82

Reprinted from

TRANSPORTATION RESEARCH

Vol. 13B, *No.* 2, *pp.* 105–111

THE DOGIT MODEL

MARC J. I. GAUDRY and MARCEL G. DAGENAIS

Centre de recherche sur les transports, Université de Montréal, Montréal, Quebec, Canada

PERGAMON PRESS
NEW YORK · OXFORD · FRANKFURT · PARIS
1979

Transpn Res.-B Vol. 13B, pp. 105–111.
Pergamon Press Ltd., 1979. Printed in Great Britain

THE DOGIT MODEL

MARC J. I. GAUDRY and MARCEL G. DAGENAIS
Département de sciences économiques et Centre de recherche sur les transports, Université de Montréal, Quebec, Canada

(*Received* 21 *November* 1977; *in revised form* 25 *April* 1978)

Abstract—This paper presents the dogit model. That model is flexible enough to permit the choice among specific pairs of alternatives to be consistent with the independence from irrelevant alternatives axiom, as in a logit model, but it simultaneously allows the choice among other pairs not to be. Dogit parameters add an "income effect" to the "substitution effect" already built into the logit model; alternatively, they allow for the joint presence of compulsive and discretionary elements in consumer behavior, or for the identification of captive markets. Eventual estimation of the values of the parameters of the dogit model appears simpler than for the probit model.

INTRODUCTION

Recent developments of probabilistic models (e.g. Theil, 1967 and 1969), notably in the field of travel demand analysis (e.g. Domencich and McFadden, 1975, Stopher and Meyburg, 1976 and Hensher and Stopher, 1978), have centered mostly on the logit model because of its relative simplicity. At the same time, numerous authors (e.g. McFadden, Tye and Train, 1976, and McFadden and Talvitie, 1977) have been concerned with the consistency of that model with the independence from irrelevant alternatives (henceforth IIA) axiom of choice theory, a property which it shares with numerous market share models. Formulations which circumvent the problem have either made the logit a special case of the generalized extreme value model (McFadden, 1978a) or developed an altogether new distribution (McLynn, 1973) which collapses to the logit as a special case. Because it is free from the IIA restrictions, the probit model has also been the object of renewed interest directed particularly at constructing efficient estimation algorithms (Albright, Lerman and Manski, 1977).

This article will introduce a new model, called the dogit model, which obviates the IIA difficulty without losing the intuitive and practical appeal of the logit format which it allows to hold for some pairs of alternatives but without simultaneously destroying itself as a distinct model for the rest of the alternatives considered.† We will provide interpretations of the additional parameters of our model in terms of analogies with micro-economic consumer demand theory and, as a special case, suggest a complementary rationale inspired by the theory of mathematical transformations.

THE MODEL AS A DISTRIBUTION

Consider the following distribution

$$p_i = \frac{e^{V_i} + \theta_i \sum_j e^{V_j}}{\left(1 + \sum_j \theta_j\right) \sum_j e^{V_j}}, \quad i,j = 1, \ldots, N, \tag{1}$$

where p_i: probability of choosing the ith of N alternatives; V_i: a function of K independent attributes X_{ik} associated with the ith alternative, for instance $V_i = \sum_k \beta_k X_{ik}$; θ_i: a non negative parameter associated with the ith alternative: $\theta_i \geq 0$.

Equation (1) is defined in such a way that

$$0 \leq p_i \leq 1, \qquad i = 1, \ldots, N. \tag{2}$$

†The model avoids or dodges the researcher's dilemma of choosing *a priori* between a format which commits to IIA restrictions and one which excludes them—whence its name.

and

$$\sum_i p_i = \frac{\sum_i e^{V_i} + \sum_i \theta_i \sum_j e^{V_j}}{\left(1+\sum_j \theta_j\right) \sum_j e^{V_j}} = 1. \tag{3}$$

This distribution has a number of interesting properties.

(a) *Unconstrained by* IIA *axiom in general*

The IIA property of a model requires that the relative probabilities of any two alternatives be independent of the attributes of the remaining alternatives whether the latter are present or absent from the set of available choices. This property, introduced by Luce (1959) to explain the psychology of choice, was used by McFadden (1968) to derive the multinomial logit model within a different theoretical framework. In our case, the relative probabilities of any two alternatives i and j are simply:†

$$\frac{p_i}{p_j} = \frac{e^{V_i} + \theta_i \sum_j e^{V_j}}{e^{V_j} + \theta_j \sum_j e^{V_j}}, \tag{4}$$

where the ratio obviously depends on the attributes of all alternatives; this implies that, if θ_i and θ_j differ from zero for all alternatives, the introduction of a new alternative will in general change the relative probabilities of alternatives. For this ratio to remain constant after a change in the value of characteristics of the lth alternative (for $l \neq i, j$), or following the introduction of a new alternative, it is necessary that

$$\frac{e^{V_i} + \theta_i \sum_j e^{V_j}}{e^{V_i} + \theta_i \sum_j e^{V'_j}} = \frac{e^{V_j} + \theta_j \sum_j e^{V_j}}{e^{V_j} + \theta_j \sum_j e^{V'_j}}, \tag{5}$$

or that the proposed change, represented by using V'_j instead of V_j, result in

$$\theta_i e^{V_j} \sum_j e^{V'_j} + \theta_j e^{V_i} \sum_j e^{V_j} = \theta_j e^{V_i} \sum_j e^{V'_j} + \theta_i e^{V_j} \sum_j e^{V_j}. \tag{6}$$

This requirement can be met if any of the following occurs:

$$\theta_i = \theta_j = 0, \text{ or} \tag{7}$$

$$\frac{\theta_i}{\theta_j} = \frac{e^{V_i}}{e^{V_j}}, \text{ or} \tag{8}$$

$$\sum_j e^{V_j} = \sum_j e^{V'_j}. \tag{8'}$$

If condition (7) holds for *all* pairs of alternatives, the specification (1) collapses to the logit specification. If (8) holds for *some* pairs (including the special case of equal θ_i's and V_i's) or if (8′) holds, the dogit remains distinct from the logit but modifying the attributes of the lth alternative, or introducing a new alternative, will leave the relative probabilities of those pairs constant. Our formulation therefore allows for *subsets* of relative probabilities to be determined consistently with the IIA axiom while others are not: the model is flexible in the sense that it does not have to collapse and require that *all* relative probabilities be simultaneously determined in the same way.

†The use of j both as a specific mode index and as a running index may force the reader to rely on the context in order to clarify some ambiguous expressions, notably (22). This was realized too late for the modifications to be incorporated into the final text.

(b) *Differences with fully competitive and GEV models*

Because the ratio p_i/p_j is in general a function of more than just the variables associated with alternatives i and j, the dogit model could be called "fully competitive" although it does not belong to the class of models wich McLynn (1973) formally specified when he coined that expression. McLynn started with a class of models of the form

$$p_i = \frac{w_i}{\sum_j w_j}, \qquad \sum_i p_i = 1, \tag{9}$$

and constructed a new model of the form

$$P_i = p_i(1 + \lambda u_i), \tag{10}$$

where $\lambda \geq 0$ and u_i is a function of the attributes of all alternatives such that $\sum_i p_i = 1$. This required $\sum_i u_i p_i = 0$. He used for that purpose the following class of u_i:

$$u_i = f(p_i) - \sum_j p_j f(p_j) \tag{11}$$

which implies for (10) the general form

$$P_i = p_i + \lambda p_i f(p_i) - \lambda p_i \sum_j p_j f(p_j). \tag{12}$$

Had one started with a multinomial logit in (9) and used $f(p_i) = p_i$ in (11), eqn (12) would be

$$P_i = \frac{e^{V_i}}{\sum_j e^{V_j}} + \lambda \left(\frac{e^{V_i}}{\sum_j e^{V_j}}\right)^2 - \lambda \frac{e^{V_i}}{\sum_j e^{V_j}} \sum_j \left(\frac{e^{V_i}}{\sum_j e^{V_j}}\right)^2 \tag{13}$$

which cannot be reduced to (1) except in the special case in which $(\theta_1, \ldots, \theta_N)$ in (1) and λ in (13) are zero. McFadden (1975) has shown that the fully competitive model satisfies "simple scalability" or "order independence", which is closely related to the IIA property of the logit model.

The dogit model also differs from McFadden's (1978a) generalized extreme value (GEV) model which can be summarized as

$$p_i = \frac{e^{V_i} G_i(e^{V_1}, \ldots, e^{V_i}, \ldots, e^{V_N})}{G(e^{V_1}, \ldots, e^{V_i}, \ldots, e^{V_N})}, \tag{14}$$

where $G(.)$ is a function satisfying certain conditions and $G_i(.)$ is the derivative of that function with respect to its ith argument. If $G(.)$ is, in a particular case, equal to $\sum_j e^{V_j}$, the GEV model reduces to the multinomial logit model. If instead, it is equal to $(1 + \sum_j \theta_j)\sum_j e^{V_j}$, it also reduces to the multinomial logit model:

$$p_i = \frac{e^{V_i}\left(1 + \sum_j \theta_j\right)}{\left(1 + \sum_j \theta_j\right) \sum_j e^{V_j}} = \frac{e^{V_i}}{\sum_j e^{V_j}}. \tag{15}$$

The dogit model is therefore different from both fully competitive and GEV classes of models.

(c) *Admission of an "income effect"*

The θ_i parameters in (1) can be interpreted as indexes of the extent to which the attributes of all alternatives influence the choice of the ith alternative. They function as a sort of "income effect" which can reinforce or weaken the automatic "substitution effect" embodied in the logit model. We will however show that the former cannot be stronger than the latter effect and lead to inferior and to complementary alternatives. One way to see both of these points is to compare the elasticities of demand computed from (1) to those of the logit model.

A certain percentage change in the value of one of the attributes of the ith alternative, X_{ik}, will induce the following percentage change in the probability p_i, if V_i is linear both in parameters β_k and in variables X_{ik}:

$$\eta_{ii} = \frac{\partial p_i}{\partial X_{ik}} \cdot \frac{X_{ik}}{p_i} = \beta_k X_{ik} \left[\frac{e^{V_i}(1+\theta_i)}{e^{V_i} + \theta_i \sum_j e^{V_j}} - \frac{e^{V_i}}{\sum_j e^{V_j}} \right] \tag{16}$$

which, if $\theta_i = 0$, reduces to the well known elasticity in the logit model

$$\eta_{ii} = \beta_k X_{ik} \left[1 - \frac{e^{V_i}}{\sum_j e^{V_j}} \right]. \tag{17}$$

To isolate the impact of θ_i, rewrite (16) as

$$\eta_{ii} = \beta_k X_{ik} \left[\frac{(1+\theta_i)}{1 + \theta_i \left(\sum_j e^{V_j}/e^{V_i} \right)} - \frac{e^{V_i}}{\sum_j e^{V_j}} \right]. \tag{18}$$

Because the expression $\sum_j e^{V_j}/e^{V_i}$ in the denominator of (18) is always positive and larger than one, the size of θ_i will determine the size of the "income effect" or the extent to which the first term in the parenthesis of (18) is smaller than one. Because that term obviously cannot be larger than one, it will *reduce* the effect which would otherwise have occured if it had been equal to one. However, we will not find a situation in which this difference is such as to imply that an increase in the price of an alternative will lead to an increased demand for it. Indeed, define

$$a = \frac{e^{V_i}}{\sum_j e^{V_j}}, \quad \text{where } 0 < a < 1, \tag{19}$$

and rewrite the square parenthesis of (18) as

$$d = \left[\frac{(1+\theta_i)}{1 + \frac{\theta_i}{a}} - a \right] = \left[\frac{(1+\theta_i)a - a^2 - a\,\theta_i}{a + \theta_i} \right]. \tag{20}$$

Equation (20) implies

$$\frac{d(a+\theta_i)}{a} = 1 - a, \quad \text{or} \quad d > 0. \tag{21}$$

Because the parenthesis is always positive, alternatives in the dogit model still behave as superior alternatives.

An analogous situation arises if cross elasticities are examined. We have

$$\eta_{ij} = \beta_k X_{jk} \left[\frac{\theta_i}{\frac{e^{V_i}}{e^{V_j}} + \theta_i \left(\sum_j e^{V_j}/e^{V_j} \right)} - \frac{e^{V_j}}{\sum_j e^{V_j}} \right] \tag{22}$$

which, if $\theta_i = 0$, reduces to the cross elasticity in the logit model,

$$\eta_{ij} = \beta_k X_{jk} \left[-\frac{e^{V_i}}{\sum_j e^{V_j}} \right]. \tag{23}$$

In this case, the first term in the square parenthesis of (22) is always greater than zero and will *increase* the effect which would otherwise have occurred if it had been equal to zero. It could further be shown, in a way analogous to (19)–(21), that the square parenthesis in (22) is always negative. In consequence, alternatives in the dogit model cannot behave as complements and must, as in the logit, behave as substitutes.

There are numerous situations in which the presence of this "income effect" might make a significant difference to the results. In fact dogit results should differ more from logit results when the price of a dominant alternative is changed than when the price of a minor alternative is changed: considering (20), we note that d will fall, or that the impact of the "income effect" will increase, as $a \to 1$. A similar result could be brought out from the cross elasticity formula. It is reasonable that the "income effect" impact should increase with the importance of the modified alternative.

(d) *Reconciliation of compulsive and discretionary behavior*†

It is possible to view the model presented in (1) in another way. Suppose that one observes aggregate market shares $y_1, \ldots, y_N$ of N alternatives or commodities. Then one can specify the following market share model:

$$y_i = \frac{e^{V_i} + \theta_i \sum_j e^{V_j}}{\left(1 + \sum_j \theta_j\right) \sum_j e^{V_j}} = \frac{e^{V_i}}{\left(1 + \sum_j \theta_j\right) \sum_j e^{V_j}} + \frac{\theta_i}{\left(1 + \sum_j \theta_j\right)}. \tag{24}$$

If consumers, in the spirit of Stone's (1954) approach to demand systems, spend a given share $[\theta_i/(1 + \sum_j \theta_j)]$ of their income to satisfy a compulsive or basic need for the ith commodity *irrespective of its price*, then $1 - \sum_i [\theta_i/(1 + \sum_j \theta_j)]$ or $[1/(1 + \sum_j \theta_j)]$ is the share of discretionary income left over after such needs are satisfied. Assuming further that the logit model correctly explains how discretionary income is spent, the ith commodity share of expenditure will be the sum of compulsive and discretionary elements or

$$y_i = \frac{1}{1 + \sum_j \theta_j} \cdot \frac{e^{V_i}}{\sum_j e^{V_j}} + \frac{\theta_i}{1 + \sum_j \theta_j}, \tag{25}$$

which is clearly identical to (24).

The dogit reconciles elements of constrained choice with elements of free choice in a simple formula. Alternatively, the dogit can be said to allow for a certain "captivity" of consumers to specific alternatives, as this term is used in travel mode choice contexts. The second term of (24) or (25) is then taken to represent the basic need for, or the captive market share of, a given alternative. Ben-Akiva (1977) has recently formally investigated and extended this interpretation of the dogit. McFadden (1978b) has recently interpreted the dogit as a special case of mixed probability random utility models.

Seen in the light of constrained choice theory, lower dogit own price market share elasticities derived above reflect the fact that, if some consumers are insensitive to price, a greater change in relative prices will be required to modify market shares by a given amount than if all consumers are unconstrained. The size of the basic or captive market puts an absolute limit on such modifications: the share of an infinitely unattractive alternative tends to

†In November 1977, Daniel Leblanc of *L'Université de Montréal* and Daniel McFadden of the University of California at Berkeley respectively suggested the analogy with Stone's expenditure system and the similar captivity interpretation as comments on an earlier version of this paper.

its captivity level $\theta_i/(1+\sum_j \theta_j)$. Similarly, equally attractive alternatives may have different market shares if they have different captivity levels.

SPECIAL CASE: AN INTERPRETATION IN TERMS OF FUNCTIONAL FORM

Share model (24) can also be viewed in a different perspective. Assume that $\theta_i = \theta_j = \theta$ and rewrite it as

$$y_i - \mu = \frac{e^{V_i}}{\frac{\theta}{\mu}\sum_j e^{V_j}}, \quad \text{where} \quad \mu = \theta/(1+N\theta). \tag{26}$$

Considering the relative shares of alternatives i and j yields

$$\frac{y_i - \mu}{y_j - \mu} = \frac{e^{V_i}}{e^{V_j}}, \tag{27}$$

where μ is simply a parameter which displaces the mean of the observations on market shares: it is in effect the location parameter studied by Anscombe and Tukey (1954), referred to in Tukey (1957), in the context of the power family of transformations. One could estimate μ and, knowing N, the number of alternatives considered, obtain an estimated value for θ. The Tukey location parameter applied to the dependent variable of a logit market share model can therefore sometimes be interpreted as a way of transforming a logit model into a particular case of the dogit model or as a way of introducing "income" or "compulsion" effects in a logit model.

By contrast, the use of the μ parameter to transform the *right-hand side* variables of a logit model does not change the fundamental properties of that model. Neither does the use of Box–Cox (1964) transformations on those same independent variables: all logit variants so obtained (Gaudry and Wills, 1977) will have the basic format defined in (9) because these transformations simply amount to complex specifications of the V_i functions used in this analysis. The application of Tukey and Box-Cox transformations to independent variables also leaves the basic format of the dogit model unaffected.

Box–Cox transformations, however, cannot be used directly on the *left-hand side* of either model type since they generally imply that $\sum_i p_i \neq 1$, as in

$$\frac{(y_i)^\lambda - 1}{\lambda} = \frac{q_i}{\sum_j q_j}, \tag{28}$$

where the right-hand side stands for any share format.

CONCLUSION

The dogit model appears promising enough to warrant empirical investigation either with disaggregate or with aggregate data: the possibility of estimating "income" or "compulsion" effects and of simultaneously relaxing the logit straightjacket wherever necessary suggests rich possibilities of application. The link between the dogit model and the literature on functional form is also of some interest. The fact that the dogit comprises no integrals might also give it a computational edge over the probit model which some would like to see replace the logit because it is free from the IIA restrictions.

Acknowledgements—This research was supported by the Research and Development Centre of Transport Canada and by the F.C.A.C. program of the *Ministère de l'éducation du Québec*. Michael Florian, Daniel Leblanc and Daniel McFadden made useful comments.

REFERENCES

Albright R. L., Lerman S. R. and Manski C. F. (1977) Report on the development of an estimation program for the multinomial probit model. Cambridge Systematics, Inc., Cambridge, Mass.

Anscombe F. J. and Tukey J. W. (1954) The criticism of transformations. Paper presented before the American Statistical Association and Biometric Society, Montreal.

Ben-Akiva M. (1977) Choice models with simple choice set generating processes. Department of Civil Engineering, M.I.T., Cambridge, Mass.

Box G. E. P. and Cox D. R. (1964) An analysis of transformation. *J. Roy. Statis. Soc.* Series B. 211–243.

Domencich T. A. and McFadden D. (1975) *Urban Travel Demand.* North-Holland, Amsterdam.

Gaudry M. J. I. and Wills M. J. (1977) Estimating the Functional Form of Travel Demand Models. Publication 63, Centre de recherche sur les transports, Université de Montréal. *Transpn Res.* to be published.

Hensher D. and Stopher P. R. (1978) *Behavioural Travel Modelling.* Croom Helm, London.

Luce R. D. (1959) *Individual Choice Behavior.* Wiley, New York.

McFadden D. (1968) The revealed preferences of a government bureaucracy. *Technical Rep.* No. 17, Department of Economics, University of California.

McFadden D. (1975) Economic applications of psychological choice models. *Working paper* No. 7519, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.

McFadden D. (1978a) A closed form multinomial choice model without the independence from irrelevant alternatives restriction. Summarized in "Quantitative methods for analyzing travel behavior: some recent developments", forthcoming in Hensher D. and Stopher P. R., *Behavioural Travel Modelling.* Croom Helm, London.

McFadden D. (1978b) Econometric models for probabilistic choice. Paper presented at the NBER-NSF Conference on Decision Rules and Uncertainty, Carnegie-Mellon University. Mimeographed. p. 19.

McFadden D. and Talvitie A. (1977) The independence from irrelevant alternatives property of the multinomial logit model. Chap. 1 of *Demand Model Estimation and Validation*, Urban Travel Demand Forecasting Project, University of California, Berkeley. *Final Rep.* 263–231.

McFadden D., Tye W. and Train K. (1976) Diagnostic tests for the independence from irrelevant alternatives property of the multinomial logit model. *Working paper* No. 7616, Urban Travel Demand Forecasting Project, University of California, Berkeley.

McLynn J. M. (1973). *Disaggregate Modal Choice Models of Fully Competitive Type.* U.S. Department of Transportation Contract DOT-UT-200021. Washington, D.C.

Stone R. (1954) Linear expenditures systems and demand analysis: an application to the pattern of British demand. *Econ. J.* **64**, 511–527.

Stopher P. R. and Meyburg A. H. (1976) *Behavioral Travel-Demand Models.* D.C. Heath, Lexington, Mass.

Theil H. (1967) *Economics and Information Theory*, Chap. 3. North-Holland, Amsterdam.

Theil H. (1969) A multinomial extension of the linear logit model. *Int. Econ. Rev.* **10**(3), 251–259.

Tukey J. W. (1957) On the comparative anatomy of transformation. *Ann. Math. Statist.* **28**, 602–632.