial logit model, for which the choice probabilities are defined by the utility functions

$$\begin{aligned} U_{i,j,t} &= \beta_p \text{ Price}_{i,j,t} \\ &\quad +\beta_q \text{Quality}_{i,j,t} \\ &\quad +\beta_s \text{ Style}_{i,j,t} \\ &\quad +\sigma_1 K_{i1} + \varepsilon_{i,j,t} \\ U_{i,N,t} &= \alpha_N + \sigma_2 K_{i2} + \varepsilon_{i,N,t} \end{aligned}$$

in which "i" is the individual $[i = 1, \ldots, N = 250]$, "j" is the alternative $[j = 1, 2, 3]$, "N" is the NONE

```
Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Geometric (Loglinear) Regression Model      |
| Maximum Likelihood Estimates                |
| Model estimated: Jan 27, 2006 at 09:31:32AM.|
| Dependent variable                DOCVIS    |
| Weighting variable                  None    |
| Number of observations             27326    |
| Iterations completed                  23    |
| Log likelihood function        -58428.85    |
| Number of parameters                  17    |
| Info. Criterion: AIC =           4.27767    |
|   Finite Sample: AIC =           4.27767    |
| Info. Criterion: BIC =           4.28278    |
| Info. Criterion:HQIC =           4.27932    |
| Restricted log likelihood      -62871.17    |
| Chi squared                     8884.642    |
| Degrees of freedom                    16    |
| Prob[ChiSqd > value] =          .0000000    |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Parameters in conditional mean function
 Constant      2.91068238      .13793951    21.101   .0000
 FEMALE         .34548207      .01637507    21.098   .0000     .47877479
 AGE          -.03492189      .00592814    -5.891   .0000    43.5256898
 AGESQ          .45208919      .06761539     6.686   .0000    2.02285549
 HSAT         -.22252983      .00328141   -67.815   .0000    6.78542607
 HANDDUM        .04123176      .02147213     1.920   .0548     .21401539
 HANDPER        .00581661      .00047008    12.374   .0000    7.01228548
 MARRIED        .04971995      .01892647     2.627   .0086     .75861817
 EDUC         -.01061506      .00380751    -2.788   .0053    11.3206310
 HHNINC       -.01399059      .00454970    -3.075   .0021    3.52083617
 HHKIDS       -.09511910      .01773246    -5.364   .0000     .40273000
 SELF         -.23924209      .03374704    -7.089   .0000     .06217522
 BEAMT          .04204388      .03692631     1.139   .2549     .07469077
 BLUEC        -.00520478      .02056494     -.253   .8002     .24376052
 WORKING        .02331361      .01979648     1.178   .2389     .67704750
 PUBLIC         .11120968      .03001372     3.705   .0002     .88571324
 ADDON          .03626686      .05349417      .678   .4978     .01880992
```

Code 8. Results of LIMDEP's built-in estimator.

choice, "t" is the choice situation $[t = 1, \ldots, 5]$, and the "kernels" $[K_1$ and $K_2]$ are individual random effects. For convenience, let the collected utility functions be written

$$U_{i,j,t} = \boldsymbol{\beta}'\mathbf{x}_{i,j,t} + C_1\sigma_1 K_{i1} + C_2\sigma_2 K_{i2} + \varepsilon_{i,j,t},$$

where $C_1 = 1$ for $j\,1,2,3$ and 0 otherwise and $C_2 = 1$ for $j = N$ and 0 otherwise. Ignoring the individual effects, with an assumption of independent, homoscedastic, type 1 extreme value distributions for the four random terms, the choice probabilities will follow the multinomial logit form

$$\begin{aligned}\text{Prob}(\text{Choice}j|i,t) &= \frac{\exp(U_{i,j,t})}{\sum_{j=1}^{j}\exp(U_{i,j,t})}\\ &= \frac{\exp(\boldsymbol{\beta}'\mathbf{x}_{i,j,t})}{\sum_{j=1}^{j}\exp(\boldsymbol{\beta}'\mathbf{x}_{i,j,t})}.\end{aligned}$$

This is a standard multinomial logit model (MNL). Estimates of the MNL model are given in Code 9. (A variety of other options and post estimation tools are not shown here. More extensive description is given in the program documentation.) The code is:

Results 1.

```
          Tree Structure Specified for the Nested Logit Model
          Sample proportions are marginal, not conditional.
          Choices marked with * are excluded for the IIA test.
----------------+----------------+----------------+----------------+------+---
Trunk    (prop.)|Limb     (prop.)|Branch   (prop.)|Choice   (prop.)|Weight|IIA
----------------+----------------+----------------+----------------+------+---
Trunk{1} 1.00000|Lmb[1:1] 1.00000|B(1:1,1) 1.00000|B1        .25360| 1.000|
                |                |                |B2        .29840| 1.000|
                |                |                |B3        .28560| 1.000|
                |                |                |NONE      .16240| 1.000|
----------------+----------------+----------------+----------------+------+---
Model Specification: Utility Functions for Alternatives
Table entry is the attribute that multiplies the indicated parameter.
              Parameter
      Row  1  PRICE    QUALITY  STYLE    ASCNONE
Choice
B1         1  PRICE    QUALITY  STYLE    ASCNONE
B2         1  PRICE    QUALITY  STYLE    ASCNONE
B3         1  PRICE    QUALITY  STYLE    ASCNONE
NONE       1  PRICE    QUALITY  STYLE    ASCNONE
Normal exit from iterations. Exit status=0.

+----------------------------------------------+
| Discrete choice (multinomial logit) model    |
| Maximum Likelihood Estimates                 |
| Model estimated: Jan 29, 2006 at 10:07:23PM.|
| Dependent variable               Choice      |
| Weighting variable                 None      |
| Number of observations             1250      |
| Iterations completed                  5      |
| Log likelihood function       -1496.927      |
| Number of parameters                  4      |
| Info. Criterion: AIC =          2.40148      |
|   Finite Sample: AIC =          2.40151      |
| Info. Criterion: BIC =          2.41790      |
| Info. Criterion:HQIC =          2.40766      |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj  |
| Constants only   -1702.3704  .12068  .11974  |
| Response data are given as ind. choice.      |
| Number of obs.=  1250, skipped   0 bad obs.  |
+----------------------------------------------+
+----------------------------------------------+
| Notes No coefficients=> P(i,j)=1/J(i).       |
|       Constants only => P(i,j) uses ASCs     |
|         only. N(j)/N if fixed choice set.    |
|         N(j) = total sample frequency for j  |
|         N    = total sample frequency.       |
|       These 2 models are simple MNL models.  |
|       R-sqrd = 1 - LogL(model)/logL(other)   |
|       RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)    |
|         nJ   = sum over i, choice set sizes  |
+----------------------------------------------+

+---------+--------------+----------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+--------------+----------------+--------+---------+
 PRICE       -5.44375589      .49576792   -10.980   .0000
 QUALITY      2.20901909      .12202398    18.103   .0000
 STYLE         .02909512      .08007411      .363   .7163
 ASCNONE      -.99273686      .17481569    -5.679   .0000
```

Code 9. Estimates of the MNL model.

Results 2

```
Normal exit from iterations. Exit status=0.
+------------------------------------------------+
| FIML Nested Multinomial Logit Model            |
| Maximum Likelihood Estimates                   |
| Model estimated: Jan 29, 2006 at 10:17:50PM.|
| Dependent variable              CHOICE         |
| Weighting variable                None         |
| Number of observations            5000         |
| Iterations completed                14         |
| Log likelihood function       -1496.252        |
| Number of parameters                 5         |
| Info. Criterion: AIC =          .60050         |
|   Finite Sample: AIC =          .60050         |
| Info. Criterion: BIC =          .60702         |
| Info. Criterion:HQIC =          .60278         |
| Restricted log likelihood     -2016.681        |
| Chi squared                    1040.858        |
| Degrees of freedom                   5         |
| Prob[ChiSqd > value] =        .0000000         |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| No coefficients  -2016.6810  .25806  .25707 |
| Constants only   -1702.3704  .12108  .11990 |
| At start values  -1732.8680  .13655  .13539 |
| Response data are given as ind. choice.        |
+------------------------------------------------+
+------------------------------------------------+
| FIML Nested Multinomial Logit Model            |
| The model has 2 levels.                        |
| Random Utility Form 2: IV parms = muj|i, gi |
| IVs for degenerate branches have muj|i=1.      |
| p(alt=k|b=j,l=i)=exp[muj*bX_k|j,i]/Sum..       |
| p(b=j|l=i)=exp[gi(aY_j|i+IVj|i/muj|i)]/Sum     |
| p(l=i)=exp[cZ_j+IVi/gi]/Sum...                 |
| Number of obs.=  1250, skipped   0 bad obs. |
+------------------------------------------------+

+---------+--------------+----------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+--------------+----------------+--------+---------+
          Attributes in the Utility Functions (beta)
 PRICE       -4.46881857      .42378957   -10.545   .0000
 QUALITY      1.81100254      .10138314    17.863   .0000
 STYLE         .01603177      .06543834      .245   .8065
 ASCNONE     -1.12436863      .14919766    -7.536   .0000
          IV parameters, RU2 form = mu(j|i),gamma(i)
 B(1|1,1)      .79895238      .04908424    16.277   .0000
 B(2|1,1)     1.00000000     ......(Fixed Parameter).......
          Underlying standard deviation = pi/(IVparm*sqr(6))
 B(1|1,1)     1.60528941      .09862217    16.277   .0000
 B(2|1,1)     1.28254980     ......(Fixed Parameter).......
```

Code 10. NLOGIT example.

```
NLOGIT ; Lhs = choice
 ; choices=b1,b2,b3,none
    ; Rhs = Price,Quality,Style,ASCNone
         ; Show tree $
```

The MNL model is known to assume the objectionable IIA property. This is an undesirable restriction on behavior. Many alternative models have been suggested to extend the MNL model into more realistic forms. The multinomial probit model, which assumes that the joint distribution of the random terms is multivariate normal rather than independent type 1 extreme value is an early competitor. The multinomial probit model does relax the undesirable IIA assumption. (It is supported in NLOGIT, although I do not recommend it. First, it is difficult to extend it to any more interesting model forms—it is quite inflexible. Second, estimation is excruciatingly slow. Simulation of the multivariate normal probabilities even for moderate sized models such as this one requires an extraordinary amount of time consuming computation.)

The nested logit model has been used for many years to relax the IIA assumption. It is a step in that direction, though a somewhat restrictive one. We will group the three brands in one branch of the tree and the NONE choice in a second branch. With only a single alternative, the second branch is "degenerate," In the first form shown below, the model is actually not identified. This would be clearly seen in the results for this model, where the estimated standard error on the inclusive parameter for the degenerate branch would be essentially infinite. It is necessary for the user to be attentive to this model failure in their specification. The problem with the model in form 1 is that it is not properly normalized. By appropriately normalizing the scale parameters in the branches, NLOGIT recognizes the specification failure internally, and fixes the second IV parameter, as seen in Code 10. NLOGIT provides two specific forms of normalization for inclusive value parameters. One of them is shown in the output below.

Model form 1, unnormalized. (Results not shown.)

```
NLOGIT ; lhs = choice
 ; choices=b1,b2,b3,none
 ; Rhs = Price,Quality,Style,ASCNone
 ; Tree = (b1,b2,b3),(none) $
```

Model form 2, normalized so scale parameters appear at the branch level

```
nlogit ; lhs = choice
 ; choices=b1,b2,b3,none
 ; rhs = Price,Quality,
         Style,ASCNone
 ; tree = (b1,b2,b3),(none)
 ; Ru2$
```

These choice data embody some additional features that the analyst would want to capture in the model. First, this is a repeated choice situation, so there should be an element of the model which captures the common effects of the same individual making all five choices. Second, there is considerable heterogeneity across individuals in terms of preference weighting and in appropriate scaling of the utility functions. We capture these aspects with a "kernel logit," random parameters model. (These features of choice modeling are discussed in a series of recent papers by Hensher and Greene.) We modify the utility functions as follows:

$$
\begin{aligned}
U_{i,1,t} &= \beta_{p,i}\,\text{Price}_{i,1,t} + \beta_{qi}\,\text{Quality}_{i,1,t} \\
&\quad + \beta_s\,\text{Style}_{i,1,t} + \varepsilon_{i,1,t} + \sigma_1 K_{i,1} \\
U_{i,2,t} &= \beta_{p,i}\,\text{Price}_{i,2,t} + \beta_{qi}\,\text{Quality}_{i,2,t} \\
&\quad + \beta_s\,\text{Style}_{i,2,t} + \varepsilon_{i,2,t} + \sigma_1 K_{i,1} \\
U_{i,3,t} &= \beta_{p,i}\,\text{Price}_{i,3,t} + \beta_{qi}\,\text{Quality}_{i,3,t} \\
&\quad + \beta_s\,\text{Style}_{i,3,t} + \varepsilon_{i,3,t} + \sigma_1 K_{i,1} \\
U_{i,N,t} &= \alpha_N + \varepsilon_{i,N,t} + \sigma_2 K_{i,2}
\end{aligned}
$$

Normally distributed "kernel" effects

$$K_{i,1} \sim N[0, 1^2], K_{i,2} \sim N[0, 1^2].$$

Results 3

```
NLOGIT ; lhs = choice
       ; choices=b1,b2,b3,none
       ; Model: U(b1,b2,b3)=bp*Price+bq*Quality
         +bs*Style / U(none)    = aN*ASCNone
       ; rpl = age,sex
       ; fcn = bp(n),bq(n)
       ; hfr=Income
       ; pds=nchoice
       ; kernel=(b1,b2,b3),(none) ; pts=50
       ; Halton
 ; Parameters $

Normal exit from iterations. Exit status=0.
+--------------------------------------------+
| Random Parameters/Kernel Logit Model       |
| Maximum Likelihood Estimates               |
| Model estimated: Jan 29, 2006 at 10:33:15PM.|
| Dependent variable               CHOICE    |
| Weighting variable                 None    |
| Number of observations             5000    |
| Iterations completed                 25    |
| Log likelihood function       -1480.486    |
| Number of parameters                 14    |
| Info. Criterion: AIC =           .59779    |
|   Finite Sample: AIC =           .59781    |
| Info. Criterion: BIC =           .61604    |
| Info. Criterion:HQIC =           .60419    |
| Restricted log likelihood     -1732.868    |
| Chi squared                    504.7639    |
| Degrees of freedom                   14    |
| Prob[ChiSqd > value] =         .0000000    |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| No coefficients  -1732.8680  .14564  .14244 |
| Constants only   -1702.3704  .13034  .12708 |
| At start values  -1496.9268  .01098  .00728 |
| Response data are given as ind. choice.    |
+--------------------------------------------+
+--------------------------------------------+
| Random Parameters/Kernel Logit Model       |
| Replications for simulated probs. =  50    |
| Halton sequences used for simulations      |
| ----------------------------------------   |
| RPL model with panel has  250 groups.      |
| Fixed number of obsrvs./group=       5     |
| Random parameters model was specified      |
| ----------------------------------------   |
| Heteroscedastic random parameters          |
| Hessian was not PD. Using BHHH estimator.  |
| Number of obs.=  1250, skipped   0 bad obs. |
+--------------------------------------------+

+---------+-------------+--------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+-------------+--------------+--------+---------+
          Random parameters in utility functions
 BP          -4.94450710      .71198009    -6.945   .0000
 BQ           1.80323891      .21730037     8.298   .0000
          Nonrandom parameters in utility functions
 BS            .02029871      .08740124      .232   .8163
 AN          -1.21868271      .20289608    -6.006   .0000
          Heterogeneity in mean, Parameter:Variable
 BP:AGE       -.57851538      .72145313     -.802   .4226
 BP:SEX       -.88364588      .71869620    -1.230   .2189
 BQ:AGE        .58965843      .26742452     2.205   .0275
 BQ:SEX        .48602090      .26325952     1.846   .0649
          Derived standard deviations of parameter distributions
 NsBP         1.66217662     4.02149543      .413   .6794
 NsBQ          .49242429      .57063382      .863   .3882
          Heteroscedasticity in random parameters
 sBP|IN      -1.18163073     5.67526652     -.208   .8351
 sBQ|IN       -.12298899     2.18504349     -.056   .9551
          Standard deviations of latent kernel effects
 SigmaK01      .59533271      .70655694      .843   .3995
 SigmaK02      .42996678     1.00366776      .428   .6684
```

continued

```
 Kernel Logit Model
 Appearance of Latent Kernel Effects in Utilities
 Alternative   K01 K02
+-------------+---+---+
|  B1         | * |   |
+-------------+---+---+
|  B2         | * |   |
+-------------+---+---+
|  B3         | * |   |
+-------------+---+---+
|  NONE       |   | * |
+-------------+---+---+

Parameter Matrix for Heterogeneity in Means.

Matrix Delta    has  2 rows and  2 columns.
          AGE           SEX
        +----------------------------
BP      |     -.57852      -.88365
BQ      |      .58966       .48602
```

Code 11. Estimates of the random parameters model.

Random parameters with heterogeneous means and heteroscedasticity

$$\begin{aligned}\beta_{p,i} &= \beta_p^0 + \delta_{p,\text{Sex}}\,\text{Sex}_i + \delta_{p,\text{Age}}\,\text{Age}_i\\ &\quad +\gamma_p\exp(\theta_p\,\text{Income}_i)w_{i,p}, w_{i,p}\sim N[0,1],\\ \beta_{q,i} &= \beta_q^0 + \delta_{q,\text{Sex}}\,\text{Sex}_i + \delta_{q,\text{Age}}\,\text{Age}_i\\ &\quad +\gamma_q\exp(\theta_q\text{Income}_i)w_{i,q}, w_{i,q}\sim N[0,1].\end{aligned}$$

The kernel functions play the role of individual (random) effects in the choice model. They also allow the modeler to build cross utility correlation into the model. Note that the preceding produces a stochastic form of the nested logit model, with the three brands in one branch and the no choice alternative in another, degenerate branch. The heterogeneity across individuals is also built into the taste weights through the observable, individual specific means that vary with age and sex and variances that change with income.

The random parameter model is estimated by maximum simulated likelihood. Estimates of the random parameters model are shown in Code 11. Note that although these results are estimated by maximum simulated likelihood, they are exactly replicable because we use Halton sequences rather than pseudorandom draws to do the integration.

The random parameters model has the useful feature that, in similar fashion to Bayesian estimation, one can, after estimation, compute conditional estimates of $E[\beta_i|$all information on individual $i]$. NLOGIT will retain the computed conditional (posterior) estimates of the mean and standard deviation of the conditional distribution for each person. (That is the function of the `;Parameters` switch in the command.) The estimates are saved as a matrix which can be accessed later. The command set shown below extracts these estimates and plots estimated confidence limits for the individual conditional means. The bars

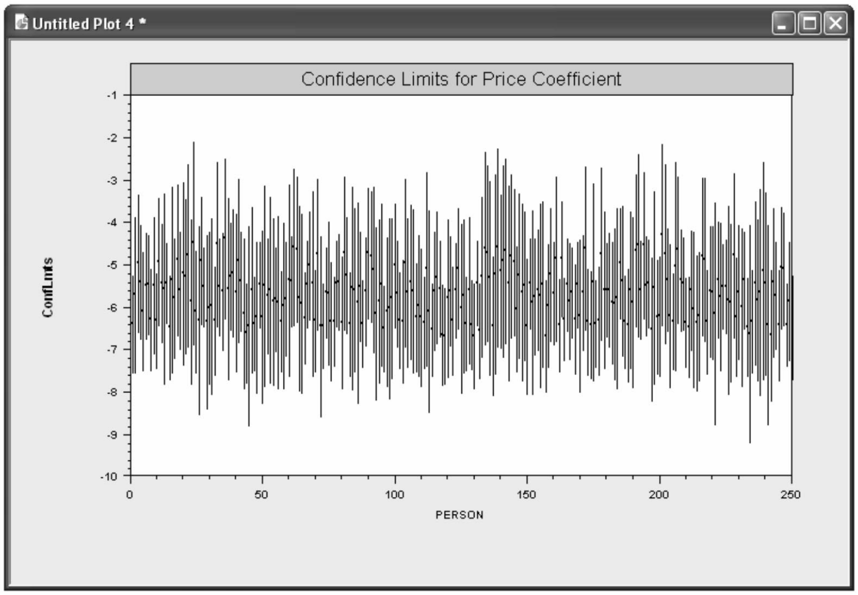

Figure 5. Conditional distributions for estimated parameters.

in Figure 5 show the individual specific estimates and a range of the mean plus and minus 1.96 standard deviations. The kernel density estimator shows how the means of the distributions vary across the individuals in the sample.

```
Matrix ; bp=beta_i(1:250,1:1) $
Matrix ; sbp=sdbeta_i(1:250,1:1) $
Sample ; 1 - 250 $
Create ; bpi=bp $

Create ; sdi = sbp$
Create ; person = trn(1,1) $
 Create
 ; bpupper = bpi + 1.96*sdi
 ; bplower=bpi - 1.96*sdi $
Plot
 ; Lhs = person ; Rhs = bplower,bpupper
 ; Centipede ; endpoints=0,250
 ; Yaxis=ConfLmts
 ; Title=Confidence Limits for
   Price Coefficient $
Kernel; Rhs = bpi $
```

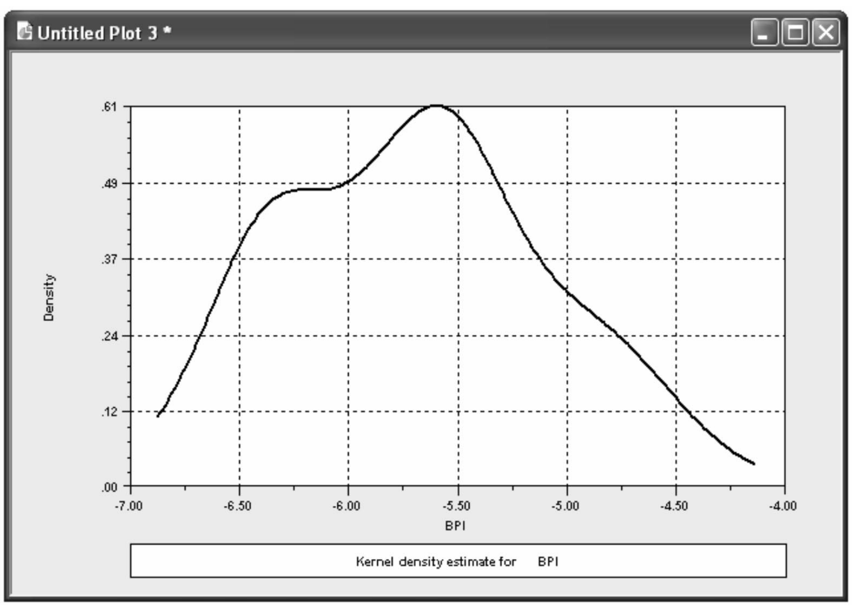

Figure 6. Kernel density estimator for conditional means.

```
nlogit ; lhs = choice
       ; choices=b1,b2,b3,none
       ; Model: U(b1,b2,b3)=bp*Price+bq*Quality+bs*Style /
                U(none)    = aN*ASCNone
       ; rpl = age,sex
       ; fcn = bp(n),bq(n)
       ; hfr=Income
       ; pds=nchoice
       ; kernel=(b1,b2,b3),(none) ; pts=50 ; Halton
       ; Simulation ; Scenario: price=[*]1.25 $

+---------------------------------------------+
| Random Parameters/Kernel Logit Model        |
| Model Simulation Using Previous Estimates   |
| Number of observations             1250     |
+---------------------------------------------+
+-----------------------------------------------------+
|Simulations of Probability Model                     |
|Model: Random Parameters Logit Model                 |
|Simulated choice set may be a subset of the choices. |
|Number of individuals is the probability times the   |
|number of observations in the simulated sample.      |
|Column totals may be affected by rounding error.     |
|The model used was simulated with   1250 observations.|
+-----------------------------------------------------+
-------------------------------------------------------------------------
Specification of scenario 1 is:
Attribute  Alternatives affected            Change type           Value
---------  ------------------------------  ------------------- ---------
PRICE      B1       B2       B3       more  Scale base by value     1.250
-------------------------------------------------------------------------
The simulator located   1250 observations for this scenario.
Simulated Probabilities (shares) for this scenario:
+----------+--------------+--------------+------------------+
|Choice    |     Base     |   Scenario   | Scenario - Base  |
|          |%Share Number |%Share Number |ChgShare ChgNumber|
+----------+--------------+--------------+------------------+
|B1        | 28.361   355 | 25.756   322 | -2.605%      -33 |
|B2        | 27.660   346 | 25.301   316 | -2.359%      -30 |
|B3        | 27.818   348 | 25.334   317 | -2.483%      -31 |
|NONE      | 16.161   202 | 23.609   295 |  7.447%       93 |
|Total     |100.000  1251 |100.000  1250 |   .000%       -1 |
+----------+--------------+--------------+------------------+
```

Code 12. Experiment using earlier model with simulated 25% increase in prices of all brands.

Finally, NLOGIT contains a model simulator that can be used with any estimated discrete choice model, and with any sample, whether it was used to estimate the model or not. The simulator computes the models predictions of the allocation of choices across the alternatives. It then allows the user to do experiments to see how the allocations would change if the attributes of the choices changed. For example, the experiment in Code 12 and Figure 6 takes the model just estimated, and simulates a 25% increase in the prices of all brands. The model construction is quite price sensitive. As the simulator shows, under the experimental change, many of the choosers would opt for no brand, rather than choose one of the three brands.

I should note that the experiment is a bit dubious, since "Brands" 1, 2, and 3 are different for each individual.

6. SUMMARY REMARKS

As the reader can likely observe from reading through the examples, LIMDEP is extremely thorough in its output and provides numerous options to effect the standard—as well as the not so standard—fit statistics. It is a, if not *the*, premiere econometrics package on the market. I have attempted to demonstrate the scope and feel of the package in presenting the examples. I also want to point out to those in other disciplines—for example, biostatistics, noneconomic social sciences, and even to those in the physical sciences—LIMDEP has many unique capabilities that you may find valuable to your own work. LIMDEP and NLOGIT have been marketed to, and primarily used by, those in the economic sciences. This need not be the case. I have found applications in health outcomes analysis that are currently unavailable in other commercial packages. There is a moderate learning curve, to be sure, but I found it to be no more than most other packages. In fact, when I first obtained the package, I read the manual on how to perform a simple Poisson regression model, with displayed standard goodness-of-fit statistics, imported data from another format, and modeled it without problem in about 10 minutes. Using the menu facility makes the modeling task quite easy; using the command line as we did in the examples, takes more effort and learning. The most difficult task for many may be in learning how to interpret the host of statistics that are displayed following modeling.

There are packages on the market that provide the user with limited options and only a few, if any, GOF tests. LIMDEP and NLOGIT are the antithesis to this type of statistical package. Perhaps some may be intimidated with the comprehensiveness of options and output, but the professional researcher and statistician should find this fact to be a welcome plus.

The statistical models found in LIMDEP and NLOGIT are written by a single person, Professor William Greene. When a possible problem is identified, he is quick to respond, and if need be, to correct the code. He continually is enhancing the software to include the most recent statistical advances in econometrics. Both programs can, without a doubt, be considered as state of the art packages with respect to statistical models, in particular discrete response models. Before commencing this review I had doubts as to my future interest in using the package in my own research. However, I have come to appreciate its depth and capability, and intend to use it in my work, and classes, whenever it is appropriate. I suspect that others who engage in statistical modeling may come to the same conclusion.

REFERENCES

Hardin, J. W., and Hilbe, J. M. (2003), *Generalized Estimating Equations*, Boca Raton, FL: Chapman & Hall/CRC Press.

Hensher, D. D., Rose, J. M., and Greene, W. H. (2005), *Applied Choice Analysis: A Primer*, New York: Cambridge University Press.

Hosmer, D., and Lemeshow, S. (2001), *Applied Logistic Regression* (2nd ed.), New York: Wiley.

McCullagh, P., Nelder, J. (1989), *Generalized Linear Models* (2nd ed.), London: Chapman and Hall.

Riphahn, R. T., Wambach, A., and Million, A. (2003), "Incentive Effects in the Demand for Health Care: A Bivariate Panel Coun Data Estimation," *Journal of Applied Econometrics*, 18, 387–405.