Identifying all alternative-specific constants in Multinomial Logit models by Inverse Power Transformation Capture

by

Marc Gaudry
Cong-Liem Tran

Agora Jules Dupuit (AJD)
Université de Montréal
Montréal, Canada

This paper, written at the suggestion of Göran Tegnér, benefited from direct support from the National Sciences and Engineering Research Council of Canada (NSERCC) and from the *Fonds québécois de la recherche sur la nature et les technologies* (FQRNT) which both contributed over time to improvements of the TRIO freeware algorithms used for model tests and downloadable from the Agora Jules Dupuit website (www.e-ajd.net). Corresponding author: marc.gaudry@umontreal.ca.

Version 6, September 3, 2011. Revised June 11, 2012

Publication AJD-141
Agora Jules Dupuit, Université de Montréal

Abstract

We recall the multiple roles of constants in Classical regression analysis and the corresponding ones of ***alternative-specific constants*** (ASC) in Multinomial Logit models and, for the latter, demonstrate a roundabout method of identification of the missing M^{th} constant based on the Linear-Inverse Power Transformation-Logit (LIN-IPT-L) model. Tests of this method, called ***IPT Capture***, are carried out on two existing transport mode choice models, aggregate and discrete respectively, with Standard Linear and Standard Box-Cox specifications of their representative utility functions (RUF). We also discuss potential application of ***IPT Capture*** to the identification of all M ***alternative-generic constants*** (AGC) associated to path alternatives of a single mode (*e.g.* air, road or public transit) using either Linear or Box-Tukey IPT-L models, and we present numerical illustrations of this possibility based on the former with M alternative-generic constants.

Key words: alternative-specific constant (ASC), alternative-generic constant (AGC), Multinomial Logit (MNL), Box-Cox transformation, Box-Tukey transformation, Linear-Inverse Power Transformation-Logit (LIN-IPT-L), Box-Tukey-Inverse Power-Transformation-Logit (BT-IPT-L), mode choice, path choice, network assignment, IPT Capture, Canada, Santiago de Chile.

Résumé

Nous rappelons qu'en régression classique la constante joue des rôles multiples qui ont des registres correspondants pour les M constantes ***spécifiques*** en régression logistique multinomiale où nous proposons d'identifier la célèbre constante manquante par une astuce inspirée du modèle Logit-Transformation Puissance Inverse-Linéaire (L-TPI-LIN). Nous appliquons cette méthode, appelée ***Captation TPI***, à deux modèles de choix modal, respectivement agrégé et discret, aux fonctions d'utilité représentative Linéaires Standard et Box-Cox Standard. Nous discutons par ailleurs de l'identification par ***Captation TPI*** des M constantes ***génériques*** associées aux choix entre itinéraires d'un mode particulier donné (*e.g.* la route, l'avion ou le transport en commun), par des modèles L-TPI des familles Box-Tukey et Linéaire dont la dernière sert ici de véhicule à des estimations de constantes génériques présentées pour illustrer numériquement la méthode par Captation TPI.

Mots clés: constantes spécifiques, constantes génériques, Logit Multinomial, transformation Box-Cox, transformation Box-Tukey, Logit-Transformation Puissance Inverse-Linéaire (L-TPI-LIN), Logit-Transformation Puissance Inverse-Box-Tukey (L-TPI-BT), choix modal, choix d'itinéraire, affectation dans un réseau, Captation TPI, Canada, Santiago de Chile.

Table of contents

List of tables

1. Introduction: the missing constant[1]

We recall here the multiple roles of alternative-specific constants (ASC) in Multinomial Logit models (MNL) and demonstrate a roundabout method of identification of the M^{th} constant based on the Linear-Inverse Power Transformation-Logit (LIN-IPT-L) model. The tests are carried out on existing transport mode choice models, both aggregate and discrete, using comparable linear and Box-Cox specifications of the representative utility functions (RUF).

In transport mode choice practice, discussions of ASC have centered on how many of the identifiable *M-1* coefficients should be retained (Tardiff, 1978) but the unidentified M^{th} remaining one has consistently nagged researchers (*e.g.* Bierlaire *et al.*, 1997). We maintain the primary perspective and language of transport mode choice modeling but also marginally touch upon multiple-path choice analysis within a single mode, a problem where the identification of all *M* generic constants (AGC) should be of as much relevance as the identification of all *M* specific constants in multimodal choice analysis.

2. Modeling context

To recall the multiple functions of regression constants and raise the question of the identification of all Logit alternative-specific ones, it is convenient both to start from the familiar grounds of Classical regression and to state our intended itinerary. Writing in vector notation (without observation subscripts), the first of three needed models of interest is the Box-Tidwell (1962) extension of Ordinary Least Squares (OLS) regression:

(1)
$$y = \beta_0 + \sum_k \beta_k X_k^{(\lambda_k)} + u \text{ ,}$$

where the dependent variable is untransformed. The second and third models of interest are the MNL model explaining *p(i)*, the market share or choice probability of the i^{th} of *M* modes:

(2)
$$p\,(i) = \frac{\exp(V_i)}{\sum_{j \in C} \exp(V_j)} \qquad , i, j = 1,..., M,$$

and the corresponding Linear-Inverse Power Transformation-Logit model (LIN-IPT-L):

(3-A)
$$p(i) = \frac{\left[\phi_i \exp(V_i) + 1\right]^{1/\phi_i} - \mu_i}{\sum_{j \in C_n} \left\{ \left[\phi_j \exp(V_j) + 1\right]^{1/\phi_j} - \mu_j \right\}} \qquad , \phi_i \geq 0 \text{ and } \mu_i \leq 1,$$

where the RUF are in both cases specified as:

(4-A)
$$V_i = \beta_{i0} + \sum_k \beta_{ik} X_{ik}^{(\lambda_k)} + e_i \text{ .}$$

Explicating the four above expressions requires precisions concerning fluctuating uses of W_f , the so-called “Box-Cox Transformation” (BCT), originally formulated in direct form by Box & Cox (1964) and later introduced in inverse form by Gaudry (1981) for model (3)[2], among other models. *Stricto sensu*, the *BC* specification is applied to single level-of-service (LOS) variables in (1) and

[1] This paper, written at the suggestion of Göran Tegnér, benefited from direct support from the National Sciences and Engineering Research Council of Canada (NSERCC) and from the *Fonds québécois de la recherche sur la nature et les technologies* (FQRNT) which both contributed over time to improvements of the TRIO freeware algorithms used for model tests and downloadable from the Agora Jules Dupuit website (www.e-ajd.net). Corresponding author: marc.gaudry@umontreal.ca.

[2] The model has been used outside of transportation, for instance by Montmarquette & Mahseredjian (1985) to explain sigmoid achievement scores in mathematics, but the authors used no constant.

(4) but the *BTG* specification is applied to complete functions, the Logit quantities $\exp(V_m)$, in (3). These strict designations of BCT practices are derived from the foursome:

(5)

Direct Power Transformations (DPT)		Inverse Power Transformations (IPT)	
$W_f\left(\lambda_f,\mu_f\right)$	Conditions[3]	$W_f\left(\lambda_f,\mu_f\right)^{-1}$	Conditions[4]
BC		BCG[5]	
$\frac{W_f^{\lambda_f}-1}{\lambda_f}$	$\lambda_f \neq 0,\ W_f > 0$	$\left(\lambda_f W_f + 1\right)^{1/\lambda_f}$	$\lambda_f \neq 0,\ \left(W_f+1\right) > 0$
$\ln W_f$	$\lambda_f \to 0,\ W_f > 0$	$\exp W_f$	$\lambda_f \to 0,\ W_f > 0$
BT[6]		BTG	
$\frac{\left(W_f+\mu_f\right)^{\lambda_f}-1}{\lambda_f}$	$\lambda_f \neq 0,\ \left(W_f+\mu_f\right) > 0$	$\left(\lambda_f W_f + 1\right)^{1/\lambda_f} - \mu_f$	$\lambda_f = 0,\ \left(\lambda_f W_f + 1\right) > 0$
$\ln\left(W_f+\mu_f\right)$	$\lambda_f \to 0,\ \left(W_f+\mu_f\right) > 0$	$\exp\left(W_f\right) - \mu_f$	$\lambda_f \to 0,\ W_f > 0$

We could, without loss of generality, limit our identification tests to Standard Linear RUF cases nested in (4), but such an imposition of constant marginal utility would be as dubious as it is popular: empirical credibility requires that we allow flexible RUF forms and provide identification results by IPT Capture under both Standard Linear and Standard Box-Cox conditions.

But we shall for the moment ignore a comparable road to the identification of all alternative-specific constants based on the Box-Tukey (BT-IPT-L) model option also available in our computer programs used below for the multimodal aggregate[7] and discrete applications:

(3-B)
$$p(i) = \frac{\exp\left\{\dfrac{\left[\exp(V_i)+\tilde{\mu}_i\right]^{\tilde{\phi}_i}-1}{\tilde{\phi}_i}\right\}}{\sum_{j\in C_n}\left\{\exp\left\{\dfrac{\left[\exp(V_j)+\tilde{\mu}_j\right]^{\tilde{\phi}_j}-1}{\tilde{\phi}_j}\right\}\right\}}\qquad,\ \tilde{\mu}_i \geq 0\,.$$

And we shall likewise neglect the Generalized Linear and Box-Cox RUF options available for all Logit and IPT-Logit families implemented in these algorithms namely:

(4-B)
$$V_i = \beta_{i0} + \sum_n \beta_{in}^i X_n^{(\lambda_{in}^i)} + \sum_{j\neq i}\sum_n \beta_{in}^j X_n^{(\lambda_{in}^j)} + \sum_s \beta_{is} X_s^{(\lambda_{is})}\,,$$

where the X_n variables, standing for the LOS variables of a given mode (*e.g.* Cost, Time, Frequency), may appear in any number of modal RUF of the same model[8].

---

[3] Under certain conditions, it is also possible to transform an explanatory variable that contains some zero observations but is not a Boolean (dummy) variable. This will be done in some of the discrete test examples below and explained further in Appendix 1.

[4] This development of IPT families in early 1978 was partly aimed at transforming the dependent variables of probability and share models such as (2) in the same way as one can easily transform the dependent variable in (1), a predicament if probabilities (or shares) must directly sum to one. But this constraint can be indirectly met if the transformations are applied to right hand side quantities, such as Logit quantities in (3). They embody in probability and share models the possibility that, in addition to belonging to say RUF levels of service variables included in (4), particular curvatures might also belong to the modal Logit quantities themselves.

[5] The label was invented by students as a pun on the name of a famous tuberculosis vaccine given to children.

[6] The label « Box-Tukey » was defined by Gaudry & Wills (1978) to distinguish between the common use of the label "Box-Cox" and the same with a shift parameter.

[7] Although all three include the Logit as a special case, neither (3-A) nor (3-B) generalizes the binary Pregibon (1980) link for aggregate data, studied recently by Koenker & Yoon (2009), which constitutes an independent nesting structure.

3. On the variable roles of constants

It is useful for our purposes to first isolate, among the multiple roles played by regression constants, the particular one related to the reproduction of sample means, a role we shall call "the link".

The link between constants and sample means. Starting with linear-in-variable specifications of the three model types, we first recall the differing status of the link between constants and sample values and extend its analysis to the presence of BCT, applied here only to explanatory variables:

Role 1: in Ordinary Least Squares regression (1), the basic link between the constant and sample values is simple: the constant β_o equals the mean of the dependent variable y when there are no explanatory variables X_k and, when there are, can be deduced from the β_k coefficients if the regressors and regressand are set at their sample means $\bar{X}_k$ and $\bar{y}$, given[9] that $\bar{u}=0$:

(6-A) $$\beta_0 = \bar{y} - \sum_k \beta_k \bar{X}_k \, ,$$

an expression that, by OLS properties, carries over to the means of transformed variables:

(6-B) $$\beta_0 = \bar{y} - \sum_k \beta_k \overline{X_k^{(\lambda_k)}} \, .$$

Role 2: in Logit regression (2) with a Standard Linear RUF, this basic link is modified. What is the context of this changed role? Recall first that the well known coefficient identification problem arises for any variable that is the same in all alternatives, such as a socio-economic variable X_s or a constant proper β_0: such common variables cannot then influence the market shares of the alternatives, as demonstrated by the fact that they can be "cancelled out" of the choice equations.

The usual solution adopted is then to "set one at 0", *i.e.* to select an alternative *r* as reference for those kinds of variables, thereby obtaining (by multiplication of all RUF by the relevant terms) estimates of differences of coefficients for the remaining alternatives:

(7) $$\begin{aligned} V_i &= \left(\beta_{i0} - \beta_{r0}\right) + \sum_n \beta_{in}^i X_n + \sum_s \left(\beta_{is} X_s - \beta_{rs} X_s\right) \\ V_i &= \beta_{i0}^{\nabla} + \sum_n \beta_{in}^i X_n + \sum_s \beta_{is}^{\nabla} X_s \end{aligned}$$

In this case, when alternative-specific BCT are applied to the socioeconomic variable X_s, the problem of identification of its M^{th} coefficient β_{M0} is solved and there remains only one coefficient written in terms of differences, that of the constant β_{i0}^{∇}.

But, in the presence of such *M-1* alternative-specific constants, the estimated choice probabilities evaluated at the mean of the X_k are not, as in (5), equal to the means of observed choices d_i, given that d_i is the [0,1] observed value of the dependent variable for the i^{th} alternative. If the model is discrete and we introduce observation subscripts, we have at the point of maximum likelihood[10], if $\hat{p}_{it}$ is the estimated probability of the i^{th} alternative for observation *t*, that the number of observations in the sample for which $d_{it}=1$ is equal to the sum of estimated probabilities:

(8) $$\sum_{t=1}^{T} d_{it} = \sum_{t=1}^{T} \hat{p}_{it} \, ,$$

---

[8] In the Logit model, some conditions apply: the associated BC powers must differ in appropriate manner. In IPT-L models, the BCT can take any value, including 1: no identification issue arises, even with the Generalized Linear RUF. Used in (2), the Generalized Box-Cox RUF (4-B) provides a reasonable incarnation of the Universal Logit Model (McFadden, 1975a) which "can describe any pattern of cross-elasticities" (McFadden, 1984, p. 1415).

[9] See any econometrics textbook, for instance Johnston (1984, p. 172 and Equation 5-48 p. 176).

[10] See for instance Maddala (1983, p. 26) or Ben-Akiva & Lerman (1985, p. 118-119).

which only holds for a specific set of observations — but not for its subsamples, an important issue in mixtures of Stated and Revealed Preference data and the establishment of ASC estimates for a new alternative (Cherchi & Ortuzar, 2006) — and from which it indeed clearly follows that:

(9) $$\left[\frac{1}{T}\sum_{t=1}^{T} d_{it} = \frac{1}{T}\sum_{t=1}^{T} \hat{p}_{it}\right] \neq \hat{p}_{i\,|\,X=\bar{X}},$$

where $\hat{p}_{i\,|\,X=\bar{X}}$ is the probability of the i^{th} alternative evaluated at point $\bar{X}$. But, if the ***observed frequency*** in Logit models equals the predicted frequency, does any linkage remain between the ***sample means of observed frequencies*** and estimated constants?

Yes, but the link now exists in a new garb: differences between constants β_{i0}^{∇} equal the mean of differences between natural logarithms of sample shares of the alternatives if there are no explanatory variables and, if there are, equal differences between such means and the variable systematic parts of the RUF (excluding the constant) evaluated at the means of the regressors. To see this new linkage, consider the natural logarithm of the odds ratio in a binomial model with Standard Linear RUF, called the "logistic transform" or "log odds" (Cox, 1970, p. 18):

(10) $$\ln\left[\frac{p_1}{p_2}\right] = \left[V_1 \quad - \quad V_2\right] + \left[e_1 - e_2\right]$$

which is of familiar[11] format, namely:

(11) $$\left[y\right] = \left[(\beta_{10} - \beta_{20}) + \sum_k \beta_k (X_{1k} - X_{2k})\right] + \left[u_1\right]$$

and which, given that $\bar{u}_1 = 0$, may be rewritten as in (6) with generic regression coefficients:

(12-A) $$\beta_{10}^{\nabla} = (\beta_{10} - \beta_{20}) = \left[\overline{\ln(p_1) - \ln(p_2)}\right] - \left[\sum_k \beta_k \left(\overline{X_{1k} - X_{2k}}\right)\right].$$

Mutatis mutandis, this expression can be adapted to the presence of BCT, to the explicit use of a reference mode in a multinomial context and to specific regression coefficients, to yield:

(12-B) $$\beta_{i0}^{\nabla} = (\beta_{i0} - \beta_{r0}) = \left[\overline{\ln(p_i) - \ln(p_r)}\right] - \left[\sum_k (\beta_{ik}\overline{X_{ik}^{(\lambda_{ik})}} - \beta_{rk}\overline{X_{rk}^{(\lambda_{rk})}})\right],$$

which applies[12] to our all of our forthcoming test cases[13].

Unfortunately, such an expression of "the link" does not much help to understand intuitively why, in choice based sampling, estimation with different sample rates across alternatives merely biases the regression constants (Manski & Lerman, 1977) and why this bias can be redressed *ex post* if population values are known by subtracting $\ln(H_i/W_i)$ from the exogenous sample estimates, where H_i is the fraction of users of the mode in the sample and W_i is the corresponding fraction in the population (Ben-Akiva & Lerman, 1985, p. 238).

Role 3: What now of the link in a LIN-IPT-L model? Note first that, by contrast with a MNL model where the addition of a constant to all alternatives has no effect on them (or on the Log-Likelihood of the sample), the addition of a constant to all alternatives in a LIN-IPT-L model will increase all RUF by an equal amount and all shares by an equal amount if parameters φ_i and μ_i are constrained equal across alternatives. As a rule, this should increase the shares of alternatives that hold less than $1/M^{th}$ of the market and decrease the shares of alternatives that have more, as one might surmise

[11] It has to be remembered that the Logit can be derived as a random or as a deterministic utility model (Anderson *et al.*, 1988): the hypotheses concerning alternative-specific errors required for the logistic transform are weak.

[12] The statement "The alternative-specific constant reflects the mean of $\varepsilon_{it} - \varepsilon_{rt}$, that is the difference in the utility between these alternatives if all else is constant" (Ben-Akiva & Lerman, 1985, p. 75) is therefore meant to be understood as "the alternative-specific constant should be specified so that the sample mean of u_1 be equal to zero".

[13] In fact the Share algorithm used below for tests with aggregate data implements an iterative maximum likelihood version of the Berkson-Theil estimator of (12-B), as recommended by Wills (1982), not the original OLS version found in Berkson (1944) and Theil (1969).

from the fact that all alternatives would have shares equal to 1/M if a constant common to all reached the limit value $(\beta_{i0} \to \infty, \forall i)$, as pointed out by Laferrière (1988, p. 107).

Can the two models be compared? Although (2) looks nested in (3), this is an illusion because it is not possible to multiply all RUF in (3) by a particular term drawn from a reference alternative and obtain differences between M-1 alternative-specific constants and a reference one. But the two models can be brought arbitrarily close to each other by choosing suitable limit values of "outer" parameters φ_i and μ_i (*i.e.* close to 1) and by forcing the "reference" constant in (3) to be extremely small (*i.e.* close to 0). This means that the link established in (12-B) can only hold at that "near-Logit" limit, where the two models are ε-close, *i.e.* are practically equivalent in spite of the fact that the M-1 constants of the LIN-IPT-L cannot be specified in terms of differences with any arbitrarily chosen M^{th} reference constant.

Consequently when, in a LIN-IPT-L model, we merely set all φ_i and μ_i at values close enough to those of the Logit but keep all M alternative-specific constants unrestricted, as we do below, we ***"capture" the M^{th} constant*** on its own, as we do the other M-1, but we remain at some distance from (12-B) to the extent that the captured constant now differs from zero. In the absence of explanatory variables, all such differences between the now identified constants would then be ***more or less close*** to the sample means of differences between the natural logarithms of actual frequencies of the pairs of alternatives; in the presence of explanatory variables, (12-B) would again hold ***more or less*** if the outer parameters φ_i and μ_i were set at near-Logit limit values. But the link (12-B) has apparently broken down, more or less: the structure of this breakdown and the search for a broader link for the LIN-IPT-L model with M constants is a task beyond our current limited perspective.

Other functions of constants. Because of this demonstrated existence of a link between constants and sample means of dependent variables in all of the above models under certain conditions, regression constants cannot, strictly speaking, be said to be estimated like other parameters and should not really count in exact tallies of degrees of freedom used up in models.

In addition, it becomes very dubious to test for the equality of constants in Logit (and even LIN-IPT-L) models if they reflect (at least partially in the presence of explanatory variables) sample means of differences in the natural logarithms of observed frequencies: what is the meaning of the test? And their simultaneous other roles would not redeem such a gesture but make it less wise still, and downright foolish if they are set equal to zero, as we presently argue.

Role 4: the modeler ignorance role of regression constants is intuitive: they reflect the mean of unobserved variables in (1) and the presence of presumably unchanging but unobserved factors, like the "qualitative" comfort of transport modes, in (2) and (3). A special case of hidden influence, which prevents from understanding the alternative-specific constants as "pure car, pure bus, *etc*." comfort dummies occurs with categorical variables coded as 0 or 1 (*e.g*. Race, Sex, Seasons) because the effect of the arbitrarily chosen reference category (*e.g*. "Male") is absorbed by the regression constant (which, shifted, becomes $\beta_0^* = \beta_0 + \beta_{Male}$) and the coefficients of the complementary dummy variable ("Female", in this case) are again implicitly transformed into estimates of the difference with the reference one (here $\beta_{Female}^* = \beta_{Female} - \beta_{Male}$).

Role 5: when variables are transformed by BCT, as in (6-B) and (12-B), the constant guarantees the invariance of the form estimate to changes in the scale of the transformed variables, a point made by Schlesselman (1971) and explicated at length in Appendix 1 below.

Who would indeed want the curvature and properties of a demand function to depend on whether prices were measured in dollars rather than cents or population in thousands rather than millions? Econometric share models with BCT where the constants have been set at zero (*e.g*. Berndt & Khaled, 1979) therefore yield BCT form estimates that depend on chosen units of measurement of the transformed variables and are consequently of severely limited validity.

Because of this ***scalar invariance*** requirement of BCT estimation, all TRIO algorithms (Gaudry et al., 1997) applied below prevent the user of any maximum likelihood estimator, be it of the Classical or Logit regression type, from setting constants at zero, and this independently from the Box-Cox form parameter value enrichments assumed or estimated by the analyst. But such ***scalar invariance*** conditions for BCT estimates should not be confused with ***power invariance*** conditions: BCT estimates are in fact invariant to a simple power transformation of variables even in the absence of a regression constant (Gaudry & Laferrière, 1989), but this is another point altogether.

The scalar invariance requirement of the presence of alternative-specific constants should not be confused either with the *prima facie* similar ***additional dummy variable*** requirement imposed by applying BCT to variables that contain observed zeroes, as also documented in Appendix 1. Models used in transportation are sometimes guilty of this neglect (*e.g.* Caves *et al*., 1980a, 1980b, 1985) and then yield again invalid scale-dependent estimates, at least of their BCT power parameters.

Role 6: the regression constant is also the intercept, and it is well known that forcing the intercept through the origin can change the signs of some regressors. This fact alone is sufficient to justify the ***no-0-intercept*** baseline rule wisely applied by many, but ignored by some (*e.g.* Kessel *et al*., 1986), in Logit mode choice analysis.

None of the multiple roles of constants examined above affects the fundamentally non-stochastic nature of the constant or turn it into a random variable, at least in the models with non stochastic regression parameters assumed here; and we will not discuss “random parameter” regression or the Mixed Logit here.

4. The missing ASC constant recovered in mode choice models

Mindful of the critical role of ASC in MNL regression, and of practitioners’ interest in identifying all *M* values (for instance to forecast the demand for new alternatives), we provide here estimates of such values obtained from two LIN-IPT-Logit mode choice cases based on different kinds of data and compare them with MNL estimates under Standard linear and Box-Cox specifications of their RUF. These models are representative in that their specifications include both network and socio-economic variables, as seen in Table 1, and in that their contents and results, presented in detail in Appendices 2 and 3, have been documented or published elsewhere, at least for the Logit cases.

Table 1. Summary of key features of test models

Model	Nature of demand	Number of modes	Trip purpose	Variables		BCT	T
				X_n	X_s	λ_k	
Canada 1976	Yearly domestic intercity flows	4	Total	3	4	3	120
Santiago 1983-85	Urban CBD corridor trips	9	Work	4	2	3	622

The results for the share model, extracted from Appendix 2, are found in Table 2. The data are found in Tran & Gaudry (2011a), the algorithm used is documented in full in Tran & Gaudry (2008) and Logit results were first published in Gaudry (1993). We note that the reference constant for the auto mode is positive (equal to 2,12 in the linear and to 3,45 in the Box-Cox case), large and significant and that estimation with Box-Cox transformations increases the difference between the largest and smallest constants. As there is a difference of only one degree of freedom between Logit and LIN-IPT-Logit results (the outer parameters being fixed at near-Logit values), the new auto constant yields a statistically significant improvement in model fit, at least for the linear model. Student’s *t*-statistics, admittedly of marginal interest when Likelihood ratio tests are at hand, are computed conditionally upon the values of the BCT.

The results for the discrete choice model, extracted from Appendix 3, are found in Table 3, where the reader may verify the presence of some *additional* dummy variables, called “associated dummies”, duly used for variables that are transformed by BCT but contain some zero observations, namely *Cost/wage rate* and *Walk time*. The data are found in Tran & Gaudry (2011b), the algorithm used is documented in full in Tran & Gaudry (2009) and Logit results were first published in

Gaudry *et al.* (1989). We note that the reference constant for the auto-passenger mode is positive (equal to 0,10 in the linear and to 1,13 in the Box-Cox case) and smaller than the auto-driver constant, their relative difference decreasing with the introduction of Box-Cox transformations on the network variables. But the small gains in Log Likelihood mean that statistical gains in model fit from the identification of the new constant are small in this particular example.

Table 2. Alternative-specific *vs* generic constants and their *t*-statistics; intercity model for Canada

Modes	LOGIT Linear	LIN-IPT-L Linear	LOGIT 3 Box-Cox	LIN-IPT-L 3 Box-Cox
1. Air	-5,50	-3,37	-11,90	-10,3
	(-2,66)	(-1,62)	(-1,53)	(-1,36)
2. Rail	-7,39	-4,47	-10,70	-7,71
	(-3,66)	(-2,20)	(-1,51)	(-1,11)
3. Bus	-6,09	-2,84	-2,59	1,27
	(-3,58)	(-1,57)	(-0,45)	(0,22)
4. Auto	**0,00**	**2,12**	**0,00**	**3,45**
		-5,89		(5,02)
Log Likelihood	**-527,46**	**-510,22**	**-499,35**	**-496,99**
1-4. Equal	-	2,01		3,56
		-5,76		(4,94)
Log Likelihood		**-517,79**		**-498,67**

Table 3. Alternative-specific *vs* generic constants: urban model for Santiago de Chile

Modes	LOGIT Linear	LIN-IPT-L Linear	LOGIT 3 Box-Cox	LIN-IPT-L 3 Box-Cox
1. Auto-driver	0,0382	0,2995	-0,5420	1,7174
2. Auto-passenger	**0,0000**	**0,1024**	**0,0000**	**1,1327**
3. Collective taxi	0,6660	0,8545	-0,2757	0.8626
4. Metro	4,2936	4,4505	3,2591	4,4222
5. Bus	2,0508	2,7005	1,1194	2,4721
6. Auto-driver & metro	-0,1175	-0,0294	-0,7571	0,3593
7. Auto-passenger & metro	0,2280	0,1979	-0,6443	0,4502
8. Collective taxi & metro	0,2066	0,1284	-0,5038	0,5034
9. Bus & metro	1,5561	1,7644	0,7216	1,8578
Log Likelihood	**-855,96**	**-853,90**	**-837,93**	**-837,86**
1-9. Equal		-0,9267		4,6164
Log Likelihood		**-986,92**		**-941,19**

5. The missing AGC constant recovered in path choice analysis

Alternative-specific common factors readmitted. Clearly, IPT Capture makes it possible to obtain relief from the tyranny of Logit *differences* summarized in (7) and to identify contributions of common factors, be they constant or socio-economic in nature[14], affecting all alternatives in a mode choice problem *even if they intervene linearly* in the RUF[15].

In mode choice, such constants must be alternative specific, *i.e.* ASC, if the linkage with sample means of (12-B) is to remain possible; furthermore, setting the ASC at zero is foolish because, in addition to rupturing any linkage with sample means of choice frequencies, it puts at risk other critical functions of constants such as their garbage can/modeler ignorance, scalar invariance preservation (in the presence of BCT) and ***no-0-intercept*** roles.

But could there be sense in attempting to use M constants in the analysis of "assignments" or path choices within a network, *i.e.* in problems where the alternatives have no recognizable[16] natural labels? We argue that, in this case, where the use of ASC is naturally impossible, the insertion of M ***alternative generic constants*** (AGC), *i.e.* of constants constrained equal, is both possible with IPT Capture and in principle preferable to the existing practice of setting all ASC to 0. The reasons include both the fact that a common constant effect could be at work across paths (something of course impossible to test without IPT Capture) and the various roles of constants explored above.

The practice in Logit path choice: four examples. Path assignment based on a random utility formulation is 50 year old. It started with Abraham (1961) who, assuming that each road path had a certain utility level composed of fixed and random components (*op. cit.* p. 66), first derived a binomial Probit choice probability formulation under Gaussian error assumptions — exactly as found later in CRA (1972, Ch. 5) or in Domencich & McFadden (1975, S. 4.4). To simplify calculations in multipath assignment problems (notably with both fully distinct and partially overlapping paths), they simultaneously proposed a probabilistic choice variant derived under Rectangular error distribution assumptions: their simulations showed that, with ratios of generalized (distance and time) costs by itinerary $GC_i = \sum_k \beta_k X_{ik}$, results under Gaussian and Rectangular error assumptions were extremely close.

In further applications, multipath road assignment in France under "Abraham's Law", as it came to be known, has in practice been effected proportionately to $\beta \cdot \ln GC_i$, rather than to say $\beta \cdot GC_i$, a single-variable formula without path constants often referred to as a "Logarithmic Logit" (*e.g.* Leurent, 1999, p. 4)[17].

---

[14] Whether that is a good idea is a different point. For instance, allowing the *Income* variable with a specific coefficient and a common BCT to be included in all (four) alternatives in the model of Column 4 of Table 2 raises the Log Likelihood from -496.99 to -489.63, but allowing the *Sex dummy* variable with a generic coefficient to belong to all (nine) alternatives in the model of Column 4 of Table 3 lowers it from -837.86 to -838.68. In passing, the latter case demonstrates that a socio-economic variable in linear form with a generic coefficient can be included in all M alternatives, just like a constant can.

[15] For a binomial (car *vs* transit) example using data from Winnipeg in Canada, see Laferrière & Gaudry (1992) who, after a trinomial Monte Carlo study of the properties of the SHARE estimator developed by Tran & Gaudry (2008), not only estimate the two ASC constants but, more importantly, demonstrate the ability of the LIN-IPT-L to simultaneously admit variable levels of captivity or modeler ignorance by mode (with distinct μ_i) and alternative-specific asymmetry of response curves (with distinct φ_i), properties that are studied both with Standard Linear specifications of (4-A) and with Generalized Linear specifications of (4-B). The emphasis of the paper is not on IPT Capture as narrowly defined here but on the estimation of the Lin-IPT-L and on the considerable improvements in fit made possible even with generalized linear RUF form enrichments.

[16] Labelling alternatives according to a criterion that looks relevant (*e.g.* shortest, cheapest, *etc.*) will not do: the issue is precisely to determine the weight of such factors.

[17] In practice, the β coefficient is high, with representative values between 8 and 10.

A comparable problem concerned the choice between road *tracés* (McFadden, 1968 or 1975b & 1976) for 65 routing decisions in California from 1958 until 1966, a problem of similar structure because *tracé* alternatives, chosen or not, again have no natural labels. McFadden's tests were made either with a multinomial or with a binomial choice set (the chosen *tracé* and a refused one chosen at random). A linear form of the RUF was used without apparent concern for best fit form (of variables like Cost, Benefit, or their ratio), in spite of Warner's anterior careful finding on the inferiority of linear specifications in his mode choice problem (Warner, 1962), a choice of form methodology also just applied by Thomas (1967) in a car commuter path choice context. The *tracé* model variants had no constants and the paper contained no reference to any other path choice examples, anterior or posterior to 1968.

The fourth "pure" case in transportation, where any labelling based for instance on in-vehicle time or on price would inject a clear *petitio principii* bias, is again a road path choice model. Between an origin and a destination separated by the shortest length L*, Dial (1971) assigned road vehicles on each path of length L "proportionately to $\exp\left[\beta\left(L^{*}-L\right)\right]$" (*op. cit.* p. 91), again without constants.

Can path constants be random variables? In these four examples, not only is there of course no use of ASC, but there might be a further problem in that these constants could have become random variables because, for instance in McFadden's model where there is no more reason to think that all rejected projects had identical alternative-specific constants than there are reasons to think that all accepted ones did. Setting differences of random variables at zero could then affect the assumed properties of the error terms, to say nothing of effects related to the other roles of constants.

For one, forcing all β_{i0} in (12-B) to equal 0 means that, in the absence of explanatory variables, $\overline{\ln(p_i)-\ln(p_r)}=0$ holds, a condition that might only be true in special cases, for instance in a binomial case where the numbers of chosen and rejected alternatives are paired and per force equal (*e.g.* McFadden, 1976, Table 2, Model 6 only). It also means that the other roles of constants cannot be fulfilled: for instance, using BCT would result in form parameter estimates that depended on units of measurement of the variables.

A case of recognizably specific alternatives. These 3 "pure" road path choice cases should not be confused with the air path model formulated by Laferrière (1988, 1989) who was the first to use IPT Capture, as defined here, in practice, but for a very special path choice problem formulation.

He first defined air paths as sets of itineraries with four identical characteristics (Duration, Frequency, Fare and Fare class) and then estimated a binomial choice model between such paths that offered only two itineraries and those that offered more than two. LIN-IPT-L Capture duly yielded the two alternative-specific path constants estimated for all OD pairs in a Canada-wide model of air demand built with a huge sample (12 million individual trips) of all domestic trips made on Air Canada and Canadian Pacific Airlines flights (representing 65% of all air services in Canada) in 1983. His explicit hypothesis was that markets with two principal paths differ from markets with more than two in a way that the 4-characteristic RUF might have failed to Capture: his constants, related to the number of itineraries, embody a sort of latent "itinerary scope" or path richness dimension. Such recognizable paths[18] obtain natural labels that resemble those of modes: it then makes sense to use ASC or even to test for AGC (the last of which Laferrière did[19] not try).

Abstract alternatives: from ASC to AGC. How would one proceed to exploit LIN-IPT-L Capture in a "pure" path choice case? In any intra-modal competition application to air or road choices (or

[18] The same holds for Lapparent's (2004, 2010) model of choice between two specific flights linking Paris and London.
[19] In his 5 principal models (Laferrière, 1988, pp. 80-81), the constant for paths with more than two itineraries is always smaller and less significant than the complementary ASC.

to transit networks), one could first select for each OD pair a number of paths sufficient to accommodate most observed choices and then use as many ASC constrained equal, *i.e.* AGC, a stand which altogether avoids deciding on labels. It would however not be possible later to relax this equality constraint and estimate all M alternative-specific constants because the alternatives have no natural labels other than having been chosen[20]; but at least some of the multiple roles of constants could still be fulfilled by AGC. Selecting the right number of paths of course depends on the problem at hand: to account for 98% of current air trip path choices by Germans, 3 paths will do the job within Germany but some 16 are needed between Germany and the United States of America. But would it work?

Estimating AGC: two numerical examples. We now want to show that the estimation of AGC is *numerically feasible* by carrying out tests on the mode choice problems above[21]. The AGC results are found in the lower parts of the tables. As compared with ASC cases, one naturally gets a reduction in the Log Likelihood values.

In Table 2, imposing an AGC instead of 4 ASC lowers the Log likelihood by 7 points to -517,79 in the Linear case and by almost 2 points to -498,67 in the Box-Cox case. In Table 3, the similar imposition of an AGC to the 9 ASC results in Log Likelihood values reduced by more than 100 points in both Linear and Box-Cox cases and in values of the AGC that have little to do with the original ASC. But our point here is numerical.

There clearly remains some way to go to improve on this AGC solution and define the detailed requirements for its proper use, but LIN-IPT-L Capture offers a fruitful way to address the issue, which at least allows for common effects in Logit models to the extent that the LIN-IPT-L option can be construed as the nearest acceptable enrichment. A similar argument could be made with the BT-IPT-L family (3-B), but obtaining closely resembling results from the Canada and Santiago examples with this model is not useful at this point. An extensive comparison of the full LIN and BT families of IPT-L approaches (with the estimation of the extra "outer" parameters) is carried out in the referenced documents[22] on the databases.

6. Conclusion

Because IPT Capture makes the identification of the coefficients of variables common to all alternatives of a Multinomial Logit model possible, we have estimated all such M alternative-specific constants (ASC) in real aggregate and discrete mode choice examples and demonstrated the possible use of the same method with alternative generic constants (AGC) for Logit path choice problems.

There are many circumstances where the knowledge of the level of impact of constants on the utility of alternatives might not just improve model fit as occurs here, a fundamental point if one is hoping to measure the utility of alternatives, but also usefully free us from the tyranny of differences built into the Logit specification, a desirable result because there is just so much one can do with differences, for both constants and any socio-economic variable common to all alternatives.

---

[20] That is just a practical device: one could also add to the choice set a certain number of alternatives not chosen and maintain for those the assumption that all ASC are equal and non-stochastic in nature. The assumption of a common constant does not introduce a *petitio principii*.

[21] For the BCT tests, the transformed variables are defined in the same units across all RUF. For the maximization of the Likelihood function, the initial values of the constants are arbitrarily set at 1/M in both ASC and AGC cases. Note that any initial value, including zero, would imply, in the absence of other regressors, market shares all equal to 1/M.

[22] Using Standard Linear and Box-Cox RUF (4-A), Logit, Standard and Generalized Dogit, as well as Linear and Box-Tukey Inverse Power Transformation-Logit estimates are provided for the aggregate Canada 1976 model (Tran & Gaudry, 2011a, pp.90-95) and for the discrete Santiago de Chile 1983-1985 model (Tran & Gaudry, 2011b, pp. 125-137). Examples based on Generalized Linear and Box-Cox RUF (4-B) are not provided but, as demonstrated above for both models with estimation of AGC constants, the SHARE S-1 to S-5 and the PROBABILITY P-2 to P-6 programs also allow for those specifications.

7. Appendix 1. Constants and scalar invariance of Box-Cox transformations

To understand the role of dummy variables associated to variables that contain zero values but are subjected to a BCT, start with a simple linear regression model with dependent variable ***y*** regressed on a constant ***C***, on a variable ***Q*** that is not strictly positive, *e.g.* snowfall, and on its associated binary variable ***D*** (where $D_t = 1$ if $Q_t \neq 0$ and $D_t = 0$ otherwise):

$$y_t = \beta_0 C_t + \beta Q_t + \beta_d D_t + u_t . \qquad (1)$$

In (1), the variable $\boldsymbol{D_t}$ can be understood as a variable of interaction capturing the qualitative effect of a *state-of-the world* (*e.g.* there is snow) on $\boldsymbol{y_t}$ beyond the quantitative effect of $\boldsymbol{Q_t}$ (*e.g.* how much snow), or as a test of whether the "first" unit of snowfall has the same impact as the next units. Replacing the state variable by its complement ($D_t = 0$ if $Q_t \neq 0$ and $D_t = 1$ otherwise) simply reverses the sign of β_d and adding a state variable consumes a degree of freedom, modifies the meaning of the model and might improve the adjustment, or not.

But, if the model is instead:

$$y_t = \beta_0 C_t + \beta Q_t^{(\lambda)} + \beta_d D_t + u_t , \qquad (2)$$

where the BCT is applied only to the positive values of $\boldsymbol{Q_t}$ (to avoid numerical problems with 0 observations when $\lambda \to 0$), the associated binary variable will play an additional role: it will preserve the invariance of the $\boldsymbol{\lambda}$ estimate to changes in the units (scale) of measurement of $\boldsymbol{Q_t}$. The analytic proof of this affirmation is beyond our reach but a numerical proof is accessible if we have OLS regression software: (i) set $\boldsymbol{\lambda}$ in (2) equal to an arbitrary value, such as 2, apply the OLS estimator while conserving $\boldsymbol{Q_t}$ in original units and obtain the vector of coefficients $[\hat{\beta}_0, \hat{\beta}, \hat{\beta}_d]$; (ii) apply scale factor ***s*** to $\boldsymbol{Q_t}$ and re-estimate with $\tilde{Q}_t = sQ_t$ to obtain $[\tilde{\beta}_0, \tilde{\beta}, \tilde{\beta}_d]$; (iii) observe that $(\hat{\beta}_0 = \tilde{\beta}_0)$, $(\hat{\beta} = s^{\lambda}\tilde{\beta})$, $(\hat{\beta}_d = \tilde{\beta}_d + s^{(\lambda)}\tilde{\beta})$ and that the error sum of squares is invariant. The model (2) is invariant because the coefficient of the constant $\boldsymbol{C_t}$ does not change, the coefficient of $\boldsymbol{Q_t}$ adjusts in inverse proportion to the factor $\boldsymbol{s^{\lambda}}$ and the scale change compensation (needed by the absence of transformation applied to the zero observations of rescaled $\boldsymbol{Q_t}$) is effected by the coefficient of the associated Boolean dummy variable which has acquired a second role: it still represents a *state-of-the world*, as in (1), but also functions as a scale compensation buffer that makes $\boldsymbol{\lambda}$ invariant to the rescaling of $\boldsymbol{Q_t}$.

We must not confuse this numerical proof of the invariance of $\boldsymbol{\lambda}$ in (2) with the invariance of the $\boldsymbol{\lambda}$ of a BCT applied to a strictly positive ***Q*** variable, demonstrated by Schlesselman (1971) to hold only if the regression has a constant ***C***. In that case, the coefficient of the required constant ***C*** provides the necessary adjustment, after a change of scale ***s*** of the transformed variable ***Q***, as demonstrated in the following simple before and after sequence:

$$\beta_0 C + \beta (Q^{\lambda} - 1)/\lambda) \qquad (3\text{-}1)$$

$$\beta_0 C + \beta s^{\lambda} \left(Q/s\right)^{\lambda} / \lambda - \beta/\lambda \qquad (3\text{-}2)$$

$$\beta_0 C + \tilde{\beta} \tilde{Q}^{\lambda} / \lambda - \beta/\lambda + \tilde{\beta}/\lambda - \tilde{\beta}/\lambda \qquad (3\text{-}3)$$

$$\beta_0 C + \tilde{\beta} (\tilde{Q}^{\lambda} - 1)/\lambda - \beta/\lambda + \tilde{\beta}/\lambda \qquad (3\text{-}4)$$

$$\tilde{\beta}_0 C + \tilde{\beta} (\tilde{Q}^{\lambda} - 1)/\lambda \qquad (3\text{-}5)$$

where, after the rescaling of ***Q***, its new coefficient $\tilde{\beta} = s^{\lambda}\beta$, the same as in (2), and the new coefficient of the required constant $\tilde{\beta}_0 = (\beta_0 - \beta/\lambda + \tilde{\beta}/\lambda)$, have been *analytically* derived.

8. Appendix 2. Results for the aggregate intercity model (Canada 1976)

```
================================================================================================
PART I. Presented statistics (*)
  Model type                                               LOGIT     LIN-IPTL    LOGIT     LIN-IPTL
  Name and version number of the variant                   LIN:1     LINIPTL:8   BC3:2     LINIPTL:11
================================================================================================
                                                    ALTERNATIVE  1 :  AIR
  P     =   PRICES
  ----------------
    FARE(AIR) / FARE(CAR)                       RFANLSC1
      BETA coefficient                                   -0.256D+00 -0.104D+01 -0.998D+00 -0.134D+01
      Elast. of S(AIR      ) - at mean(X)                  -0.6695    -2.9482    -0.7159    -0.9694
      T-statistic conditional on LAMBDA(X)               (  -6.90)  (  -6.69)  (  -2.55)  (  -3.09)
                                                            (GE)       (GE)    L 2  (GE)  L 2  (GE)
  N     =   NETWORKS
  ------------------
    TIME(CAR) / TIME(AIR)                       RETNLSC1
      BETA coefficient                                    0.568D+00  0.262D+00  0.197D+01  0.205D+01
      Elast. of S(AIR      ) - at mean(X)                   1.7821     0.8943     1.4097     1.4639
      T-statistic conditional on LAMBDA(X)               (   7.93)  (   3.28)  (   7.05)  (   7.18)
                                                            (GE)       (GE)    L 2  (GE)  L 2  (GE)
    UTILITARY FREQUENCY (AIR)                   FREQU1
      BETA coefficient                                    0.142D-01  0.161D-01  0.104D-01  0.900D-02
      Elast. of S(AIR      ) - at mean(X)                   0.2468     0.3034     0.2176     0.2275
      T-statistic conditional on LAMBDA(X)               (   4.95)  (   5.35)  (   5.39)  (   5.54)
                                                            (GE)       (GE)    L 3  (GE)  L 3  (GE)
  S     =   SOCIOECONOMIC
  ----------------------
    LANGUAGE SIMILARITY O-D                     LANGFR
      BETA coefficient                                    0.594D-03  0.586D-02  0.131D+00  0.104D+00
      Elast. of S(AIR      ) - at mean(X)                   0.0175     0.3533     0.2272     0.1953
      T-statistic conditional on LAMBDA(X)               (   0.13)  (   1.35)  (   1.65)  (   1.39)
                                                            (SP)       (SP)    L 1  (SP)  L 1  (SP)
    O-D WITHIN THE SAME PROVINCE                ODREG
                                                =====
      BETA coefficient                                   -0.209D+00 -0.114D+00 -0.133D+00 -0.275D-01
      Elast. of S(AIR      ) - at mean(X)                  -0.1435    -0.0402    -0.1247    -0.0239
      T-statistic conditional on LAMBDA(X)               (  -0.57)  (  -0.32)  (  -0.41)  (  -0.08)
                                                            (SP)       (SP)       (SP)       (SP)
    INCOME                                      REVE
      BETA coefficient                                    0.461D-03  0.113D-02  0.319D+00  0.386D+00
      Elast. of S(AIR      ) - at mean(X)                   1.3097     3.8664     1.4318     2.0082
      T-statistic conditional on LAMBDA(X)               (   0.89)  (   2.20)  (   0.98)  (   1.29)
                                                            (SP)       (SP)    L 1  (SP)  L 1  (SP)
  ET    =   ET CETERA
  ------------------
    REGRESSION CONSTANT                         CONSTANT
      BETA coefficient                                   -0.550D+01 -0.337D+01 -0.119D+02 -0.103D+02
      T-statistic conditional on LAMBDA(X)               (  -2.66)  (  -1.62)  (  -1.53)  (  -1.36)
                                                            (SP)       (SP)       (SP)       (SP)
------------------------------------------------------------------------------------------------
                                                   ALTERNATIVE  2 :  RAIL
  P     =   PRICES
  ----------------
    FARE(RAIL) / FARE(CAR)                      RFANLSC2
      BETA coefficient                                   -0.256D+00 -0.104D+01 -0.998D+00 -0.134D+01
      Elast. of S(RAIL     ) - at mean(X)                  -0.2812    -1.1004    -0.9493    -1.2674
      T-statistic conditional on LAMBDA(X)               (  -6.90)  (  -6.69)  (  -2.55)  (  -3.09)
                                                            (GE)       (GE)    L 2  (GE)  L 2  (GE)
  N     =   NETWORKS
  ------------------
    TIME(CAR) / TIME(RAIL)                      RETNLSC2
      BETA coefficient                                    0.568D+00  0.262D+00  0.197D+01  0.205D+01
      Elast. of S(RAIL     ) - at mean(X)                   0.4154     0.1853     1.8782     1.9614
      T-statistic conditional on LAMBDA(X)               (   7.93)  (   3.28)  (   7.05)  (   7.18)
                                                            (GE)       (GE)    L 2  (GE)  L 2  (GE)
    FREQUENCY (RAIL)                            FREQ2
      BETA coefficient                                    0.142D-01  0.161D-01  0.104D-01  0.900D-02
      Elast. of S(RAIL     ) - at mean(X)                   0.1828     0.1997     0.1690     0.1645
      T-statistic conditional on LAMBDA(X)               (   4.95)  (   5.35)  (   5.39)  (   5.54)
                                                            (GE)       (GE)    L 3  (GE)  L 3  (GE)
  S     =   SOCIOECONOMIC
  ----------------------
    LANGUAGE SIMILARITY O-D                     LANGFR
      BETA coefficient                                   -0.689D-03 -0.133D-02  0.256D-01  0.235D-01
      Elast. of S(RAIL     ) - at mean(X)                  -0.0765    -0.1663    -0.0481    -0.0282
      T-statistic conditional on LAMBDA(X)               (  -0.16)  (  -0.32)  (   0.34)  (   0.33)
                                                            (SP)       (SP)    L 1  (SP)  L 1  (SP)
```

```
    O-D WITHIN THE SAME PROVINCE          ODREG
                                          =====
      BETA coefficient                             -0.404D+00 -0.144D+00  0.278D+00  0.368D+00
      Elast. of S(RAIL    ) - at mean(X)             -0.3383    -0.0521     0.2865     0.1468
      T-statistic conditional on LAMBDA(X)          (  -1.41)  (  -0.50)  (   1.05)  (   1.36)
                                                       (SP)       (SP)       (SP)       (SP)
    INCOME                                REVE
      BETA coefficient                              0.945D-03  0.102D-02  0.315D+00  0.316D+00
      Elast. of S(RAIL    ) - at mean(X)              3.3689     3.3171     1.4054     1.4879
      T-statistic conditional on LAMBDA(X)          (   1.95)  (   2.11)  (   1.07)  (   1.17)
                                                       (SP)       (SP)   L 1  (SP)  L 1  (SP)
  ET    =   ET CETERA
  -------------------
    REGRESSION CONSTANT                   CONSTANT
      BETA coefficient                             -0.739D+01 -0.447D+01 -0.107D+02 -0.771D+01
      T-statistic conditional on LAMBDA(X)          (  -3.66)  (  -2.20)  (  -1.51)  (  -1.11)
                                                       (SP)       (SP)       (SP)       (SP)
-----------------------------------------------------------------------------------------------
                                       ALTERNATIVE  3 :  BUS
  P     =   PRICES
  ----------------
    FARE(BUS) / FARE(CAR)                 RFANLSC3
      BETA coefficient                             -0.256D+00 -0.104D+01 -0.998D+00 -0.134D+01
      Elast. of S(BUS     ) - at mean(X)             -0.2569    -0.9859    -0.9685    -1.2908
      T-statistic conditional on LAMBDA(X)          (  -6.90)  (  -6.69)  (  -2.55)  (  -3.09)
                                                       (GE)       (GE)   L 2  (GE)  L 2  (GE)
  N     =   NETWORKS
  ------------------
    TIME(CAR) / TIME(BUS)                 RETNLSC3
      BETA coefficient                              0.568D+00  0.262D+00  0.197D+01  0.205D+01
      Elast. of S(BUS     ) - at mean(X)              0.3972     0.1737     1.9153     1.9938
      T-statistic conditional on LAMBDA(X)          (   7.93)  (   3.28)  (   7.05)  (   7.18)
                                                       (GE)       (GE)   L 2  (GE)  L 2  (GE)
    FREQUENCY (BUS)                       FREQ3
      BETA coefficient                              0.142D-01  0.161D-01  0.104D-01  0.900D-02
      Elast. of S(BUS     ) - at mean(X)              0.4333     0.4642     0.4300     0.4328
      T-statistic conditional on LAMBDA(X)          (   4.95)  (   5.35)  (   5.39)  (   5.54)
                                                       (GE)       (GE)   L 3  (GE)  L 3  (GE)
  S     =   SOCIOECONOMIC
  -----------------------
    LANGUAGE SIMILARITY O-D               LANGFR
      BETA coefficient                              0.881D-02  0.908D-02  0.233D+00  0.231D+00
      Elast. of S(BUS     ) - at mean(X)              0.6201     0.5617     0.4957     0.5361
      T-statistic conditional on LAMBDA(X)          (   2.46)  (   2.43)  (   3.81)  (   3.85)
                                                       (SP)       (SP)   L 1  (SP)  L 1  (SP)
    O-D WITHIN THE SAME PROVINCE          ODREG
                                          =====
      BETA coefficient                             -0.926D-02  0.965D-01  0.522D+00  0.572D+00
      Elast. of S(BUS     ) - at mean(X)              0.0560     0.0484     0.5303     0.2335
      T-statistic conditional on LAMBDA(X)          (  -0.04)  (   0.36)  (   2.36)  (   2.54)
                                                       (SP)       (SP)       (SP)       (SP)
    INCOME                                REVE
      BETA coefficient                              0.248D-03  0.204D-03 -0.110D+00 -0.118D+00
      Elast. of S(BUS     ) - at mean(X)              0.4030    -0.0684    -1.3701    -1.6025
      T-statistic conditional on LAMBDA(X)          (   0.61)  (   0.48)  (  -0.46)  (  -0.53)
                                                       (SP)       (SP)   L 1  (SP)  L 1  (SP)
  ET    =   ET CETERA
  -------------------
    REGRESSION CONSTANT                   CONSTANT
      BETA coefficient                             -0.609D+01 -0.284D+01 -0.259D+01  0.127D+01
      T-statistic conditional on LAMBDA(X)          (  -3.58)  (  -1.57)  (  -0.45)  (   0.22)
                                                       (SP)       (SP)       (SP)       (SP)
-----------------------------------------------------------------------------------------------
                                       ALTERNATIVE  4 :  CAR
  N     =   NETWORKS
  ------------------
    NUMBER OF NIGHTS                      NUITA
                                          -----
      BETA coefficient                             -0.463D+00 -0.405D+00 -0.412D+00 -0.365D+00
      Elast. of S(CAR     ) - at mean(X)             -0.4042    -0.1259    -0.4321    -0.1659
      T-statistic conditional on LAMBDA(X)          (  -7.43)  (  -6.43)  (  -6.60)  (  -5.57)
                                                       (SP)       (SP)       (SP)       (SP)
  ET    =   ET CETERA
  -------------------
    REGRESSION CONSTANT                   CONSTANT
      BETA coefficient                                         0.212D+01             0.345D+01
      T-statistic conditional on LAMBDA(X)                     (   5.89)             (   5.02)
                                                                  (SP)                  (SP)
```

PART II. Parameters

T-statistic unconditional (=0) [=1]

Model type	LOGIT	LIN-IPTL	LOGIT	LIN-IPTL
Name of the variant	LIN	LINIPTL	BC3	LINIPTL
Version number of the variant	1	8	2	11
BOX-COX TRANSFORMATIONS				
LAMBDA(X) 1			0.2245	0.2355
			(0.42)	(0.44)
			[-1.45]	[-1.42]
LAMBDA(X) 2			-0.0080	-0.0416
			(-0.09)	(-0.42)
			[-10.67]	[-10.64]
LAMBDA(X) 3			1.0884	1.1353
			(2.97)	(3.16)
			[0.24]	[0.38]
LIN-IPTL				
PHI 1-4		0.9900		0.9900
		FIXED		FIXED
MU 1-4		0.9900		0.9900
		FIXED		FIXED

PART III. General statistics

LOG-LIKELIHOOD	-527.46	-510.22	-499.35	-496.99
DEGREES OF FREEDOM	13	13	16	16
RHO-SQUARED - OVERALL	0.9289	0.9453	0.9530	0.9531
.Alternative 1 : AIR	0.9441	0.9611	0.9653	0.9652
.Alternative 2 : RAIL	0.0647	0.0822	0.2049	0.2009
.Alternative 3 : BUS	0.2344	0.2420	0.3421	0.3462
.Alternative 4 : CAR	0.9271	0.9428	0.9523	0.9526
MEAN SHARES OBSERVED / ESTIMATED				
.Alternative 1 : AIR	0.36/0.35	0.36/0.37	0.36/0.36	0.36/0.37
.Alternative 2 : RAIL	0.04/0.03	0.04/0.03	0.04/0.03	0.04/0.03
.Alternative 3 : BUS	0.03/0.02	0.03/0.02	0.03/0.02	0.03/0.02
.Alternative 4 : CAR	0.57/0.60	0.57/0.59	0.57/0.58	0.57/0.58
SAMPLE - NUMBER OF ALTERNATIVES	4	4	4	4
- NUMBER OF OBSERVATIONS	120	120	120	120
- AVAILABLE OBSERVATIONS:				
.Alternative 1 : AIR	120	120	120	120
.Alternative 2 : RAIL	120	120	120	120
.Alternative 3 : BUS	120	120	120	120
.Alternative 4 : CAR	120	120	120	120
TOTAL NUMBER OF FIXED OR ESTIMATED PARAMETERS:				
- BETA .Estimated	13	13	13	13
.CONSTANTS	3	4	3	4
- LAMBDA(X)				
.Fixed (always in G-DOGIT)	0	0	0	0
.Estimated	0	0	3	3
- EXTRA PARAMETERS				
.Fixed	0	8	0	8
.Estimated	0	0	0	0
- TOTAL				
.Fixed	0	8	0	8
.Estimated	16	17	19	20
COVARIANCE MATRIX SIGMA	FULL	FULL	FULL	FULL

* Under each t-statistic, one finds flags for the presence of a Box-Cox transformation (L) and an indication of whether the regression coefficient is generic (GE) or specific (SP).

9. Appendix 3. Results for the discrete urban model (Santiago 1983-1985)

```
===================================================================================================
PART I. Presented statistics[23]
  Model type                                                   LOGIT     LIN-IPTL    LOGIT     LIN-IPTL
  Name and version number of the variant                       LIN:1       LIN:15    BC3:3      BC3:14
===================================================================================================
                                               ALTERNATIVE  1 :  auto-driver
  P     =     PRICES
  ------------------
    TRAV.COSTS/NETWAGE/MIN.(CAR DRIVER)    cosinc1
      BETA coefficient                                         -0.0233    -0.0432    -0.1806    -0.1841
      Elast. of P(autc    ) - at mean(X)                       -0.4259    -0.7517    -0.4884    -0.5292
      T-statistic conditional on LAMBDA(X)                   (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                                (GE)       (GE)    L 3  (GE)  L 3  (GE)
  N     =     NETWORKS
  --------------------
    WALKING TIME (CAR DRIVER)              itcam1
      BETA coefficient                     ------              -0.1616    -0.1752    -0.7832    -0.7838
      Elast. of P(autc    ) - at mean(X)                       -0.8921    -0.9204    -0.8830    -0.8887
      T-statistic conditional on LAMBDA(X)                   (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                                (GE)       (GE)    L 1  (GE)  L 1  (GE)
    IN-VEHICLE TIME (CAR DRIVER)          tdv1m
      BETA coefficient                                         -0.0823    -0.1348    -0.0112    -0.0166
      Elast. of P(autc    ) - at mean(X)                       -1.0915    -1.7021    -0.7866    -1.0142
      T-statistic conditional on LAMBDA(X)                   (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                                (GE)       (GE)    L 2  (GE)  L 2  (GE)
  Y     =     CONSUMER CHARACTERISTICS
  ------------------------------------
    NO. OF CARS/NO. OF DRIVING LICENCES    autlic2
      BETA coefficient                                          2.1947     2.4028     2.1418     2.2135
      Elast. of P(autc    ) - at mean(X)                        1.2421     1.3096     1.2066     1.2195
      T-statistic conditional on LAMBDA(X)                   (   4.99)  (   5.01)  (   4.83)  (   4.80)
                                                                (GE)       (GE)       (GE)       (GE)
  AD    =     ASSOCIATED DUMMIES
  ------------------------------
    ASSOCIATED DUMMY FOR itcam1            dumtcam1
      BETA coefficient                     ========                                  -1.2703    -1.2498
      Elast. of P(autc    ) - at mean(X)                                             -1.2101    -1.1589
      T-statistic conditional on LAMBDA(X)                                         (  -2.63)  (  -3.33)
                                                                                      (GE)       (GE)
  ET    =     ET CETERA
  ---------------------
    REGRESSION CONSTANT                    constant
      BETA coefficient                                          0.0382     0.2995     0.5680     1.7174
      T-statistic conditional on LAMBDA(X)                   (   0.09)  (   0.43)  (   0.94)  (   1.16)
                                                                (SP)       (SP)       (SP)       (SP)
---------------------------------------------------------------------------------------------------
                                              ALTERNATIVE  2 :  auto-passenger
  P     =     PRICES
  ------------------
    TRAV.COSTS/NETWAGE/MIN.(CAR PASS.)     cosinc2
      BETA coefficient                     -------             -0.0233    -0.0432    -0.1806    -0.1841
      Elast. of P(auta    ) - at mean(X)                       -0.0797    -0.1220    -0.3357    -0.3166
      T-statistic conditional on LAMBDA(X)                   (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                                (GE)       (GE)    L 3  (GE)  L 3  (GE)
  N     =     NETWORKS
  --------------------
    WALKING TIME (CAR PASSENGER)           itcam2
      BETA coefficient                     ------              -0.1616    -0.1752    -0.7832    -0.7838
      Elast. of P(auta    ) - at mean(X)                       -0.9689    -0.8672    -0.9151    -0.8521
      T-statistic conditional on LAMBDA(X)                   (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                                (GE)       (GE)    L 1  (GE)  L 1  (GE)
    IN-VEHICLE TIME (CAR PASSENGER)       tdv2m
      BETA coefficient                                         -0.0823    -0.1348    -0.0112    -0.0166
      Elast. of P(auta    ) - at mean(X)                       -1.1001    -1.4882    -0.7773    -0.9275
      T-statistic conditional on LAMBDA(X)                   (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                                (GE)       (GE)    L 2  (GE)  L 2  (GE)
  Y     =     CONSUMER CHARACTERISTICS
  ------------------------------------
    SEX AT ORIGIN (1=MALE,0=FEMALE)        sexo
      BETA coefficient                     ====                -0.3430    -0.3760    -0.3391    -0.3637
      Elast. of P(auta    ) - at mean(X)                       -0.1928    -0.1759    -0.1901    -0.1840
      T-statistic conditional on LAMBDA(X)                   (  -1.52)  (  -1.28)  (  -1.48)  (  -1.39)
                                                                (GE)       (GE)       (GE)       (GE)
```

[23] Under each t-statistic, one finds flags for the presence of a Box-Cox transformation (L) and an indication of whether the regression coefficient is generic (GE) or specific (SP.

```
  AD    =     ASSOCIATED DUMMIES
  ------------------------------
    ASSOCIATED DUMMY FOR cosinc2           dumcos2
      BETA coefficient                     =======                          0.2029     0.2218
      Elast. of P(auta    ) - at mean(X)                                    0.2001     0.1968
      T-statistic conditional on LAMBDA(X)                               (   0.48)  (   0.48)
                                                                            (GE)       (GE)
    ASSOCIATED DUMMY FOR itcam2            dumtcam2
      BETA coefficient                     ========                        -1.2703    -1.2498
      Elast. of P(auta    ) - at mean(X)                                   -1.2528    -1.1090
      T-statistic conditional on LAMBDA(X)                               (  -2.63)  (  -3.33)
                                                                            (GE)       (GE)
  ET    =     ET CETERA
  ---------------------
    REGRESSION CONSTANT                    constant
      BETA coefficient                                         0.1024                1.1327
      T-statistic conditional on LAMBDA(X)                  (   0.16)             (   0.81)
                                                               (SP)                  (SP)
---------------------------------------------------------------------------------------------------
                                  ALTERNATIVE  3 :  collective taxi
  P     =     PRICES
  ------------------
    TRAV.COSTS/NETWAGE/MIN.(TAXI)          cosinc3
      BETA coefficient                     -------      -0.0233    -0.0432    -0.1806    -0.1841
      Elast. of P(taxi    ) - at mean(X)                -0.3589    -0.5319    -0.4695    -0.4712
      T-statistic conditional on LAMBDA(X)            (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                         (GE)       (GE)    L 3  (GE)  L 3  (GE)
  N     =     NETWORKS
  --------------------
    WALKING TIME (TAXI)                    itcam3
      BETA coefficient                                  -0.1616    -0.1752    -0.7832    -0.7838
      Elast. of P(taxi    ) - at mean(X)                -1.0301    -0.8925    -0.8649    -0.8139
      T-statistic conditional on LAMBDA(X)            (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                         (GE)       (GE)    L 1  (GE)  L 1  (GE)
    WAITING TIME (TAXI)                    ites3m
      BETA coefficient                                  -0.1484    -0.1351    -0.3216    -0.2642
      Elast. of P(taxi    ) - at mean(X)                -0.1737    -0.1265    -0.3199    -0.2392
      T-statistic conditional on LAMBDA(X)            (  -1.68)  (  -1.22)  (  -1.40)  (  -1.10)
                                                         (GE)       (GE)    L 1  (GE)  L 1  (GE)
    IN-VEHICLE TIME (TAXI)                 tdv3m
      BETA coefficient                                  -0.0823    -0.1348    -0.0112    -0.0166
      Elast. of P(taxi    ) - at mean(X)                -1.1473    -1.5024    -0.8310    -0.9980
      T-statistic conditional on LAMBDA(X)            (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                         (GE)       (GE)    L 2  (GE)  L 2  (GE)
  Y     =     CONSUMER CHARACTERISTICS
  ------------------------------------
    SEX AT ORIGIN (1=MALE,0=FEMALE)        sexo
      BETA coefficient                     ====         -0.3430    -0.3760    -0.3391    -0.3637
      Elast. of P(taxi    ) - at mean(X)                -0.1928    -0.1698    -0.1901    -0.1853
      T-statistic conditional on LAMBDA(X)            (  -1.52)  (  -1.28)  (  -1.48)  (  -1.39)
                                                         (GE)       (GE)       (GE)       (GE)
  AD    =     ASSOCIATED DUMMIES
  ------------------------------
    ASSOCIATED DUMMY FOR cosinc3           dumcos3
      BETA coefficient                     =======                          0.2029     0.2218
      Elast. of P(taxi    ) - at mean(X)                                    0.1997     0.1980
      T-statistic conditional on LAMBDA(X)                               (   0.48)  (   0.48)
                                                                            (GE)       (GE)
  ET    =     ET CETERA
  ---------------------
    REGRESSION CONSTANT                    constant
      BETA coefficient                                   0.6660     0.8545    -0.2757     0.8626
      T-statistic conditional on LAMBDA(X)            (   2.34)  (   1.28)  (  -0.48)  (   0.61)
                                                         (SP)       (SP)       (SP)       (SP)
---------------------------------------------------------------------------------------------------
                                       ALTERNATIVE  4 :  metro
  P     =     PRICES
  ------------------
    TRAV.COSTS/NETWAGE/MIN.(METRO)         cosinc4
      BETA coefficient                                  -0.0233    -0.0432    -0.1806    -0.1841
      Elast. of P(met     ) - at mean(X)                -0.0169    -0.0242    -0.0601    -0.0593
      T-statistic conditional on LAMBDA(X)            (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                         (GE)       (GE)    L 3  (GE)  L 3  (GE)
  N     =     NETWORKS
  --------------------
    WALKING TIME (METRO)                   itcam4
      BETA coefficient                                  -0.1616    -0.1752    -0.7832    -0.7838
      Elast. of P(met     ) - at mean(X)                -0.3311    -0.2766    -0.1895    -0.1845
      T-statistic conditional on LAMBDA(X)            (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                         (GE)       (GE)    L 1  (GE)  L 1  (GE)
```

```
  WAITING TIME (METRO)                    ites4m
    BETA coefficient                                 -0.1484    -0.1351    -0.3216    -0.2642
    Elast. of P(met     ) - at mean(X)               -0.0434    -0.0304    -0.0690    -0.0531
    T-statistic conditional on LAMBDA(X)           (  -1.68)  (  -1.22)  (  -1.40)  (  -1.10)
                                                      (GE)       (GE)    L 1  (GE)  L 1  (GE)
  IN-VEHICLE TIME (METRO)                 tdv4m
    BETA coefficient                                 -0.0823    -0.1348    -0.0112    -0.0166
    Elast. of P(met     ) - at mean(X)               -0.1118    -0.1411    -0.0557    -0.0707
    T-statistic conditional on LAMBDA(X)           (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                      (GE)       (GE)    L 2  (GE)  L 2  (GE)
ET    =     ET CETERA
---------------------
  REGRESSION CONSTANT                     constant
    BETA coefficient                                  4.2936     4.4505     3.2591     4.4222
    T-statistic conditional on LAMBDA(X)           (  10.99)  (   7.29)  (   4.74)  (   3.39)
                                                      (SP)       (SP)       (SP)       (SP)
-------------------------------------------------------------------------------------------------
                                               ALTERNATIVE  5 :  bus
P     =     PRICES
------------------
  TRAV.COSTS/NETWAGE/MIN.(BUS)            cosinc5
    BETA coefficient                                 -0.0233    -0.0432    -0.1806    -0.1841
    Elast. of P(bus     ) - at mean(X)               -0.1247    -0.2109    -0.3187    -0.3282
    T-statistic conditional on LAMBDA(X)           (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                      (GE)       (GE)    L 3  (GE)  L 3  (GE)
N     =     NETWORKS
--------------------
  WALKING TIME (BUS)                      itcam5
    BETA coefficient                                 -0.1616    -0.1752    -0.7832    -0.7838
    Elast. of P(bus     ) - at mean(X)               -1.0369    -1.0248    -0.8497    -0.8467
    T-statistic conditional on LAMBDA(X)           (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                      (GE)       (GE)    L 1  (GE)  L 1  (GE)
  WAITING TIME (BUS)                      ites5m
    BETA coefficient                                 -0.1484    -0.1351    -0.3216    -0.2642
    Elast. of P(bus     ) - at mean(X)               -0.3613    -0.3000    -0.3286    -0.2639
    T-statistic conditional on LAMBDA(X)           (  -1.68)  (  -1.22)  (  -1.40)  (  -1.10)
                                                      (GE)       (GE)    L 1  (GE)  L 1  (GE)
  IN-VEHICLE TIME (BUS)                   tdv5m
    BETA coefficient                                 -0.0823    -0.1348    -0.0112    -0.0166
    Elast. of P(bus     ) - at mean(X)               -1.9612    -2.9295    -2.0074    -2.4937
    T-statistic conditional on LAMBDA(X)           (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                      (GE)       (GE)    L 2  (GE)  L 2  (GE)
ET    =     ET CETERA
---------------------
  REGRESSION CONSTANT                     constant
    BETA coefficient                                  2.0508     2.7005     1.1194     2.4721
    T-statistic conditional on LAMBDA(X)           (   5.69)  (   3.67)  (   1.68)  (   1.96)
                                                      (SP)       (SP)       (SP)       (SP)
-------------------------------------------------------------------------------------------------
                                      ALTERNATIVE  6 :  aut0-driver and metro
P     =     PRICES
------------------
  TRAV.COSTS/NETWAGE/MIN.(CAR D.-MET)     cosinc6
    BETA coefficient                                 -0.0233    -0.0432    -0.1806    -0.1841
    Elast. of P(autcmet ) - at mean(X)               -0.2318    -0.4033    -0.3925    -0.4178
    T-statistic conditional on LAMBDA(X)           (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                      (GE)       (GE)    L 3  (GE)  L 3  (GE)
N     =     NETWORKS
--------------------
  WALKING TIME (CAR DRIVER-METRO)         itcam6
    BETA coefficient                                 -0.1616    -0.1752    -0.7832    -0.7838
    Elast. of P(autcmet ) - at mean(X)               -0.8281    -0.8422    -0.8267    -0.8322
    T-statistic conditional on LAMBDA(X)           (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                      (GE)       (GE)    L 1  (GE)  L 1  (GE)
  WAITING TIME (CAR DRIVER-METRO)         ites6m
    BETA coefficient                                 -0.1484    -0.1351    -0.3216    -0.2642
    Elast. of P(autcmet ) - at mean(X)               -0.2133    -0.1823    -0.3138    -0.2531
    T-statistic conditional on LAMBDA(X)           (  -1.68)  (  -1.22)  (  -1.40)  (  -1.10)
                                                      (GE)       (GE)    L 1  (GE)  L 1  (GE)
  IN-VEHICLE TIME (CAR DRIVER-METRO)      tv6
    BETA coefficient                                 -0.0823    -0.1348    -0.0112    -0.0166
    Elast. of P(autcmet ) - at mean(X)               -1.1245    -1.7286    -0.8135    -1.0487
    T-statistic conditional on LAMBDA(X)           (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                      (GE)       (GE)    L 2  (GE)  L 2  (GE)
Y     =     CONSUMER CHARACTERISTICS
------------------------------------
  NO. OF CARS/NO. OF DRIVING LICENCES     autlic2
    BETA coefficient                                  2.1947     2.4028     2.1418     2.2135
    Elast. of P(autcmet ) - at mean(X)                1.2421     1.2937     1.2066     1.2201
    T-statistic conditional on LAMBDA(X)           (   4.99)  (   5.01)  (   4.83)  (   4.80)
                                                      (GE)       (GE)       (GE)       (GE)
```

```
ET      =      ET CETERA
---------------------
  REGRESSION CONSTANT                      constant
    BETA coefficient                                    -0.1175    -0.0294    -0.7571     0.3593
    T-statistic conditional on LAMBDA(X)             (  -0.27)  (  -0.04)  (  -0.99)  (   0.25)
                                                        (SP)       (SP)       (SP)       (SP)
-------------------------------------------------------------------------------------------------
                                        ALTERNATIVE  7 :  auto-passenger and metro
P       =      PRICES
------------------
  TRAV.COSTS/NETWAGE/MIN.(CAR P.-MET)      cosinc7
    BETA coefficient                                    -0.0233    -0.0432    -0.1806    -0.1841
    Elast. of P(autamet ) - at mean(X)                  -0.1047    -0.1491    -0.3038    -0.2912
    T-statistic conditional on LAMBDA(X)             (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                        (GE)       (GE)    L 3  (GE)  L 3  (GE)
N       =      NETWORKS
--------------------
  WALKING TIME (CAR PASSENGER-METRO)       itcam7
    BETA coefficient                                    -0.1616    -0.1752    -0.7832    -0.7838
    Elast. of P(autamet ) - at mean(X)                  -0.8307    -0.6915    -0.8543    -0.7934
    T-statistic conditional on LAMBDA(X)             (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                        (GE)       (GE)    L 1  (GE)  L 1  (GE)
  WAITING TIME (CAR PASSENGER-METRO)       ites7m
    BETA coefficient                                    -0.1484    -0.1351    -0.3216    -0.2642
    Elast. of P(autamet ) - at mean(X)                  -0.2197    -0.1537    -0.3248    -0.2418
    T-statistic conditional on LAMBDA(X)             (  -1.68)  (  -1.22)  (  -1.40)  (  -1.10)
                                                        (GE)       (GE)    L 1  (GE)  L 1  (GE)
  IN-VEHICLE TIME (CAR PASSEN.-METRO)      tv7
    BETA coefficient                                    -0.0823    -0.1348    -0.0112    -0.0166
    Elast. of P(autamet ) - at mean(X)                  -1.2584    -1.5832    -0.9642    -1.1433
    T-statistic conditional on LAMBDA(X)             (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                        (GE)       (GE)    L 2  (GE)  L 2  (GE)
Y       =      CONSUMER CHARACTERISTICS
------------------------------------
  SEX AT ORIGIN (1=MALE,0=FEMALE)          sexo
    BETA coefficient                       ====         -0.3430    -0.3760    -0.3391    -0.3637
    Elast. of P(autamet ) - at mean(X)                  -0.1928    -0.1629    -0.1901    -0.1836
    T-statistic conditional on LAMBDA(X)             (  -1.52)  (  -1.28)  (  -1.48)  (  -1.39)
                                                        (GE)       (GE)       (GE)       (GE)
ET      =      ET CETERA
---------------------
  REGRESSION CONSTANT                      constant
    BETA coefficient                                     0.2280     0.1979    -0.6443     0.4502
    T-statistic conditional on LAMBDA(X)             (   0.82)  (   0.29)  (  -1.04)  (   0.31)
                                                        (SP)       (SP)       (SP)       (SP)
-------------------------------------------------------------------------------------------------
                                       ALTERNATIVE  8 :  collective taxi and metro
P       =      PRICES
------------------
  TRAV.COSTS/NETWAGE/MIN.(TAXI-METRO)      cosinc8
    BETA coefficient                                    -0.0233    -0.0432    -0.1806    -0.1841
    Elast. of P(taximet ) - at mean(X)                  -0.2906    -0.3454    -0.4374    -0.4011
    T-statistic conditional on LAMBDA(X)             (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                        (GE)       (GE)    L 3  (GE)  L 3  (GE)
N       =      NETWORKS
--------------------
  WALKING TIME (TAXI - METRO)              itcam8
    BETA coefficient                                    -0.1616    -0.1752    -0.7832    -0.7838
    Elast. of P(taximet ) - at mean(X)                  -1.2118    -0.8418    -0.8794    -0.7635
    T-statistic conditional on LAMBDA(X)             (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                        (GE)       (GE)    L 1  (GE)  L 1  (GE)
  WAITING TIME (TAXI - METRO)              ites8m
    BETA coefficient                                    -0.1484    -0.1351    -0.3216    -0.2642
    Elast. of P(taximet ) - at mean(X)                  -0.4725    -0.2758    -0.3425    -0.2401
    T-statistic conditional on LAMBDA(X)             (  -1.68)  (  -1.22)  (  -1.40)  (  -1.10)
                                                        (GE)       (GE)    L 1  (GE)  L 1  (GE)
  IN-VEHICLE TIME (TAXI - METRO)           tv8
    BETA coefficient                                    -0.0823    -0.1348    -0.0112    -0.0166
    Elast. of P(taximet ) - at mean(X)                  -1.1254    -1.1816    -0.8017    -0.8866
    T-statistic conditional on LAMBDA(X)             (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                        (GE)       (GE)    L 2  (GE)  L 2  (GE)
ET      =      ET CETERA
---------------------
  REGRESSION CONSTANT                      constant
    BETA coefficient                                     0.2066     0.1284    -0.5038     0.5034
    T-statistic conditional on LAMBDA(X)             (   0.48)  (   0.17)  (  -0.67)  (   0.32)
                                                        (SP)       (SP)       (SP)       (SP)
-------------------------------------------------------------------------------------------------
```

```
                         ALTERNATIVE  9 :  bus and metro
  P     =     PRICES
  ------------------
    TRAV.COSTS/NETWAGE/MIN.(BUS-METRO)     cosinc9
      BETA coefficient                                  -0.0233    -0.0432    -0.1806    -0.1841
      Elast. of P(busmet  ) - at mean(X)                -0.2091    -0.3357    -0.3847    -0.3956
      T-statistic conditional on LAMBDA(X)            (  -2.80)  (  -4.35)  (  -2.87)  (  -3.47)
                                                         (GE)       (GE)    L 3  (GE)  L 3  (GE)
  N     =     NETWORKS
  --------------------
    WALKING TIME (BUS - METRO)             itcam9
      BETA coefficient                                  -0.1616    -0.1752    -0.7832    -0.7838
      Elast. of P(busmet  ) - at mean(X)                -1.2057    -1.1311    -0.8648    -0.8500
      T-statistic conditional on LAMBDA(X)            (  -7.82)  (  -7.42)  (  -7.95)  (  -8.07)
                                                         (GE)       (GE)    L 1  (GE)  L 1  (GE)
    WAITING TIME (BUS - METRO)             ites9m
      BETA coefficient                                  -0.1484    -0.1351    -0.3216    -0.2642
      Elast. of P(busmet  ) - at mean(X)                -0.6050    -0.4769    -0.3421    -0.2728
      T-statistic conditional on LAMBDA(X)            (  -1.68)  (  -1.22)  (  -1.40)  (  -1.10)
                                                         (GE)       (GE)    L 1  (GE)  L 1  (GE)
    IN-VEHICLE TIME (BUS - METRO)          tv9
      BETA coefficient                                  -0.0823    -0.1348    -0.0112    -0.0166
      Elast. of P(busmet  ) - at mean(X)                -1.3467    -1.9095    -1.0810    -1.3422
      T-statistic conditional on LAMBDA(X)            (  -3.59)  (  -4.41)  (  -3.39)  (  -3.73)
                                                         (GE)       (GE)    L 2  (GE)  L 2  (GE)
  ET    =     ET CETERA
  ---------------------
    REGRESSION CONSTANT                    constant
      BETA coefficient                                   1.5561     1.7644     0.7216     1.8578
      T-statistic conditional on LAMBDA(X)            (   3.58)  (   2.61)  (   0.99)  (   1.34)
                                                         (SP)       (SP)       (SP)       (SP)
================================================================================================
PART II. Parameters and T-statistic unconditional (=0) [=1]
================================================================================================
  BOX-COX TRANSFORMATIONS
    LAMBDA(X)  1                                                             0.0617     0.0809
                                                                          (   0.30)  (   0.51)
                                                                          [  -4.53]  [  -5.76]
    LAMBDA(X)  2                                                             1.6301     1.5887
                                                                          (   2.74)  (   2.39)
                                                                          [   1.06]  [   0.89]
    LAMBDA(X)  3                                                             0.3529     0.3828
                                                                          (   1.46)  (   1.23)
                                                                          [  -2.68]  [  -1.98]
LIN-IPTL: all 9 PHI and MU FIXED at                                0.9900                0.9900
================================================================================================
PART III. General statistics
================================================================================================
  LOG-LIKELIHOOD  - FINAL VALUE (LF)                    -855.96    -853.90    -837.93    -837.86
                  - INITIAL VALUE                       -924.69    -924.70    -924.69    -924.70
                  - DEGREES OF FREEDOM                        6          6         11         11
  RHO-SQUARED                                            0.0743     0.0766     0.0938     0.0939
  RHO-SQUARED BAR - AKAIKE                               0.0592     0.0603     0.0733     0.0723
                  - HOROWITZ                             0.0668     0.0685     0.0835     0.0831
                  - HENSHER AND JOHNSON                  0.0703     0.0723     0.0885     0.0883
  PERCENT RIGHT                                           47.11      47.59      49.52      49.84
  SAMPLE - NUMBER OF ALTERNATIVES                             9          9          9          9
         - NUMBER OF OBSERVATIONS                           622        622        622        622
           .ALTERNATIVE  1 : autc                           420        420        420        420
           .ALTERNATIVE  2 : auta                           466        466        466        466
           .ALTERNATIVE  3 : taxi                           560        560        560        560
           .ALTERNATIVE  4 : met                            141        141        141        141
           .ALTERNATIVE  5 : bus                            613        613        613        613
           .ALTERNATIVE  6 : autcmet                        375        375        375        375
           .ALTERNATIVE  7 : autamet                        421        421        421        421
           .ALTERNATIVE  8 : taximet                        435        435        435        435
           .ALTERNATIVE  9 : busmet                         456        456        456        456
  TOTAL NUMBER OF FIXED OR ESTIMATED PARAMETERS:
         - BETA .Estimated                                    6          6          8          8
                .CONSTANTS                                    8          9          8          9
         - LAMBDA(X)
                .Fixed (always in G-DOGIT)                    0          0          0          0
                .Estimated                                    0          0          3          3
         - EXTRA PARAMETERS
                .Fixed                                        0          9          0          9
                .Estimated                                    0          0          0          0
         - TOTAL
                .Fixed                                        0          9          0          9
                .Estimated                                   14         15         19         20
================================================================================================
```

10. References

Abraham, C. (1961). La répartition du trafic entre itinéraires concurrents: réflexions sur le comportement des usagers, application au calcul des péages. *Revue générale des routes et aérodromes* 357, pages 57-60, 65-72, 75-76.

Anderson, S.P., Palma, A. de, Thisse, J.-F. (1988). A representative consumer theory of the logit model. *International Economic Review* 29, 3, 461-466, August.

Ben-Akiva, M., Lerman, S. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA, The MIT Press.

Berkson, J. (1944). Applications of the Logistic Function to Bioassay. *Journal of the American Statistical Association* 39, 357-365.

Berndt, E.R., Khaled, M.S. (1979). Parametric Productivity Measurement and Choice among Flexible Functional Forms. *Journal of Political Economy* 87, 6, 1220-1245, December.

Bierlaire, M., Lotan, T., Toint, P. (1997). On the Overspecification of Multinomial and Nested Logit Models Due to Alternative-specific Constants. *Transportation Science* 31, 4, 397-371.

Box, G.E.P., Cox, D.R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society B*, 26, 211-243.

Box, G.E.P., Tidwell, P.W. (1962). Transformation of the Independent Variables. *Technometrics* 4, 531-550.

Caves, D.W., Christensen, L.R., Swanson, J.A. (1980a). Productivity in U.S. Railroads, 1951-1974. *Bell Journal of Economics* 11, 1, 168-181.

Caves, D.W., Christensen, L.R., Tretheway, M. W. (1980b). Flexible Cost Functions for Multiproduct Firms. *Review of Economics and Statistics*, 477-481 (Notes Section).

Caves, D.W., Christensen, L.R., Tretheway, M.W., Windle. J. (1985). Network effects and the measurement of returns to scale and density for U.S. railroads. Ch. 4 in A. F. Daughety (ed.), *Analytical Studies in Transport Economics*, 97-120, Cambridge University Press, Cambridge.

Cherchi, E., Ortúzar, J. de D. (2006). On fitting mode-specific constants in the presence of new options in RP/SP models. *Transportation Research A* 40, 1, 1-18.

Cox, D.R. (1970). *An analysis of binary data*. 142 p., Methuen and Co Ltd, London.

CRA (1972). *A disaggregated behavioural model of urban travel demand*. Report No. CRA-156-2 prepared under contract No. FH-11-7566 "Development of a behavioural modal split model" for the Federal Highway Administration. Charles River Associates Incorporated, Cambridge, Massachusetts, March.

Dial, R.B. (1971). A Probabilistic Multipath Traffic Assignment Model which Obviates Path Enumeration. *Transportation Research* 5, 83-111.

Domencich, T., McFadden, D. (1975). *Urban Travel Demand: A Behavioral Analysis*. North Holland, Amsterdam.

Gaudry, M. (1981). The Inverse Power Transformation Logit and Dogit Mode Choice Models. *Transportation Research B* 15, 2, 97-103.

Gaudry, M. (1993). Asymmetric Shape and Variable Tail Thickness in Multinomial Probabilistic Responses to Significant Transport Service Level Changes. In Comtois, C., F. Bingquan, F., eds. *Proceedings of the '92 International Conference on Urban Land Use and Transport System ICULTS '92*, Publication CRT-910, Centre de recherche sur les transports, Université de Montréal, avril.

Gaudry, M. et al. (1997). *Cur cum TRIO?* Publication CRT-901, Centre de recherche sur les transports, Université de Montréal, 20 pages.

Gaudry, M., Laferrière, R. (1989). The Box-Cox Transformation: Power Invariance and a New Interpretation. *Economics Letters* 30, 27-29.

Gaudry, M.J.I., Wills, M.J. (1978). Estimating the Functional Form of Travel Demand Models. *Transportation Research* 12, 4, 257-289, 1978.

Gaudry, M., Jara-Diaz, S.R., Ortuzar, J. de D. (1989). Value of Time Sensitivity to Model Specification. *Transportation Research B* 23, 2, 151-158.

Johnston, J. (1984). *Econometric Methods*. Third Edition, 568 p., McGraw-Hill Book Company.

Kessel, P., Röling, W., Kienzler, H.-P. (1986). *Untersuchungsbericht zum Forschungsprojekt der MVP Versuchs- und Planungsgesellschaft für Magnetbahnsysteme, Modellentwicklung zur Prognose der Verkehrsnachfrage auf Magnetbahnsystemen*. Beratergruppe für Verkher+Umwelt (BVU), Freiburg im Bresgau.

Koenker R., Yoon, J. (2009). Parametric links for binary choice models: A Fisherian-Bayesian colloquy. *Journal of Econometrics* 152, 2, 120-130, October.

Laferrière, R. (1988). *Une Agrégation Nouvelle des Itinéraires de Transport Aérien (ANITA).* Publication CRT-574, Centre de recherche sur les transports, Université de Montréal, 106 p., février.

Laferrière, R. (1999). The Air Network Itinerary and Trip Aggregation (ANITA) Model. Ch. 14 in Gaudry, M. and R. R. Mayes (eds), *Taking Stock of Air Liberalization*, Kluwer Academic Press, pages 223-228.
Laferrière, R., Gaudry, M. (1992). Testing the Linear Inverse Power Transformation Logit Mode Choice Model. *Demand, Traffic and Network Modelling,* Selected Proceedings of the Sixth World Conference on Transport Research, Volume II, 739-750.
Lapparent, M. de (2004). *La valeur du temps dans les déplacements réguliers professionnels*. Thèse d'économie, Université Paris I –Panthéon-Sorbonne. Publication AJD-81, Agora Jules Dupuit, Université de Montréal, 180 p., juin. www.e-ajd.net.
Lapparent, M. de (2010). Attitude towards Risk of Time Loss in Travel Activity and Air Route Choices. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, forthcoming. Formerly: Optimism, Pessimism toward Risk of Time Losses in Business Air Travel Demand, and Insurance Premium. Publication AJD-77, Agora Jules Dupuit, Université de Montréal, 29 p., (www.e-ajd.net).
Leurent, F. (1999). *Calcul de deux modèles d'affectation*. Rapport de convention SETRA-INRETS, Référence INRETS: 18-98 SETRA, 41 p., février.
Maddala, G.S. (1983). *Limited-dependent and qualitative variables in econometrics*. Econometric Society Monographs in Quantitative Economics, Cambridge University Press.
Manski, C.F., Lerman. S.R. (1977). The Estimation of Choice Probabilities from Choice Based Samples, *Econometrica* 45, 1977-1988.
McFadden, D. (1968). *The Revealed Preferences of a Government Bureaucracy*. Technical Report No 17, Institute of International Studies, University of California at Berkeley, 52 p.
McFadden, D. (1975a). *On independence, structure, and simultaneity in transportation demand analysis*. Working paper 7511, Institute of Transportation and Traffic Engineering, University of California, Berkeley, CA.
McFadden, D. (1975b). The Revealed Preferences of a Government Bureaucracy: Theory. *Bell Journal of Economics* 6, 401-416.
McFadden, D. (1976). The Revealed Preferences of a Government Bureaucracy: Empirical Evidence, *Bell Journal of Economics* 7, 55-72.
McFadden, D. L. (1984). Econometric analysis of qualitative response models. Ch. 24 in Griliches, Z., Intriligator, M.D., *Handbook of Econometrics*, Volume II, Elsevier Science Publishers BV, 1395-1457.
Montmarquette, C. and S. Mahseredjian (1985). Functional form and educational production functions, *Economic Letters* 19, 291-294.
Pregibon, D. (1980). Goodness of link tests for generalized linear models. *Applied Statistics* 29, 15-24.
Schlesselman, J. (1971). Power families: a note on the Box and Cox transformation. *Journal of the Royal Statistical Society* Series B 33, 307-31.
Tardiff, T.J. (1978). *The Use of Alternaive-specific Constants in Choice Modelling.* Working Paper UCI-ITS-SP-78-6, Institute of Transportation Studies, University of California, Irvine, 28 p., December.
Theil, H. (1969). A Multinomial Extension of the Linear Logit Model. *International Economic Review* 10, 251-259.
Thomas, T.C. (1967). *The value of time for passenger cars: an experimental study of commuters'* values. Report PB-175-731, Stanford Research Institute, May.
Tran, C.-L., Gaudry, M. (2008). *SHARE: The S-1 to S-5 estimation procedures 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 AJD-114, Agora Jules Dupuit, Université de Montréal, 76 p., October.
Tran, C.-L., Gaudry, M. (2009). *PROBABILITY: The P-2 to P-6 estimation procedures for the Standard and Generalized BOX-COX LOGIT and DOGIT and for the Linear and Box-Tukey INVERSE POWER TRANSFORMATION-LOGIT models with disaggregate data.* Publication AJD-111, Agora Jules Dupuit, Université de Montréal, 72 p., January.
Tran, C.-L., Gaudry, M. (2011a). *SHARE: The S-1/S-5 Program user guide with databases, examples and Tablex outputs*. Publication AJD-115, Agora Jules Dupuit, Université de Montréal, 190 p., June.
Tran, C.-L., Gaudry, M. (2011b). *PROBABILITY: The P-2/P-6 program user guide with databases, examples and Tablex outputs*. Publication AJD-112, Agora Jules Dupuit, Université de Montréal, 221 p., June.
Warner, S.L. (1962). *Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice*, Northwestern University Press, Evanston, Illinois.
Wills, M.J. (1982). On the Equivalence of Common Rearrangements for Sum Constrained Travel Estimation Problems. *Transportation Research B* 16B, 339-359.