Attitudes to Distance, Time and Cost in Logit Transport Choice Models

Marc Gaudry\textsuperscript{1}
Matthieu de Lapparent\textsuperscript{2}

\textsuperscript{1}Département de sciences économiques et Agora Jules Dupuit (AJD), Université de Montréal, Canada
\textsuperscript{2}Institut Français des Sciences et Technologies des Transports, de l’Aménagement et des Réseaux (IFSTTAR), Université Paris-Est, France
Abstract

During the last 20 years, a distance variable has sometimes been added to specifications of Logit utility functions that already included travel time and cost variables long assumed in theory and practice to embody the impeding effect of distance on transport demand. We briefly recall some successful milestones of this seemingly superfluous enrichment of modeling practice and propose to understand it as allowing for the expression of an “Attitude to Distance” distinct from the “Attitude to Time or Cost” level of service outcomes.

The framework adopted to document this split role of distance, a duality not entirely absent from common language, is the Multinomial Box-Cox Logit model where, in the hope of improving estimates of Values of Travel Time Savings, distance raised to a simple power has often been introduced in interactions with the time and cost terms. The new power parameters of distance then betray an optimistic, pessimistic or neutral behavior towards distance distinct from any attitudes to the level of service variables themselves. In other Logit specifications of utility functions based from the start on price, speed and distance variables, the self-standing distance term can also be re-interpreted as comprising a “Distance Attitude” component in addition to an impedance effect.

Key Words. Multinomial Logit, Attitude to outcome, Attitude to distance, Power functions, Box-Cox transformations, Value of travel time savings, constancy of marginal utility in Logit models.

Table of contents

The absence of distance from transport mode and itinerary choice models .................................................................3
Some occurrences of the introduction of distance in Logit utility functions .................................................................4
  Econometric framework: the Box-Cox MNL and its derived values of time savings ..............................................4
  The intuitive introduction of interactions between distance and LOS variables ......................................................5
A survey of Box-Cox Logit models and the multiple roles of distance ........................................................................6
  A Distance Attitude (DA) ..............................................................................................................................................6
  The changed meaning of rate-specified RUF ..............................................................................................................7
Conclusion ........................................................................................................................................................................8
References ........................................................................................................................................................................8

Journal of Economic Literature Classification: C-35, C-53, D-12, R-41.
The absence of distance from transport mode and itinerary choice models

The workhorse of transport demand analysis has long been the Logit model. Ever since early Binomial applications 55 years ago, for instance as an itinerary choice model to produce road traffic forecasts for the Channel Tunnel (Setec, 1959) or as an urban mode choice model (Warner, 1962), it was implicitly assumed that key level of service (LOS) variables, primarily travel time and travel cost, duly incorporated the role of distance. Other variables potentially missing from the $V_i$ utility functions defined by alternative $(i=1,\ldots,M)$, such as comfort or even risk and uncertainty, were often seen as inherent qualities of the alternatives reflected in the constant (intercept), other specific coefficients, or orthogonal error terms. In practice, a distance term was deemed unnecessary, even as potential stand-in for the missing variables.

What about theory? The first proper deductions of the Probit and Linear Probability model types from utility functions defined by a combination of systematic and random terms (of Normal and Uniform distributions) were based only on the time and cost determinants of systematic utility and on their errors (Abraham, 1961). And later derivations of the Logit and Arctan model types, based on Weibull and Cauchy distributions (CRA, 1972) of the random terms, did not break step on this point. Until Ramjerdi (1993), distance was always absent from the systematic utility components of Logit models of any character — Multinomial or Nested (e.g. McFadden, 1978); with fixed or random regression coefficients (e.g. Hensher & Johnson, 1979; Johnson, 1979) — containing time and cost.

Although we limit our discussion to the Logit model type, which rapidly displaced its early Random Utility Maximization (RUM) competitors — such as the Binary Probit favored in some urban modal choice problems (e.g. Barbier, 1966) — from the mainstream of theoretical and applied transport demand work, minority streams did not differ from the dominant Logit stream in this respect: they did not use distance as an explanatory variable in addition to the LOS variables themselves.

But minority stream Probit models, by contrast with majority stream Logit models, were for many years seldom estimated with non linear forms of their time and cost variables because of the numerical difficulty of the task. Consequently, to the extent that firstly used interactions between distance and LOS variables aimed precisely at modifying the constant values of travel time savings (VTTS) implied by linear LOS specifications, it was easier to introduce them, and to further question those linear restriction as well, in Logit models. The framework of the Box-Cox Logit specification is then very fitting for us to present both a brief history of the introduction of distance over the last 20 years and our interpretation of its new role. Similarly, we adopt the available general terminology of rank dependent utility even if distance is not here shown to act as a rough measure of risk and if LOS outcomes are assumed certain. This avoids inventing a new expression for the interpretation proposed.

That interpretation, as the expression of an “attitude to distance” distinct from the “attitude to time or to cost”, with parameters perhaps jointly estimated by co-monotonic non linear transformations of distance and LOS variables (assumed known with certainty), has implications for other transport demand contexts, such as trip distribution and destination choice. We do not discuss those here but simply adopt the language of mode (and itinerary) choice applications, where the new role of distance was first introduced with increasing success, as our chronological focus briefly documents.

---

1 The author even associated pairs of time and cost specific errors to road itineraries in California.
2 Bolduc (1999), for instance, stated that: “to implement a Box-Cox technique within a MNP setting represents a too formidable task”.
3 The present authors first derived the «Attitude to Distance» interpretation in April 2010 as a device to classify some of the 50 models reported in a survey of Box-Cox Logit mode choice models (Gaudry, 2010, Section 5) that notably included a prospect theory path choice application (Lapparent, 2004). Here, we draw extensively upon this work, without further due reference.
Some occurrences of the introduction of distance in Logit utility functions

We are principally interested in how the contents of typical random utility functions (RUF) progressively changed from \( V_i = f_i(T, C, \text{etc.}) \) to \( V_i = f_i(T, C, D, \text{etc.}) \), but we shall also briefly discuss implications of this development for the understanding of the less frequently found RUF that had included distance from the start, e.g. \( V_i = f_i(P, S, D, \text{etc.}) \). For any origin to destination (O-D) pair, RUF belonging to the latter class are written directly in terms of rates and distance, rather than in terms of time and money expenditures, of the modes.

We discuss neither cases of “universal” or “mother” utility functions (i.e. Marshallian ones), where the service characteristics of one mode might also appear in the utility functions of other modes (McFadden, 1975), nor cases of random regression coefficients or powers of variables. For the sake of clarity, we also neglect all O-D and observational subscripts and will not directly report in our table on the t-statistics (or on the Likelihood ratio tests) associated with selected parameter results because these statistics are not always available from the referenced papers.

Econometric framework: the Box-Cox MNL and its derived values of time savings

In the Standard Box-Cox MNL, the choice probabilities for the \( i^{th} \) mode are assumed given by

\[
p(i) = \frac{\exp(V_i)}{\sum_j \exp(V_j)} , \quad i, j = 1, \ldots, M
\]

\[
V_i = \beta_{i0} + \sum_k \beta_k X_k^{(\lambda_k)} , \quad k = 1, \ldots, K
\]

\[
X_k^{(\lambda_k)} = \begin{cases} (X_k)^{\lambda_k} - 1, & \text{if } \lambda_k \neq 0, \\ \ln(X_k), & \text{if } \lambda_k = 0. \end{cases}
\]

where the Box-Cox transformation (BCT) defined in (1-C) without Tukey’s shift parameter is for simplicity assumed to be generic and applicable only to strictly positive explanatory variables \( X_k \). It is a particularly convenient way\(^4\) to tests consumers’ responsiveness to these characteristics, linearity implying constant marginal utility.

A key statistic derived from such models is the marginal rate of substitution (MRS) between time and cost, or Value of Travel Time Savings (VTTS), defined\(^5\) for instance for the rail mode as:

\[
VTTS_{\text{rail}} = \frac{\partial p}{\partial X_{\text{rail}, \text{Time}}} \frac{\partial X_{\text{rail}, \text{Time}}}{\partial V_{\text{rail}, \text{Time}}} = \frac{\partial V}{\partial X_{\text{rail}, \text{Cost}}} \frac{\partial X_{\text{rail}, \text{Cost}}}{\partial V_{\text{rail}, \text{Cost}}} = \frac{\beta_{\text{rail}, \text{Time}} X_{\text{rail}}^{(\lambda_{\text{rail}} - 1)}}{\beta_{\text{rail}, \text{Cost}} X_{\text{rail}}^{(\lambda_{\text{rail}} - 1)}}
\]

which, after estimation of the vectors of parameters \( (\beta_{\text{rail}}, \lambda) \), may simply be rewritten in terms of the implicit price \( P \), speed \( V \) and distance \( D \) terms (presuming C=PD and \( T = V^{-1} D \)) as:

\[
VTTS_{\text{rail}} = \frac{\beta_{\text{rail}, \text{Time}} V^{-1}_{\text{rail}}}{\beta_{\text{rail}, \text{Cost}} P_{\text{rail}}^{(\lambda_{\text{rail}} - 1)}} D_{\text{rail}}^{\lambda_{\text{rail}} - \lambda_{\text{cost}}},
\]

where \( \lambda_{\text{rail}} - \lambda_{\text{cost}} > 0 \) is a necessary and sufficient condition for the VTTS to increase with distance, traditionally interpreted as impeding spatial interaction.

\(^4\) There exist several ways to test for non-linearities such as, to cite a few, power series expansions or ad hoc mixtures of pre-determined non-linear forms, but there is no parametric function combining flexibility and parsimony like the BCT. Non-parametric formulations of the problem often amount to «super-parametric» ones!

\(^5\) In models with non separable «Marshallian» utility, the derivative with respect to the choice probability differs from the derivative with respect to the utility because any given modal time of cost variable may appear in many utility functions.
If one allows in (1-B) for interactions between distance raised to a simple power $\gamma$ and LOS variables, namely for new $D^{\gamma}_{\text{Time}}$ and $D^{\gamma}_{\text{Cost}}$ terms, the corresponding expression for the VTTS becomes, after re-estimation of the vectors of regression and power parameters $(\beta_{\text{rail}}, \gamma, \lambda)$:

$$
(2-C) \quad \text{VTTS}_{\text{rail}} = \frac{\beta_{\text{rail, Time}} \left[ (\gamma_{\text{Time}})^{-1} \right]}{\beta_{\text{rail, Cost}} \left[ (\lambda_{\text{Cost}})^{-1} \right]} \cdot D^{\gamma_{\text{Time}}} - \lambda_{\text{Cost}} + D^{\gamma_{\text{Cost}}} - \lambda_{\text{Cost}},
$$

where a new identifiable role appears for distance, apparently independent from its previous impedance role. Summing the simple and BCT powers $\gamma_{\text{Time}} - \lambda_{\text{Cost}} + \lambda_{\text{Time}} - \lambda_{\text{Cost}}$ then yields the direction of the total effect of distance on the VTTS.

The intuitive introduction of interactions between distance and LOS variables

For the VTTS to increase with distance in this last sum if $\gamma_{\text{Cost}} = 0$ and the BCT are set to 1,00, it suffices that $\gamma_{\text{Time}} > 0$. This is precisely the seminal specification used by Ramjerdi (1993), with the following unique interaction with time:

$$
(3-A) \quad f(D,T) = \beta_{T} \left[ D \right]^{\gamma_{T}} \left[ T \right],
$$

which yielded the positive value 0,40 reported in table 1. This simple interaction was introduced on an intuitive basis at the end\(^6\) of an extensive study of intercity travel patterns and related VTTS in Norway.

This success and various other indices prompted five experienced researchers (Mackie et alii, 2003) to perform an extensive meta-analysis of the link between distance and VTTS estimates. It showed that:

(i) relatively to cost, the value of in-vehicle time systematically increased with distance, but more for car users than for users of public transport modes (train and bus). This property of models estimates is called « cost damping » by the UK Department for Transport (DfT);

(ii) relatively to in-vehicle time, the value of walking and waiting times systematically decreased with distance while that of vehicle headway increased, again more for car users than for users of public modes.

These authors’ conclusive results and correspondingly strong recommendations stimulated many to introduce interactions between distance and LOS variables. Notably, a group of six co-authors (Axhausen et alii, 2008) even extended interactions jointly to distance and income (I), with a view to obtaining for Switzerland VTTS that varied continuously with both distance and income. Their specification for time, similar to that adopted for cost, remained linear in the LOS variables, e.g.:

$$
(3-B) \quad f(D,I,T) = \beta_{I} \left[ D \right]^{\gamma_{I}} \left[ I \right]^{\rho_{I}} \left[ T \right],
$$

and their paper title made it clear that interactions with distance were now part of the state-of-the-art, a point of practice confirmed also by Hess et al. (2008) with 4 distinct data sets considered separately and jointly. Soon enough, triply co-monotonic functions, specified for instance as:

$$
(3-C) \quad f(D,I,T) = \beta_{I} \left[ D \right]^{\gamma_{I}} \left[ I \right]^{\rho_{I}} \left[ T \right]^{\lambda_{T}},
$$

were successfully estimated for the first time (Lapparent et alii, 2009), and continue to be estimated (e.g. Lapparent, 2014), despite their technical difficulty\(^7\). The negative tally of -0,42 obtained for

\(^6\) It does not appear in the first part of her study (Ramjerdi & Rand, 1992).

\(^7\) For reasons that remain unclear, here thinking about strictly numerical challenges that still need to be overcome despite the availability of computational power, products of BCT are notably more difficult to estimate that products of simple and BCT powers, or products of simple powers.
\[ \gamma_{Time} - \gamma_{Cost} + \lambda_{Time} - \lambda_{Cost}, \]

as reported in the darkly framed square of column 9 of table 1, may be due to the very unusual nature of the sample pertaining to «occasional» trips because such a tally is usually positive.

In these three intercity models (A, B, and C in table 1), the justification for the use of distance remains intuitive, by contrast with the justification for the inclusion of income, which has a long micro-economic history almost requiring its presence in some form or other (Train & McFadden, 1978). For consumer surplus calculations, the simplest form for the inclusion of income is linearity, lest «the calculation of the log-sum be perturbed» (McFadden, 1998). Unfortunately, linearity in income is no more generally tenable empirically than the assumption that the marginal utilities of time and cost are constant in (3-C). This points to the existence of a major scientific gap still to be filled between the search of a parcimonious and a parametrized non-linear form of some utility functions and a more general cost-benefit approach.

**A survey of Box-Cox Logit models and the multiple roles of distance**

As part of an examination of «cost damping» carried out for the DfT, Daly (2008, 2009) asked in particular whether damping claims were consistent with evidence provided by the class of Box-Cox Logit models, thereby prompting the survey of such models mentioned above (Gaudry, 2010).

On the basis of a sample of about 50 models where distinct BCT had been estimated for time and cost in 10 countries, it was found in intercity freight\(^8\) and in intercity or urban passenger models that:

(i) as expected, in the overwhelming majority of cases, \(\lambda_T - \lambda_C > 0\);

(ii) simultaneously, \(\lambda_T > 1\) in urban markets and \(\lambda_T < 1\) in intercity markets.

This empirical finding of a potential structural difference in consumer preferences with respect to LOS between urban and intercity markets, whereby demand appeared to fall faster with distance in urban than in intercity markets, made it imperative to make sense of the new role of distance interaction terms, perhaps competing with the simple impedance function of distance built into time and cost terms and giving rise, in any case, to an applied calculus of component parameters such as

\[ \gamma_{Time} - \gamma_{Cost} + \lambda_{Time} - \lambda_{Cost}. \]

**Table 1. Estimates of simple powers of distance (\(\gamma\)) and of Box-Cox powers (\(\lambda\)) of time or cost**

<table>
<thead>
<tr>
<th>Study</th>
<th>Market</th>
<th>Modes</th>
<th>Purpose</th>
<th>Power parameter estimates</th>
<th>Differences</th>
<th>Mean Distance (km)</th>
<th># obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Ramjerdi (1993), T. 9.5.28</td>
<td>Norway</td>
<td>5 modes</td>
<td>Work</td>
<td>(\gamma_T) = 0.40, (\lambda_T) = 1.00, (\gamma_C) = 1.40, (\lambda_C) = 1.00</td>
<td>(\gamma_T - \gamma_C) = 0.40, (\lambda_T - \lambda_C) = 0.00</td>
<td>108.79, 15824</td>
<td>10</td>
</tr>
<tr>
<td>B. Axhausen et alii (2008), T. 4</td>
<td>Switzerland</td>
<td>4 modes</td>
<td>All</td>
<td>(\gamma_T) = 0.26, (\lambda_T) = 0.74, (\gamma_C) = 0.60, (\lambda_C) = 1.00</td>
<td>(\gamma_T - \gamma_C) = 0.04, (\lambda_T - \lambda_C) = 0.34</td>
<td>42.89, 15870</td>
<td>10</td>
</tr>
<tr>
<td>C. Lapparent et alii (2009), T. 6</td>
<td>Czech Rep.</td>
<td>4 modes</td>
<td>Occasional</td>
<td>(\gamma_T) = 0.66, (\lambda_T) = 0.74, (\gamma_C) = 0.08, (\lambda_C) = 0.21</td>
<td>(\gamma_T - \gamma_C) = -0.87, (\lambda_T - \lambda_C) = 0.45</td>
<td>145.47, 2044</td>
<td>10</td>
</tr>
<tr>
<td>D. 'Houre et alii (2013), p. 30</td>
<td>France</td>
<td>3 modes</td>
<td>All</td>
<td>(\gamma_T) = 0.44, (\lambda_T) = 0.56, (\gamma_C) = 0.92, (\lambda_C) = 1.00</td>
<td>(\gamma_T - \gamma_C) = 0.48, (\lambda_T - \lambda_C) = 0.08</td>
<td>n.c., 18039</td>
<td>10</td>
</tr>
</tbody>
</table>

\[ \gamma = \text{Attitude towards Distance} \]

**A Distance Attitude (DA)**

In the survey, the present authors had noted that, as in columns 1, 4 and 7 of table 1, estimated powers of distance associated with time or cost and their differences were always smaller than unity, which implied a «contraction» of actual distances. In these conditions, perhaps might distance, as revealed in everyday language, have a role in consumer utility functions independently from that of its implicit presence in LOS variables. This split was then implemented by distinguishing formally between the

---

\(^8\) The sample contained no urban freight model.
attitude to LOS and the attitude to distance, per force linked in (2-B). The adopted mechanism, to be described presently, was that of a product of functions which assumed that LOS service variables (the network) are provided under certainty and taken as given\(^9\) by the traveller.

Consider then a random variable \( x \) characterizing transport run distance, \( x \in [D, \bar{D}], D > 0, \bar{D} < D \), its associated distribution excluding the null value, \( F : \mathbb{R} \setminus \{0\} \to [0, 1] \), and a transformation function \( \psi(\cdot) \) maintaining the capacity of that distribution, i.e. with the property \( \psi : [0, 1] \to [0, 1] \). Assume further that \( x \) is linked to an exogenously determined network supply \( n \) drawn from a finite set of discrete possibilities \( N(x) \). The traveller may then be assumed to maximize utility on this predetermined distance interval offering a given fixed Supply:

\[
(4-A) \quad \int_D^\infty \max_{x \in \mathbb{R}[x]} \left[ aou \left( x, g'(x) \right) \right] d\psi \left( F(x) \right)
\]

where, if the supply actually consists in \( L \) service attributes, the interaction of attitude to distance and attitude to attribute LOS terms might be taken to be

\[
(4-B) \quad aou \left( x, g'(x) \right) = \beta_0 + \sum_{l=1}^{L} \beta_{l} \left[ h_{los}(x, D, \bar{D}) \right] \left[ aou_{los}(g_{l}(x)) \right]
\]

and where, selecting \( aou_{los}(g_{l}(x)) = T^{(\lambda_{los})} \) as expressing the attitude to time outcome, obvious specifications for the attitude to distance term certainly include

\[
(4-C) \quad h_{los}(x, D, \bar{D}) = \begin{cases} 
\left( \frac{x - D}{\bar{D}} \right)^{\gamma_{los}} \\
\left( x^{\gamma_{los}} - 1 \right) \lambda_{los}
\end{cases}
\]

We then chose the simple power candidate to match and account for past practice, say with time:

\[
(3-D) \quad f(D, T) = \beta_T \left[ D \right]^{\gamma_T} \left[ T^{(\lambda_T)} \right],
\]

where the Distance Attitude (DA) power parameter is readily interpreted: convexity \((\gamma < 1)\) contracts objective distance, indicating optimism; concavity \((\gamma > 1)\) amplifies it, indicating pessimism; and a neutral pivot \((\gamma = 1)\) designates “untwisted” neutrality. Estimates found in table 1 all show optimism, including the most recent ones based on (3-A) to explore a huge sample of intercity trips in France (L’Hour et alii, 2013).

**The changed meaning of rate-specified RUF**

If RUF are of exactly logarithmic form, the specification \( V_i = f_i(\text{Time}, \text{Cost}, \text{etc.}) \) is indistinguishable from the alternate \( V_i = f_i(\text{Price}, \text{Speed}, \text{Distance}, \text{etc.}) \); but, as this form restriction almost never holds in reality, the latter perfectly distinct and identified variant is sometimes preferred to the former because of reduced collinearity and better fit.

But rate-specified RUF that include distance from the start can also be reinterpreted under DA because the new term (4-C) need not appear as an interaction with LOS variables but can stand alone:

\[^9\text{We therefore excluded from the “Distance Attitude” formulation both LOS uncertainty and any potential endogeneity resulting, at least in aggregate models, from equilibration between demand and supply, but we did not exclude the possibility that distance be observed with error, notably of the Berksonian kind (Berkson, 1950) where the true value } X^{\text{true}} \text{ is distributed around observed values } X^{\text{obs}}, \text{ so that } (X^{\text{true}} = X^{\text{obs}} + u). \text{ In the classical case, the measured value is distributed around the true value and the observed value is equal to the true value plus an error } (X^{\text{obs}} = X^{\text{true}} + u).\]
the self-standing distance term will then fulfill two roles, that of pure impedance and that of the DA, but some effort will be required to distinguish them.

If M mode-specific pairs of parameters by alternative \((\beta_{i,d}, \lambda_{i,d})\) for all \(i\) are estimated, they will reflect both DA and pure impedance roles and effects. To remove the DA effect and be left only with the impedance component, it is necessary to assume that the impedance role is common to all alternatives. This implies the estimation of a generic effect with a reference variable \(D_r\) common to all modes; but, as only M-1 of the coefficients of any variable common to all alternatives are then identified, estimation of M generic coefficients and powers \((\beta_{i}, \lambda_{i})\) requires roundabout methods.

Existing models with M-1 pairs of generic or specific parameters therefore provide an implicit and constrained measurement of both effects. For instance, the very strongly negative generic coefficient estimates of \((\beta_{i} - \beta_{j})D^{(\lambda_{j})}\) obtained for the classic and container rail modes (as opposed to the trucking mode) in Gaudry et alii (2008) combine pure impedance and DA effects that reveal strongly negative attitudes towards classical and container rail freight distances, relative to road freight distance, in France.

**Conclusion**

If the “Distance Attitude” (DA) interpretation makes good sense of the new functional roles of distance that are progressively emerging, and if distance-enriched practice continues to receive successful confirmation in empirical work, the question of distance as a stand-in for missing variables may need to be explored further and the dual roles of distance, built into time or cost factors but also used alongside these LOS variables, accepted in both RUM and other types of transport models.

**References**


\[\text{For an example of such methods based on Inverse-Power Transformation capture, see Gaudry & Tran (2012).}\]