A Review On Some Alternative Specifications Of The Logit Model

Md. Zakir Hossain, Sultan Qaboos University, Oman

ABSTRACT

The logit model, perhaps the simplest and the best possible probabilistic choice model in the discrete choice modeling literature. As a matter of fact, it has been extensively used in many statistical and economic applications. Unfortunately, a very unattractive property of this model in its multinomial situation is independence of irrelevant alternative (IIA) property. Due to such limitation, a number of alternative possible specifications have been proposed in the literature. This paper investigates some of the important alternative specifications of the logit model along with their merits, demerits, estimation techniques, testing procedures and attempts to advocate about the superiority of the existing specifications for the users. Interestingly, we found no such absolutely superior model to be used as an alternative to the logit model. However, generalized extreme value (GEV) model and multinomial probit model have been found to be very promising and much better than other models.

Keywords: Logit model; independence of irrelevant alternative property; generalized extreme value; nested logit; multinomial probit and dogit.

1.   INTRODUCTION

$\mathcal{P}$robably the simplest and best-known probabilistic choice model among the discrete choice behavior is the logit model. During the last few decades many studies have been conducted using this model. See for example, Theil (1969), Crag and Uhler (1970), Schmidt and Strauss (1975), McFadden (1973, 1978, 2001), Hosmer and Lemeshow (2000) and Cramer (2003). The logit model arises from the assumption that the random component $\in_{ij}$ (where i=1,2,…,n choosing alternatives/individual and j=0,1,…,m-1 categories) of the model follows the double exponential distribution - a member of the extreme value (EV) family of distributions. An important but very unappealing - property of the logit model in its multinomial situation is the independence of irrelevant alternatives (IIA). This property implies that adding another alternative or changing the characteristics of a third alternative does not affect the relative odds between the two alternatives considered. This implication is not realistic for applications with similar alternatives. Many examples have been constructed to illustrate this problem, see for example, Red Bus/Blue Bus problem in McFadden 1974. However, this property no longer holds if we relax the assumption that the random component $\in_{ij}$ is identically and independently distributed (iid). Due to such limitations of the widely used logit model, alternative specifications have become a natural development.

As an alternative specification, Tversky (1972) proposed the elimination-by-aspect (EBA) model and showed that the model is consistent with random preference maximization. But the model becomes computationally infeasible for large choice sets. Moreover, the model does not have the latent-variable characterization. So, the EBA model has not been used in any econometric applications and hence we have not considered this model in this paper. Tversky and Sattath (1979) specialized the EBA model to the situation in which the alternatives are represented by a tree graph. When aspects have a tree structure, the EBA model reduces to the hierarchical elimination-by-aspects (HEBA) model. Although it involves fewer parameters than the EBA model, the HEBA model has also not been used in any econometric field, see for example, Maddala (1983).

The conditional probit model of Hausman and Wise (1978) and the negative exponential distribution model can also be used as alternatives to the logit model as they overcome the restriction of IIA assumption. But these models can be represented by the multinomial probit model, which we have considered in our paper. In this paper we mainly concentrate on the alternative specifications of the logit model as a consequence of its well IIA assumption. The alternative model specifications presented subsequently in this paper will, it is hoped, shed some light on a critical review of the literature that may play a key role to the interested researchers in future. The organization of the paper is as follows. Generalized extreme value (GEV), nested logit, multinomial probit and dogit models are discussed in sections 2, 3, 4 and 5, respectively. Testing IIA and estimation of nested logit and dogit models are also explained in sections 3 and 5, respectively. Concluding remarks are provided at the final section of the paper.

2. GENERALIZED EXTREME VALUE MODEL

The Generalized Extreme Value (GEV) model has been introduced by McFadden (1978) in the context of residential location. The model provides a nice theoretical framework for the development of new discrete choice model, like Koppelman and Wen (1997) and Vovsha (1997). It is a class of models that involve generalizing the distributional assumption to be that of extreme value for the random error term from the random utility model (RUM)[2] model. This generalization is concerned with the assumption that the extreme value disturbances are not statistically independent and the model as such, is operationalized by the following distribution function of the error term

$$F(\epsilon_0, \epsilon_1, ..., \epsilon_{m-1}) = \exp[-G\{\exp(-\epsilon_0), ..., \exp(-\epsilon_{m-1})\}] \tag{1}$$

where G is a non-negative homogeneous function of degree 1 i.e., G satisfies the following conditions

(i) $\quad G(\gamma_0, \gamma_1, ..., \gamma_{m-1}) \geq 0$ and $\gamma_0, \gamma_1, ..., \gamma_{m-1} \geq 0$

(ii) $\quad G(\alpha\gamma_0, \alpha\gamma_1, ..., \alpha\gamma_{m-1}) = \alpha G(\gamma_0, \gamma_1, ..., \gamma_{m-1})$

(iii) $\quad G_j = \dfrac{\alpha n G(.)}{\alpha\gamma\frac{n}{j}} \geq 0$ if n is odd $\forall j$

where, $G(.) = \sum\limits_{j=0}^{m-1} \gamma_j, \ \forall \ j = 0, 1, 2, ..., m-1.$

[2]According to Haneman (1984) a random utility model arises when an individual assumes that, although a consumer's utility function is deterministic to him, it contains some components which are unobservable to the econometric investigator and are treated by the observer as random variables. The unobservable components could be characteristics of the consumer and/or attributes of the commodities. This concept, therefore, combines two ideas that have a long history in economics - the idea of a variation in choices among individuals in a population and the idea of unobserved variables in econometric models.

For a given G, McFadden (1978) has proved that the probabilistic choice model of GEV is consistent with RUM, giving a choice probability of the form

$$P_{ij} = \dfrac{\exp(V_{ij})G_j\{\exp(V_0), ..., \exp(V_{m-1})\}}{G\{\exp(V_0), ..., \exp(V_{m-1})\}} \tag{2}$$

where G_j is the partial derivative with respect to the jth argument and the utility

$$U_{ij} = V_{ij} + \in_{ij} \tag{3}$$

is the chosen alternative with highest utility. Here V_{ij} is the deterministic component and $\in_{ij}$ is the random component. As can be seen, the G function is the pivotal component of the GEV model as it provides the functional form of the random component. With the form of the choice probability as given by (2), it is clear that the odds ratio,

$\dfrac{\exp(V_{ij})}{\exp(V_{ik})}$, is now determined not only by the attributes of j and k but also by the characteristics of the entire

choice set, where $j \neq k = 0, 1, \ldots, m - 1.$ Thus, the GEV model is no longer characterized by the IIA property.

3.   NESTED LOGIT MODEL

The nested logit model, first derived by Ben-Akiva (1973), is an extension of the multinomial logit (also called conditional logit in econometrics) model designed to capture correlations among alternatives. It is based on the partitioning of the choice set into several nests. The nested logit model and the higher-level nested logit model are the special cases of the GEV model. This model is attributed to McFadden (1978) and is developed in greater detail in an article by McFadden (1981). The nested logit model can be considered to possess a branching structure. The model is, in fact, analogous to the one-way analysis of variance partitioning a choice set into mutually exclusive subsets. Here, the random utilities of similar alternatives within the same subgroup are correlated but the utilities of alternatives in different subgroups are independent.

Here we consider the well-known red bus-blue bus example to illustrate a two-level nested logit model. Commuters initially face a decision between two modes of transportation: car and red bus. Suppose a consumer chooses between these two options with equal probability, 0.5, so that the odds equal 1. Now suppose a third mode, blue bus, is added. Assuming the bus commuters do not care about the color of the bus, consumers are expected to choose between bus and car still with equal probability. But IIA implies that this is not the case, the probability of commuters that make each of the three modes equals one third. Here, the choice set is split between two alternatives in such a way that the relative probability of the choice of two alternatives, the red bus and the car, is made dependent on the characteristics of the blue bus.

This sort of two-level nested logit models is, in fact, based on psychological studies of preference determination and is consistent within individual utility mappings in a random utility scheme. Within the framework of the RUM model, we consider P_{ij} as the probability that j is selected by individual i conditional on subsets being chosen. Mathematically, the choice probability P_{ij} is given by the following equation

$$P_{ij} I_r = \frac{\exp\left(\dfrac{V_{ij}}{p_s}\right)}{\sum\limits_{\kappa \in B_r} \exp\left(\dfrac{V_{ij}}{p_s}\right)} \times \frac{\exp(p_s I_s)}{\sum\limits_{r-1}^{R} \exp(p_r I_r)} \tag{4}$$

where $I_r = \log\left[\sum\limits_{\kappa \in B_r} \exp\left(\dfrac{V_{ik}}{p_r}\right) \right]$ is the inclusive value which summarizes the attributes of alternatives

below a node. It is clear from equation (4) that the nested logit model is essentially a product of multiple logits,

where the number of the multiple is equivalent to the level of nesting within the model. It should be pointed out that this two-level nested logit model can be generalized to more than two levels, see for example, Chow (1983).

3(a)   Testing IIA

The fundamental role for alternative modeling specifications (such as the nested logit model) is allowed for traditional by hypothesis testing of null models (in our case the logit model). The test is as follows.

$$H_0 : p_r = 1 \rightarrow \text{logit model i.e. IIA holds}$$

$$H_1 : \text{at least one } p_r \neq 1 \text{ logit model specification is incorrect i.e. IIA does not hold.}$$

Where p_r is the coefficient on the inclusive value term for the rth subset. Notice that among the traditional testing procedures that could be employed to test H_0 against its alternative, each of the classical testing strategies could be used. Among three classical tests, however, the Lagrange Multiplier (LM) test may be referred in practice as it requires only the estimation of the null model.

3(b)   Estimation of Nested Logit Model

Maximum likelihood (ML) method can be employed to carry out estimation of the nested logit model because it provides desirable statistical properties in large samples. But the model is computationally tractable for large problems and involves the sequential use of the logit program. Maddala (1983) outlines the sequential estimation procedure for the model. The problems associated with the sequential estimation procedure may be as follows.

1.   Sequential estimates are consistent but inefficient,
2.   A simple computer program often provides unappealing results for the model,
3.   The model is complicated to specify and exploit, see for example, Daly (1987).

Despite the problems mentioned above, sequential estimation has been used in a number of studies. But in practice, the sequential estimation does not seem to be appreciated because of the following reasons.

1.   It is often different to derive acceptable results (e.g. coefficient with the right sign),
2.   The calculation of the `logsums' is somewhat tedious and difficult to check,
3.   It is time consuming and creates problem imposing constraints on coefficients that may be desirable,
4.   The results obtained by the process are less accurate in terms of statistical efficiency.

Many studies e.g., Daly and Zachary (1978), Small and Brownstone (1982) suggest that full-information maximum likelihood (FIML) estimation give better results than sequential estimation, by being statistically more efficient, time saving and accurate calculation. Econometric software packages, such as LIMDEP, Greene (1991), are capable of both of these estimations for nested logit model. As an alternative specification of logit model nested logit seems to be quite reasonable. Ortuzar (1983) also indicated that nested logit gives better result than multinomial logit model. Furthermore, for the vast majority of econometric applications, nested logit model could be a better specification than other tree models, such as the hierarchical elimination-by-aspects (HEBA) model, see for example, Maddala (1983).

4.   MULTINOMIAL PROBIT MODEL

Multinomial probit (MNP) model has been considered as one of the most appealing and promising alternatives to the logit model in the literature. The random components $\in_{ij}$ in the random utility model for such a case have a multivariate normal distribution. The main alternative feature of the MNP model is that it allows the

general covariance structure for the alternative specific errors, see for example, Bunch (1991). The MNP model was first proposed by Thurstone (1927) and has been applied to psychological-choice data by Bock and Jones (1968). Hausman and Wise (1978) applied the MNP model to the transit-choice problem. In addition to these, a large number of applications of the model has been done by Daganzo (1979), Johnson and Hensher (1982), Dansie (1985) and Kamakura (1989), etc. Although, the model has wide applications in various fields but the main disadvantage of the model is that it can be applied for a limited number (e.g. at most 4) of alternatives because of its computational difficulties, see for example, Maddala (1983).

In recent times, several alternative computational methods for the model have been suggested by different experts in the discipline. Lerman and Manski (1982) proposed a Monte Carlo method which starts with given values of the expected utility V_j (which is linear function of the attributes X_i) and draws vectors $(\in_0, \in_1, ..., \in_{m-1})$ from a multivariate normal distribution. Clark (1961) used an approximation method which is good for non-negatively correlated varieties of equal variances but is poor for negative correlations or unequal variances. It is also doubtful that the MNP model is worth for all the computational trouble when the number of choices will be greater than 4.

McFadden (1989) and Pakes and Polland (1989) suggest simulation methods which may lead to practical probit estimation codes for more than 4 alternatives. Kamakura (1989) demonstrates using a simulation study that the Mendal Elston approximation is more accurate than the Clark method. Bunch and Kitamura (1989) support Kamakura's results in a study using empirical data and discuss improve algorithms for maximum likelihood estimation. However, before conclusion, formulation of model is of prime importance. Following McFadden (1981, 1989), Bunch (1991) has given a random utility model as

$$U_n = X_n (\theta + \delta_n) + \mu + \in_n \tag{5}$$

where $U_n, \mu, \in_n \in R^m, 0, \delta_n \in R^k$ and $X_n \in R^{k \times m}$, and also

$$\delta_n \to MVN(0, s); s \in R^{k \times k}$$

and $\in_n \to MVN(0, t); t \in R^{m \times m}$

In this formulation, the terms δ and $\in$ are assumed to be independent. This framework is, however, consistent with Hausman and Wise (1978). Furthermore, it is flexible and consistent with many other specifications in the literature, see for example, Bunch (1991).

A serious limitation of Bunch's (1991) multinomial probit specification is that the model is usually incapable to identify all the estimable parameters. Users of MNP model frequently make modeling assumptions which are analogous to choosing among various alternatives structures in the nested multinomial logit or extreme value model. Nevertheless, we must say that the MNP model discussed here has some new insights into the model's behavior related to its specification, estimation and performance. This work may be remained for those researchers seeking to apply multinomial probit model as an alternative to the logit model.

5. DOGIT MODEL

The Dogit model once again, is a generalization of the logit model. The Dogit model can remove the IIA difficulty without losing the intuitive and practical appeal of the logit format. The model can be considered as a distinct model for the rest of the alternatives discussed in this paper. Because, it is quite different from both fully competitive (i.e. which satisfies simple salability or order independence - closely related to the IIA property of the logit model) and GEV classes of models, see for example, Gaudry and Dagenais (1979).

The Dogit model, however, does so by making no use of the assumption of underpinning the derivation of GEV models. Instead, Gaudry and Dagenais (1979) derived the Dogit specification by arbitrarily expanding the form of the choice probability. Within the framework of this specification the relative probability of selection of alternatives j and k is in terms of more than just the characteristics of these two alternatives, which gives

$$P_{ij} = \frac{\exp(V_{ij}) + \theta_j \sum_{k=0}^{m-1} \exp(V_{ik})}{\left(1 + \sum_{k=0}^{m=1} \theta_k\right) \sum_{k=0}^{m-1} \exp(V_{ij})} \tag{6}$$

$$i = 1, 2, ..., n; \; j = 0, 1, ..., m-1$$

An important characteristic of the Dogit model is that it does not impose the IIA property a priory on all pairs of alternatives. The odds ratio in the Dogit model is

$$\frac{P_{ij}}{P_{ik}} = \frac{\exp(V_{ij})0_i \sum_{k=0}^{m-1} \exp(V_{ij})}{\exp(V_{ik})\theta_j \sum_{t=0}^{m-1} \exp(V_{it})} \tag{7}$$

where $j \neq k = 0, 1, ..., m-1$. As the ratio clearly depends upon all of the alternatives in the choice set and the Dogit model does not show the IIA property and hence the model is seen as a useful specification. But on the other hand, it should be noted here that the Dogit model is not generated from the principles of random utility maximization theory. So, the main disadvantage of the model is that its applicability in qualitative response modeling is substantially limited as it is almost impossible to attach to it a plausible economic justification.

5(a). Testing IIA

As has been mentioned, the Dogit model is intended to be a behavioral model for individual decision making problem. It actually appears in the literature to provide usual diagnostic checking of the logit model. In this way the Dogit model gives a simple test of IIA by the following parameter restrictions

$H_0 : \theta_j = \mathbf{0,}$ where $j = 0, 1, ..., m-1 \rightarrow$ logit model (i.e. IIA holds)

$H_1 :$ at least one non-zero $\rightarrow$ logit model is the incorrect specification (i.e. IIA does not hold)

Thus, as the logit model is nested within the Dogit model, any classical tests (e.g. Wald, LR and LM) may be employed to undertake the test of H_0 against $H_1.$ Tse (1987) has outlined a lager sample diagnostic test (using the LM principle) for the IIA property using the Dogit model as the alternative (non-IIA) specification. The test is straightforward, where the power of the test depends on how satisfactorily the model represents the underlying probabilities of discrete choice behavior.

5(b). Estimation of Dogit Model

The software packages by which the Dogit model can be estimated are limited. For example, Greene's LIMDEP, which has various applications in econometric model estimation, does not allow Dogit estimation. This is

one of the main reasons why we use the LM test procedure as the only viable method by which IIA can be tested against the alternative hypothesis of the Dogit model.

The Dogit model has, however, wide scopes to empirical investigation either with aggregate or disaggregate data set. Moreover, the model is associated with its various functional forms that may be of future interest for the researchers. Among the alternative specifications of the logit model mentioned above, the nested logit model seems to be more sensible, popular and widely used model in the literature. Perhaps, its popularity in many areas is due to the fact that full information maximum likelihood (FIML) estimates for the model can be computed very easily. In fact, the FIML estimation gives better results by being statistically efficient, time saving and error-prove calculation. The LIMDEP, see for example, Green (1991), a widely used econometric software package, is capable of estimating the nested logit model. Ortuzar (1983) has indicated that the nested logit model gives better result than the multinomial logit model. Maddala (1983) also reported that in most of econometric applications, the nested logit model is a better specification than many other alternative models. Thus we can conclude that the nested logit model may be considered as one of the most appealing and promising alternatives to the logit model.

6. CONCLUDING REMARKS

In this paper we briefly reviewed some important alternative specifications of the logit models namely, generalized extreme value, nested logit, multinomial probit and dogit indicating their formulation, estimation, testing procedures and several directions for future research. The recent developments in logit models are quite encouraging but literature reveals that much yet to be done. In particular, there is a pressing need for development of more practical models for selection probabilities which do not have the IIA assumption. This is because the use of logit models has the potential of being largely affected by the characteristic axiom of this IIA assumption.

Amongst the alternative specification, we first identified the GEV model as it is one of the most promising models in the literature, see for example, McFadden (1981). It is a generalization of the logit model taking into account non-independence of alternatives. The model can be specialized to preference trees to accommodate the patterns of inter-alternative substitution found in many choice situations and can be expressed as a nested sequence of multinomial logit. Although, McFadden (1981) reports that the sequential estimation of such models is not too cumbersome even for relatively large and complex trees. But on the other hand, many studies, see for example, Daly and Zachary (1978) and Small and Brownstone (1982) suggest that the full information maximum likelihood estimation can give better results than sequential estimation by being statistically more efficient and time saving.

Multinomial probit model has also been considered as one of the most attractive alternatives of the logit model. A number of studies have been conducted using this model in the past years, see for example, Daganzo (1979), Johnson and Hensher (1982), Dansie (1985) and Kamakura (1989). The main disadvantage of the model is that its applicability is, to some extent, limited due to its computational difficulties, see for example, Maddala (1983). To overcome this problem, several alternative computational methods have also been suggested by researchers, see for example, Clark (1961), McFadden (1989), Pake sand Polland (1989), Kamakura (1989), Bunch and Kitamura (1989). The methods reveal that parsimonious estimation and prediction can be done by the multinomial probit model. Furthermore, the model has some new insights into its behavior related to its specification, estimation and performance. This work remains for the interested researcher in future.

As has already been mentioned, in our paper we have considered the nested logit model as one of the most reasonable alternatives to the logit model due to its desirable statistical properties. Although, a number of studies reveal that the nested logit model can provide better results than many other models, but the model is computationally tractable for various problems. The well-known FIML method has, however, given some solutions to its computational difficulties. However, if we wish to have a more successful and practical alternative specification to the logit model, we need to give more concentration to this problem in the near future.

In section 5, we presented our last port of all: Dogit model - a generalization (albeit arbitrary) of the logit model. Unlike the GEV models which are generated by distributional assumptions of the residuals of logit model, it, on the other hand, makes no use of such assumptions. Instead, Gaudry and Dagenais (1979) derived the model

specification by arbitrarily expanding the form of the model but the use of such model seems very limited. Nevertheless, it may have some scopes in empirical investigation considering its various functional forms for the future.

ACKNOWLEDGEMENTS

AUTHOR INFORMATION

Dr. Hossain has 23 years of teaching experience in 5 different universities in Bangladesh, Australia and now in Oman. Dr. Hossain's academic records are very bright. He obtained first class in all the examinations throughout his academic life. Dr. Hossain obtained his Ph.D. in Econometrics and Business Statistics in 1998 from Monash University under the guidance of Sir John Monash Distinguished Professor Maxwell L. King. Dr. Hossain worked as an Assistant Lecturer at Monash University from 1994 to 1997. He served Statistics Department, Jahangirnagar University, Bangladesh for more than a decade.

Dr. Hossain has published about 30 research papers in the field of econometrics, financial econometrics and business statistics in internationally reputed journals including *Communications in Statistics* (USA), *Journal of Statistical Theory and Application* (USA), *Journal of Applied Statistical Science* (USA), *International Journal of Social Economics* (UK), *Managerial Auditing Journal* (UK), *Managerial Finance* (UK), *Asian Profile* (Canada), *Genus* (Italy), *International Journal of Information and Management Sciences* (Taiwan), *Pakistan Journal of Statistics* (Pakistan), *Journal of Accounting, Business and Management* (Indonesia), *Journal of Statistical Studies* (Bangladesh). He received the "***Outstanding Research Award***" from the SQU in 2006.

REFERENCES

1.   Ben-Akiva, M.E (1973). Structure of passenger travel demand models. PhD thesis, Department of Civil Engineering, MIT, Cambridge, Ma, USA.
2.   Bock, R. and L. Jones (1968). *The measurement and prediction of judgement and choice*. San Francisco: Holden day.
3.   Bunch, D.S. (1991). Estimability in the multinomial probit model. *Transportation Research*, 1, 112.
4.   Bunch, D.S. and K. Kitamura (1989). Multinomial probit estimation revisited: Testing new algorithms and evaluation of alternative model specifications for trinomial models of household car ownership. Transportation Research Group, Research Report UCD-TRC-RR-4 University of California.
5.   Clark, C. (1961). The greatest of a fiite set of random variables. *Operation Research*, 9, 145-162.
6.   Chow, G.C. (1983). *Econometrics*. McGraw-Hill, New York.
7.   Cragg, J.G. and R. Uhler (1970). The demand for auto-mobiles. *Canadian Journal of Economics*, 3, 386-406.
8.   Cramer, J.S. (2003). The origins and development of the logit model. Working Paper. University of Amsterdam and Tinbergen Institute, Amsterdam.
9.   Daganzo, C. (1979). *Multinomial probit: The theory of its application to demand forecasting*. Academic Press, New York.
10.   Daly, A.J. (1987). Estimating tree logit models. *Transportation Research*, 4, 251-267.
11.   Daly, A.J. and S. Zachary (1978). Improved multiple choice models. In David A.H. and Q. Dalvi (eds.). *Determinants of Travel Choice*, Praeger Publishers, New York.
12.   Dansie, B.R. (1985). Parameter estimatimability in the multinomial probit model. *Transportation Research*, 19, 526-528.
13.   Gaudry, M.J.I and M.G. Dagenais (1979). The dogit model. *Transportation Research*, 13, 105112.
14.   Green, W.H. (1991). LIMDEP, version 6.0. Econometric Software, New York.
15.   Hanemann, W.M. (1984). Discrete/Continuus models of consumer demand. *Econometrica*, 52, 544-561.

16. Hosmer, D.W. and S. Lemeshow (2000). *Applied Logistic Regression* (2nd ed.) New York: Wiley. First edition in 1989.

17. Housman, J.A. and D. Wise (1978). A conditional probit model for qualitative chic, discrete decisions, recognizing interdependence and heterogeneous performance. *Econometrica*, 46, 403-426.

18. Johnson, L. and D. Hensher (1982). Application of multinomial probit to a two-period panel data set. *Transportation Research*, 16, 457-464.

19. Kamakura, W. (1989). The estimation of multinomial probit models: a non calibration algorithm. *Transportation Science*, 23, 253-265.

20. Koppelman F.S. and Chieh-Hua Wen (1997). The paired combinatorial logit model: properties, estimation and application. Transportation Research Board, 76th Annual Meeting, Washington DC, Paper #970953.

21. Lerman, S. and C. Manski (1982). On the use of simulated frequencies to approximate choice probabilities. In C. Manski and D. McFadden (eds.). *Structural Analysis of Discrete Data: With Econometric Applications*. Cambridge, MIT Press.

22. Maddala, G.S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge.

23. McFadden, D. (1973). Conditional logit analysis of qualitative chic behavior. In P. Zarembka (ed.) *Frontiers in Econometrics*, Academic Press, New York.

24. McFadden, D. (1978). Modeling the choice of residential location. In A. Kurlquist (ed.). *Theory and Residential Location*, North-Holland, Amsterdam.

25. McFadden, D. (1981). Econometric Models of Probabilistic Choice. In C.F. Manski and D. McFadden (eds.). *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, M.A.

26. McFadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration. *Econometrica*, 57, 995-1026.

27. McFadden, D. (2001). Economic Choices. *American Economic Review* 91, 352-370. Nobel prize acceptance speech.

28. Ortugar, J.D. (1983). Nested logit models for mixed travel in urban corridors. *Transportation Research*, 17, 283-299.

29. Pakes, A. and D. Polland (1989). Simulation and the asymptotic of optimization estimators. *Econometrica*, 57, 1027-1057.

30. Schmidt, P. and R.P. Strauss (1975). Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach. *Econometrica*, 43, 745-755.

31. Small, K.A. and D. Brownstone (1982). Efficient estimation of nested logit models: an application to trip training. Economic Research Program, Research memo. No. 296, Princeton University.

32. Theil, H. (1969). A Multinomial Extension of the Liner Logit Model. *International Economic Review*, 10, 251-259.

33. Thurstone, L. (1927). A law of comparative judgement. *Psychological Review*, 34, 273-286.

34. Tse, Y.K. (1987). A diagnostic test for the multinomial logit model. *Journal of Business and Economic Statistics*, 5, 283-286.

35. Tversky, A. (1972). Elimination by aspects: a theory of choice. *Psychology Review*, 79, 281-299.

36. Tversky, A. and S. Sattath (1979). Preference trees. *Psychology Review*, 86, 542-573.

37. Vovsha P. (1997). Cross-nested logit model: an application to mode choice in the Tel-Aviv metropolitan area. Transportation Research Board, 76th Annual Meeting,Washington DC, Paper # 970387.

NOTES