CRT 893

CENTRE DE RECHERCHE SUR LES TRANSPORTS
(C.R.T.)
CENTRE FOR RESEARCH ON TRANSPORTATION

UNIVERSITÉ DE MONTRÉAL
ÉCOLE DES HAUTES ÉTUDES COMMERCIALES
ÉCOLE POLYTECHNIQUE

CRT-893

Asymmetric Shape and Variable Tail Thickness in Multinomial Probabilistic Responses to Significant Transport Service Level Changes

by

Marc Gaudry

CENTRE DE RECHERCHE SUR LES TRANSPORTS
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCCURSALE A
MONTRÉAL
CANADA H3C 3J7

TÉLÉPHONE : (514) 343-7575
EMAIL : crt@crt.umontreal.ca
TÉLÉCOPIEUR (FAX) : (514) 343-7121

CRT-893

Asymmetric Shape and Variable Tail Thickness in Multinomial Probabilistic Responses to Significant Transport Service Level Changes

by

Marc Gaudry

Centre de recherche sur les transports et Département de sciences économiques, Université de Montréal, C.P. 6128, succursale A, Montréal, Canada H3C 3J7.

This ongoing work and previous research on the models have been supported by Transport Canada, by the *Ministère de l'éducation supérieure, de la science et de la technologie* (M.E.S.S.T.) of Québec and by the Social Sciences and Humanities Research Council (S.S.H.R.C.) and the Natural Sciences and Engineering Research Council (N.S.E.R.C.) of Canada. Yvan Taillefer contributed as research assistant.

Queen's University
John Deutsch Institute for the study of Economic Policy,
Discussion paper # 13

Université de Montréal
Centre de recherche sur les transports, Publication #893
Département de sciences économiques, Cahier #9311

June 1989
Revised July 1990 — Revised April 1993

Dépôt légal – Bibliothèque nationale du Québec,
Bibliothèque nationale du Canada, 1993

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Tables

Figures

Abstract

In this paper, we consider five types of mode choice models, all of which contain the LINEAR LOGIT as an embedded special case. The model types belong to the BOX-COX LOGIT, DOGIT and INVERSE POWER TRANSFORMATION families. We focus on particular features of the probabilistic response curve that can make a difference in the evaluation of changed transport conditions, such as fares or speeds. We find, with a sample of four intercity passenger modes for 120 Canadian city pairs in 1976, evidence of both modal captivity and asymmetric response that would make a difference to the evaluation of significant service improvements. From a modeling point of view, the LINEAR LOGIT form is easily rejected in favour of any of the five types of generalizations explored.

Key-words: Logit, Box-Cox Logit, Standard Dogit, Generalized Dogit, Linear Inverse Power Transformation Logit, Box-Tukey Inverse Power Transformation Logit, High Speed Rail, Intercity Mode Choice, Canada.

Résumé

Nous considérons cinq genres de modèles de choix du mode de transport qui tous comprennent le Logit linéaire comme cas particulier. Les genres de modèles considérés appartiennent aux familles BOX-COX LOGIT, DOGIT et TRANSFORMATION PUISSANCE INVERSE. Nous nous attardons à des caractéristiques de la courbe de réaction qui peuvent avoir une influence sur l'évaluation de nouvelles conditions du transport, dont les tarifs et les vitesses. À l'aide d'un échantillon interurbain canadien portant, pour les quatre modes de transports de voyageurs, sur 120 paires origine-destination, nous montrons que la présence de captivité et d'une courbe asymmétrique influencerait l'évaluation de l'impact de changements significatifs du niveau de service. Il est facile de rejeter le Logit Linéaire si on a le choix entre lui et n'importe lequel des cinq autres genres de modèles considérés.

Mots-clés: Logit, Logit Box-Cox, Dogit Standard, Dogit Généralisé, IPT-Logit Linéaire, IPT-Logit Box-Tukey, Choix du mode de transport interurbain, train à grande vitesse, Canada.

1. INTRODUCTION: MODELING THE RESPONSE TO SIGNIFICANT SERVICE LEVEL CHANGES

In this paper, we are concerned with the adequacy of travel demand model design features in giving adequate treatment to the introduction of new transportation services, such as high speed rail. We focus on features that can make a difference in the evaluation of significant changes in fares and other transport conditions, notably speeds.

Our emphasis is on five types of mode choice models, all of which contain the **LINEAR LOGIT as an embedded nested special case.** The model types belong to the BOX-COX LOGIT, DOGIT AND INVERSE POWER TRANSFORMATION families. The mode choice models are situated in a Quasi-Direct Format (QDF) in order to give an idea of their limitations as measures of total demand.

Using an improved specification of the mode choice and total demand components found in Gaudry and Wills (1978), we effectively generalize the mode choice component of that paper by considering five model types instead of considering only the STANDARD BOX-COX LOGIT. Two of the mode choice model types considered, the STANDARD DOGIT and the LINEAR INVERSE POWER TRANSFORMATION-LOGIT, have been the object of published empirical work by this author or others[1] but the other two, the GENERALIZED DOGIT AND THE BOX-TUKEY INVERSE POWER TRANSFORMATION-LOGIT[2], are just beginning to be applied.

We find, with a sample of four intercity modes for 120 Canadian city pairs in 1976, evidence of both modal captivity and asymmetric response that would make a difference to the evaluation of significant service improvements.

2. IMPLICIT MODELING FORMAT

2.1. Issue: from introducing new modes to improving old modes

Twenty-five years ago, the introduction of new transportation technology, such as the implementation of STOL (Short Take-Off and Landing) aircraft, appeared to require the development of a new generation of demand models. These models would naturally be specified independently from the existing "institutional" modes because the demand functions required for forecasting and evaluation of new modal services could clearly not be estimated from revealed preferences for goods that did not yet exist: even perfect knowledge of the Slutsky matrix for all existing goods in the world economy would have said nothing about forthcoming new goods. The solution to **Slutsky irrelevance** seemed to lie in the Lancastrian recognition of the vectorial nature of goods and services and in the abandonment of the very productive scalar simplifications made in microeconomics one hundred years ago.

In principle, this vision of objects as vectors of characteristics should have led to fully mode- and itinerary-abstract transportation models (imposing identical weights to a given characteristic in all

[1] In mode choice applications, the STANDARD DOGIT has been used for instance by Gaudry and Wills (1979) and Swait and Ben-Akiva (1987); the LINEAR INVERSE POWER TRANSFORMATION-LOGIT has been used by Montmarquette and Mahseredjian (1985) to estimate educational production functions.

[2] The present author is preparing with M. Dagenais a paper containing detailed tests of a mode-specific version of the multinomial GENERALIZED DOGIT used here and with R. Laferrière a paper containing detailed tests of the LINEAR INVERSE POWER TRANSFORMATION-LOGIT on a binomial urban example.

alternatives). In practice, a number of factors have limited the usefulness of the fully abstract vision to itinerary choice. Prudent modelers have not used abstract specifications to forecast demand for "new modes" represented as yet another abstract alternative: service innovations are more convincingly treated as additional services (or paths) of existing modes. Of course, mode choice specifications do not typically use abstract specifications for socioeconomic variables: the option of defining a significant innovation as a new modal alternative does not really arise in that case because one cannot decide which of the modal "colors" (mode-specific weights) are appropriate to the new mode.

It is not altogether understood why so much mode specificity remains in mode choice models. Relevant factors pertain to: i) the specification of variables, such as incompleteness in the description of alternatives, other error in the selection of socio-economic variables (including their different roles in mode and itinerary choice), and errors of observation (or of construction when they are built as aggregations of paths) of the network characteristics; ii) the implicit presence of a selection bias for the problem at hand because, even in random samples of individuals, trips not made are not represented.

2.2. Model features of interest

Once it is clear that the new technology will be represented as an improvement of an existing institutional mode, the focus shifts from the determination of the appropriate level of "abstractness" to the provision of an **adequate representation of the effects of fares and service levels** of existing modes.

We want to isolate here two features of mode choice models that are of specific relevance when significant changes in fares or service levels are under consideration: the first is that of the **form of the monotonically increasing sigmoïd response curve**, particular its potential asymmetry, and the second is that of the **probability limits of this curve**, notably its thickness in the various directions.

These features are of interest because strong asymmetry might mean that a certain large improvement in modal speeds will lead to stronger modal shifts if one is in the appropriate range[3], and to more trivial shifts outside of this range, than if a symmetric shape had been assumed: this would influence the optimal dispersion of investments among links of the network. The presence of thick tails, indicative of captivity, might influence pricing of the services because fare elasticities obtained **after** captivity has been taken into account might be higher (or lower) in the relevant range than those obtained without explicit recognition of the presence of captivity: then new services will not (or will) be priced much higher than existing ones without fear of market share loss[4].

2.3. Implicit context: aggregate cross-sectional origin-destination models

One expects that mode choice models that allow for tail thickness, asymmetry of response, or both, would **also have different "trip generation" or "induced traffic" effects** than models that do not have these features. The implicit context of our exercise is that of the quasi-direct format, originally proposed by McLynn et *al.* (1968, 1969), that explains trip demand by mode $\mathbf{T_m}$ as the product of a market share function and a total demand function for travel irrespective of mode. Neglecting the origin-destination subscripts:

[3] A case in point might be that of the addition of fast TGV trains in the Paris-Lyon corridor where total passenger traffic on all trains increased by 60% between 1980 and 1987 while air traffic fell by 48%.

[4] TGV train prices are only about 20% higher than regular trains prices.

$$T_m = T \cdot p_m \,, \tag{1}$$

in which the dot separates total demand and market share parts, explicitly

$$T = g^d([X_S], U) \,, \tag{2}$$

and

$$p_m = U_m / U \,, \; m = 1, \ldots, M; \tag{3}$$

where

$$U_m = u_m([X_N], [X_S]) \,, \; m = 1, \ldots, M; \tag{4}$$

$$U = (U_1 + \ldots + U_M) \,, \tag{5}$$

and the sets $[\mathbf{X_s}]$ and $[\mathbf{X_N}]$ respectively denote socioeconomic (s = 1,...,S) and network (n = 1,...,N) variables, and $\mathbf{g^d}(.)$ and $\mathbf{u_m}(.)$ are functions. We assume that the latter set contains, in particular, indices of modal service levels (such as fare, frequency or travel time by air, train, bus or car) by origin-destination pair, obtained by aggregating somehow over itineraries or paths[5].

This quasi-direct format (QDF) is an efficient compromise, in the current state of knowledge of origin-destination demand models, between the easily inconsistent classical sequential format and the nearly forgotten[6] direct format in which the modal demand by origin-destination pair is obtained by a single equation effecting in one step the classical generation, distribution and mode choice steps.

---

5 The requirement that demand functions by origin-destination pair contain unique values of service components is a problem that can be resolved on an ad hoc basis (say using the mean or "representative" trip characteristics) or, better still, by using other functions of the characteristics of paths or itineraries that weigh them more formally. In the context of the LIN-LOGIT model, for instance, the inclusive value or natural logarithm of the denominator (U in eq. (3)) represents, under certain conditions, the expected value of the maximum of the utility over all possible alternatives and has often been used to aggregate modes. To aggregate itineraries, the major objection to the use of the model is that it does not make sense to define a "reference" itinerary, as one defines a reference mode in a mode choice model, in order to estimate the (differences of) coefficients of the constants for the remaining modes. To get around this problem, Laferrière (1988), in a model of air demand, used the LIN-IPT-LOGIT specification in the neighborhood of the LIN-LOGIT point of interest, because the IPT-LOGIT makes it possible to identify the coefficients of all alternative-specific constants and even to constrain them equal if desired. It was then possible to estimate a meaningful choice model among itineraries (and to aggregate over them) defined for each origin-destination pair as sets of paths having the same fare, frequency and travel time (controlling as well for the number of competing itineraries and the presence of multiple fare classes in the same itinerary). We are preparing a paper with Laferrière on this problem. More generally, the denominator U can be seen as a nonlinear index of the utility of alternatives.

6 The direct demand format has been extremely successful when the model is not explicitly specialized and pertains for instance to the total travel demand over a certain area. In the case of origin-destination models, the direct format may yet recover if one can be sure that the sign problems that have always plagued it have been solved (Dagenais and Gaudry, 1986).

The QDF format means that the effect of any variable $\mathbf{X_k}$ on $\mathbf{T_m}$, trip demand by mode, **can be decomposed** between its impact on $\mathbf{p_m}$, mode share, and it's impact on **T**, total demand irrespective of mode. If this effect is expressed in terms of elasticities ϵ, it is easy to show that, because QDF is a product, we have:

$$[\epsilon \text{ of Mode}] \equiv [\epsilon \text{ of Total}] + [\epsilon \text{ of Share}] \tag{6}$$

or

$$\epsilon\,(T_m, X_k) \;=\; \epsilon\,(T, X_k) \;+\; \epsilon\,(p_m, X_k). \tag{7}$$

And we are sometimes especially interested in the "inclusive value" or impedance term U for which (7) may be written more explicitly as

$$\epsilon\,(T_m, X_k) = \epsilon(T, U) \cdot \epsilon(U, X_k) + \epsilon(p_m, X_k). \tag{8}$$

2.4. Emphasis and relevance outside of transportation

In this paper, we shall concentrate exclusively on a discussion of share model family types and resulting elasticities of share $\epsilon(\mathbf{p_m}, \mathbf{X_k})$. Although our perspective will be limited to transportation mode choice, one should remember that choice models are applicable outside of transportation. For instance, thick tails and asymmetry of response may occur in studies of the resistance of materials and in toxicology where the underlying probability of occurrence of events may well possess a shape that is problem-dependent and should not be imposed on a *priori* grounds.

3. THREE SHARE MODEL TYPE FAMILIES EMBEDDING THE LINEAR-LOGIT

We want to give a balanced view of theoretical economic or econometric specification issues, of estimation program capabilities and of illustrative example characteristics for the problem at hand.

3.1. Statistical model and main results presented

To estimate share model parameters, we use a multinomial nonlinear generalization of the Berkson-Theil estimator. For a random sample of T observations on M market shares, the statistical model is:

$$\ln\,(p_{it}/p_{rt}) = \ln\,(U_{it}/U_{rt}) + u_{it}\,,\ {\scriptstyle (i=1,\ldots,M-1;t=1,\ldots,T),} \tag{9}$$

where the subscript r denotes an arbitrary (model neutral) reference mode.

Assuming that the vector of residuals $\mathbf{u_t} = (\mathbf{u_{1,t}},\ldots,\mathbf{u_{m-1,t}})$ is (M-1)-variate normal $\mathbf{N}(\mathbf{0}, \underline{\sum})$, the log-likelihood $\boldsymbol{L}$ can be written:

$$L = -\frac{1}{2}T(M-1)\ln(2\pi) - \frac{1}{2}T[\ln(\det \underline{\Sigma})] - \frac{1}{2}\sum_{t=1}^{T}(u_t'\underline{\Sigma}^{-1}u_t) \tag{10}$$

and, after concentration on $\underline{\sum}$, as:

$$L_c = -\frac{1}{2}T(M-1)[1+\ln(2\pi)] - \frac{1}{2}T\left[\ln\left(\det\left\{T^{-1}\sum_{t=1}^{T} u_t' u_t\right\}\right)\right]. \tag{11}$$

The Davidon-Fletcher-Powell algorithm (Fletcher and Powell, 1963) that makes use of analytical derivatives is employed to maximize the concentrated log-likelihood. Asymptotic t-statistics are computed for all estimated parameters with the Berndt, Hall et *al.* (1974) technique, both unconditionally and conditionally upon the estimated values of Box-Cox transformation parameters when those are part of the specification.

In addition to the maximized **log-likelihood value** and **t-statistics,** we present an intuitive coefficient of multiple correlation R^2 and **share elasticities**, defined, using ∂ to denote partial derivatives, as

$$\epsilon(p_m, X_k) \equiv [\partial p_m / \partial X_k][X_k / p_m] \tag{12}$$

and evaluated at the sample means. Because of the large number of specifications studied, we will only show own elasticities. Our fully documented SHARE program (Liem *et al.*, 1993) produces many other statistics of interest, for instance **percentage points** π, which express the number of market share points gained (or lost) due to a one percent change in the variable considered

$$\pi(p_m, X_k) \equiv [\partial p_m / \partial X_k] X_k \ , \tag{13}$$

partial derivatives $[\partial p_m / \partial X_k]$ and **marginal rates of substitution** σ among variables, often used to evaluate values of travel time, defined as

$$\sigma(p_m, X_k X_{k^1}) \equiv [\partial p_m / \partial X_k] / [\partial p_m / \partial X_{k^1}], \quad k \neq k^1 \ . \tag{14}$$

3.2. LINEAR-LOGIT reference model type: specification and emphasis

A. Building blocks

To specify the form of the U_m functions in (4), consider for the i^{th}mode:

$$U_i = \sum_j w_{ij}\{\exp[V_j]\} \ , \ i, j = 1, \ldots, m, \ldots, M, \tag{15}$$

$$V_j = \beta_{jo} + \sum_k \beta_{jk} X_k \ , \tag{16}$$

$$(X_1, \ldots, X_k) = \left([{}^iX_n], [{}^jX_n], [X_s]\right) \ , \ i \neq j, n = 1, \ldots, N; s = 1, \ldots, S; N + S = K, \tag{17}$$

where the w_{ij} and β_{jk} are parameters and the left superscripts associated with the set of network variables partition this set between characteristics that belong to the i^{th} mode and those of other modes.

When the latter set is empty, the former contains variables that, for any observation t, vary across modes, in contrast with the set of "socioeconomic" variables that do not vary across alternatives.

To specify our models, we also need to define the direct and inverse Box-Tukey transformation of a variable or function $\mathbf{Z_f}$ as, respectively

$$Z_f(\lambda_f, \mu_f) = \begin{bmatrix} \left((Z_f + \mu_f)^{\lambda_{f-1}}\right)/\lambda_f \,, & \lambda_f \neq 0, \\ \ln\left(Z_f + \mu_f\right) & , \lambda_f \to 0. \end{bmatrix} \tag{18}$$

and

$$Z_f(\lambda_f . \mu_f)^{-1} = \begin{bmatrix} \left(\lambda_f Z_f + 1\right)^{1/\lambda_f} - \mu_f \,, & \lambda_f \neq 0 \\ \exp\left(Z_f\right) - \mu_f & , \lambda_f \to 0 \end{bmatrix} \tag{19}$$

where μ is the shift parameter introduced by Tukey (1957) in the context of the power model: if it is set equal to zero, (18) and (19) yield the direct and inverse Box-Cox (1964) transformations as commonly designated.

These transformations will be used to transform the X_k variables in (16) and the $\{\exp[V_j]\}$ functions in (15) as follows:

$$f_j[g_j] \tag{20}$$

and

$$g_j\{\exp\left[V_j\right]\}, \tag{21}$$

and

$$h_{jk}(X_k), \tag{22}$$

One could then rewrite (15) – (16) in a very general but obscure nesting fashion

$$U_i = \sum_j w_{ij} \left\{ f_j \left[g_j \left\{ \exp \left[\beta_{jo} + \sum_k \beta_{jk}\left(h_{jk}(X_k)\right) \right] \right\} \right] \right\} \tag{23}$$

that would include (some normalizations excepted) all model types of interest.

B. The LINEAR LOGIT: classical and universal types

To specify the LINEAR **CLASSICAL** LOGIT from (15) – (17) and (22), one requires:

[L-CL]

- (i) : that all cross weights be equal to zero (w_{ij}=0, for all $i \neq j$) and all own weights be equal to one (w_{ij}=1, i = j);
- (ii) : linearity in variables (in (22), h_{jk}=1, for all j and k);
- (iii) : the exclusion of characteristics of other alternatives from the representative utility of the i[th] one $\left(\left[{}^{j}X_n\right] = 0, i \neq j\right)$;
- (iv) : equal "abstract" coefficients for the network characteristics (i.e. $\beta_{jk} = \beta_{mk}$ for all j and m), a constraint that is not necessary but is frequently imposed.

The principal characteristics of this model for our purposes are:

Additive separability of utility

Because (i) and (iii), the model implies

a) equal cross elasticities of demand: this means that setting up a bicycle path between two cities will draw the same percentage of clients from the plane and the bus, in conformity with Luce's Independence from irrelevant alternatives (IIA) axiom;
b) the exclusion of complementarity among alternatives;
c) only differences in the level of characteristics matter: this means that the function is not homogeneous of degree zero. In consequence, doubling all fares and income will change market shares. Of course homogeneity is often imposed in the estimation of demand equations systems and generally rejected by the data.

Symmetry of response curve

Because of (ii), the model further implies that

(d) the effect of a given difference is independent of the level of characteristics: for instance a 10 minute air service improvement has the same impact on choice probabilities for the Montreal-Quebec origin-destination pair as for the Montreal-Mexico pair. Stated otherwise, this also means that the response curve to changes in service characteristic $[{}^{i}X_1]$ is symmetric with respect to its inflexion point, as for the LIN-LOGIT curves in Fig. 1.

Zero probability limits

(e) The choice probabilities go to zero (one) when the representative utility V_i goes to $-\infty\,(+\infty)$: this means that one cannot model thick tails due to specification error, modeler ignorance, compulsive consumption or captivity to modes. This is the thin-equal-tails property.

Underidentification of modal-specific effects

(f) Identical (linear) applications due to (ii) or (22), and the implicit use of linear forms for (20)and (21) impose a form of abstractness that implies the underidentification of alternative-specific β_k coefficients for the $[X_s]$ variables common to all alternatives and prevent the use of both $[{}^{i}X_n]$ and $[{}^{j}X_n]$ in all V_i at the same time;

(g) implicit use of linear forms for (20) and (21) also implies the underidentification of alternative-specific β_{oj} coefficients for the constants.

To specify the LINEAR UNIVERSAL LOGIT, one needs to modify two assumptions:

[L-UL]

(ii') : introduce sufficient nonlinearities in variables so as to:

(iii') : be able to use of both $[{}^{i}X_n]$ and $[{}^{j}X_n]$ at the same time in some or all V_i in order to give a particular usable form to McFadden's (1975) unspecified universal logit.

C. Actual specification and chosen reference model

We formulated the classical and universal LIN-LOGIT specifications shown in Table 1, where the coefficients of the three service variables are clearly generic. The specifications differ for the Fare and Time[7] variables. The classical-type formulation is sometimes referred to as the "difference model" and the universal-type formulation as the "ratio model". Note that the ratio model is not invariant to the chosen reference mode[8] even in a binomial case. Using ratios of variables modifies the classical properties in the following ways:

a) the model is not consistent with the IIA axiom, so the cross elasticities of demand are not equal;
b) the function is homogeneous of degree zero in prices, so doubling all fares has no influence on choice probabilities;
c) the function is still symmetric with respect to its inflexion point but this means that a certain **percentage** improvement service is independent of the relative level of service, so a 10% change in air service has the same effect on the choice probabilities whether the plane is twice as fast as the car or four times as fast as the car.

The maximized log-likelihood was -560.64 for the classical difference model and, as shown in the first column of Table 4, -527.45 for the ratio model. As the log-likelihood for a model containing both the difference and ratio specifications simultaneously was only -526.55, **the ratio specification was retained,** although results for both models appeared reasonable[9]. Results for Model L-CL will not be shown in Table 4.

---

7 Tests of a relative frequency variable made under linearity assumptions did not improve the results.

8 We tested an alternate specification in which the bus was used as reference mode: the results appeared as good.

9 These tests were performed under linearity assumptions. It remains an open question, to be addressed in another paper, whether the ratio model dominates the difference model irrespective of the functional form used for the variables (indeed, one can be shown to be a special case of the other if the logarithm of all variables is used). One could also wonder whether specific coefficients for the fare and service variables could alter our conclusion. The classical model is known to be extremely fragile in this respect, a fact that almost certainly hides an inherent weakness of models that assume additive separability.

Table 1. LINEAR LOGIT specifications, 4 modes, 120 Canadians O-D city pairs, 1976

Model		L-CL	L-UL
Variables in V_m	β_k		
	β_{0m}	Constant for mode m	Constant for mode m
	β_1	Fare by mode m	Fare by mode m/ fare by car
Network	β_2	Time[a] by mode m	Time by car/time by mode m
	β_3	Frequency[b] by mode m	Frequency by mode m
	β_{4m}	Number of nights for car	Number of nights for car
	β_{5m}	O-D within same province[c]	O-D within same province
Socioeconomic	β_{6m}	Language similarity O-D[d]	Language similarity O-D
	β_{7m}	Income at O-D[e]	Income at O-D

[a]Excluding overnight stays but including access, egress and transfer.
[b]For air, "utilitary frequency" ≡ {min. [air frequency, max. (train frequency, bus frequency)]}.
[c]Dummy variable.
[d]{100 - [(%French speaking at origin)-(%French speaking at destination)]}.
[e]Geometric mean.

D. Emphasis

Our emphasis will be three-fold: firstly, we will point out how the model families under examination can in principle have more general properties than the **classical** difference LINEAR LOGIT model; secondly, we will note the capabilities of our SHARE program in this respect; thirdly, we will examine the impact on results of using those model families with our chosen universal ratio reference model, with special emphasis on fare and travel time service level own elasticities of market share. To focus our comments, we will pretend that we are considering the implementation of high speed trains services. To visualize the effects that we are looking for, the reader may consult the four figures.

Although our discussion of formal model properties of the three families of generalizations will be in terms of the classical form (16), our example uses a special case of the universal branch as root. Table 2 contains a summary of the structural properties of all model types considered, with a middle section reserved for the two properties of particular interest. Note that the LINEAR CLASSICAL LOGIT of column (1) has none of the desirable general features but that, as one moves to the right, each family has more desirable features than the previous one, the IPT-LOGIT having all of them except homogeneity of degree zero in prices and income because the general properties are defined with respect to the classical specification in which the variables appear as levels (or differences).

Our SHARE program to estimate the parameters of these families of models is extremely general: for instance, it allows the analyst to estimate direct Box-Cox transformation ((18) with μ=0) on all explanatory variables of all model types except the GENERALIZED DOGIT (for which they cannot be estimated but may be fixed a *priori*), which means that one can specify a GENERALIZED BOX-COX LOGIT model as the nested root of the STANDARD DOGIT, and of the LINEAR and BOX-TUKEY INVERSE POWER TRANSFORMATION-LOGIT. Our example, however, will use the LINEAR LOGIT as the root of all generalizations in order to focus on the **distinguishing**

parameters of each family: λ Box-Cox parameters in the BOX-COX LOGIT **—asymmetry only—**; Θ weights in the DOGIT **—thick tails only—**; and BOX-TUKEY ϕ and μ parameters in the INVERSE POWER TRANSFORMATION LOGIT **—asymmetry and thick tails—**.

3.3. BOX-COX LOGIT family: STANDARD and GENERALIZED model types

A. The Standard BOX-COX LOGIT

To specify the STANDARD BOX-COX LOGIT, modify the [L-CL] specification with:

[BC-L-S] [(ii′) : $h_{jk}(X_k) \equiv (X_k)^{(\lambda_{jk})}$ defined in (18);

This introduces the following features:

(d) **asymmetry of response**, as shown in Fig. 1 (noting X_i, not V_i space), because the effect of a unit change in service will depend on the level of X_{jk} for all values of λ_{jk} other than 1. To see how a unitary improvement in one mode will have this variable effect, consider in Table 3 the first and second partial derivatives of the representative utility of the j^{th} mode, namely of

$$V_j = \sum_k \beta_{jk}\left(\left(X_{jk}^{\lambda_{jk}-1}\right)/\lambda_{jk}\right), \tag{24}$$

and notice that the effect of additional service will be smaller at higher service levels than at lower ones if λ_{jk} is smaller than 1: these diminishing returns mean that given absolute (or relative, in a ratio model) **reductions** in travel time have more impact when travel times are low than high. Conversely, increasing returns exist if λ_{jk} is larger than one. There is no doubt that one expects diminishing returns with respect to service and fare changes: then, in a classical model, a gain of 5 minutes will mean less on a long trip than on a short one; in a ratio model, a gain of 5% will mean less or on a relatively long trip than on a relatively short one;

(f) **identification** of all coefficients of socioeconomic variables that are common to all alternatives as long as the λ_{jk} differ across the modes.

The function actually specified was:

[BC-L-S]

Use a Box-Cox transformation on the strictly positive socioeconomic variables, one on the relative fare and travel time variables and one on departure frequency.

Constrain each to be equal across modes: $\lambda_{mk}=\lambda_{jk}$ for all j.

The results found in the second column of Table 4 show that the additional three parameters cause the log-likelihood value to increase considerably from -527.456 to -499.345. As expected, the estimated transformation (k = -.008) for fare and time indicates decreasing returns: in fact, one cannot reject the logarithmic form for these variables. In consequence, **the presence of asymmetry of response is clear** — a result that also held when each variable was allowed its own transformation, but the gains (not shown here) in log-likelihood were trivial. Note also in Table 4 that the **fare and travel**

Table 2. Properties of general, SHARE program and chosen example* model types

Column #	1	2	3	4	5	6	7	8
Family	LIN-LOGIT		BC-LOGIT		BC-DOGIT		IPT-LOGIT	
Type	C**	U	S	G	S[a]	G[aa]	LIN[a]	BT[a]
Additivity of utility								
a. IIA inconsistency								
General	**no**	**yes**	**no**	**yes**	**yes**	**yes**	**yes[b]**	**yes[b]**
SHARE	no	yes	no	yes	yes	yes	yes[b]	yes[b]
Example	no**	yes	no	. . .	yes	yes	yes[b]	yes[b]
b. Complements possible								
General	**no**	**yes**	**no**	**yes**	**no**	**yes**	**yes[c]**	**yes[c]**
SHARE	no	yes	no	yes	no	yes	yes[c]	yes[c]
Example	no**	yes	no	. . .	no	yes	no	no
c. Homogeneity of degree zero								
General	**no**	**no**	**no**	**no**	**no**	**no**	**no**	**no**
SHARE	no	no	no	no	no	no	no	no
Example	no**	yes	yes	. . .	yes	yes	yes	yes
Asymmetry of response in V_j space								
d. Level of X_k matters								
General	**no**	**no**	yes	yes	yes	yes[aa]	yes	yes
SHARE	no	no	yes	yes	yes	yes[aa]	yes	yes
Example	no**	no	yes	...	no	no	yes	yes
			in X_k	space				
Probability limits of response								
e. Thick tails possible								
General	**no**	**no**	**no**	**no**	**yes**	**yes**	**yes**	**yes**
SHARE	no	no	no	no	yes	yes	yes	yes
Example	no**	no	no	. . .	yes	yes	yes	yes
Identification of coefficients								
f. Common variables								
General	**no**	**no**	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**
SHARE	no	no	yes	yes	yes	yes	yes	yes
Example	no**	no	no	. . .	no	no	no	no
g. Constants								
General	**no**	**no**	**no**	**no**	**no**	**no**	**yes**	**yes**
SHARE	no	no	no	no	no	no	yes	yes
Example	no**	no	no	. . .	no	no	no	no

*As our chosen reference model belongs to the **Universal** LOGIT type, one could formally say that we are testing its BC-U-LOGIT, BC-U--DOGIT and IPT-U-LOGIT generalizations. In that sense, the Generalized BOX-COX LOGIT is not a model type tested with our example, a fact that is denoted by (...) in column 4.

**The classical form example would require a detailed study.

[a] With Box-Cox transformations allowed on all explanatory variables.

[aa] With Box-Cox transformation values on explanatory variables fixed a *priori*.

[b] With general U specification (or as example) but not with C specification.

[c] With general U specification but not in example or with C specification.

time own elasticities of market share are generally higher than in the LIN-LOGIT case: the rail travel time elasticity is 4.6 times higher, which implies that high speed rail would have a greater impact on modal share than if one had used the root model.

Table 3. Non constant returns in a BOX-COX LOGIT model

Returns	$\partial V_j / \partial X_{jk} = \beta_{jk} X_{jk}^{(\lambda_{jk}-1)}$		$\partial V_j^2 / \partial X_{jk}^2 = \beta_{jk}(\lambda_{jk}-1) X_{jk}^{(\lambda_{jk}-2)}$
Decreasing	λ=-1	β_{jk}/X_{jk}^2	$-2\beta_{jk}/X_{jk}^3$
Decreasing	λ=0	β_{jk}/X_{jk}	$-\beta_{jk}/X_{jk}^2$
Constant	λ=1	β_{jk}	0
Increasing	λ=2	$\beta_{jk} X_{jk}$	β_{jk}

B. The Generalized BOX-COX LOGIT

To specify the GENERALIZED BOX-COX LOGIT, modify specification [L-CL-S] with:

[BC-L-G] [(iii') use [^iX_n] and [^jX_n] at the same time in some or all V_i.

This model is easy to specify and study theoretically (Gaudry, 1978) because one simply puts Box-Cox transformations on all variables. The resulting system is very close to those found in microeconomics where all "prices" appear in all demand functions. As long as the values of the transformations on a variable present in all V_j differ, there is no coefficient underidentification problem. The model should dominate the practice of adding in very *ad hoc* fashion the logarithm of X_{mk} in V_j $(m \neq j)$, and only in one place, because trials have shown it to be relevant. And relevant it should be: who would formulate demand (representative utility) functions with only own prices? One expects LOGIT results to be very sensitive to such gross misspecification error based on the assumption of additive separability, an expectation once confirmed by Joël Horowitz in Monte Carlo tests.

In part because of the large number of service level variables involved in typical mode choice problems, the model remains untested. We have performed preliminary tests that showed the importance of constraining the Box-Cox parameters, lest there be too many, or to prevent those used on the variable common to all modal functions from sometimes converging to a common value, thus leading again to underidentification[10]. To circumvent such problems, Wills (1981) has defined a special case of the GENERALIZED BOX-COX LOGIT by grouping into an index the variables pertaining to other modes in each V_j function; unfortunately the model remains untested. Although our SHARE program would make it possible to explore this model type, it is premature to do so. Note however that, if our reference model is log linear (λ_k=0 for all k), it is nested within a GENERALIZED BOX-COX LOGIT structure.

[10] The econometric reasons for this, which are not clear, are currently under investigation.

Figure 1. LIN-LOGIT vs BC-LOGIT

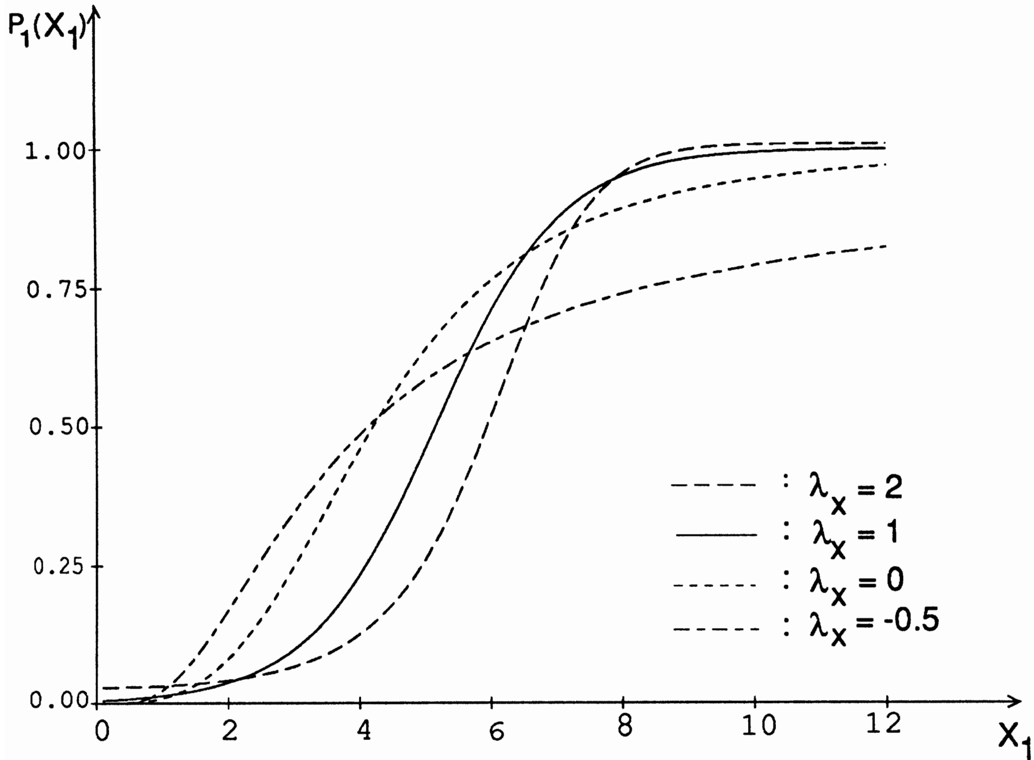

Figure 2. LIN-LOGIT vs STANDARD-DOGIT

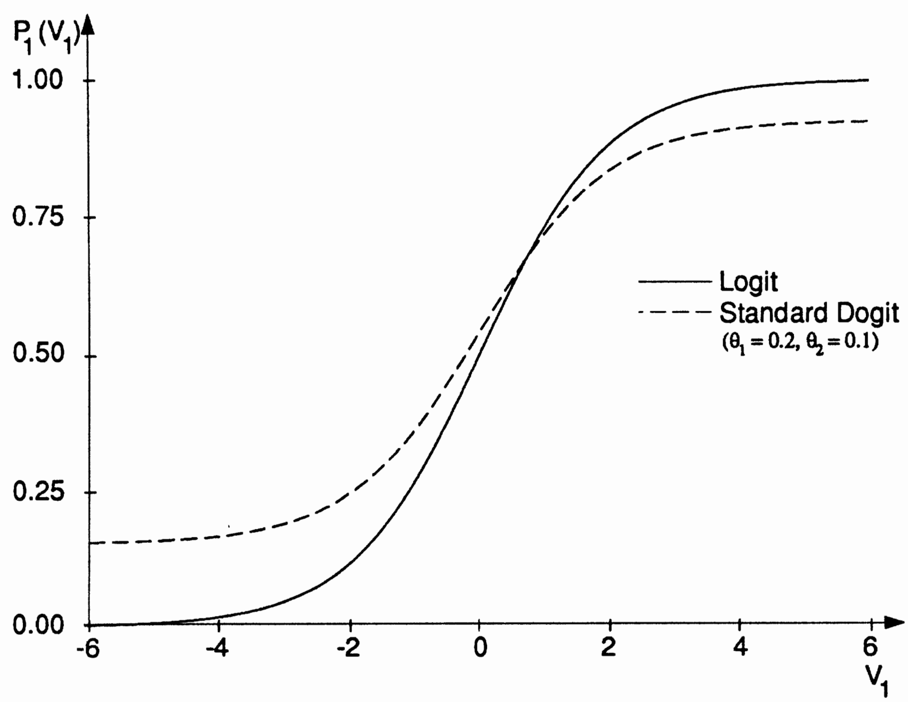

Figure 3. LIN-LOGIT vs LIN-IPT-LOGIT

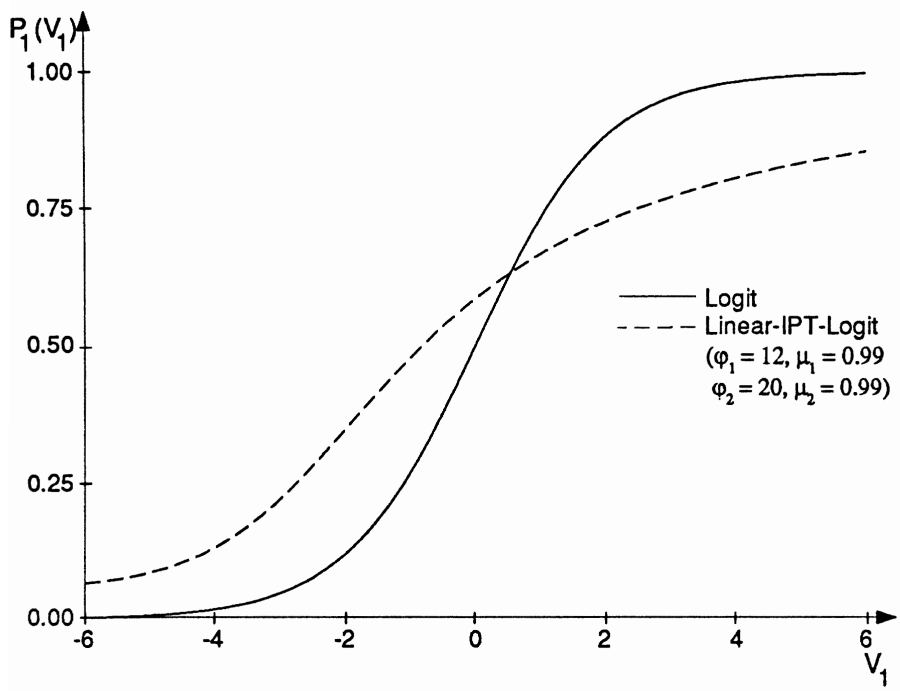

Figure 4. LIN-LOGIT vs BT-IPT-LOGIT

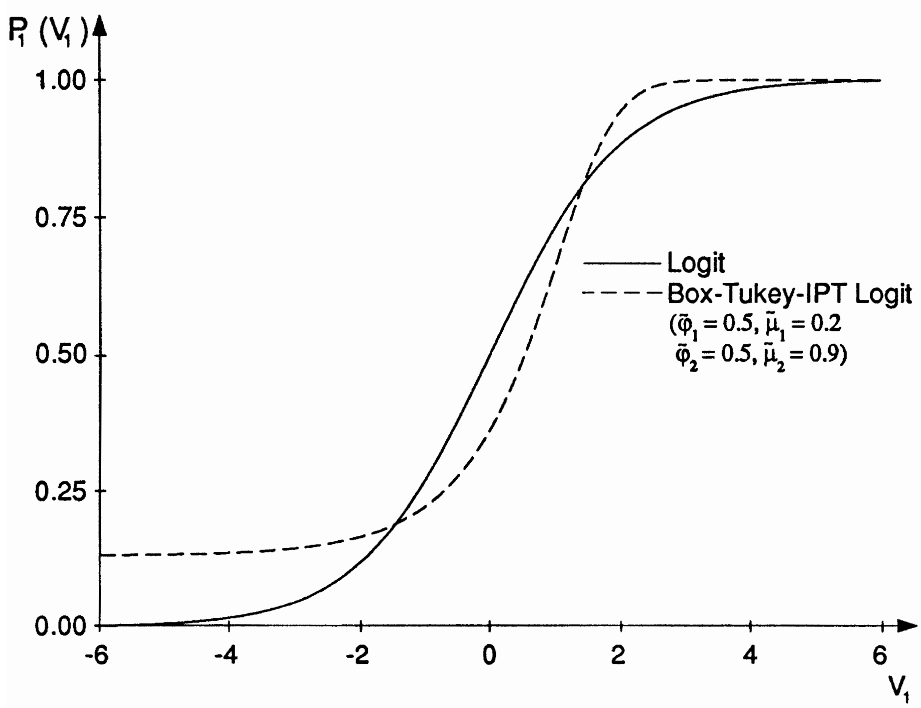

Table 4. Own SHARE Elasticities and Own Statistics for Six Model Types

ELASTICITIES AT MEANS (COND. T-STATISTICS)							
	TYPE =	LOGIT	LOGIT	S-DOGIT	G-DOGIT	LIN-IPTL	BT-IPTL
	VARIANT =	logit	bclogit	sdogit	gdogit	linipt	btipt
	VERSION =	0	1	1	1	1	1
	DEP.VAR. =	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE
ALTERNATIVE 1 : AIR							
Network Variables							
PLANE FARE / CAR FARE	RFANLSC1	-.670	-.716	-2.931	-3.291	-2.436	-.482
		(-6.90)	(-2.55)	(-6.36)	(-4.91)	(-5.57)	(-2.51)
	(1)	(GE)	L 2(GE)	(GE)	(GE)	(GE)	(GE)
TIME BY CAR / TIME BY PLANE	RETNLSC1	1.782	1.410	.894	1.146	1.434	1.736
		(7.93)	(7.05)	(3.22)	(2.80)	(2.25)	(5.65)
	(1)	(GE)	L 2(GE)	(GE)	(GE)	(GE)	(GE)
PLANE UTILITARY FREQUENCY	FREQU1	.247	.218	.307	.485	.257	.061
		(4.95)	(5.39)	(4.90)	(3.95)	(5.01)	(3.67)
	(1)	(GE)	L 3(GE)	(GE)	(GE)	(GE)	(GE)
Socioeconomic Variables							
LANGUAGE SIMILARITY O-D	LANGFR	.018	.227	.401	.829	.682	.117
		(.13)	(1.65)	(1.46)	(2.02)	(1.72)	(.55)
	(1)	(SP)	L 1(SP)	(SP)	(SP)	(SP)	(SP)
O-D WITHIN THE SAME PROVINCE	ODREG	-.144	-.125	-.080	-.335	-.331	-.043
	=====	(-.57)	(-.41)	(-.26)	(-.28)	(-.97)	(.04)
	(1)	(SP)	(SP)	(SP)	(SP)	(SP)	(SP)
INCOME	REVE	1.310	1.432	3.535	4.482	4.805	1.949
		(.89)	(.98)	(1.91)	(1.76)	(2.38)	(1.56)
	(1)	(SP)	L 1(SP)	(SP)	(SP)	(SP)	(SP)
REGRESSION CONSTANT	CONSTANT	-	-	-	-	-	-
		(-2.66)	(-1.53)	(-2.38)	(-2.21)	(-2.87)	(-3.68)
	(1)						
ALTERNATIVE 2 : RAIL							
Network Variables							
TRAIN FARE / CAR FARE	RFANLSC2	-.281	-.949	-1.127	-1.220	-1.059	-.903
		(-6.90)	(-2.55)	(-6.36)	(-4.91)	(-5.57)	(-2.51)
	(2)	(GE)	L 2(GE)	(GE)	(GE)	(GE)	(GE)
TIME BY CAR / TIME BY FARE	RETNLSC2	.415	1.878	.191	.236	.346	1.805
		(7.93)	(7.05)	(3.22)	(2.80)	(2.25)	(5.65)
	(2)	(GE)	L 2(GE)	(GE)	(GE)	(GE)	(GE)
TRAIN FREQUENCY	FREQ2	.183	.169	.208	.317	.197	.203
		(4.95)	(5.39)	(4.90)	(3.95)	(5.01)	(3.67)
	(2)	(GE)	L 3(GE)	(GE)	(GE)	(GE)	(GE)
Socioeconomic Variables							
LANGUAGE SIMILARITY O-D	LANGFR	-.077	-.048	-.141	-.272	-.172	-.127
		(-.16)	(.34)	(-.17)	(-.52)	(.05)	(-.05)
	(2)	(SP)	L 1(SP)	(SP)	(SP)	(SP)	(SP)
O-D WITHIN THE SAME PROVINCE	ODREG	-.338	.287	-.066	.117	-.159	.513
	=====	(-1.41)	(1.05)	(-.27)	(.55)	(-.94)	(2.13)
	(2)	(SP)	(SP)	(SP)	(SP)	(SP)	(SP)
INCOME	REVE	3.369	1.405	2.974	.479	2.998	-.102
		(1.95)	(1.07)	(1.86)	(.18)	(2.24)	(.59)
	(2)	(SP)	L 1(SP)	(SP)	(SP)	(SP)	(SP)
REGRESSION CONSTANT	CONSTANT	-	-	-	-	-	-
		(-3.66)	(-1.51)	(-3.04)	(-.96)	(-3.65)	(-4.53)
	(2)						

Table 4. (Continued) Own SHARE Elasticities and Own Statistics for Six Model Types

ELASTICITIES AT MEANS (COND. T-STATISTICS)		TYPE =	LOGIT	LOGIT	S-DOGIT	G-DOGIT	LIN-IPTL	BT-IPTL
		VARIANT =	logit	bclogit	sdogit	gdogit	linipt	btipt
		VERSION =	0	1	1	1	1	1
		DEP.VAR. =	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE
ALTERNATIVE 3 : BUS								
Network Variables								
BUS FARE / CAR FARE	RFANLSC3		-.257	-.968	-1.034	-.969	-.959	-1.128
			(-6.90)	(-2.55)	(-6.36)	(-4.91)	(-5.57)	(-2.51)
	(3)		(GE)	L 2(GE)	(GE)	(GE)	(GE)	(GE)
TIME BY CAR / TIME BY BUS	RETNLSC3		.397	1.915	.183	.196	.328	2.361
			(7.93)	(7.05)	(3.22)	(2.80)	(2.25)	(5.65)
	(3)		(GE)	L 2(GE)	(GE)	(GE)	(GE)	(GE)
BUS FREQUENCY	FREQ3		.433	.430	.494	.653	.462	.657
			(4.95)	(5.39)	(4.90)	(3.95)	(5.01)	(3.67)
	(3)		(GE)	L 3(GE)	(GE)	(GE)	(GE)	(GE)
Socioeconomic Variables								
LANGUAGE SIMILARITY O-D	LANGFR		.620	.496	.578	-.761	.532	.730
			(2.46)	(3.81)	(2.52)	(-1.15)	(2.77)	(2.90)
	(3)		(SP)	L 1(SP)	(SP)	(SP)	(SP)	(SP)
O-D WITHIN THE SAME PROVINCE	ODREG		.056	.530	.144	1.705	.090	.808
			(-.04)	(2.36)	(.45)	(2.17)	(-.02)	(3.67)
	(3)		(SP)	(SP)	(SP)	(SP)	(SP)	(SP)
INCOME	REVE		.403	-1.370	-.171	-4.532	-.383	-2.189
			(.61)	(-.46)	(.37)	(-1.68)	(.62)	(-.59)
	(3)		(SP)	L 1(SP)	(SP)	(SP)	(SP)	(SP)
REGRESSION CONSTANT	CONSTANT		-	-	-	-	-	-
			(-3.58)	(-.45)	(-2.71)	(.43)	(-3.15)	(-6.09)
	(3)							
ALTERNATIVE 4 : CAR								
Socioeconomic Variables								
NUMBER OF NIGHTS	NUITA		-.404	-.432	-.315	-.235	-.282	-.904
			(-7.43)	(-6.60)	(-4.46)	(-1.25)	(-4.48)	(-4.68)
	(4)		(SP)	(SP)	(SP)	(SP)	(SP)	(SP)

P A R A M E T E R S UNCOND.[T-STATISTICS=0] UNCOND.[T-STATISTICS=1]		TYPE =	LOGIT	LOGIT	S-DOGIT	G-DOGIT	LIN-IPTL	BT-IPTL
		VARIANT =	logit	bclogit	sdogit	gdogit	linipt	btipt
		VERSION =	0	1	1	1	1	1
		DEP.VAR. =	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE
BOX-COX TRANSFORMATIONS								
LAMBDA(X) - GROUP 1	LAM 1			.224				
				[.42]				
				[-1.45]				
LAMBDA(X) - GROUP 2	LAM 2			-.008				
				[-.09]				
				[-10.67]				
LAMBDA(X) - GROUP 3	LAM 3			1.088				
				[2.97]				
				[.24]				

Table 4. (Continued) Own SHARE Elasticities and Own Statistics for Six Model Types

P A R A M E T E R S		TYPE =	LOGIT	LOGIT	S-DOGIT	G-DOGIT	LIN-IPTL	BT-IPTL
UNCOND.[T-STATISTICS=0]		VARIANT =	logit	bclogit	sdogit	gdogit	linipt	btipt
UNCOND.[T-STATISTICS=1]		VERSION =	0	1	1	1	1	1
		DEP.VAR. =	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE
EXTRA PARAMETERS								
THETA 1	AIR				.12E-02 [2.77]			
THETA 2	RAIL				.99E-13 [.00]			
THETA 3	BUS				.22E-13 [.00]			
THETA 4	CAR				.97E-02 [.60]			
PHI 1	AIR						.38E+01 [1.31] [.97] GRP 1	-.13E+01 [-1.76] [-3.10]
PHI 2	RAIL						.38E+01 [1.31] [.97] GRP 1	-.25E+01 [-2.19] [-3.05]
PHI 3	BUS						.38E+01 [1.31] [.97] GRP 1	-.30E+01 [-2.50] [-3.33]
PHI 4	CAR						.38E+01 [1.31] [.97] GRP 1	-.28E+00 [-.61] [-2.78]
MU 1	AIR						.10E+01 [2710.90] [-1.74] GRP 1	.15E+00 [1.48] [-8.34] GRP 1
MU 2	RAIL						.10E+01 [2710.90] [-1.74] GRP 1	.15E+00 [1.48] [-8.34] GRP 1
MU 3	BUS						.10E+01 [2710.90] [-1.74] GRP 1	.15E+00 [1.48] [-8.34] GRP 1
MU 4	CAR						.10E+01 [2710.90] [-1.74] GRP 1	.15E+00 [1.48] [-8.34] GRP 1
THETAS 1	AIR							
	1 , 1 AIR					.00E+00 FIXED		
	1 , 2 RAIL					.89E-01 [1.84]		
	1 , 3 BUS					.13E-01 [.83]		
	1 , 4 CAR					.00E+00 [.00]		
THETAS 2	RAIL							
	2 , 1 AIR					.22E-01 [2.80]		
	2 , 2 RAIL					.00E+00 FIXED		
	2 , 3 BUS					.86E-01 [.74]		
	2 , 4 CAR					.00E+00 [.00]		

Table 4. (Continued) Own SHARE Elasticities and Own Statistics for Six Model Types

PARAMETERS UNCOND.[T-STATISTICS=0] UNCOND.[T-STATISTICS=1]	TYPE = LOGIT	LOGIT	S-DOGIT	G-DOGIT	LIN-IPTL	BT-IPTL
VARIANT =	logit	bclogit	sdogit	gdogit	linipt	btipt
VERSION =	0	1	1	1	1	1
DEP.VAR. =	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE
THETAS 3 BUS						
3 , 1 AIR				.14E-01 [4.64]		
3 , 2 RAIL				.20E-09 [.00]		
3 , 3 BUS				.00E+00 FIXED		
3 , 4 CAR				.43E-12 [.00]		
THETAS 4 CAR						
4 , 1 AIR				.49E-01 [2.51]		
4 , 2 RAIL				.71E-10 [.00]		
4 , 3 BUS				.64E+01 [1.23]		
4 , 4 CAR				.00E+00 FIXED		

GENERAL STATISTICS	TYPE = LOGIT	LOGIT	S-DOGIT	G-DOGIT	LIN-IPTL	BT-IPTL
VARIANT =	logit	bclogit	sdogit	gdogit	linipt	btipt
VERSION =	0	1	1	1	1	1
DEP.VAR. =	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE	FOURMODE
LOG-LIKELIHOOD -FINAL VALUE	-527.456	-499.345	-509.271	-499.224	-508.794	-490.465
RHO-SQUARED OVERALL	.929	.953	.946	.951	.952	.946
1 AIR	.944	.965	.962	.967	.967	.959
2 RAIL	.065	.205	.082	.092	.109	.183
3 BUS	.234	.342	.244	.232	.247	.326
4 CAR	.927	.952	.943	.949	.950	.946
MEAN SHARES OBSERVED / ESTIMATED						
1 AIR	.36/.35	.36/.36	.36/.37	.36/.38	.36/.37	.36/.37
2 RAIL	.04/.03	.04/.03	.04/.03	.04/.03	.04/.03	.04/.03
3 BUS	.03/.02	.03/.02	.03/.02	.03/.02	.03/.02	.03/.02
4 CAR	.57/.60	.57/.58	.57/.58	.57/.58	.57/.58	.57/.58
SAMPLE -NUMBER OF ALTERNATIVES	4	4	4	4	4	4
-NUMBER OF OBSERVATIONS	120	120	120	120	120	120
-FIRST OBSERVATION	1	1	1	1	1	1
-LAST OBSERVATION	120	120	120	120	120	120
-AVAILABLE OBSERVATIONS	120	120	120	120	120	120
1 AIR	120	120	120	120	120	120
2 RAIL	120	120	120	120	120	120
3 BUS	120	120	120	120	120	120
4 CAR	120	120	120	120	120	120
NUMBER OF ESTIMATED PARAMETERS						
-BETAS .VARIABLES	13	13	13	13	13	13
.CONSTANTS	3	3	3	3	3	3
.ASSOCIATED DUMMIES	0	0	0	0	0	0
-BOX-COX TRANSFORMATIONS	0	3	0	0	0	0
-THETA	0	0	4	12	0	0
-PHI	0	0	0	0	1	4
-MU	0	0	0	0	1	1
-VARIANCE MATRIX OF RESIDUALS	6	6	6	6	6	6

Under each t-statistic of the first section of the table, the symbols denote a generic (GE) or specific (SP) regression coefficient, as well as (L), the presence of a BCT and its group identification.

3.4. DOGIT family: STANDARD and GENERALIZED model types

Starting from a LINEAR LOGIT form, the DOGIT family is obtained by relaxing the constraints on the w_{ij} weights in (15). Normalizing the w_{ij} weights in such a way that we have the overidentified form

$$U_i = \left\{ \exp [V_i] + \theta_{ij} \sum_j \exp [V_j] \right\} , \tag{25}$$

from which the GENERALIZED and STANDARD and types are obtained, respectively:

[L-D-G]

(i′) : by setting all $\theta_{ii} = 0$, which yields the identified form

$$U_i = \left\{ \exp [V_i] + \theta_{ij} \sum_{j \neq i} \exp [V_j] \right\} , \tag{26}$$

and

[L-D-S]

(i′) by setting $\theta_{i1} = \theta_{i2} = \ldots = \theta_{iM} = \theta_i$ for m = $1, \ldots, M$ as follows :

$$U_i = \left\{ \exp [V_i] + \theta_i \sum_j \exp [V_j] \right\} , \tag{27}$$

Although both (26) and (27) are obtained from (25), **(27) is not nested in (26)** — for the special case of two modes, they have the same log-likelihood value but estimated values Θ_1 and Θ_{12} differ, as do Θ_2 and Θ_{21}. These models, which respectively add M(M-1) and M parameters, introduce the following features:

a) they have unequal cross elasticities of demand because the relative probabilities of choice between any two alternatives depend on the characteristics of all alternatives;
b) the generalized form (26) admits of complements;
c) tails of variable thickness are possible.

If the Θ_i are positive in the STANDARD DOGIT (Gaudry and Dagenais, 1979), they may be interpreted as measures of compulsive behavior to satisfy basic consumption needs independently from prices, in the spirit of Stone's (1954) linear expenditure system: one would say that a proportion of income $\theta_i/(1 + \Sigma_i \theta_i)$ is preassigned to satisfy this minimum need for good i. But they also lead to the same ambiguities. Indeed, in a number of empirical studies of linear expenditure systems, Solari (1971) found that the best fits were obtained when certain of these parameters were negative. In these cases, he reinterpreted them as indicative of "superior" goods. At first, our SHARE program did not constrain the Θ_i to be positive, but we realized that this makes the model too flexible in that one can forecast negative probabilities. So we now constrain the Θ_i to be nonnegative.

The most popular interpretation if the Θ_i are positive is that of captivity (Ben-Akiva and Swait, 1987) reflecting the hypothesis that one may be drawing from a mixed distribution in which some consumers choose according to a LOGIT and the remaining ones are captive to each mode i in the proportion $\theta_i/(1+\Sigma_i\theta_i)$. This interpretation closely fits the practice of transportation planners who preassign a certain proportion of a market to a mode and assign the rest according to a choice model of some sort. In effect, a number of models that put an upper limit on the tail of the LOGIT and "compress" it are special cases of the STANDARD DOGIT with positive Θ_i (Small, 1981). The general idea is always to have a thick tail (or many), as in Figure 2.

Other derivations of the STANDARD DOGIT exist that need not always assume non negative values of the Θ_i and certainly broaden these two interpretations. Tse (1989) derives it as a special case of his proportional random utility model. In the marketing literature, Colombo and Morrison (1989) have developed a model describing an individual's probability of buying a product conditioned on the product they owned previously, a model which Bordley (1989) has reinterpreted as a conditional LOGIT model which allowed for varying degrees of loyalty across the population. He then showed (Bordley, 1990) both that, for perfectly loyal or unloyal buyers, it reduces to the STANDARD DOGIT and also that, because the Colombo-Morrison model allows for varying degrees of loyalty across the population and does not really presuppose perfectly loyal buyers, the STANDARD DOGIT can also be applicable even without perfectly loyal buyers: it provides the equilibrium choice probabilities when buyer choices from period to period follow the two-period marketing science loyalty model.

Our specification was:

[L-D-S] [allow one distinct and unconstrained Θ_j per mode.

The results shown in Table 4 indicate that the additional four parameters cause the log-likelihood value to increase considerably from -527.456 to -509.271. But the unconstrained values for rail and bus are negative —a Solari-like result for these modes— and significantly so for the bus mode. The results shown incorporate the nonnegativity requirement. In either case, only the air and car modes exhibit some captivity, but **the presence of captivity is clear only for the air mode**. Note also that the effect of the additional four parameters is to increase the fare elasticity and reduce the travel time elasticity of market share for all modes, relatively to the LIN-LOGIT values. Using STANDARD DOGIT values would therefore yield a smaller impact of high speed rail than using LIN-LOGIT values.

In the GENERALIZED DOGIT (Gaudry and Dagenais, 1978; Dagenais et *al.*, 1982), the economic interpretation of the Θ_{ij} over and above the fact that they make the utility of each alternative dependent on all others –as in systems of consumer demand equations– is not as clear. Meyer and Eagle (1982) interpret them as "dissimilarity parameters ... weighing the utilities according to the degree to which they are differentially competitive"; because they appear related to the costs of transition among states, we unsuccessfully tried with Giorgio Leonardi to derive them as limits of dynamic stochastic processes, in the light of the asymptotic theory of extremes (Leonardi, 1984), but obtained $-\Theta_{ij}$ instead of Θ_{ij}. They are at least a way of making all modes ubiquitous in all utility functions. The interested reader will find further discussions of their meaning in Liem *et al.* (1993).

Our specification was:

[L-D-G] [allow M(M-1) unconstrained Θ_{ij} parameters.

However, although the log likelihood reached -480.646, 5 out of 12 were negative: the imposition of nonnegativity requirements lowered this value to that shown in Table 4: -499.224, which is still a very large increase from -527.456. It is interesting to note that the most significant parameters are Θ_{21}, Θ_{31} and Θ_{41}, indicating that **air (m=1) has the most influence on the other modes.** Fare share elasticities increase but time elasticities decrease relatively to LINEAR-LOGIT values.

3.5. INVERSE-POWER TRANSFORMATION-LOGIT family: LINEAR and BOX-TUKEY types

Always using the LINEAR-LOGIT as the nested root, we turn to a family that has both the asymmetry and the thick tails of interest. But now transformations are applied to functions –so asymmetry is a property of the whole V_i, as shown in Figure 3 and Figure 4, rather than a property associated with individual variables–, and thick tails depend on a second parameter of the transformation[11]. As they involve at most 2M additional parameters, the model types considered are very economical in the way in which they introduce these forms of modal specificity. In both cases the ϕ parameters produce asymmetry when they differ across alternatives, and the μ parameters can be interpreted as indicators of captivity or of "modeler ignorance" as we suggested in the original paper on the LINEAR INVERSE POWER TRANSFORMATION (Gaudry, 1981).

To specify the LINEAR and BOX-TUKEY INVERSE POWER TRANSFORMATION LOGIT model types, we need to modify the [L-CL] formulation with, respectively:

$$
\text{[LIN-IPT-L]} \left[\begin{array}{ll} (v) & : \quad f_j, [g_j] \equiv [g_j]^{(\phi_f, \mu_f)^{-1}}, \text{ defined in (19),} \\ (vi) & : g_j\{\exp[V_j]\} \equiv \{\exp[V_j]\} \text{ ; namely} \\ & \quad U_i = \left\{ [\phi_i \exp[V_i] + 1]^{1/\phi_i} - \mu_i \right\} \end{array} \right. \tag{28}
$$

$$
\text{[BT-IPT-L]} \left[\begin{array}{ll} (v) & : \quad f_j, [g_j] \equiv [g_j]^{(\phi_f)^{-1}}, \text{ defined in (19), and } \phi_f\text{=0,} \\ (vi) & : g_j\{\exp[V_j]\} \equiv \{\exp[V_j]\}^{(\phi_f, \mu_f)} \text{ ; namely} \\ & \quad U_i = \left[\exp\{\exp[V_j]\}^{(\phi_i, \mu_i)} \right] \end{array} \right. \tag{29}
$$

where it is clear that (29) is a *conditional* IPT form that belongs to the extreme value class.

[11] In the case of the LINEAR type, Jean-Marie Claudius has pointed out to us that certain combinations of values of ϕ and μ define a critical region in which the probability limits when V_i goes to $-\infty$ are hyperbolic (the reaction curve is not monotonically increasing anymore). Although one can (and we do) avoid these regions during estimation, there is no guarantee that forecasts would preserve the property that the sum of shares sum to one: this situation is analogous to that of the linear probability model which can generate forecasts of probabilities outside of 0 and 1 when values of the explanatory variables, or combinations of them, take us out of the sample range, or to obtaining negative Θ parameters in the DOGIT family.

It appears from a number of our experiments carried out since 1982 with both urban and intercity data that the LINEAR model type shown in Figure 3 has some difficulty generating a thick right-hand side tail without simultaneously flattening out to the left, which shifts the inflexion point too far left in the sense that it linearizes the function in the range of interest shown in Figure 3, which is too costly in terms of fit in that region (even if the overall fit improves). It also appears that using as many μ_i parameters as there are modes sometimes leads to very slow convergence, as some tails are in part "substitutes" for others. To simplify the task and have numerically robust results, we chose here specifications less general than some of those that we succeeded in estimating:

[LIN-IPT-L] [use a common $\phi_i=\phi$ and a common $\mu_i=\mu$.

[BT-IPT-L] [allow one ϕ_i per mode but use only a common tail $\mu_i=\mu$.

Results shown in Table 4 indicate that in both cases the additional parameters caused the log-likelihood to increase significantly from -527.456 to 508.794* and -490, respectively. Detailed examination of the estimated values for the additional parameters shows that **asymmetry and the common thick tail are more significant** in the LINEAR case than in the BOX-TUKEY case. In these cases, the BT-IPT-LOGIT fare and time elasticities were more affected than the LIN-IPT-LOGIT ones (except for air), as compared to the LIN-LOGIT values. Results for the former shown in column 6 would then produce much stronger effects of either fares or travel time for high speed trains than the symmetric LIN-LOGIT, as was the case for the asymmetric BOX-COX LOGIT.

4. CONCLUSION

The LINEAR LOGIT model is the workhorse of transportation mode choice modeling and is applied in many other areas as well. We have shown that its theoretical properties are not generally credible and that one can obtain far superior results with relatively simple generalizations that all include it as a nested special case. The particular branches of the broad families that we have considered here easily do much better that the workhorse at explaining mode choice. They also do this in ways that make sense and have major implications for the analysis of the potential response to significant service level changes, such as those associated with high speed rail. It is quite clear that, instead of segmenting samples in order to model the phenomena of interest, it is much easier to introduce explicit nonlinearities.

* Using one ϕ_i per mode and a common tail $\mu_i=\mu$. increased the likelihood to -493.700 but was a less convenient stepping stone to more advanced formulations under study, which allow for unavailable modes. So it was neglected to allow better comparability with these future results.

References

Ben-Akiva, M. and Swait, J. (1987) Incorporating random constraints in discrete models of choice set generation.Transportation Research B 21, 2, 91-112.

Berndt, E.K., Hall, B.H., Hall, R.E., and Hausman, J.A. (1974) Estimation and inference in nonlinear structural models.Annals of Economic and Social Measurement 653-665.

Bordley, R.F. (1989) Relaxing the loyalty condition in the Colombo/Morrison model. Marketing Science 8, 1.

Bordley, R.F. (1990) The DOGIT model is applicable even without perfectly captive buyers. Operating Sciences Department, General Motors Research Labs, Warren, Michigan 48090.Forthcoming inTransportation Research B.

Box G.E.P. and Cox D.R. (1964) An analysis of transformations.Journal of the Royal Statistical Society Series B, 26, 211-214.

Colombo, R. and Morrison, D. (1989) A brand-switching model with implications for marketing strategies. Marketing Science, 8, 1.

Dagenais, M., Gaudry, M. and Liem, T.C. (1982) Testing the GENERALIZED DOGIT mode choice model (Abstract). Atlantic Economic Journal 60.

Dagenais, M. and Gaudry, M. (1986) Can aggregate direct travel demand models work?. Proceedings of the World Conference on Transport Research, The Centre for Transportation Studies, University of British Columbia, Vancouver, 1669-1676.

Fletcher, R. and Powell, M.J.D. (1963) A rapidly convergent descent method for minimization. Computer Journal 6, 163-168.

Gaudry, M. (1978) A generalized LOGIT mode choice model. Publication 98, Centre de recherche sur les transports, et Cahier #7804, Département de sciences économiques, Université de Montréal, 20 p.

Gaudry, M. (1981) The INVERSE POWER TRANSFORMATION LOGIT AND DOGIT mode choice models. Transportation Research B 15, 2, 97-103.

Gaudry, M. and Dagenais, M. (1978) The GENERALIZED DOGIT mode choice model. Publication 107, Centre de recherche sur les transports, et Cahier #7826, Département de sciences économiques, Université de Montréal, 20 p.

Gaudry, M. and Dagenais M. (1979) The DOGIT model.Transportation Research B 13, 2, 105-112.

Gaudry, M.J.I. and Wills, M.J. (1978) Estimating the functional form of travel demand models. Transportation Research 12, 4, 257-289.

Gaudry, M. and Wills, M. (1979) Testing the DOGIT model with aggregate time series and cross sectional travel data.Transportation Research B 13, 2, 155-166.

Laferrière, R. (1988) Une agrégation nouvelle des itinéraires du transport aérien (ANITA). Publication # 574, Centre de recherche sur les transports, Université de Montréal.

Leonardi, G. (1984) The structure of random utility models in the light of the asymptotic theory of extremes. In M. Florian, ed.Transportation Planning Models, North Holland, 107-133.

Liem, T.C., Gaudry, M. and Laferrière, R. (1993) SHARE: the S-1 to S-5 programs for the Standard and Generalized Box-Cox LOGIT and

DOGIT and for the Linear and Box-Tukey INVERSE POWER TRANSFORMATION-LOGIT models with aggregate data. Publication # 899, Centre de recherche sur les transports, Université de Montréal.

Meyer, J.M., and Eagle, T.C. (1982) Context-induced parameter instability in a disaggregate-stochastic model of store choice. Journal of Marketing Research 19, 62-71.

McFadden, D. (1975) On independence, structure and simultaneity in transportation demand analysis. Working paper 7511, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.

McLynn, J.M., Goldman, A.J., Meyers, P.R. and Watkins, R.H.(1968) Analysis of a market split model. Journal of Research of the National Bureau of Standards 72-B, # 1.

McLynn, J.M. and Woronka, T. (1969) Passenger demand and modal split models. Northeast Corridor Project, U.S. Department of Transportation, Report # 230.

Montmarquette, C. and Mahseredjian, S. (1985) Functional forms and educational production functions. Economics Letters 19, 291-294.

Small, K.A. (1981) Review of G.D. Westley "Planning the location of urban-suburban rail lines: an application of cost-benefit and optimal path analysis". Transportation Research A 15, 1, 111-112.

Solari, L. (1971) Théorie des choix et fonctions de consommation semi-agrégées Librairie Droz, Genève, Paris.

Stone, R. (1954) Linear expenditure systems and demand analysis: an application to the pattern of British demand. Economic Journal 64, 511-527.

Swait, J. and Ben-Akiva, M. (1987) Empirical tests of a constrained choice discrete model: mode choice in So Paulo, Brazil. Transportation Research B 21, 2, 103-115.

Tse, Y.K. (1989) A proportional random utility approach to qualitative response models. Journal of Business and Economic Statistics 7, 1, 61-65.

Tukey, J.W. (1957) On the comparative anatomy of transformations. Annals of Mathematical Statistics 28, 602-632.

Wills, M.J. (1981) A generalized extreme value model admitting substitution and complementarity between alternatives. Strategic Planning Group, Transport Canada.