Notes

Chapter 1

1. The possibility of ties (i.e., $U_{in} = U_{jn}$) is ignored in this discussion; in practice, ties never occur.

2. An alternative way of viewing the probabilities is to suppose that the researcher observes one decisionmaker facing the same choice repeatedly. In each repetition, the observed component of utility is the same, but the unobserved component changes (due, perhaps, to randomly varying tastes of the decisionmaker). P_{in} is then the proportion of times the decisionmaker chooses alternative i as the number of repetitions becomes large. With this interpretation, the probability still arises from the researcher's lack of knowledge. At each time, the decisionmaker's choice is deterministic, and if the researcher knew the varying component of utility with every repetition, then the researcher could perfectly predict each choice.

A third interpretation of the choice probabilities arises from the concept that probabilities are not necessarily limits of proportions for repeated events but rather reflect subjective views of an uncertain world. With this viewpoint, P_{in} is the subjective probability assigned by the researcher to the event that the unobserved utility components of person n are such that, given the observed components, the person will choose alternative i. Again, the probability reflects the researcher's uncertainty, not the decisionmaker's.

Chapter 2

1. Also called the Weibull distribution. See section 2.9 for the density function for this distribution.

2. Shoulder room is the width of the passenger cabin of a car measured at the height of a seated passenger's shoulders. Cars with greater shoulder room carry more passengers and, for a given number of passengers, allow more room per passenger than cars with less shoulder room.

3. Usually the planner's goal can be achieved consistently with the fact that derivatives of choice probabilities sum to zero by defining the choice set appropriately. For example, if an airport commission utilizes a qualitative choice model describing passengers' choice of airline (with the alternatives being the various airlines), increasing demand for one airline will necessarily imply reduced demand for other airlines. By expanding the choice set to include the alternative of traveling by nonair modes (auto, bus, rail), then one airline's demand can be increased without decreasing other airlines' demand as long as the additional demand is drawn from the nonair modes.

Unfortunately, standard logit models cannot be used for this purpose, since, as stated, increasing the representative utility of one alternative in a logit model decreases the probabilities for all other alternatives in proportion to their probabilities before the change. Consequently, drawing additional demand from nonair modes but not other airlines is not possible with logit; other qualitative choice models, such as GEV, handle this situation more appropriately.

4. The stratification cannot be on the basis of the individual's choice of alternative; if it is, more complex methods are required.

Chapter 3

1. A homebuyer's perception and concern about the risk of a variable rate is probably related to the maximum allowable increase in the interest rate for a loan. If this is the case, then we can write

$R_{in} = \theta_n M_{in} + m_{in},$

such that

$e_{in} = -\theta_n M_{in} - m_{in} + \eta_{in},$

which could be handled in a manner similar to that for taste variation. However, for the present example, we assume R_{in} is unrelated to M_{in}.

2. As in the taste variation example, a convenient normalization that reflects the fact that multiplicative transformations of utility do not affect choices is $\omega = 1$.

Chapter 4

1. The parameter λ_k is usually assumed to be between zero and one. If the parameter is within this range, then the resulting choice probabilities are consistent with utility maximization for all possible levels of observed data. If the parameter is above one (and hence $1 - \lambda_k$ is below zero), then the choice probabilities are consistent with utility maximization, if at all, only within a range of observed data. Values of λ_k below zero are inconsistent. See McFadden (1978).

2. That the product of these marginal and conditional probabilities equals the joint probability in (4.1) is verified as follows:

$$\begin{aligned}
P_{in} = P_{in|B_n^k} \cdot P_{B_n^k} &= \frac{e^{Y_{in}^k}}{\sum_{j \in B_n^k} e^{Y_{jn}^k}} \cdot \frac{e^{W_n^k + \lambda_k I_k}}{\sum_{l=1}^K e^{W_n^l + \lambda_l I_l}} \\
&= \frac{e^{Y_{in}^k}}{\sum_{j \in B_n^k} e^{Y_{jn}^k}} \cdot \frac{e^{W_n^k}(\sum_{j \in B_n^k} e^{Y_{jn}^k})^{\lambda_k}}{\sum_{l=1}^K e^{W_n^l}(\sum_{j \in B_n^l} e^{Y_{jn}^l})^{\lambda_l}} \\
&= \frac{e^{Y_{in}^k} e^{W_n^k}(\sum_{j \in B_n^k} e^{Y_{jn}^k})^{\lambda_k - 1}}{\sum_{l=1}^K e^{W_n^l}(\sum_{j \in B_n^l} e^{Y_{jn}^l})^{\lambda_l}} \\
&= \frac{(e^{Y_{in}^k + W_n^k/\lambda_k})(\sum_{j \in B_n^k} e^{Y_{jn}^k + W_n^k/\lambda_k})^{\lambda_k - 1}}{\sum_{l=1}^K (\sum_{j \in B_n^l} e^{Y_{jn}^l + W_n^l/\lambda_l})^{\lambda_l}} \qquad \text{since} \quad e^{W_n^k} = (e^{W_n^k/\lambda_k})^{\lambda_k} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = e^{W_n^k/\lambda_k} \cdot (e^{W_n^k/\lambda_k})^{\lambda_k - 1} \\
&= \frac{e^{V_{in}/\lambda_k}(\sum_{j \in B_n^k} e^{V_{jn}/\lambda_k})^{\lambda_k - 1}}{\sum_{l=1}^K (\sum_{j \in B_n^l} e^{V_{jn}/\lambda_l})^{\lambda_l}} \qquad \text{since} \quad V_{in} = W_n^k + \lambda_k Y_{in}^k, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad V_{in}/\lambda_k = W_n^k/\lambda_k + Y_{in}^k.
\end{aligned}$$

3. The term "inclusive price" is also used occasionally. Actually, however, the **negative** of I_k more closely resembles a price.

4. Recall from section 2.3 that only differences in representative utility affect choice probabilities in logit models. Consequently, each of the two conditional submodels can be estimated with actual time and cost for each mode entering as explanatory variables rather than with deviations from the average. Since the average is constant over the two alternatives in each submodel, it drops out automatically.

5. Sequential estimation takes more steps than standard maximum likelihood estimation. When the labor involved in these steps is considered, sequential estimation is not necessarily less expensive.

Chapter 5

1. The indirect utility function must satisfy certain criteria, most of which are very intuitive, to be consistent with a direct utility function. For example, the indirect utility must be nonincreasing in price (i.e., if the price of a good rises, the utility that a consumer obtains after maximizing utility at the new price cannot rise). See Varian (1978) for a full discussion of these criteria.

2. This is simply a restatement of the fact that utility is maximized at a point of tangency between an indifference curve and the budget constraint. It can be demonstrated as follows:

$$(\partial U(x_1, (y - x_1 p_1)/p_2)/\partial x_1) = \mathrm{MU}_1 - (p_1/p_2)\,\mathrm{MU}_2 = 0,$$

or

$$\mathrm{MU}_1/\mathrm{MU}_2 = p_1/p_2.$$

3. If the choices are independent, a joint modeling approach is unnecessary. A precise meaning of the word "interrelated" in this context is given at the conclusion of this subsection.

4. Example: a person chooses a type of car and decides how many miles to drive. The price of driving (i.e., the cost per mile) that the person faces will be different for different types of cars.

5. For some situations, it might be useful to generalize the choice situation described in section 5.3 to allow for the possibility of some of the variables entering V_i being functions of the continuous good. Consider, for example, the choice of which long distance service to acquire (AT&T, Sprint, Allnet, etc.) and the choice of how much to use the long distance service. Given that volume discounts are given by some carriers, the marginal price of a long distance call depends on the number of calls made, such that p_i entering V_i is a function of x_i. A stepwise method for estimating choice probabilities in these types of situations is usually feasible. For the case of long distance service choice, (1) estimate the demand equation for the number of calls as a function only of variables that do not depend on the long distance service chosen; (2) use this equation to estimate the number of calls, and calculate a marginal price for alternative service given this number of calls; and then (3) estimate the choice model with this constructed price entering as an explanatory variable.

6. For choice situations with more than two alternatives, expression (5.10) is generalized as follows. For any alternative i in J, where J is the set of available alternatives,

$$E(e_i) = (\sqrt{6\sigma^2}/\pi)\left[\left(\sum_{j\in J} \rho_j P_j \ln P_j/(1 - P_j)\right) - (\rho_i \ln P_i)/(1 - P_i)\right],$$

where σ^2 is the variance of e in the entire population, and ρ_j is the correlation of e with the unobserved utility associated with alternative j, for all j in J.

7. For choice situations with more than two alternatives, expression (5.11), and the correction mechanism are generalized as follows. Using the notation of the previous note, we know (for the same reasons that $\rho_c = -\rho_q$ in the binary case) that $\sum_{j\in J} \rho_j = 0$, or, stated equivalently,

$$\rho_i = -\sum_{\substack{j\in J\\ j\neq i}} \rho_j.$$

Substituting into the expression for $E(e_i)$ in the previous note, we have

$$E(e_i) = \sum_{\substack{j\in J\\ j\neq i}} (\sqrt{6\sigma^2}/\pi)\rho_j\left(\frac{P_j \ln P_j}{1 - P_j} + \ln P_i\right),$$

which is the generalized form of (5.11). Therefore, with N alternatives in the set J, there are $N - 1$ selectivity correction terms to be added to the regression equation, with coefficients of $(\sqrt{6\sigma^2}/\pi)\rho_j$ for each $j \neq i$.

8. That is, the coefficients in the structural equations (5.7) and (5.8) are not common. The coefficients of the selectivity correction terms that are added to these equations are equal in magnitude and opposite in sign.

Chapter 6

1. Often a researcher is unable to predict changes in the number of decisionmakers in each segment. Suppose instead that the researcher can predict the row totals and column totals in the segmentation (i.e., the marginals) but not the individual cell counts (i.e., the joint distribution). For example, for figure 6.2, the researcher might be able to predict changes in the proportion of males and females and changes in the proportion of decisionmakers at each education level, but is unable to predict changes in the proportion of each sex **and** education level. A method for estimating cell counts from predicted values of marginals, called iterative proportional fitting, can be used in these cases. See Bishop, Fienberg, and Holland (1975) for details.

2. The values of z_i in the forecast area will also differ from those in the estimation area. However, since z_i consists of input variables, no reestimation is required to accommodate these changes. The values of β, which reflect the behavioral importance, or weight, attached to each element of z_i, will be the same even if z_i changes, provided only that the process by which households make decisions is not area-specific.

Chapter 7

1. A note is needed regarding terminology. Throughout part II the word "automobile," or "auto" for short, designates both cars and trucks. "Truck" refers to pickups, vans, and utility vehicles such as jeeps, while "car" denotes any automobile that is not a truck. Even though "vehicle" encompasses, in common parlance, objects other than autos, the word is used herein interchangeably with "automobile," primarily to provide a possibility for textual variety when needed.

2. Hocherman, Prashker, and Ben-Akiva estimate the probability that a household will make a particular transaction (e.g., purchase a vehicle, or sell a currently held vehicle), rather than the probability that the household will hold a particular number of vehicles. Since the number of vehicles held at any time can be calculated by knowing the holdings in some base period and all transactions since the base period, this model is grouped with those of vehicle holdings per se.

3. The last four of these studies utilize a nested logit model, with number of autos (or the purchase decision, in the case of Hocherman, Prashker, and Ben-Akiva) as the "upper level" choice, and make and model of autos as the "lower level" choice of each of these models. Auto cost variables enter significantly in the lower level choice; they enter indirectly in the upper level choice through the inclusive value term, which itself enters significantly in each of these studies. Consequently, it is appropriate to say that auto cost enters these quantity models significantly.

4. Hensher and Le Plastrier do not enter price per se, but enter sales tax, which depends on purchase price.

5. The Cambridge Systematics, Inc., model includes age through the scrappage probability variable, which is a function of age.

6. Cambridge Systematics, Inc.; Booz, Allen, and Hamilton, Inc.; and Mannering and Winston reflected the effect of the number of autos owned by constructing separate models for one- and two-auto households.

7. Of the twenty studies using disaggregate compensatory models based on real choice situations, only four have examined both the number and type of vehicles owned: Hocherman, Prashker, and Ben-Akiva; Booz, Allen, and Hamilton, Inc.; Hensher and Le Plastrier; and Mannering and Winston.

8. Lave and Train; Booz, Allen, and Hamilton, Inc.; and Hensher and Le Plastrier enter current-year miles traveled. The number of miles traveled in the previous year enters the models of Mannering and Winston, and Winston and Mannering; these latter models are subject to the simultaneity bias (to be discussed) only if there is serial correlation in the number of miles traveled.

9. Lave and Bradley do not include vehicle characteristics in their model, but the concept of grouping into classes (in this case, foreign and domestic) and forecasting demand for each class, independent of variations of make/model characteristics within each class, still applies.

10. See section 2.2 for a full discussion of this property.

11. If a constant term were included in the representative utility of each make and model, then the logit model could possibly produce consistent estimates of the model parameters. However, none of the studies included constants for each make and model because the number of makes and models is so large.

12. The Cambridge Systematics, Inc., model forecasts quite large year-to-year fluctuations in demand for makes and models. The procedure of assigning vehicles can easily produce such a pattern. Consider, for example, a case of one household facing a choice between two vehicles that, in the eyes of the household, are exactly the same. The household is assigned in each year one of the two vehicles on the basis of a flip of the coin, reflecting the true probabilities of 50 : 50. For the first year, heads appears, and the first vehicle is assigned to the household. In the second year, tails appears, and the second vehicle is assigned. Since this household is the entire sample, the prediction of market share for the first vehicle is 100% in year 1 and 0% in year 2; for the second vehicle, predicted shares are 0% in year 1 and 100% in year 2. This phenomenon can occur whenever the sample size is sufficiently small with respect to the number of vehicles.

13. Weight is valued, at least partially, because of its real or perceived relation to crash worthiness.

14. It is worth noting that noncompensatory models can (depending on their exact form) be consistent with utility maximization by the consumer. Therefore, the distinction between the models discussed in this section and those in sections 7.2 and 7.3 is not whether utility maximization is assumed. Rather, the difference lies in the form of the utility function, with noncompensatory models having lexicographic preferences.

15. The researchers at Charles River Associates have called their technique "hedonic demand models." This is an unfortunate term since the word "hedonic" has already been established for a type of price analysis made popular by Griliches (1961). Charles River Associates's model is no more related to Griliches's method than other demand analyses are, and so it does not seem that their approach should be called "hedonic" to emphasize any inherent connection. The term can only cause confusion.

Chapter 8

1. The precise definitions of these categories will be given in the detailed discussion of the submodel.

2. Households that chose more than two vehicles were not included in the estimation sample. Given that the full choice set includes the alternatives of owning three vehicles, four vehicles, etc., eliminating households that own more than two vehicles is equivalent to estimating the vehicle quantity model on a subset of alternatives. With logit models, estimating on a subset of alternatives is consistent due to the independence of irrelevant alternatives property (see sections 2.2 and 2.6). Households were also eliminated if data were missing for any relevant variable.

3. This formula is based on the normalization $V_0 = 0$ such that $\exp(V_0) = 1$, which is the third term in the denominator. See the discussion "Differences in Representative Utility" in section 2.3 for an explanation of this normalization.

4. Note that the term representing the average utility in the class/vintage choice (i.e., I_n defined in equation (8.8)) depends on the parameters of the class/vintage submodels. Since these are estimated parameters rather than the true parameters, the term used in estimating the vehicle submodel is an estimate of I_n rather than of the true I_n. Amemiya (1978) has shown that using an estimate of I_n in logit estimation results in a downward bias in the standard errors (though the parameters themselves are estimated without bias). Consequently, the t-statistics given in the third column of table 8.1 should be viewed as upper limits on the true t-statistics.

5. To show this, let the indirect utility of a household be

$$V = \alpha P + \beta \mathrm{OC},$$

where P is purchase price of the class/vintage, OC is operating cost, and α and β are parameters. To calculate the change in price required to keep a household at the same level of utility with a one-unit change in operating cost, totally differentiate V to obtain

$$\Delta V = \alpha \Delta P + \beta \Delta \mathrm{OC}.$$

Constrain this total derivate to zero (that is, keep the household's utility unchanged), and solve for the change in price required to offset a change in operating cost:

$\Delta V = \alpha \Delta P + \beta \Delta \mathrm{OC} = 0,$

$\Delta P = -\beta/\alpha \, \Delta \mathrm{OC}.$

That is, the change in price that keeps utility constant with a one-cent change in operating cost (i.e., $\Delta \mathrm{OC} = 1$) is the (negative of the) ratio of the operating cost and price coefficients.

6. The transaction cost variable introduces dynamic effects into the model such that a household's choice in one period affects its choice in the next period. Dynamics have only recently played a role in qualitative choice models of auto demand. Cambridge Systematics, Inc. (1980b), included a transaction cost variable in its model. However, since data on households' previous auto holdings were not available, the coefficient of this variable was not estimated within the model. Rather, a model excluding this variable was estimated, and the variable was added afterward, with its coefficient chosen so that the model produced the observed aggregate turnover rate of vehicles. The model presented in this chapter is the first to include a variable representing dynamics with its coefficient estimated statistically along with the other model parameters. More recently, Mannering and Winston (1983) have examined alternative variables for representing dynamics.

A transaction cost variable, and variables used by Mannering and Winston, are essentially lagged dependent variables, the inclusion of which raise econometric questions concerning the consistency and efficiency of the standard maximum likelihood estimation. Heckman (1981) has examined the situation in the context of probit models. However, the analysis is more complex for logit models since the convenient convolution properties of the normal distribution cannot be utilized.

7. Significance in this situation was defined as having a t-statistic exceeding 0.5.

8. Since W^2 in equation (8.13) necessarily increases with the variance of each characteristic, any variances that approximate W^2 must enter with a positive coefficient.

9. The author thanks D. McFadden for deriving this expression.

10. Since most 1976–1978 vintage vehicles are designated as prestigious, including this prestige dummy in the submodel reduces the coefficient of the vintage 1976–1978 variable below that of the vintage 1972–1975 variable. Perhaps the best way to interpret the coefficient of the vintage 1976–1978 variable is that it represents the extra utility that a household obtains from having two 1976–1978 vintage vehicles over that which it obtains with only one 1976–1978 vintage vehicle. This interpretation reflects the fact that the prestige dummy takes the value of one if either or both of the vehicles is prestigious, while the vintage 1976–1978 variable takes the value of one if one of the two vehicles is vintage 1976–1978 and two if both of the vehicles are vintage 1976–1978. This interpretation is not valid, however, for 1976–1978 vintage vehicles that are not classified as prestigious.

11. See section 5.4 for a full explanation of the potential bias and alternative methods for correcting it.

12. One further note is required concerning the estimation of the VMT submodel. Equation (8.10) includes the operating cost of the actual make and model of vehicle that a household owns. However, as stated the household's choice of make and model of vehicle is not predicted in our system of submodels. Consequently, the operating cost of the household's chosen class/vintage was used as a proxy for the operating cost of the particular make and model of vehicle chosen within the class/vintage.

13. A truly behavioral model of VMT by category would take as data for each household the cost and time of travel by each mode to work and nonwork destinations. Such data are unavailable for statewide or nationwide samples.

14. Gas price was entered rather than operating cost since operating cost is endogenous, and an instrumental variable approach to endogenous variables is not appropriate with logit.

Chapter 9

1. For cars, the Commission relied primarily on projections developed by Energy and Environmental Analysis, Inc.

2. The model is capable of handling a different income growth for each household in each forecast year, but this was not specified by the commission.

3. An elasticity is defined as the percent change in one variable resulting from a percent change in another variable. For example, the elasticity of VMT with respect to income is the percent change in VMT that results from a percent change in income (calculated in this case as 13.7%/48% = .29.)

Appendix B

1. Provided by Energy and Environmental Analysis, Inc. (1982), and the California Energy Commission.