The Problem of Congestion Unreliability Cost: The Certain Error of a Mean and the Measure of Travel Time Utility Gains from Public Works

Marc J.I. Gaudry
Université de Montréal

This summary-for-students presents the claim of some recent analyses of home-based daily urban trips made in France or in the USA that the first four moments of travel time, not just the first, matter and that travelers react to changed service on road or transit networks by adjusting all four of their constructed trip duration moments, including the three higher moments characterizing time unreliability. Sole reliance on mean time changes to evaluate improved or worsened trip conditions then short-changes the benefit-cost analysis and notably fails to explain fast fill-ups and super-peaks that are part of the demanded trip time profile.

Keywords: trip duration demand, determinate moments of transport service, determinate moments of trip time duration, expected transport service, precautionary time margin, mass transit, private car, rail transit line automation, transport service improvement effects, fast highway fill-up, transit super-peak, travel time reliability echo, paris, USA SMA, france

EXECUTIVE SUMMARY

This introductory note for senior students summarizes a stream of recent microeconomic analyses of surveyed home-based daily urban trips made in France or in the United States of America.

A Foursome Parcel

It is claimed in these reports that: (i), in transport demand models, emphasis on the first moment of travel time as an explanatory variable, to the exclusion of its next three higher moments, fails to capture the true nature of travel time and its full disutility, which also depends on these ignored higher moments characterizing the unreliability of travel time; (ii) the first four moments of travel time, the only moments of a random variable to have proper names in our daily languages, should be used to specify and define endured transport service S and actual trip duration T.

A New Equation Explaining Constructed Journey Durations

First, we show that travelers plan their intended trip durations T^*, identifiable as observed T plus an outcome error, by combining an expected value of endured transport service S^p with a constructed safety margin I^* offsetting the uncertainty of that transport service. There must then exist, complementing the structural trip demand equation, a tandem demand equation for constructed durations proper whereby observed duration T is explainable by endured transport service itself as well as other variables. In it, the

issue of the existence and nature of the safety margin turns on the elasticity of duration T with respect to perceived endured service S^p: the existence of I^* is established by an elasticity estimate different from unity and its role as offset to the uncertainty of transport service by a value of that estimate smaller than unity.

Relative Moment Valuations and Absolute Moment Demands

That elasticity value secured, one extracts from the flexibly fitted trip duration demand equation (formulated in tandem with a trip frequency equation) the relative and absolute effects of explanatory variables on duration moments. The former, the empirical marginal rates of substitution among the moments, are assumed to reveal relative moment utility valuations; the latter, the absolute impacts of service on duration moments, are shown to modify the time profile distribution of trip demand and to close the demand model. Notably, elasticities of duration moments with respect to mean service jointly imply that service betterment contracts the time profile of demand and leads to higher utility choices due not only to shorter mean durations but also to net changes in the upper three unreliability moments of duration. We carry out such duration moment analyses by trip purpose (work and shopping), mode (mass public transit and car), region (Paris, large US Standard Metropolitan Areas and Metropolitan France) as well as sex, and to some extent by socio-cultural group. Key findings of these analyses can be summarized under two headings: (A) relative moment valuations; (B) absolute changes in moment demands. Consider them in turn.

Bettered or New Infrastructure and its Fast Fill-Ups

Invariant across explanatory variables, as in Poisson models, the marginal rates of substitution among trip duration moments vary, contrary to the constant ones of Poisson models (with unit coefficients of variation, for instance). Actual estimates reveal risk aversion in all markets and show that: (a) women tend to value the second moment more, and the third less, than men; (b) for work trips, the second moment is more valued in large US Standard Metropolitan Areas and in Metropolitan France than in the Greater Paris Region, a trade-off found to be socio-culturally sensitive and to imply different local "fast fill-up" effects for improved infrastructure; (c) trips to French shopping centers reveal lower valuations of reliability moments than do work trips.

Bettered or New Infrastructure and its Traffic Super-Peaks

With respect to absolute effects of mean travel time service $S^p \equiv \overline{t_e}$ on traveler duration constructs T, elasticities clearly imply that, when average service on the network improves (or worsens), (i) not only does mean trip duration fall (or rise) significantly less than proportionately (ii) but the three upper unreliability moments also adjust, each more or less, and jointly create a complementary net "echo" that amplifies the benefits (or costs) of changes in mean chosen duration. This amplified "error of the mean" notably implies higher (or lower) super-peaks, the 4th moment of duration acting as a key component of the net echo reaction.

Fixed-Form Models Found Wanting

Maximum Likelihood parameter estimates are obtained from a Box-Cox model enriched to calculate adjustments (model fits) of moments higher than the usual first, $E(y_t)$, but otherwise unchanged. Popular Linear and Log-Log cases are rejected as inferior to flexible Box-Cox estimates, infinitely so for the former and uncontrovertibly so for the latter.

INTRODUCTION: IS CONGESTION REALLY JUST MEAN TRAVEL TIME?

The Unease of Feeling Cheated

There are strong reasons to think that urban transport slowness, defined primarily as long average travel time, is in fact much more disliked than "standard" transport Quantity & Mean Travel Time

(Q&MTT) model estimates would let on. It is not just that rider altercations aboard transit vehicles and driver road rage outbursts cannot be meaningfully linked to modifications of average suffered service run time $S^p \equiv \overline{t_e}$, but that Q&MTT specification emphasis on the first moment of travel time t_e to the exclusion of higher moments clearly fails to capture the true nature and full disutility of travel time, an economic modeling shortfall yielding structurally biased net measures of total time disutility, or generalized time cost, of dully hated urban travel time.

Average Travel Time Blinkers

What do models miss? Car riders do dislike crowded roads with low speeds (*i.e.* congested, technically speaking) and mass transit users do shun long in-vehicle times (*i.e.* slow service −as distinct from infrequent or uncomfortably jammed vehicles −), but their daily grumblings imply stronger repulsion than average speed effects in Q&MTT models can account for. This seeming paradox can be solved if *first-moment-of-travel-time-blinkers* cause the unease by ignoring (give or take feeble glimpses of wider moment horizons) the unreliability of travel time, a critical part of endured transport service characterized by the next three moments of travel duration.

Two Foursomes

Net responses of this n-tuple of unreliability-defining moments to modified mean service $S^p \equiv \overline{t_e}$ may in fact be as important as the response of mean duration itself. Blinker removal therefore matters to all urban time empirical results. Since Dupuit's (1844) bridge with an implicitly constant crossing time, service **S** defined as mean time t_e has been added to price in lists of demand function regressors but its standard error is still not a regular attribute of service **S**, itself still formally devoid of unambiguous skewness or kurtosis dimensions. We criticize wishy-washy two-moment service specifications and propose to enrich both mean service and mean trip duration times by their three higher unreliability moments. These two foursomes should both be shown to matter, one day.

Starting with Trip Duration Analysis

But first steps first. We summarize here a first effort to study the effects of changes in mean road or mass transit transport service, dubbed endured perceived service $S^p \equiv \overline{t_e}$, on the level and concatenation of user-constructed trip duration **T** moments. The more complete original arguments and demonstrations were developed during the last three years to evaluate the user benefits of the Grand Paris Express (GPE) automated metro. Under construction since June 2016, it will double the current length of Paris metro lines, from 200km to 400km, all the while adding some 68 new stations or interchanges with current metro and/or regional rail lines.

Applications to France and to the USA

On this topic, the core arguments were developed in three series of source documents, all listed in a special list of references (*cf.* Section 9), about home-based daily (quotidian) trips. The first batch (Gaudry 2015, 2016a, 2016b, 2016c, 2018a) recalls previous uses of the methodology for the analysis of constructed moments of road accidents and applies it anew to work trip constructs by car and transit in the Greater Paris Region (Île-de-France, with 12 million inhabitants).

The second effort formalizes the trip duration elasticity test with respect to **S^p** previously applied and carries out new applications to car trips for the same work purpose in Metropolitan France and in US Standard Metropolitan Areas (SMA) of more than 3 million inhabitants (Gaudremeau & Gaudry, 2017a). The third wave notably adds analyses of car trips for non-work purposes in Metropolitan France (Gaudremeau & Gaudry, 2017b) and provides summaries of the work stream since 2015.

An Introduction Aimed at Senior Undergraduates

This introductory paper, with slight modifications to the manuscript (Gaudry, 2018c) and taking after a summary paper in French (Gaudry, 2018b), concentrates as much technical material as possible in appendices, but without yielding an inch on the substance, namely a theoretical and empirical rejection, as inconsistent with the data, of the triplet of inherited *penchants* of those who are (a) firmly against the existence of safety margins; (b) wobbly against use of moments of travel time higher than the second; (c) fixed-form lazy or obtuse on the estimation of flexible mathematical regression forms.

Genesis of the Foursome Approach

We first recall how making sense of the success of an automated Paris metro line required adopting new modeling postures to account for the key role of regularity. We then explain the economic and econometric approaches taken to explain the asymmetric trip duration data of interest and finally summarize key results of the source papers from which we copy and paste extensively without shame or quotation marks.

AUTOMATED METRO LINE SUCCESS AND THE ISSUE OF TRAVELER PREFERENCES

Explaining the Success of Paris Metro Line M14 Running Automated Trains
The Abraham-McFadden RUM Approach and Terminology

Since Abraham (1961) initiated, and McFadden (CRA, 1972) expanded[1] the random utility Maximization (RUM) discrete choice approach, Logit attractiveness functions V_i in (1-A) have increasingly been called "utility functions", a terminology also proposed by Rassam *et al.* (1970, 1971) in aggregate mode share applications.

RATP, the main operator of bus, metro and regional train lines in the Île-de-France (Paris) region, which has developed such Logit mode choice models containing in their utility functions level-of-service (LOS) variables for public transit (PT) and private car (PC) networks since 1982, recently decided to secure LOS variables from the origin-destination (O-D) paths of Logit itinerary choice models.
A RUM Path Choice Model

For PT, its chosen RUM discrete multinomial Logit (MNL), considered below as typical of "demand" models, has path choice probabilities p_i for individual t given by:

$$p_t(i) = \frac{\exp(V_{it})}{\sum_{j=1}^{j=M} \exp(V_{jt})}, \quad i, j = 1, ..., M; \ t=1, ..., T, \tag{1-A}$$

where the V_i utility functions[2] of the itineraries, or "paths", may be written for individual t as:

$$V_{it} = \beta_{i0} + \sum_n \beta_{in}(X_{nt}^i)^{(\lambda_{in}^i)} + \sum_s \beta_{is}(X_{st})^{(\lambda_{is})}, \tag{1-B}$$

with a distinction made, among the explanatory variables X_k, between attributes X_n^i that vary across itineraries (*e.g.* network characteristics) and factors X_s, associated to individuals, that are common to all *M* path options (*e.g.* age, sex, or income). This formulation is referred to as a "Box-Tidwell Logit" because only it's strictly positive and continuous X_k are subjected to the Box-Cox transformation (BCT) operator, used here without Tukey's μ_v shift found in the original Box & Cox (1964) paper, namely:

$$X_{vt}^{(\lambda_v)} \equiv \begin{cases} \left[(X_{vt})^{\lambda_v} - 1\right]/\lambda_v & , \quad \lambda \neq 0, \\ \ln(X_{vt}) & , \quad \lambda \to 0, \end{cases} \tag{2}$$

Yielding a Residual Surprise

RATP modelers used this data-determined RUM form estimation with sophisticated utility functions specifications comprising crowding comfort-weighted in-vehicle time (based on passenger density), walk and wait times, transfer penalties, *etc*. (Leblond & Langlois, 2013). Their model errors, defined by differences between actual and assigned flows based on model parameters, surprisingly implied large underestimates of ridership on automated line M14. Contrary to classical lines, where link flow errors were small, ridership on M14 was underestimated by some 20% during off-peak, and 25% during peak, service periods (Prat & Leblond, 2014). An unexpected result: in particular, how could one explain that, at peak times, travelers would much rather use a crowded automated line than equally crowded man-driven ones? Might train automation itself explain this result?

The Impact of Automation on Four Explicit Moments of In-vehicle Travel Time
*A Very Discrete Nanny*

Unfortunately, the scientific transport literature totally ignores the role and value of train service automation. The first automated line, London's Victoria Line, was automated twice: at birth in 1968 and again in 2013. Much mileage was made about the impact of its opening on (average) time savings (Foster & Beesley, 1963) and about their valuation in terms of the hourly wage rate (Beesley, 1965 or 1970), but we could find no discussion of the impact of its automation as such, even when the recent second modernization increased train frequency per hour and direction from 27 to 33 trains!

A Foursome of On-Board Time Dimensions

One way to understand the M14 line error results is to look at the impact of metro line automation on the in-vehicle (on-board) time distribution between two representative stations A and B on a recently automated Paris metro line Z, shown in Figure 1. In it, transport service, defined solely by 4 measures (or notions) applied to in-vehicle time (wait-time is excluded from the calculation), has here no existence independently from that of these 4 moments.

FIGURE 1
MODERNIZATION OF RATP LINE Z OF THE PARIS METRO:
THE FIRST 4 MOMENTS OF IN-VEHICLE TIME
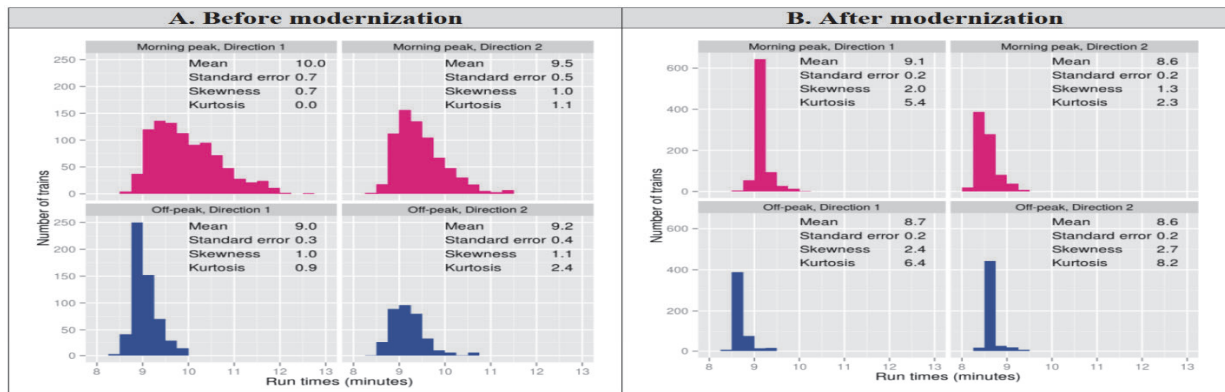
Analysis shows that both changes (from Peak to Off-peak before modernization, and from any time period and direction before modernization to any afterwards) have the following properties: (a) higher speed lowers mean in-vehicle time and its standard error; (b) its skewness increases, but less in the first set of changes (Peak/Off-peak) than in the second set; (c) but percentage increases in skewness are small relative to those of kurtosis, which can be huge, as estimated easily by regrouping the data:

Before (Z) and after (Z') modernization		Direction	Z	Z'	Direction	Z	Z'
Morning peak	Mean	1	10,0	9,1	2	9,5	8,6
	Standard error		0,7	0,2		0,5	0,2
	Skewness		0,7	2,0		1,0	1,3
	Kurtosis		0,0	5,4		1,1	2,3
Off-peak	Mean	1	9,0	8,7	2	9,2	8,6
	Standard error		0,3	0,2		0,4	0,2
	Skewness		1,0	2,4		1,1	2,7
	Kurtosis		0,9	6,4		2,4	8,2

Is there an intuitive view of these 4 moments of a random variable? The first moment, the mean, stating location, is associated in common language with fatality and provides the anchor of the next three moments, jointly defining service reliability here, characterized by successively higher powers of distances between individual observations and the mean (*cf.* Appendix 1), each result being scaled, or "normalized", by a denominator. The second moment, the standard error, a symmetric measure with similar units as the first, is associated with risk in common parlance and the third, skewness, which is without units and denotes the direction of the longest tail of the distribution, with prudence. The fourth, kurtosis, also without units, is the least intuitive because, associated with temperance in daily language, it measures the lack of extreme values, or of acuity, sometimes called "*peakedness*" (*e.g.* Kane, 1968), of the distribution. We stop at the fourth moment because higher ones, in the case of the strictly positive (or of the absolute) value of a normally distributed variable **y** of interest here, are in fact indeterminate (Berg, 1988)[3] and, perhaps for that reason, without known proper names in European languages.

Moment Performance Equilibria and the Metro as Metronome

The moment values shown in every one of the 8 Parts of Figure 1 are equilibrium performance outcomes resulting from the interaction of train supply with passenger demand flows. And, clearly, the most striking impact of train automation is its dramatic impact on distribution acuity: the metro has become a metronome. It is this neglected metronomic regularity effect that no doubt primarily explains the "overuse" by 20-25% noted above despite the small increase in skewness and the small decrease in standard error (also ignored in the RATP model, which accounts for wait time and mean speed) also due to automation. Crucially, these first four moments of on-board time must also be relevant for the appreciation of road service.

Making Sense of Four Moments: Multiple-Moment Dependent Utility (MMDU)

Consumer Preferences for the Four Moments

This led us to the view that all first four moments of transport service matter and that the 4th, regularity, may play a role relatively as important in transport time utility as does the third, characterizing jackpots, in national lotteries. But this idea that all determinate moments of the distribution as a whole matter to utility is not new: it has been stated forcefully time and again by Allais (*e.g.* 1953, 1987). It means that, if one for instance makes a financial investment, one will have an interest at least in: (i) its average yield or return $e(y)$; (ii) the standard error of this return $\sigma(y)$; (iii) the asymmetry of the return $\gamma(y)$, or relative chances of yield increases (*upside risk*) or decreases (*downside risk*); (iv) the acuity (or kurtosis) of the return $\kappa(y)$

A New Behavioral Modeling Posture: Multiple-Moment Dependent Utility (MMDU)

We therefore assume that consumer preferences for distributions are moment dependent, *i.e.* entirely and solely characterized by their empirical moments, and that all first four moments of the travel time "lottery" matter to utility, and notably the fourth (regularity) as just noted. How? If people are known to correctly identify, even on the basis of few observations, at least distributions as varied of those by Gauss, Poisson and Erlang, as well as various power functions (Griffiths & Tenenbaum, 2006), they will *a fortiori* identify mere sequences of moments, and do so more easily than they would identify (now superfluous) features of the distributions such as their exact form, perhaps not even unique (Gut, 2002). But this Multiple-Moment Dependent Utility (MMDU) view, more than good old bi-moment ($e;\sigma$) utility, needs to be completed by linking individual moments to the direction of preference in utility.

Moments of Travel Time and the Direction of Preference

Whether the prospect is a «good» like financial return or a «bad» like transport time or accident, naturally reverses the sign of its marginal effect on utility but not the sign of the MRS between any two such effects because we adopt the rule that signs of $\partial U/\partial(moment\ i)$ alternate (Scott & Horvath, 1980),

as summarized in Table 1 where signs of the first two moments are considered obvious and those of higher orders as reasonable.

TABLE 1
PRESUMED DIRECTION OF PREFERENCE (EFFECTS ON UTILITY)
OF INCREASES IN MOMENTS OF y

Moment	1st	2nd	3rd	4th
Prospect: good	+	-	+	-
Prospect: bad	-	+	-	+

As a consequence for an investment, increased yield is pleasant (+) but not increased variability (-), whereas more asymmetry (less negative on the left or more positive on the right, depending on its nature) is also pleasant (+). Analogously, and *mutatis mutandis*, the occurrence of longer transport times or more road crashes is unpleasant (-) and one wishes that the standard variability of this outcome increased (+), but not its asymmetry (by becoming less negative or more positive) because higher asymmetry (in absolute value) of a bad is unpleasant (-). The signs of kurtosis obey the alternance rule.

"Theory is Good, but it does not Prevent from Existing" (Charcot)[4]
And we also assume the direction of preference for that 4th moment, despite again the difficulty of proof arising from the logical impossibility to calculate the partial derivative of a utility function with respect to, say, the third central or fourth moment without affecting lower order ones (Brockett & Garven, 1998). Recognizing their point that moment ordering is a necessary but not sufficient condition for an ordering of utility (stochastic dominance), moment trade-offs estimates to-day must then come from actual behavior, without waiting for proof of their theoretical standing, lest we be compelled to silence on the structuring of national lotteries or on the needed valuation of 4-moment transport service. In all cases, one is interested in the weight put by consumers on each moment, and in the quantities of each they are willing to trade for a unit of another. For this, we need sharpened utility-linkage assumptions.

Attitudes to Risk, Directions of Preference and their Ratios
Notably, do *"riscophobes"* and *"riscophiles"* differ on weight signs? Limiting the field for the present to the first three moments of a variable of interest, we estimate below marginal rates of substitution (MRS), or trade-offs -[∂ (moment i)/∂ (moment j)] and their signs, shown in Table 2. The preferences of a risk-averse investor assumed on the first line of Table 1 imply in Table 2 the sign pattern of marginal rate of substitution ratios of matrix [A]; the contrasting sign pattern of matrix [B] reveals a risk-seeker. Both matrix sets are mathematically coherent and admissible (*cf.* Tran *et al.*, 2008, Table 6, p. 36).

TABLE 2
RISK-AVERSION, RISK-SEEKING, AND EXPECTED SIGNS OF THE MRS
AMONG THE FIRST 3 MOMENTS

	Two sets of marginal rates of substitution among moments: -[∂ (moment i)/∂ (moment j)]							
	[A]. *Riscophobe*			**[B].** *Riscophile*				
i\j	e	σ	γ	e	σ	γ		
e	1	+	–	1	–	+		assumed MRS signs
σ		1	–		1	–		derived MRS signs
γ			1			1		

In Table 2, risk aversion and risk seeking are defined on line *e* by assuming the existence of certain signs for the MRS between the first and higher moments; on line σ the MRS between the second and third moments are then derived from those of the first line, which explains their common negative value in matrices [A] and [B]. The matrices are in effect based on usual preference maps and ignore trade-offs equal to 1 or 0 arising from the very unusual horizontal indifference curves in Tobin's (1965) risk neutral formulation. Table 2 then states our key sharpened MMDU behavioral hypotheses.

Searching for Multiple-Moment Transport Service Trade-Off Valuations
Granted, the literature on automation does not contribute to our search for estimates of the distinct weights of the relevant 4 trip time moments presented in Figure 1. But then, what about the general literature on trip demand? Limiting ourselves to formulations using distinct moments with own weights, we presently see that it is weak in multiple-moment specifications of service by mode and very shy in taking into account the uncertainty of that service by a precautionary time margin, let alone an offsetting one...

Specification of Suffered Transport Service and of its Moments in Demand Models
From the First to the Second Moment of Endured Transport Service
Basic demand theory only uses cost and time features of modes or paths in their utility functions and, for scores of years, only their mean values mattered. Neglecting cost, the seminal RUM Abraham-McFadden discrete choice approach referred to above and independent Logit mode choice practice (Warner, 1962) were all concerned only with mean time, as were early utility-linked aggregate Logit formulations applied to Channel road path choice (Setec *et al.*, 1959) or to mode choice (Rassam *et al.*, *op. cit.*).
This changed with Jackson & Jucker (1982) who considered the 2nd moment explicitly and for whom S_n^p, endured transport service by mode **n**, is characterized by a weighted combination, in effect a vertically concatenated sum, of the mean *e* and the standard error σ of experienced travel time:

$$S_n^p = r_1 \overline{t}_{e(n)} + r_2 \overline{t}_{\sigma(n)}, \text{ } r_1 \text{ and } r_2 \text{ denoting weights,} \tag{3-A}$$

where (our) added bars indicate that, for the user, moments are given and endured – not chosen. As already mentioned, we neglect other approaches where the service variable construct S_n^p does not maintain distinctly weighted mean and variance (or standard error) terms, for instance in Senna (1994).

More on the Sign of the Variance Term
Jackson & Jucker (*op.cit.*) had stated that the sign of r_2 was positive (as in the second line of Table 1) for risk-averse individuals but not for risk-lovers, a distinction often forgotten in practice, as in the Seshadri & Srinivasan (2017) Logit route choice model. Obtaining negative weights for both r_1 and r_2 (and a barely significant r_2), they claim this result is "logical" and are closed to its implication that the 602 surveyed Chennai City road users are risk-lovers. With much more credible specifications and a sample of 13 723 trips, Prato *et al.* (2014, Table 2) find instead, for peak and off-peak periods, the three expected pairs of significant negative r_1 and positive r_2 coefficients for their variables: (i) free flow time, (ii) congested time, (iii) time "reliability" defined as the difference (P90-P50) between percentiles of the travel time distribution.

The Lurking Third Moment
We are still some distance away from using the third moment, as in

$$S_n^p = r_1 \overline{t}_{e(n)} + r_2 \overline{t}_{\sigma(n)} + r_3 \overline{t}_{\gamma(n)}, \text{ with valuation weights } r \equiv (r_1, r_2, r_3) \tag{3-B}$$

But perhaps not as much as it might seem because the third moment, of the same sign as the first on line 2 of Table 1 for risk-averse individuals dominating typical urban markets, as in our results below, has a way of barging in surreptitiously, for instance in the result just noted for (P90-P50).

The modeling posture context is one where trade-offs between the first two moments are long known in finance (Tobin, 1957/1958; 1965) and estimation of marginal rates of substitution between them, as joint determinants of the demand for asset y, $MRS_{X_k, X_\ell} = \partial y / \partial X_k / \partial y / \partial X_\ell = \partial X_\ell / \partial X_k$, is well accepted. But higher moments are ignored because of the Brockett-Garven point, noted above, of the logical impossibility of their partial derivatives of a utility function containing all 4 first moments. Transport path choice authors therefore state dogmatically *ab initio* that only mean and variance (presumably with opposite signs…) determine utility but, in tests, rapidly abandon that untenable assumption, notably by replacing mean and variance determinants by the median and a high percentile point PX0 of the time distribution (now with the same signs…). In this hocus-pocus game of double substitutions, high (PX0) quantiles arguably stand for third moments, with the same sign as the mean.

Such slippages from the first two clean analytical moments to floating PX0 quantiles, all exact multiples of 5 at that, can be found in Logit models of choice between free and tolled California State Road 91 lane options. With stated preference data on lane choices, Lam & Small (2001) finally retain the median (P50) and the 90th (P90) percentiles of the time distribution; but with revealed preference data on the same SR91 lane case, Small *et al*. (2005) end up with P50 and P80-P50. The clean mean-variance model has been transformed into a mean/median *vs* variance/skewness *smörgåsbord* of sorts…

Indeed, these manually chosen percentile point replacements "yield a better fit" than the mean and standard error pair; and the estimated signs of, say P50 and P90, coefficients match in practice, implying that latent first and third moment effects dominate those of the variance in line 2 of Table 1. Of course, such descriptive *ad hoc* P80 or P90 terms, silent about the form of the distribution, only "work" because actual SR91 lane use durations have positively skewed distributions. Making sense of casual PX0 results always requires appeal to a missing key latent third moment with the same sign on utility as the first but with an import greater than that of the replaced original variance and its sign.

Currently Available Data on Suffered Transport Network Service Moments
In our empirical part below, we will only use as an explanatory variable of trip duration demand the first moment of service $S^p \equiv \overline{t_e}$ for each of the PT or PC trips considered: the construction of multiple-moment service indicators like (3-B) is under consideration for future research, starting with PT lines. In these conditions, we rely instead on the dependent variable to estimate weights and marginal rates of substitution among relevant moments of service featured in Figure 1. They will be extracted from the very structure of trip durations, an idea now requiring some contextualization from the literature.

TRIP DURATION CONSTRUCT REVELATION OF TIME MOMENT PREFERENCES

Sources of Revealed Moment Preference Extractions
Preferences with respect to moments of travel time need not be extracted from endured service n-tuples like (3-B), concatenations rarely found among explanatory variables of modal choice models like (1) which tend to ignore the three higher moments of service time jointly characterizing its reliability. We propose their extraction instead from moments of T_n, observed trip duration by mode **n**, defined by the difference between given rendezvous H_{rdv} and chosen departure H_{dep} times. This requires modeling journey duration T_n.

Trip Duration as Comprising a Precautionary Offset to its Suffered Service Base
To model duration itself, we adopt an innovative posture. First, travelers are assumed to have preferences for the first four empirical moments of travel time, which they can weigh, compare and value

directly without any resort to analytical distributions of that travel time. Second, to plan their trip times, they evaluate endured network service moments and offset the uncertainty of their own perceived valuation of that suffered service base by a palliative safety margin construct congruent to the nature of the trip.

Precautionary Palliative Margin Existence and Structure

The relative importance of this margin in total planned durations, consisting in expected service and safety margin components, and the internal structure of such durations, notably vary with journey purpose and other factors. Its existence and palliative role need to be established by an elasticity test which opens the door to subsequent analyses of its moment structure. Unfamiliar steps are per force involved in this innovative process.

The Structure of Planned Trip Time and the New Equation Required to Explain it

Trip Durations as Endogenous Traveler Constructs

Contrary to the moments of service S^p, moments of intended (unobservable) duration T_n^*, or of (observed) duration $T_n \equiv (H_{rdv} - H_{dep}) + u_{T_n}$, may be said to consist in traveler constructs that can reveal Allais' valuations of due time moments. Why? We assume, and later test, the view that trip durations do not merely consist in an expectation of service S^p but also include a precautionary time margin *I** taken to offset the uncertainty of the expected transport service. The result, conceivable as a sum, or as a more complex interplay of those terms, is a completely endogenous multi-moment cocktail even if $T^* \equiv f(I^*, S^p)$ contains suffered service with moments that are exogenous to any single traveler. Writing out $T^* \equiv T + u_{T_n^*}$, we have:

$$\left[r_1 \hat{e}(\hat{T}_n) + r_2 \hat{\sigma}(\hat{T}_n) + r_3 \hat{\gamma}(\hat{T}_n) \right] + u_{T_n^*} \equiv f \left[(r_1 \hat{t}_{e(n}^* + r_2 \hat{t}_{\sigma(n)}^* + r_3 \hat{t}_{\gamma(n)}^*), (r_1 \overline{t}_{e(n)} + r_2 \overline{t}_{\sigma(n)} + r_3 \overline{t}_{\gamma(n)}) \right] \qquad (4\text{-}A)$$

where (i) endogenous values (with circumflex accents) contrast with endured exogenous ones (with bars), (ii) stars denote intentions and (iii) the error term $u_{T_n^*}$ is the outcome error arising from the fact that observed durations T_n (to be explained)[5] differ from unobservable intended durations T_n^*. The right-hand side of (4-A) is inspired by the distinction (Knight, 1921) between a calculable probabilistic "risk" (here S^p) and its "globally perceived uncertainty" (countered here by *I**): it combines the risk associated to $S_n^p = (r_1 \overline{t}_{e(n)} + r_2 \overline{t}_{\sigma(n)} + r_3 \overline{t}_{\gamma(n)})$, endured service perceived, to the reaction margin intended $I^* = (r_1 \hat{t}_{e(n}^* + r_2 \hat{t}_{\sigma(n)}^* + r_3 \hat{t}_{\gamma(n)}^*)$, a countermeasure offsetting the uncertainty of that service.

Both *I** and *T* are user-constructed mental moment concatenations. Concatenations, long familiar in the joint demand formulations of Demsetz (1970) and Samuelson (1969), here consist in weighted vertical sums of time moment demand schedules (*cf.* Appendix 2). Moments of time have replaced these authors' jointly supplied goods: for the former, hides and meat provided in fixed ratios by steers; for the latter (*cf.* also Samuelson, 1954, 1955), individual schedules of willingness to pay for public goods. In all cases, horizontal summation of individuals' schedules yields market demands.

The Dearth of Margin Variables in Transport Demand Models

The idea that travelers plan their travel time, reacting to the uncertainty of an expected PT or PC service by adding to their expectation of that service S^p an intended precautionary time margin *I**, is old: Polak (1987) might have stated it, according to Senna (*op. cit.*). But this would, if true, be in any case long after Gaver (1968), Starkie (1971) and Knight (1974) had, to little avail, called for use of margins under various garbs (to wit: "head-start", "idle time", "safety margin"), the latter pair of authors even

adding to the pure unreliability of service reaction a so-called "schedule delay" adjustment to cater to low frequency PT line schedule-rendezvous matching quandaries. The idea always implies that, in modal demand or choice equations such as (1), the resulting margin appear along the endured service regressor variable. If it can be identified, a margin might perhaps be added to service, as often done in models explaining for instance trips to the airport of the rail station with a reasonable (but exogenous) precautionary term complementing other path or route characteristics. The most sophisticated of all multimodal Europe-wide intercity passenger demand models (Mandel, 2014) is a case in point.

The Tandem of Structural Trip Frequency and Trip Duration Demand Equations

What may however be a newer stand to most readers is the other idea of extracting relevant service moment valuations from the dependent variable (4-A) of a distinct equation explaining duration, even without estimating the tandem trip demand equation. To see this, rewrite the mode choice equation (1) more generally to allow for explanations of demand levels (and not just of choice probabilities) of journeys Q and add to the revised formulation an estimable demand equation for their duration D. A simplified two-level (Q_n)-(D_n) system for a unique mode n serving a single O-D transport link could be:

$$\boxed{Q_n} \qquad Dem_n = q\left[P_n; T_n^c \equiv (I_n^* + S_n^p); X_s; A\right] + u_{Q_n} \qquad (4\text{-B})$$

$$\boxed{D_n} \qquad \left[T_n^c \equiv (H_{rdv} - H_{dep})_n\right] = d\left[P_n; S_n^p; X_s\right] + u_{D_n} \qquad (4\text{-C})$$

where previously undefined terms of the simple equation tandem needed to make sense of empirical work below are: (i) P_n: price of the unique O-D transport mode n; (ii) X_s: socioeconomic traveler characteristics, such as sex, age, *etc.*; (iii) A: activities ($A_1, ..., A_f, ..., A_F$) at the origin O or destination D requiring O-D travel, such as work, shopping and personal trips; (iv) u_{D_n}: modal duration own-equation error notably including $u_{T_n^*}$ from (4-A) and an observed (surveyed) duration measurement error u_{H_n}.

This simplified tandem system of two demand equations per mode fully endorses the Starkie-Knight critique noted above on the misspecification of mode choice models – that they ignore the existence of I_n^* margins and treat them as implicitly cancelling out in mode choice models or as comprised in the β_{in} alternative-specific modal coefficients found in utility functions of type (1-B). In (4-B)-(4-C), the palliative nature of time precaution offsets makes them specific to each mode (or route) considered and can for instance be used to test the credible observation made by Hendrickson & Plank (1984) that departure time to work decisions are more flexible than mode choice (or frequency) decisions.

Duration vs Departure Time Modeling

The dependent variable in (4-C) differs from that of departure time models associated principally with Vickrey's (1969) formulation. Gaver (*op.cit.*) had made departure time depend on the asymmetry of costs associated with early, as opposed to late, arrival at destination and on the distribution of route delay. Vickrey replaced that distribution by an exogenously given preferred arrival time H_{pref}, a detour tantamount to positing this duration demand:

$$(H_{pref} - H_{dep})_n = d_{pref}\left[\bar{t}_{e(n)}, \breve{a}(\bar{t}_{e(n)}, H_{pref}), \breve{r}(\bar{t}_{e(n)}, H_{pref})\right] + u_{pref} \qquad (4\text{-D})$$

where the terms $\breve{a}(.)$ and $\breve{r}(.)$ are associated with early or later arrival costs and H_{pref} is an unobserved fixed (non stochastic) *deus ex machina* which displaces the departure time pickle to that of making sense of H_{pref}, itself now begging for an explanation as a new endogenous system variable. It requires a

supplementary equation, as in Koster *et al.* (2011) who then have to explain $H_{rdv} - H_{pref}$ as well. This costly dilatory detour is avoided with (4-C) where the unknown asymmetric costs of early and late arrival may remain implicit and are subsumed within moment level and valuation time constructs.

Duration Equation Statistics of Particular Interest for the MMDU Approach
Anticipated Output Requirements for the Future Regression Method

Consider in turn three statistics needed from regression estimates of (4-C) to test our approach: (i) elasticities of duration $\eta(y, X_k)$ and (iii) marginal rates of substitution MRS^{m_τ, m_s} among moments of duration, but also (ii) marginal rates of substitution MRS_{X_k, X_ℓ} among explanatory variables X_k and X_ℓ.

    i.   ***Elasticity of demand of one or many moments of duration***. We are interested in the effects of any explanatory variable X_k on duration $y \equiv d$ and on its moments, effects measured either absolutely:

$$\partial \hat{d} / \partial X_k \, ; \, \partial \hat{d}_{e(n)} / \partial X_k \, ; \, \partial \hat{d}_{\sigma(n)} / \partial X_k \, ; \, \partial \hat{d}_{\gamma(n)} / \partial X_k \tag{5-A}$$

or as elasticities, listed in Table 3 inspired by a survey of elasticity notions (Dagenais *et al.*, 1987):

TABLE 3
FOUR NOTIONS OF THE POINT ELASTICITY η OF y
WITH RESPECT TO A CONTINUOUS VARIABLE X_k

	Name of elasticity concept with respect to X_k	Point elasticity measure* for continuous variable X_k
1.	Sample elasticity of y	$\eta^s(y, X_k) = \dfrac{\partial y}{\partial X_k} \cdot \dfrac{X_k}{y}$
2.	Elasticity of *e*(y)	$\eta^e(y, X_k) = \dfrac{\partial e(y)}{\partial X_k} \dfrac{X_k}{e(y)}$
3.	Elasticity of σ(y)	$\eta^\sigma(y, X_k) = \dfrac{\partial \sigma(y)}{\partial X_k} \dfrac{X_k}{\sigma(y)}$
4.	Elasticity of γ(y)	$\eta^\gamma(y, X_k) = \dfrac{\partial \gamma(y)}{\partial X_k} \dfrac{X_k}{\gamma(y)}$
(*) where simplified writing ignores indications that all measures have to be evaluated at some sample point *t*. For comparable expressions used in source papers if the variable is not continuous but a dummy, see App. 3.		

Elasticity notions differ in Table 3. The "sample" elasticity ignores that the relationship explaining *y* contains a random residual but other notions are "strict" and recognize the stochastic nature of *y*.

The sample notion arose in France in railway courses at the *École Nationale des Ponts et Chaussées* (ENPC) where it casually[6] appears in Sévène (1877) or Tarvernier (1889); a former Austrian student even uses its mathematical formulation in a study of rail costs (Nördling, 1886).

By contrast, all strict notions are rarely used, except for the second in nonlinear regression models, despite Goldberger's (1968) known proof that elasticities of the sample, the expected value and the median of *y* are identical in a logarithmic model, yielding for each and for any observations *t*:

$$\left[\eta^s(y_t, X_{kt}) = \frac{\partial y_t}{\partial X_{kt}} \frac{X_{kt}}{y_t}\right] = \left[\eta^e(y_t, X_{kt}) = \frac{\partial E(y_t)}{\partial X_{kt}} \frac{X_{kt}}{E(y_t)}\right] = \left[\eta^m(y_t, X_{kt}) = \frac{\partial m(y_t)}{\partial X_{kt}} \frac{X_{kt}}{m(y_t)}\right] = \beta_k \qquad \text{(5-B)}$$

In the case of our explanatory variable S_n^p, or of its components in (4-C), the effects are particularly interesting because:

a) values different from unity, *i.e.* $\eta(T_n^c, S_n^p) \neq 1$, demonstrate that duration $T_n^c = I_n^* + S_n^p$ cannot be reduced to expected service but includes a precautionary margin – otherwise the effects would be strictly proportional, *i.e.* $\eta(T_n^c, S_n^p) = 1$ – and the regression of y on S_n^p would not be legitimate, in effect regressing S_n^p on itself (and other X_k). Table 7 formalizes this test below;

b) the strict notions together account for the impact of transport service on the time profile of demand for duration, a feedback which, eventually completed by changes in the fourth moment, closes the Q-D tandem of traffic frequency and duration demand models (4-B)-(4-C);

ii. ***Said marginal rates of substitution between explanatory variables***. With respect to relative effects between explanatory variables, it is always relevant to compute the usual ratios of marginal effects $MRS_{X_k, X_\ell} = \partial y/\partial X_k \big/ \partial y/\partial X_\ell = \partial X_\ell / \partial X_k$ but our quantitative model must also allow for the ratios of effects of any X_k and X_ℓ pair (components of S_n^p included) with respect to moments of y, namely:

$$MRS_{X_k, X_\ell} = \frac{\partial E(y)/\partial X_k}{\partial E(y)/\partial X_\ell} = \frac{\partial \sigma(y)/\partial X_k}{\partial \sigma(y)/\partial X_\ell} = \frac{\partial \gamma(y)/\partial X_k}{\partial \gamma(y)/\partial X_\ell} = \frac{\partial y/\partial X_k}{\partial y/\partial X_\ell} = \frac{\partial X_\ell}{\partial X_k} \qquad \text{(5-C)}$$

which have the surprising property that the ratio $\partial X_\ell / \partial X_k$, for any pair of variables, is equal[7] for all moments of y and for the simple "sample" formula $\partial y/\partial X_k \big/ \partial y/\partial X_\ell$ that altogether ignores the error term distribution. This equality property, valid unless variables X_ℓ and X_k are also used to explain heteroskedasticity of residual u_{D_n} in (4-C), as noted in Tran *et al.* (*op.cit.*), provides below the basis of a statement concerning the robustness of our MRS estimates to potential collinearity;

iii. ***Said marginal rates of substitution among moments of duration***. After verification that $\eta(T_n^c, S_n^p) \neq 1$, one might be interested in extracting the MRS^{m_r, m_s} among the *moments* of duration T_n^c. Explicitly, with the first three moments considered, there will be only three MRS among them (their inverses being counted as identical) generated by the following expression which interestingly holds for all X_k even if, in (5-C), $[\partial(mom.\ i(y))/\partial X_k] \neq [\partial(mom.\ i(y))/\partial X_\ell]$:

$$\left[\partial(moment\ i(y))/\partial X_k\right] \big/ \left[\partial(moment\ j(y))/\partial X_k\right] = \partial(moment\ i(y))/\partial(moment\ j(y)), \forall X_k \qquad \text{(5-D)}$$

or, more explicitly, with the first MRS devoid of units but not the next two MRS:

$$MRS^{m_e,m_\sigma} = \frac{\partial e(y_t)}{\partial \sigma(y_t)} = \frac{\partial e(y_t)/\partial X_{kt}}{\partial \sigma(y_t)/\partial X_{kt}} = \frac{\partial e(y_t)/\partial X_{\ell t}}{\partial \sigma(y_t)/\partial X_{\ell t}} \tag{5-D.1}$$

$$MRS^{m_e,m_\gamma} = \frac{\partial e(y_t)}{\partial \gamma(y_t)} = \frac{\partial e(y_t)/\partial X_{kt}}{\partial \gamma(y_t)/\partial X_{kt}} = \frac{\partial e(y_t)/\partial X_{\ell t}}{\partial \gamma(y_t)/\partial X_{\ell t}} \tag{5-D.2}$$

$$MRS^{m_\sigma,m_\gamma} = \frac{\partial \sigma(y_t)}{\partial \gamma(y_t)} = \frac{\partial \sigma(y_t)/\partial X_{kt}}{\partial \gamma(y_t)/\partial X_{kt}} = \frac{\partial \sigma(y_t)/\partial X_{\ell t}}{\partial \gamma(y_t)/\partial X_{\ell t}} \tag{5-D.3}$$

From Mere Ratios of Partial Derivatives to Real Marginal Rates of Substitution

Formally speaking, substitution involves the sacrifice of a unit of X_1 traded against a unit of X_2 while maintaining the value of their combination constant. It is the slope of an isoquant, defined as **minus** the ratio of partial derivatives of the output or utility quantity held constant with respect to X_1 and X_2 (Allen, 1968, p. 42), as noted in the header of Table 2 and incorporated in its values.

Note that the above expressions (5-C) and (5-D), if used for such trades among moments of the dependent variable, neglect that required negative sign and keep the ratios of partial derivatives as they are in order to relate them with greater ease to moment utility impacts in Tables 1. But Table 2 takes the needed negative sign into account, as will future tables of results on MRS among moments.

Skipping the Toolbox

We now turn to the regression method adopted to secure all those results. Contrary to this section, which used only elementary calculus, the next one requires notions not all covered in first quantitative method courses. It may be skipped by those unfamiliar with Maximum likelihood and nonlinear regression, as long as they keep in mind that Box-Cox transformations defined in (2), with actual values determined by the data, are important because of their flexibility (and not only due to the inclusion of linear and logarithmic cases), which plays a key role in Section 5.

A PLURAL-FIT SKEWNESS-INCLUSIVE MOMENT OUTCOME REGRESSION

Rosett-Nelson, or Box-Cox, Regression Extended Beyond First Moment Fit

We adopt a Rosett-Nelson (1975) model with Gauss-distributed[8] $N(\sigma_w^2, I)$ errors w_t, collapsing to the Box-Cox model in the absence of limit observations on **y**. This familiar model can notably yield an asymmetrically fitted $E(y_t)$ that we will need because realized trip durations, as in Figures 1, 3 or 4, always obtain distributions that are skewed to the right, and are never symmetric, a seemingly universal phenomenon independent from mode, purpose, sex and socio-economic standing of the travelers:

$$y_t^{(\lambda_y)} = \beta_0 + \sum_{k=1}^{k=K} \beta_k X_{kt}^{(\lambda_k)} + w_t, \quad y_t > 0, X_{kt} \geq 0, t=1, ..., T, \tag{6-A}$$

with[9], as in (2)

$$X_{vt}^{(\lambda_v)} \equiv \begin{cases} \left[(X_{vt})^{\lambda_v} - 1\right]/\lambda_v &, \quad \lambda \neq 0, \\ \ln(X_{vt}) &, \quad \lambda \to 0, \end{cases} \tag{6-B}$$

and yielding fitted values of the first three moments of outcomes

$$E(y_t) = \varepsilon \int_{-\infty}^{w_t(\varepsilon)} \varphi(w)dw + \int_{w_t(\varepsilon)}^{w_t(v)} y_t\varphi(w)dw + v \int_{w_t(v)}^{\infty} \varphi(w)dw, \quad (\varepsilon \text{ and } v > 0) \tag{6-C}$$

$$\sigma(y_t) = \sqrt{Var(y_t)} = \sqrt{E[y_t - E(y_t)]^2} = \sqrt{E(y_t^2) - [E(y_t)]^2} \tag{6-D}$$

$$\gamma(y_t) = \frac{e_3}{\sigma^3} = \frac{E(y_t^3) - E(y_t)\{3E(y_t^2) - 2[E(y_t)]^2\}}{\sigma^3}, \text{ with } e = E(y_t); \tag{6-E}$$

where, for powers *r = 1,2,3*, one has:

when	$\varepsilon \to 0$			$\varepsilon \to 0$ and $v \to \infty$		
$E(y_t^r)$	$= \int_{-\infty}^{w_t(v)} y_t^r \varphi(w)dw + v^r \int_{w_t(v)}^{\infty} \varphi(w)dw$		does not exist	$= \int_{-\infty}^{\infty} y_t^r \varphi(w)dw$	$= \int_{w_t}^{\infty} y_t^r \varphi(w)dw$	(6-F)
if	$\lambda_y < 0$			$\lambda_y = 0$	$\lambda_y > 0$	

because, for (6-C), it has been assumed that observed y_t is, as in a two-limit Tobit model, censured both upwards and downwards: $\varepsilon \leq y_t \leq v$, where ε and v denote the strictly positive limits assumed identical for all observations[10]. Note that the measure of 1st moment fit calculated for (6-A) with a normally distributed error is then indeed (6-C) and not the following unrolled version of (6-A):

$$\hat{y}_t = \left\{1 + \hat{\lambda}_y\left[\hat{\beta}_0 + \sum_k \hat{\beta}_k X_{kt}^{(\hat{\lambda}_k)}\right]\right\}^{\frac{1}{\hat{\lambda}_y}} \tag{6-G}$$

Measure (6-C) is preferred, even in the logarithmic Case 4 ($\lambda_y = 0$) of Table 4, when it collapses to:

$$E(y_t) = \hat{k}\left\{\exp\left[\sum_k \hat{\beta}_k X_{kt}^{(\lambda_k)}\right]\right\} \tag{6-H}$$

the particulars of which require an unbiased estimate of $\hat{k}$, the sample mean of the log-normal random variable *exp(w_t)* for the sample in question. But estimation of $\hat{k}$ from the Maximum Likelihood adjustment raises further issues that make calculation of expression (6-C) preferable to that of (6-H).

Satisficing our Modeling Requirements

For linear and nonlinear cases to be reasonably nested in (6-A), it is assumed that $E(y_t)$ in (6-C) is large enough relative to σ_w (Davidson & MacKinnon, 1985, p. 501), a fair hypothesis with quotidian trip data (*cf.* the table in Figure 3). This model then allows for estimation of the derivatives of the first three moments of *y_t* with respect to any X_{kt}, as summarized in Table 4 (the analytical derivatives $D^e_{X_{kt}}$, $D^\sigma_{X_{kt}}$, $D^\gamma_{X_{kt}}$ are found in Appendix 4), and of all elasticity or MRS impacts in (5-A) and (5-B), be they strict or sample such as $\partial y_t / \partial X_{kt} = \beta_k X_{kt}^{\lambda_k - 1}$.

The only originality of Table 4 (and of Appendix 4) is to present an enrichment to higher moments of usual practice typically focused solely on calculation of the derivative of the first moment. But this

extension to higher moment fits than the first adds nothing to, and subtracts nothing from, the original Box-Cox model. In particular, it does not preclude further enrichment with $D_{X_{kt}}^{\kappa}$, the partial derivative of the 4th moment of y_t with respect to X_{kt}, in view of the possibilities evoked in Endnote 3.

TABLE 4
DERIVATIVES OF MOMENTS OF *y* EXPLAINED BY BOX-COX REGRESSION WITH GAUSS-LAPLACE ERRORS

Case	Domain of y	Box-Cox power of y	$D_{X_{kt}}^{e} = \dfrac{\partial E(y_t)}{\partial X_{kt}}$	$D_{X_{kt}}^{\sigma} = \dfrac{\partial \sigma(y_t)}{\partial X_{kt}}$	$D_{X_{kt}}^{\gamma} = \dfrac{\partial \gamma(y_t)}{\partial X_{kt}}$
1	$-\infty < y_t < \infty$	(OLS)[11]	√	0	0
2		$\lambda_y = 1$	√	√	0
3		$\lambda_y > 0$	√	√	√
4	$y_t > 0$	$\lambda_y = 0$	√	√	√
5		$\lambda_y < 0$	√	√	√
√ = the partial derivative of the moment of y in question with respect to X_{kt} is not zero.					

Case 1 is well known to all students: we have $[-\infty < y_t < \infty]$ with Gaussian errors where *y* is explained linearly [implicitly with ($\lambda_y = 1$)] by OLS and changes of X_k have no effect on $\sigma(y_t)$ and $\gamma(y_t)$, as indicated by the null values. But it is easy to forget that, in Case 2 with ($y_t > 0$) as required by the BCT on *y*, $\sigma(y_t)$ is sensitive to changes in X_k (symmetry of Gaussian errors remaining) despite the linearity of *y*. More generally, when, as in Cases 3, 4, and 5, the nonlinearity of *y* produces a variable with an asymmetric distribution, all moments of *y* react to changes in X_k, a standard and natural property of models that are "nonlinear in *y*" (there are many other than the Box-Cox model) and which we use to extract the freely determined MRS among its moments. In particular, here we in principle have $\partial(E(y))/\partial(\sigma^2(y)) \neq 1$ and can always test the Poisson equality restriction[12], for instance rejected in road accident data tests, using the above algorithm, to be recalled presently.

The idea that preferences for distributions characterized solely by their moments can extend the mean-variance approach, and the need to enrich a Box-Cox algorithm until then limited to calculation of 1st moment fit to higher moment fits and to the extraction of MRS among the first 3 moments of *y*, were prompted by the serendipitous reading of Allais (1987) in 1988. It led to a research proposal in 1989 and to a report of first results to the funding agency (FCAR, 1991) on an exploratory 2-moment application to monthly road accident frequencies by severity category in Quebec: it yielded results clearly inconsistent with the Poisson restrictions on moments of *y* (Gaudry, 1997, 2000).

Road Speed Choice Precedents
Moments of road speed, and consequently of accident frequencies by severity category, can indeed be said to be endogenously chosen and concatenated, just like those of trip durations, and there is no need to limit oneself to the first moment of speed choice, as in Ashenfelter & Greenstone (2004). Simultaneously recognizing that speed variance also matters to accident risk, as Lave (1985) had noted, they do not dare consider it as also properly endogenous. It is as if drivers had blinkers with preferences limited to picking only mean speed and its accident risk, higher moments of speed then just floating above the first, but somehow unexplainable and irrelevant.

By contrast, extensive applications of our multi-moment methodology to Quebec and West Germany, funded again by FCAR and by the *Deutsche Forschungs-Gemeinschaft* (DFG), yielded 3-moment tables of MRS (Blum & Gaudry, 2000), of the same format as Table 9 below, showing almost identical MRS among moments in the two regions. Later, revealed values of a statistical life were derived from those trade-offs among accidents frequencies of different severity and cost categories (Gaudry, 2006). At the time, the 4th moment of road safety indicators was not of concern or examined.

The Incontrovertible Superiority of the Box-Cox Power over the Simple Power Model

The likelihood for the above model (6-A) and (6-B), as in Box & Cox (*op. cit.*), is well known to be:

$$\Lambda = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left(-\frac{w_t^2}{2\sigma_w^2}\right) \left|\frac{\partial w_t}{\partial y_t}\right| \tag{6-I}$$

where $\left|\partial u_t / \partial y_t\right| = y_t^{\lambda_y - 1}$ denotes the Jacobian of the transformation from the w_t to the observed y_t, giving rise to a natural question concerning its advantages in comparison with an alternative formulation for the apparently less complicated simple power model

$$y_t^{\rho_y} = \delta_0 + \sum_{k=1}^{k=K} \delta_k X_{kt}^{\rho_k} + u_t \tag{7-A}$$

And reasons for choosing BCT powers (6-A) over simple powers (7-A) are certainly of general interest.

i. ***Continuity at 0 or degeneracy?*** The BCT specification is continuous at 0 and nests both Case 2 (linear) and Case 4 (logarithmic) of Table 4. On this continuity point, the simple power model is woefully inadequate: not only does $\rho \to 0$ make all variables tend to 1, but estimation admits the corresponding degenerate "black hole" solution $\rho_y = \rho_k = 0$, the avoidance of which requires estimation under parameter constraints. Moreover, it can be shown that the $\hat{\lambda}_y$ of (6-A) is interpretable as a non-degenerate estimate of $\hat{\rho}_y$ in (7-A) because of BCT invariance to a power transformation of y even in the absence of an intercept β_0 in (6-A), a polite demonstration inviting to abandon simple power models (Gaudry & Laferrière, 1989), as we did in 1976.

FIGURE 2
ILLUSTRATED MAINTENANCE OF THE ORDER OF VALUES
AND OF THE CONTINUITY AT 0, OF THE BCT
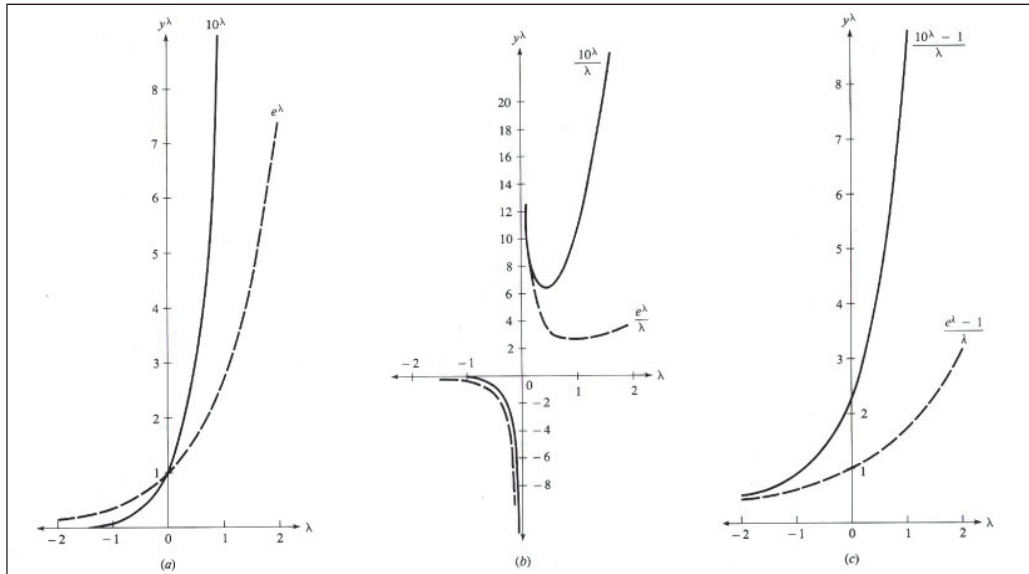
ii.   *Maintenance or risk of inversion of the order of the data?* Another risk, never mentioned, is that simple powers do not maintain the order of the data, contrary to Box-Cox transformations, as illustrated in Figure 2 (Johnston, 1984, p. 63) where the values 10 and e = 2,84128 are transformed in three ways: in (a) by a simple power y^λ, which implies an inversion of values as one goes through point (0,1); in (b) by y^λ / λ, which maintains the order but causes discontinuities at $\lambda = 0$; in (c) by $(y^\lambda - 1)/\lambda$, which preserves both continuity and the order of data values. Estimates of simple powers y^λ therefore require an ***ex post*** verification of the order of data for all transformed variables, in both their sample and forecast domains, if weird situations are to be avoided.

iii.   *Legibility of signs of derivatives, marginal rates of substitution and elasticities*. As seen in Cell 1.A of Table 5, an advantage of a BCT applied to X_k is that its own sign has no effect on the sign of the derivative with respect to that variable, which only depends on that of β_k. With simple powers, the direction of the effect of a change in X_k depends on the signs of the three parameters in $\delta_k \rho_k / \rho_y$; and the same holds for the sign of the ordinary elasticity in 3.B. Legibility is not better for marginal rates of substitution, the signs of which depend on four parameters in $\delta_k \rho_k / \delta_\ell \rho_\ell$ because the sign of the ratio of transformed variables is positive, as it is in all cells of Column A of Table 5. The legibility of signs is apparently always higher with BCT than with simple powers.

TABLE 5
SIGNS OF STATISTICS DERIVED FROM POWER MODELS:
BOX-COX *VS* SIMPLE POWERS

Definitions of statistics derived from models estimated with:		**A. Box-Cox**	**B. Simple**
1.	Partial derivative	$\dfrac{\partial y}{\partial X_k}$	$\beta_k \dfrac{X_k^{\lambda_k-1}}{y^{\lambda_y-1}}$	$\dfrac{\delta_k \rho_k}{\rho_y} \cdot \dfrac{X_k^{\rho_k-1}}{y^{\rho_y-1}}$
2.	Marginal rate of substitution (MRS) between X_k and X_ℓ	$\dfrac{\partial y/\partial X_k}{\partial y/\partial X_\ell}$	$\dfrac{\beta_k}{\beta_\ell} \cdot \dfrac{X_k^{\lambda_k-1}}{X_\ell^{\lambda_\ell-1}}$	$\dfrac{\delta_k \rho_k}{\delta_\ell \rho_\ell} \cdot \dfrac{X_k^{\rho_k-1}}{X_\ell^{\rho_\ell-1}}$
3.	Sample elasticity (*i.e.* ordinary) $\eta(y,X_k)$	$\dfrac{\partial y}{\partial X_k} \cdot \dfrac{X_k}{y}$	$\beta_k \dfrac{X_k^{\lambda_k}}{y^{\lambda_y}}$	$\dfrac{\delta_k \rho_k}{\rho_y} \cdot \dfrac{X_k^{\rho_k}}{y^{\rho_y}}$

All reasons just stated to prefer BCT also hold for the following "inverted" simple power model:

$$y_t = \left[\delta_0 + \sum_k \delta_k X_{kt}^{\rho_k} \right]^{\frac{1}{\rho_y}} + \xi_t \qquad\qquad (7\text{-}B)$$

where direct non linear estimation can yield σ, the Constant Elasticity of Substitution (CES) parameter, if and when it is recovered from $\sigma = f(-\rho)$ or from another function relating σ to ρ.

Form and the Randomisation of Regression Coefficients

The BCT is a local nonlinear approximation replaceable by splines (Rich & Mabit, 2016; Rich, 2018) or eventually by the Fourier transform (Gallant, 1981). But not by the randomization of β_k coefficients, an entirely distinct procedure seemingly successful because it mimics curvature, but at the cost of confusing form and randomness, as demonstrated by Orro *et al.* (2005, 2010) in multivariate Monte Carlo tests.

Randomness, an old innovation in classical regression models (Swamy, 1970; Johnson, 1977, 1978), including transportation ones (Hensher & Johnson, 1979) where the form implication was noted early (Johnson, 1979), appeared in Logit models under the "Mixed" label, held capable of approximating any Random Utility Model (*e.g.* McFadden & Train, 2000), a stunning claim in view of Orro's work.

In addition to the Monte Carlo results, just noted, showing that form and randomness are distinct notions, atomisation of segments effected by randomisation of coefficients poses problems of its own because distributions of the β_k are all strictly unknown. Moreover, in "Mixed Logit" models, the information matrix does not have a closed form, which implies an undefined efficiency bound (Cirillo, 2005). In spite of this built-in design ignorance and obscurity, Lapparent *et al.* (2009) have shown with data on three countries that BCT on Time, Cost and Access time were generally different from both 0 and 1, at least with their own assumed β_k distributions, a partial empirical validation of the above Monte Carlo proofs that "mixed" can just mean "mixed up" because, if β_k regression coefficients can have distributions, then β_k form parameters obviously can have them as well!

DATA, MODELS AND RESULTS: MARGINS, RELIABILITY ECHOES AND VALUATIONS

The Nature of Available American and French Surveys
Home-based Trips in Three Publicly Available Databases

To analyze durations, we draw individual home-based trips from 3 surveys: the US Nationwide Personal Transportation Survey (NPTS 2009, of 150 000 households), limiting ourselves to the large standard metropolitan areas (listed in Appendix 5); the *Enquête globale transport* (EGT 2010, of 18 000 households) of the Greater Paris Region (Île-de-France, 12M inhabitants); and the *Enquête nationale transports et déplacements* (ENTD 2008, of 22 200 households) of Metropolitan France (63M inhabitants).

Survey Information and Sample Size Limitations

Durations of 7h30-9h30 morning peak work trips (shown in Figure 3) will be studied comparatively in the 3 regions, but those of 10h00-18h00 car trips to shopping centers will be studied only in Metropolitan France (despite the relatively small samples shown in Table 4), due primarily to the fact that, in the NPTS 2009 Survey, the variable Great Circle Distance, required as the key instrument for the construction of the reference road network service variable S^p below for use in (8-B), exists only for work trips.

The Approach Adopted with Strictly Comparable Home-Based Survey Trip

In all of these surveys, *rendezvous* time H_{rdv} proposed in our duration demand equation (4-C) is unknown and must be replaced by *arrival time* H_{arr} at the activity in question. We per force assume:

$$\left[\hat{e}(T_n^c); \hat{\sigma}(T_n^c); \hat{\gamma}(T_n^c) \right]_{H_{rdv}-H_{dep}} \approx \left[\hat{e}(T_n^c); \hat{\sigma}(T_n^c); \hat{\gamma}(T_n^c) \right]_{H_{arr}-H_{dep}} \tag{8-A}$$

i.e. that trip duration moments of interest are approximately the same, a hypothesis clearly testable in the future with datasets containing both rendezvous and arrival times by activity (trip purpose).
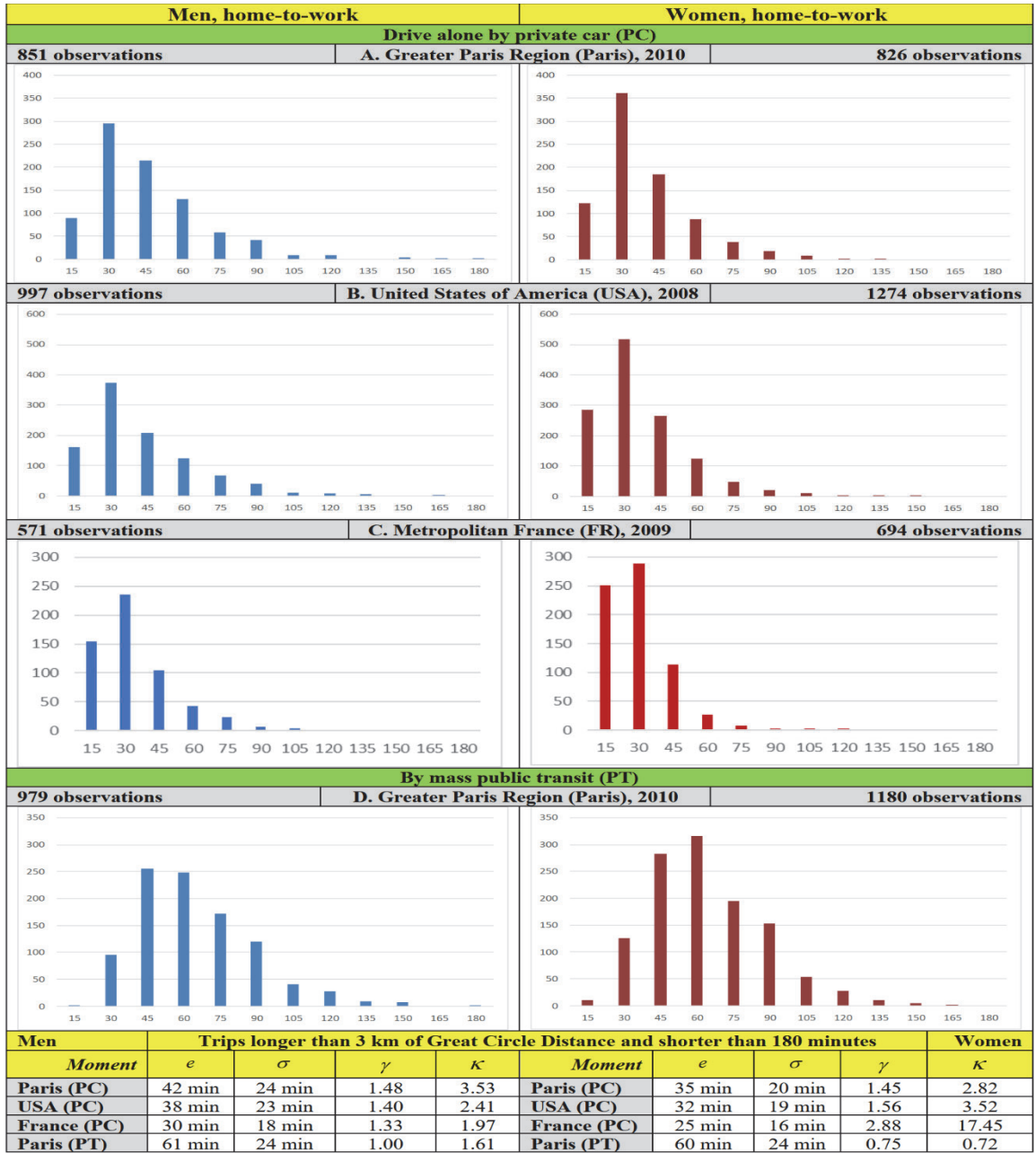
Summary of Economic and Econometric Specifications

The most extensive specifications, developed originally in source research reports III and IV for the Greater Paris Region, had 29 variables for car trip equations and 21 for transit trip equations. They can be represented by:

$$(T_n^c)^{(\lambda_T)} = f\left[(Day\ of\ week,\ Direction,\ Parking\ type);\ S_n^{(\lambda_S)};\ (Age^{(\lambda_{Age})},\ Income^{(\lambda_{Income})},\ etc.)\right], \tag{8-B}$$

as indicated in Table 6 where 4 continuous variables are to be transformed by BCT and all other variables but one (the number of persons in the household) are dummies. As **Service S^p** was the only highly significant explanatory variable and other variables were never significant (*cf.* regression details in source paper V), but for two rare exceptions (free street parking and driving towards the Paris city centre), the more recent specifications for the US and France were drastically pared down, after further tests had confirmed the core Greater Paris Region finding – that additional variables contributed almost nothing to the explanation of chosen durations after service had been taken into account.

FIGURE 3
WORK TRIP TIME BY MASS TRANSIT (PT) OR PRIVATE CAR (PC)
IN PARIS, THE USA AND FRANCE

Men	Trips longer than 3 km of Great Circle Distance and shorter than 180 minutes				Women				
Moment	*e*	*σ*	*γ*	*κ*	*Moment*	*e*	*σ*	*γ*	*κ*
Paris (PC)	42 min	24 min	1.48	3.53	**Paris (PC)**	35 min	20 min	1.45	2.82
USA (PC)	38 min	23 min	1.40	2.41	**USA (PC)**	32 min	19 min	1.56	3.52
France (PC)	30 min	18 min	1.33	1.97	**France (PC)**	25 min	16 min	2.88	17.45
Paris (PT)	61 min	24 min	1.00	1.61	**Paris (PT)**	60 min	24 min	0.75	0.72

For transit trips in the Greater Paris Region, we use for S^p an exogenous *Logsum* measure (or, rather better, its antilog subjected to a Box-Cox transformation) of multiple-path PT service extracted from the RATP Global-9 demand model. Its refined publicly documented Logit path choice model notably includes Box-Cox transformed in-vehicle time (Leblond & Langlois, *op. cit.*), weighted by tested comfort functions based on passenger densities, and dummy variables by type of line (Prat & Leblond, *op. cit.*) capturing residual effects, or "attitudes" to them as proposed in Gaudry & Lapparent (2015).

For car trips in the Greater Paris Region, the network service variable S^p is defined by a Wardrop user equilibrium assignment time, an exogenous measure available per O-D. For France and the USA,

exogeneity required building a reference time equal to Great Circle Distance (D_{GCD}) divided by mean speed on the network $V_{network\ t}$ constructed from declared Duration and calculated Distance for trip t.

$$S_t = \frac{D_{GCD}}{V_{networkt}} = \frac{D_{GCD}\ T_t}{D_{networkt}} \tag{8-C}$$

In principle, this complex reference car service time variable is both exogenous to the explanation of declared car trip durations and shorter than durations actually incurred: (a) Great Circle Distance is itself an instrument necessarily exogenous and shorter than any shortest path O-D distance; (b) also, observed O-D speed for any individual's trip is certainly exogenous at congested times but perhaps less so at other times. To verify this exogeneity, we included in the regression explanation of US trip durations the local fuel price P on the day of the surveyed trip. Given that the car mode has been chosen, that price should not be significant when congestion reigns because differences in driving style then have little room for expression, but could become more significant at other times if and where free flow uncongested conditions prevailed on some links of a path. As it turns out, the coefficient of that daily price variable is never even barely significantly different from 0, as anticipated.

TABLE 6
EXPLANATORY VARIABLES OF TRIP DURATIONS
BY REGION AND ACTIVITY (TRIP PURPOSE)

	P_n	S_n^p (unit)	Parking	Age	Income	Other	Total number of X_k excluding β_0
PC							Morning peak work trips (7h30-9h30)
Paris		√ Wardrop user equilibrium time	√	√	√	√	29
USA	√	√ built-up reference time		√			3
France		√ built-up reference time		√			2
							Day shopping centre trips (10h00-18h00)
France		√ built-up reference time		√			2
PT							Morning peak work trips (7h30-9h30)
Paris		√ log-sum of path choice model		√	√	√	21

Estimation of BCT Models of Particular Interest
For each market, 4 distinct BCT specifications were tested: the Linear, corresponding to Case 2 in Table 4; the Logarithmic, corresponding to Case 4; the BCT, with λ_k applied to each of duration T_n^c and service S_n^p (Case 3); an enriched Case 3 (Paris trips only), with BCT also applied to age and household income (Case 3+):

Case 2	Linear	BCT=1 on all transformable variables; Linear otherwise	
Case 4	Logarithmic	BCT=0 on **Duration** and **Service**; Linear otherwise	
Case 3	Box-Cox, 2 BCT	BCT estimated on **Duration** and **Service**; Linear otherwise	(8-D)
Case 3+	Box-Cox, 3 BCT	Case 3 with the addition and estimation of a BCT on **Age** and **Income**	

All specifications were tested to ensure that a global maximum of the Log-Likelihood of (6-I) had been found: it has long been known in practice, and more recently studied in theory (Kouider & Chen, 1995), that the presence of BCT in (6-I) can generate local maxima, especially with 2 or more BCT. Belsley-Welsh-Kuh condition indices (Belsley *et al.*, 1980), as re-interpreted by Erkel-Rousse (1995), were extensively used to guard against collinearity, as described in detail in source papers II and III, with a view to isolating the required specific effects of the Service variable on duration. Property (5-C) guarantees in any case the robustness of estimates of the also needed MRS among moments.

Neither Linear, Nor Logarithmic but Box-Cox Optimal Forms with $0,05 \leq \lambda \leq 0,60$

All tests showed that Case 2 was always infinitely dominated by Case 4 and that Case 3+ led to linear forms for age (and for income, when present). Also, Case 3 gave significantly better fits than Case 4 everywhere except with the very small samples shown in Figure 4. Forthcoming tables therefore document only Case 3 results (and Case 4 Log Likelihood values in Table 8 for PC trip durations and Table 10 for PT trip durations), but detailed information on all 4 cases can be found in source reports.

Meaning of Estimated BCT Parameter Values

Tables 8 and 10 show that *a priori* fixation on Linear or Log-Log forms would have been extremely damaging, as found long ago in trip demand and mode choice models (Gaudry & Wills, 1978) and confirmed more recently in a survey of tens of BCT applications to mode choice worldwide (Gaudry, 2010). But two further comments should be made on the exact values of BCT found for all work trip markets (drive alone in Table 8 and mass transit in Table 10), because they signal very significant differences in the shapes of marginal disutility curves:

    i.   **women's BCT for duration and service are roughly twice the size of men's**. These differences are statistically significant, as demonstrated by losses in Log-Likelihood from re-estimation of each equation (variants 35 *vs* 22 of T.8; 31 *vs* 17 of T.10) with the other's optimal values. In addition, unconditional *t*-tests with respect to 0 and 1 confirmed that all BCT values (in the original and restricted re-estimations) differed from 0 and 1. The results of these re-estimations are not shown;

    ii.   **public transit vs car BCT.** Comparing the first two columns of Table 8 with those of Table 10, BCT values for PT are roughly twice the size of PC values for Duration and half of their size for Service. This intermodal difference was established by the same cross-constrained re-estimation method (variants 17 *vs* 22; 31 *vs* 35; results not shown) used to establish differences between men and women.

Establishing the Existence and Palliative Nature of the Safety Margin

It would not suffice here to state that precautionary margins were long "implicitly lived but not explicitly known", as famously stated by Blondel (1904, p. 437) about certain doctrinal developments. How then is it possible to be in fact certain that properly observed durations do really contain margins, as claimed, so that their extracted moments are not just those of endured route service? A reasonable demonstration is available by an analysis of the elasticity of duration (*see* its possible definitions in Table 3) with respect to endured network service. This elasticity estimate is the foundation of the model, an argument summarized in Table 7 to which one might add the following comments:

    i.   **Nonunitary elasticity of duration with respect to endured service as a test of the existence of the safety margin.** Econometric model (8-B) is legitimate if $\eta(T_n^c, S_n^p) \neq 1$: otherwise, it amounts to a regression of service on itself (and on other variables *X*), which is of little interest beyond that very point. Were the units of T_n^c and S_n^p identical, one could perhaps imagine a test of the value of the regression coefficient but, as units differ and both variables are subjected to BCT, a test of the unitary value of the elasticity is much preferable, even if it remains intuitive. Here, it is carried out at sample means between the first moments of the duration and service variables in question.

    ii.   **The internal structure of compunction time T_n^c conditional to a non-unitary elasticity.** In the source papers, T_n^c is called "compunction time" because it is assumed to combine expected service and offsetting margin time "punctions". Once it is established that $\eta(T_n^c, S_n^p) \neq 1$, can anything interesting then be said about the internal structure of T_n^c? If

$\hat{\lambda}_y = 0$ and one further assumes that $I_n^* S_n^p$, it is possible to calculate $\eta(I_n^*, S_n^p)$ from $\eta(T_n^c, S_n^p) \neq 1 = \hat{\eta}$ because the elasticity of a product is the sum of the elasticities of its components. More generally, $\hat{\lambda}_y \neq 0$ and nothing can be said about the internal structure of T_n^c. But note that acceptance of $\hat{\eta} \neq 1$ does not require to take a position on whether the existing safety margin punction is multiplicative, additive (as many think and build into models) or entirely something else such as $(T_n^c)^{(\lambda_y)} \equiv \left[\alpha_1 (I_n^*)^{(\lambda_I)} + \alpha_2 (S_n^c)^{(\lambda_S)} \right]^{(\lambda_y)}$.

TABLE 7
SAFETY MARGIN EXISTENCE TEST BASED ON
A UNITARY ELASTICITY OF DURATION W.R.T. ENDURED SERVICE

Given $\eta(T_n^c, S_n^p) = \hat{\eta}$, an estimate of the elasticity of duration with respect to endured network service derived from a Box-Cox model with regressand $(T_n^c)^{(\lambda_y)}$ and regressors $(S_n^p)^{(\lambda_x)}$ and other variables X, (n = market). Then (in current tests, the elasticity of mean duration with respect to mean endured service is calculated):		
if $\hat{\eta} = 1$,	duration and service are identical ($T_n^c \equiv S_n^p$): no safety margin exists. The model is not legitimate;	
if $\hat{\eta} \neq 1$,	the safety margin legitimately exists ($T_n^c \equiv f(I_n^c, S_n^p)$), but its (internal) structure is unknown;	
	and $\hat{\lambda}_y = 0$	the additional assumption of a multiplicative margin $I_n^* \cdot S_n^p$ allows for the identification of its elasticity w.r.t. service, $\eta(I_n^*, S_n^p) = \hat{\eta} - 1$, but common language and all known authors envisage nothing but an additive safety margin;
	and $\hat{\eta} < 1$	one understands that I_n^* compensates for the uncertainty of S_n^p;
	for any $(\hat{\eta}, \hat{\lambda}_y, \hat{\lambda}_x)$ set, one may further calculate (i) elasticities of other moments of duration than the first with respect to mean endured service; (ii) the MRS among moments of actual duration.	
	It would not make sense, for (ii), to put our own ***ex ante*** restrictions on the levels of moments of duration or on their relative valuations, for instance by assuming a Poisson formulation imposing equality between the mean and variance v of constructed duration and a skewness equal to $v^{-1/2}$, or by other artificial constraints such as the − no doubt tongue-in-cheek − safety margin proportional to the standard deviation of endured travel time (Fosgereau, 2015). How could the effect on utility of any time margin with its own moments be ***a priori*** of the same sign as, or proportional to, that of the standard deviation of endured service, or indeed of any of its other moments?	

iii. ***The palliative interpretation of*** I_n^* ***when*** $\hat{\eta} < 1$. One expects $\hat{\eta} < 1$ because I_n^* is understood to compensate for the uncertainty of S_n^p: service improvements then reduce the safety margin, and the reverse occurs when service worsens. In the particular case of $\hat{\lambda}_y = 0$, the offsetting (negative) role of I_n^* directly follows from $\eta(I_n^*, S_n^p) = \hat{\eta} - 1$ when $\hat{\eta} < 1$.

iv. ***The link between a palliative margin and the "reliability echo".*** This palliative role of the safety margin is at the root of the "reliability echo" documented below. The echo, defined by the sum of adjustments of the 2nd, 3rd and 4th reliability moments, as measured by their

elasticities with respect to endured service S_n^p, turns out to always amplify adjustments of the 1st moment: the echo always acts as a net complement to the adjustment of the mean duration to changes in suffered service.

Car Trip Durations: Unit Elasticity Tests, Moment Demand Elasticities and MRS

The car work trip data to be analyzed are those shown in Figure 3 for the 3 regions of interest; and, despite the relatively small size of the samples, we will also use the car shopping trip data of Figure 4.

TABLE 8
ELASTICITIES OF PC TRIP DURATION W.R.T. MEAN SERVICE
AND STUDENT'S _t_ OF PARAMETERS, CASE 3 OF (8-D)

Home-based drive alone (PC) trips for:		work						shopping	
		Greater Paris Region		USA (SMA ≥ 3 M)		Metropolitan France			
(M= Men; W=Women):		1. M	2. W	3. M	4. W	5. M	6. W	7. M	8. W
Variant run number in source paper tables		22	35	H-0	F-0	H-0	F-0	H-0	F-0
Number of X_k variables except for β_0		29	29	3	3	2	2	2	2
Nomber of observations		851	826	997	1274	571	694	77	121
	BCT estimates and their unconditional Student t statistics								
BCT	on Duration λ_y	0.11	0.24	0.26	0.36	0.48	0.45	0.20	0.32
	Uncond. Student _t_ (=0)	(2.81)	(5.31)	(7.21)	(13.97)	(8.91)	(9.74)	(1.19)	1.83)
	on Service λ_x	0.21	0.29	0.32	0.46	0.52	0.60	0.38	0.55
	Uncond. Student _t_ (=0)	(2.81)	(4.09)	(9.47)	(25.46)	(8.24)	(14.21)	(1.17)	(3.14)
Elasticity of duration with respect to endured Service time S^p	*Casual sample elasticity of (y) as defined in Table 3*								
	$\eta^s(y,X_k)=\dfrac{\partial y}{\partial X_k}\cdot\dfrac{X_k}{y}$	0,760	0,750	0,907	0,869	0,836	0,849	0,837	0,739
	Student _t_ (=0) condit. on BCT	(37.59)	(9.32)	(103.16)	(108.12)	(53.08)	(76.77)	(13.32)	(16.41)
	Strict elasticities of moments of (y) as defined in Table 3								
	$\eta^e(y,X_k)=\dfrac{\partial e(y)}{\partial X_k}\dfrac{X_k}{e(y)}$	0,750	0,740	0,904	0,867	0,828	0,850	0,835	0,738
	$\eta^\sigma(y,X_k)=\dfrac{\partial \sigma(y)}{\partial X_k}\dfrac{X_k}{\sigma(y)}$	0,660	0,560	0,667	0,557	0,430	0,470	0,666	0,505
	$\eta^\gamma(y,X_k)=\dfrac{\partial \gamma(y)}{\partial X_k}\dfrac{X_k}{\gamma(y)}$	-0,091	-0,180	-0,240	-0,310	-0,400	-0,380	-0,170	-0,230
Log-Likelihood (LL) Case 4 (Log-Log)		-3305	-3023	-3169	-3979	-1824	-2104	-223.9	-357.3
Log-Likelihood (LL) Case 3 (Box-Cox)		-3300	-3008	-3136	-3887	-1783	-2040	-221.5	-351.3
LL difference (2 degrees of freedom)		5	15	33	92	41	64	2.41	5.84
Pseudo-(L)-R²		0.66	0.69	0.91	0.91	0.87	0.87	0.79	0.78

Establishing the Existence of the Precautionary Palliative Time Margin

Table 8 presents values of the two estimated BCT and of the elasticity of duration with respect to mean service, as well as the _t_-statistics of the associated BCT and of the service regression parameters. The latter are conditional on BCT estimates and consequently invariant to units of measurement of the regressors (Spitzer, 1984; Dagenais & Dufour, 1994).

In all variants, the sample measure and the strict first moment measure of the elasticity of duration with respect to mean service clearly differ from 1 and, in view of the extremely high _t_-statistics, are indeed all much smaller than 1, thereby passing the model validity test defined in Table 7. To verify this for the highest values (0,87 and 0,91 for the USA), consult their detailed analysis in Appendix 6.

Elasticities of the Moments of Duration with respect to Endured Service and Closure of the Transport Model

The feedback of endured service on the time profile of trip demand duration itself, which closes the demand model, can be calculated with the 3 strict moment elasticities highlighted in green in Table 8.

Improved service translates into *minus* these elasticity values and leads, for all regions and both sexes, to the creation of super-peaks (increased kurtosis) because the only mitigating reaction, that of asymmetry, is weak (but less in women than in men) and cannot compensate for the much stronger reduced mean and standard error (seemingly fast fill-up) effects. The consequence is in all markets an increased concentration or acuity of the time profile of demand during the period considered, namely:

$$-\left[\eta(e(Dur_n), X_k) + \eta(\sigma(Dur_n), X_k) + \eta(\gamma(Dur_n), X_k)\right] \Rightarrow SUPERPEAK\ of\ kurtosis\ \kappa(Dur_n) \quad (9\text{-}A)$$

The Reliability Echo, a Function of the Weak Elasticity of Asymmetry

The size of the elasticity of asymmetry, always producing a relatively small counter effect as compared to those of the second moment, makes possible the existence of what we have called a "reliability echo". For work trips, as well as for shopping centre trips (shown in Table 10), net changes in the three reliability moments (the 4th, which can be qualitatively estimated or guessed, included) caused by changes in mean service time always augment the utility gains or losses associated with changes in mean time. This would not be the case if the elasticity of asymmetry were much higher in absolute value: the sign of the net effect on utility of changes in reliability moments could then easily be indeterminate.

Because of this echo, it would be incorrect to evaluate, for instance, the utility of the massive investment in major urban roads and interstate highway urban portions effected in the US between NPTS dates of 1998 and 2009 by focusing only on the lower average speeds achieved in the largest SMA: lower average speeds are compatible with hugely higher overall unreliability characterized by increased standard error combined to decreased kurtosis (flatter super-peaks), both only very partially offset by lower asymmetry. Road works change mean travel time, which causes restructuration of the time profiles of demand through adjustments of the safety margins embodied in durations choices.

If mean travel time falls, commuters might sleep longer, take to the road in bunched manner (the perceived fast fill-up) and exacerbate the acuity of the super-peak, a desirable solution, not a problem to be implicitly regretted as in Dutzig *et al.* (2017); if it rises, they may have to sleep less and adjust the net moments of unreliability in the opposite direction through higher palliative safety margins.

Car Trip Purpose Levels and Mix

For a given demand time slot, the total utility variation resulting from increased lane capacity depends on changes, at equilibrium, in the first 4 moments of duration for trips for distinct purposes considered separately and summed, and not just on changed mean service time for all trip purposes together: trip purpose levels and mix affect the mean time. Valuation of moment changes (in any direction) from the initial to the new user situation will then be needed by trip purpose: one can make use of estimated MRS among them, expressed for instance per unit of the 1st moment, discussed presently.

Substitution Among Moments of Duration Choices

One finds in Table 9, for the 3 regions, marginal rates of substitution among the first three moments of duration MRS^{m_τ, m_s} expressed per unit of the 1st moment and, in that sense, comparable across regions and markets. In the source papers, original tables also present elasticities of substitution among the moments, which are always well-behaved and consequently relatively uninteresting, as well as measures of fit by moment, as one can verify in Appendix 7 which gives those supplementary statistics for Parts A, B and C of Table 9.

Concerning MRS^{m_τ, m_s} estimates in Table 9, note that: (a) their signs are compatible with those of risk-averse agents (matrix [A] in Table 2); (b) women tend to value the second moment more, and the third less, than men; (c) the greatest difference in valuations occurs between Greater Paris and the other two

regions; it pertains to the relative value of the second moment despite sensibly close elasticities of substitution of about 1.5-1.6 (*cf.* Appendix 7) across regions. Let us comment on the last two points.

i. ***differences between men and women***. Women make shorter commuting trips than men[13], a bare fact repeated *ad nauseam* even by those who fail to provide any explanation for it (*e.g.* Ng & Aker, 2018), contrary to White (1977, 1986) who first tried to make sense of it. But by far the most serious analyses of the complex issues involved are those of Picard *et al.* (2013; 2018) who, after controlling for various endogeneity biases and the effects of negotiations within couples, allow in their latter paper for comparisons of trip utility levels between men and women within a unified framework.

TABLE 9
WORK AND SHOPPING CENTRE (C.1) PC TRIP DURATION MOMENT MARGINAL RATES OF SUBSTITUTION

Marginal rates of substitution [Equations (5-D.1) to (5-D.3)]		**A. Greater Paris Region**					
		M (variant 22)			**W (variant 35)**		
Moment i\j	e	σ	γ	e	σ	γ
MRS^{m_e,m_σ} , MRS^{m_e,m_γ} MRS^{m_σ,m_γ} e	1	+ 3.5	− 400	1	+ 4.5	− 220
σ		1	− 110		1	− 49
γ			1			1

Marginal rates of substitution		**B. US Standard Metropolitan Areas**						**C. Metropolitan France**					
		M (variant H-0)			**W (variant F-0)**			**M (variant H-0)**			**W (variant F-0)**		
Moment i\j	e	σ	γ	e	σ	γ	e	σ	γ	e	σ	γ
MRS^{m_e,m_σ} , MRS^{m_e,m_γ} MRS^{m_σ,m_γ} e	1	+ 7.9	− 380	1	+ 8.6	− 250	1	+ 9.5	− 190	1	+ 9.0	− 160
σ		1	− 47		1	− 29		1	− 20		1	− 18
γ			1			1			1			1

Marginal rates of substitution		**B.1. US Standard Metropolitan Areas**						**C.1. Metropol. France shopping trips**					
		M: Whites			**W: Whites**			**M (variant H-0)**			**W (variant F-0)**		
Moment i\j	e	σ	γ	e	σ	γ	e	σ	γ	e	σ	γ
MRS^{m_e,m_σ} , MRS^{m_e,m_γ} MRS^{m_σ,m_γ} e	1	+ 8.0	− 420	1	+ 9.4	− 240	1	+ 4.8	− 140	1	+ 5.5	− 110
σ		1	− 46		1	− 26		1	− 30		1	− 19
γ			1			1			1			1

Marginal rates of substitution		**B.2. US Standard Metropolitan Areas**					
		M: Blacks/Hispanics			**W: Blacks/Hispanics**		
Moment i\j	e	σ	γ	e	σ	γ
MRS^{m_e,m_σ} , MRS^{m_e,m_γ} MRS^{m_σ,m_γ} e	1	+ 5.6	− 280	1	+ 5.0	− 730
σ		1	− 50		1	− 150
γ			1			1

With utility specifications of type (1-B) −except that their variables appear linearly−, they show that women dislike in-vehicle time by mode (including driving, then) much more than men (about 33% more) and that both sexes dislike driving alone much more than using transit (about 100% more), as one can verify in this table regrouping their latest, and most refined, results:

Pooled sample + households without a car	Table 4 model		Table 8 model		Table 9 model	
(Negative) coefficient of travel time	**Man**	**Woman**	**Man**	**Woman**	**Man**	**Woman**
-Public mass transit (PT)	1.54	2.27	1.54	2.35	0.83	1.10
-Private car drive alone (PC))	2.95	4.32	2.90	4.44	1.62	2.64
ratio PC/PT	*1.91*	*1.90*	*1.88*	*1.89*	*1.96*	*2.41*

The key point here is that differences in distances covered by sex, irrespective of mode, are explained by higher disutility per unit of mean travel time in women than in men. The previous literature on urban commuting gave that general impression but its results were often difficult to interpret because, in addition to endogeneity issues, trip distance, disutility per unit of distance and regression form (typically fixed[14]) were and remain always entangled. For instance, White (*op. cit.*) emphasized women's shorter trip lengths, but not implied disutility per minute or mile as such. Our results in Tables 9 and 11, relating higher moments of travel time to its first, say nothing about the weights of a unit of the first by sex.

This weight, in our model, is the partial derivative of the first moment of duration with respect to mean service, shown in the following table for distinct and pooled sex samples by mode. The comparison yields first indications consistent with a higher disutility of mean time in women than in men: at sample means, the partial derivative of $e(y)$ with respect to mean network service is higher for women than for men by 7,3 % in PC and 9,1% in PT, as can be verified in this table:

Mode	Private car drive alone (PC)			Mass public transit (PT)		
Sample	**Men**	**Women**	**Pooled**	**Men**	**Women**	**Pooled**
Variant number in source paper III	22	35	13	17	31	8
$\partial e(duration)/\partial\ mean\ service\ time$ *	0.82	0.87	0.84	-0.22	-0.24	-0.23
Derivatives evaluated at sample means. All specifications of (8-B) are for Case 3 optima defined in (8-D).						

Taking form into account in a formulation that neglects all endogeneity biases engendered within couples then apparently confirms the results obtained under linearity in the previous summary table (with those biases corrected): women dislike every minute of mean travel time more than men, independently of mode: measured with linear or Box-Cox forms, their marginal disutility of urban commuting travel time is just higher than that of men, which explains their famously shorter trips;
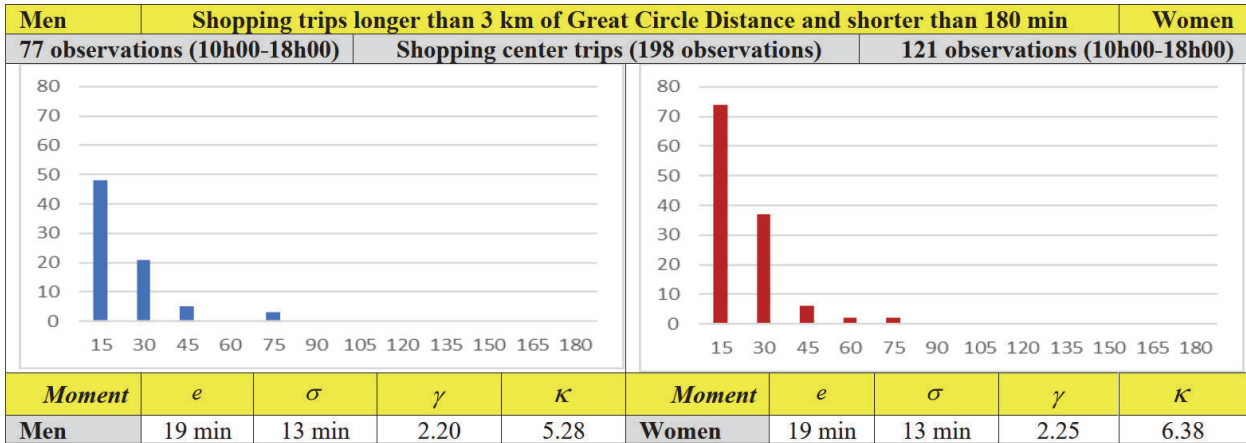
    ii.   ***relative value of arrival at work on time.*** To understand why the valuation of the second moment would in the Greater Paris Region amount to half of that found in the US and in Metropolitan France, we hypothesize that socio-cultural importance of on-time arrival at work might differ between these three regions, an interpretation inspired by the examination, in the US, of revealed valuation differences between Whites and Blacks or Hispanics reported in Table 9.

Results, shown in Parts B.1 and B.2 of that table, are consistent with the existence of cultural differences, in this case revealing Whites as stricter than Blacks and Hispanics on the importance of on-time arrival at work represented by the relative value of the second moment. But if the risk of late arrival at work is linked instead to the relative value of the 3rd moment, Black or Hispanic women are those with the greatest relative valuation of the third moment and the "strictest" attitude among regions and groups, perhaps because of their types of jobs. This cultural hypothesis should be probed further with ethnic or international data and multiple types of jobs.

On Shopping Trips

The results for trips to shopping centers ("*Grandes surfaces*") in Columns 7 and 8 of Table 8 and in Part C.1 of Table 9 must be affected by the small sample sizes (see Figure 4).

Men	Shopping trips longer than 3 km of Great Circle Distance and shorter than 180 min			Women
77 observations (10h00-18h00)	Shopping center trips (198 observations)			121 observations (10h00-18h00)

Moment	*e*	σ	γ	κ	*Moment*	*e*	σ	γ	κ
Men	19 min	13 min	2.20	5.28	**Women**	19 min	13 min	2.25	6.38

Still, the latter reveal reasonably lower relative valuations of reliability moments MRS^{m_τ,m_s} than for work trips. Samples of trips to the doctor or the hairdresser were too small for many of the statistics to be reliably calculated and their exploratory results, listed in source paper II, are not reproduced here.

On Regional Resemblances

All in all, the analysis of car trips shows great structural similarities among the 3 regions at the same time as some differences reveal the sophistication of travellers' trip planning.

The available samples in Figure 3 are all markedly asymmetric and of similar sizes clearly adequate for regression work. But, as the US values pertain to large metropolitan areas, they are in principle closer to the Greater Paris area sample than to the sample for Metropolitan France, which includes many small cities or towns, as its 30% lower average mean trip times implies. In fact, the US and Paris samples have, for both men and women, uncannily similar structures, as Figure 5 makes even clearer than do the distinct Figure 3 graphs and moment statistics.

FIGURE 5
DRIVE-ALONE A.M. PEAK WORK CAR TRIP DURATIONS
IN GREATER PARIS AND IN LARGE US STANDARD METROPOLITAN AREAS
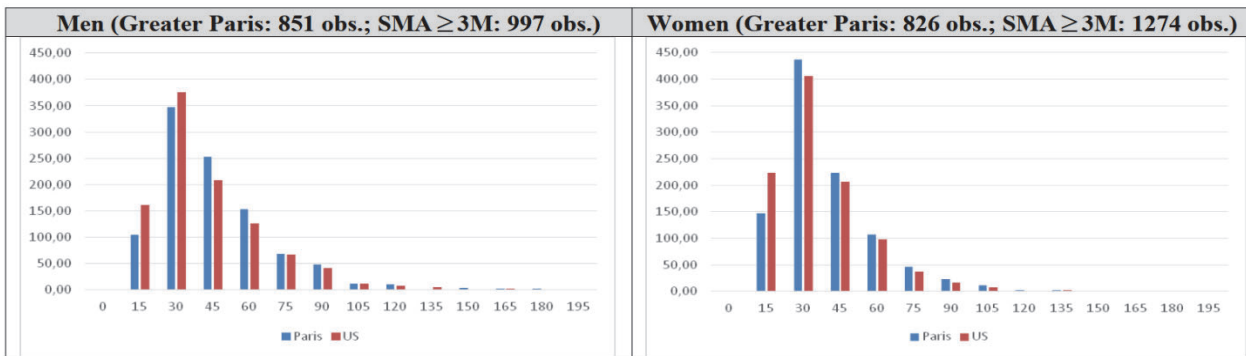
This remarkable basic fact has not prevented formal analysis of these durations from revealing statistically meaningful differences between them, either in the effects of service levels on duration choices or in their internal structures and in the MRS among the moments of duration. Both similarities

and differences concerning home-based work trips point to the importance of micro data by trip purpose, as opposed to city-wide averages, for the meaningful analysis of urban transport markets and policies in the future.

Transit Trip Durations: Unit Elasticity Tests, Moment Demand Elasticities and MRS

We used the mass transit data shown in Part D of Figure 3 for the Greater Paris Region, the only region to produce large enough mass transit use samples for work trips. This is partly due to our imposed requirement that, in the absence of face-to-face control interviews in US NPTS surveys, the retained traveler should always be the very person who answered by phone the relevant survey questions, which include those on journey departure and arrival times defining the dependent variable in (8-B).

Establishing the Existence of the Precautionary Palliative Time margin

Table 10 presents, like Table 8, selected model results: (i) the pair of BCT estimates (already commented on above) for trip duration and mean service variables; (ii) the 4 elasticities of the duration regressand with respect to mean service; (iii) *t*-statistics, unconditional for the two BCT form parameters and conditional on the BCT estimates for the β_k regression coefficient of the service regressor variable.

It has to be recalled from Table 6 that the average service variable in the equation explaining transit trips is represented by U, the denominator of a Standard strict utility Box-Cox Logit model explaining path shares p_m defined in (1), to which the BCT power $\lambda_S = \lambda_U$ is applied and which becomes a *Logsum* if $\lambda_U \to 0$. As this RATP-produced measure is differently defined and more complex than the simple path time indicator used for S^p in all variants of Table 8 car trip duration models, transit duration elasticities will be of the opposite sign and their analytical expressions different. For instance the sample elasticity of duration for a particular trip with respect to its path service, which was straightforwardly $\beta_k X_k^{\lambda_k} / y^{\lambda_y}$ (with $\hat{\beta}_k \le 0$) for car trips in Table 8, is now for public transit service a product of $\beta_k U^{\lambda_U} / y^{\lambda_y}$ (with $\hat{\beta}_k \ge 0$), the elasticity with respect to the anti-*Logsum* variable U found in Table 10, by $\beta_k X_{mk}^{\lambda_{mk}} p_m$, the elasticity of U with respect to in-vehicle time on path m.

In short, we have a sign reversal (seen by comparing Table 10 with Table 8) and the sample elasticity

$$\text{is not anymore} \quad (\beta_{X_k} X_k^{\lambda_k} / y^{\lambda_y}) \quad \text{but} \quad [(\beta_U U^{\lambda_U} / y^{\lambda_y})(\beta_{X_k} X_{mk}^{\lambda_{X_{mk}}} p_m)] \qquad (10)$$
$$\text{(i),} \qquad\qquad\qquad \text{(ii),}$$

where X_{mk} is attribute *k* (such as in-vehicle time) of itinerary *m*. The same holds for the 3 strict notions of elasticity: Table 10 only presents the first term of (10.ii) and never the required product[15].

TABLE 10
ELASTICITIES OF PT TRIP DURATION W.R.T. MEAN SERVICE
AND STUDENT'S *t* OF PARAMETERS, CASE 3 OF (8-D)

Home-based mass public transit (PT) work trips		Greater Paris Region	
(M= Men; W=Women):		**1. M**	**2. W**
Variant run number in source paper tables		17	31
Number of X_k variables except for β_0		21	21
Number of observations		979	1180
BCT	*BCT estimates and their unconditional Student t statistics*		
	on Duration λ_y	**0.19**	**0.42**
	Unconditional Student *t* (=0)	(3.44)	(14.07)
	on Service λ_x	**0.05**	**0.10**
	Unconditional Student *t* (=0)	(5.19)	(7.93)
Elasticity of duration with respect to endured Service U in Eq. (10)	*Casual sample elasticity of (y) as defined in Table 3*		
	$\eta^s(y, X_k) = \dfrac{\partial y}{\partial X_k} \cdot \dfrac{X_k}{y}$	-0,076	-0,100
	Student *t* (=0) conditional on BCT	(-31.97)	(-30.39)
	Strict elasticities of moments of (y) as defined in Table 3		
	$\eta^e(y, X_k) = \dfrac{\partial e(y)}{\partial X_k} \dfrac{X_k}{e(y)}$	-0,081	-0,110
	$\eta^\sigma(y, X_k) = \dfrac{\partial \sigma(y)}{\partial X_k} \dfrac{X_k}{\sigma(y)}$	-0,065	-0,067
	$\eta^\gamma(y, X_k) = \dfrac{\partial \gamma(y)}{\partial X_k} \dfrac{X_k}{\gamma(y)}$	0,016	0,048
Log-Likelihood (LL) Case 4 (Log-Log)		-4039	-5027
Log-Likelihood (LL) Case 3 (Box-Cox)		**-4014**	**-4909**
LL difference (2 degrees of freedom)		25	118
Pseudo-(L)-R^2		**0.57**	**0.53**

When these calculations are done in source paper III, all elasticities, precisely estimated because the *t*-statistics of underlying service variable regression coefficients are equal to about 30, are again smaller than 1. In the upper ranges of path time values, elasticities of the first two moments are of the same order of magnitude as in Table 8 but skewness is slightly more sensitive to improved (or to worsened) service for transit trip durations than for car trip durations, especially for women.

Improved Service and the Duration Choice Reliability Echo
First, it should be noted that the optimal power of the *U* term in Table 10 is close to 0 but is in fact statistically different from it with values of 0.05 or 0.10. The *Logsum* is therefore rejected here, in favor of the simple sum with a BCT power parameter, as the most adequate summary of transit path service.
The weakness of the elasticities of skewness relative to those of the first two moments again implies, as it did for car trips, that improved path service X_{mk} (requiring the minus sign) concentrates the time profiles of duration demand. The relevant expression is identical to (9-A):

$$-\left[\eta(e(Dur_n), X_{mk}) + \eta(\sigma(Dur_n), X_{mk}) + \eta(\gamma(Dur_n), X_{mk})\right] \Rightarrow SUPERPEAK \ of \ kurtosis \ \kappa \qquad \text{(9-B)}$$

but this relative super-peak effect on kurtosis derived from Table 10 is a little weaker for transit services (even if women exert some particularly strong moderating influence through skewness) than it was for car services in Table 8. This increased 4th moment of the time profile of demand, or metronomic acuity effect, might sometimes pose capacity problems for transit planning (even at constant demand for the time period considered). Also, the decreased 2nd moment of improved transit services can still be said to imply fast fill-ups, as do new urban highways.

Substitution Among Moments of Duration Choices

Table 11 presents statistics on marginal rates of substitution among moments comparable to those found in Table 9, but in this case the additional information on elasticities of substitution and fits by moment have been kept.

TABLE 11
WORK PT TRIP DURATION MOMENT RATES
AND ELASTICITIES OF SUBSTITUTION, CASE 3 OF (8-D)

Mass public transit trips (PT) Rates & elasticities of substitution		Greater Paris Region work trips					
		Men (variant 17)			Women (variant 31)		
Moments	i\j	e	σ	γ	e	σ	γ
MRS^{m_e,m_σ}, MRS^{m_e,m_γ}	e	1	+ 4.5	− 330	1	+ 5.4	− 180
MRS^{m_σ,m_γ}	σ		1	− 73		1	− 33
Eq. (5-D.1) to (5-D.3)	γ			1			1
η^{m_e,m_σ}, η^{m_e,m_γ}	e	1	+ 1.2	− 5.2	1	+ 1.7	− 2.4
η^{m_σ,m_γ}	σ		1	− 4.2		1	− 1.4
Eq. (5-D.1*) to (5-D.3*)*	γ			1			1
Sample moment value		61	24	1.00	60	24	0.75
Fitted value at the means		42	13	0.68	42	13	0.56
Mean of fitted values		61	18	0.71	60	17	− 0.08
* Defined in Appendix 7.							

Note that, although the marginal rates of substitution are extremely close to those of car trips (shown in Table 9 and Appendix 7), the PT elasticities of substitution between the first and the third moments shown in the bottom Part of Table 11 are roughly half of the size of the PC ones found in Appendix 7, namely (-5.2 ; -2.4) *vs* (-8.2 ; -4.1), as are those between the second and the third moments, namely (-4.2; -1.4) *vs* (-7.2 ; - 3.1), implying a greater sensitivity of trip planning responses in drivers despite similar valuations of time moments.

Also, again, the higher the moment, the less well it is fitted if the fit is measured by the simplistic *Fitted value evaluated at the means* of all variables (at the bottom of Table 11). Clearly, reasonable indices of fit have yet to be developed for all moments in order to extend to higher moments than the first appropriate measures for nonlinear models in the spirit of the long-used Pseudo-(L)-R^2 (Aigner, 1971, p. 85-90) calculated in our tables.

On Modal Resemblances

All in all, results obtained for transit trips are extremely close to the results obtained for car trips in the Greater Paris Region and their small differences again show, if anything, the sophistication of travelers' journey planning and time moment valuations.

Concerning car trips, we showed that network improvements apparently do little to speed commutes because they notably generate super-peaks that are part of the solution, not the problem as one simplistically assumes by wearing *first-moment-of-travel-time-blinkers*. The situation is identical here, in that transit improvements also increase the acuity of trip time distribution profiles, as travelers reduce precautionary margins. Bemoaning modal project effects, in witch-hunts for mean speed changes that ignore its higher (echo-generating) moments, commits a new "certain error of a mean".

CONCLUSION: ON STUDENTS COMMITTING A NEW "CERTAIN ERROR OF A MEAN"

Safety Margins by Urban Home-Based Journey Market

First, we have demonstrated that, given a chosen transport mode, any empirical telic urban trip duration T combines a planned palliative endogenous precautionary margin component I, a complementary offset to the uncertainty of the exogenous component, the expected endured network service time S. The existence demonstration of this inferred composition of terms, inspired by Frank Knight's distinction between uncertainty and risk here applied to the margin and the expectation of endured travel service, relied on the estimation of an infra-unitary elasticity of duration with respect to service obtained from flexible form Box-Cox regressions explaining actual duration T by service S and other variables X.

The elasticity test was readily passed by estimates secured from such duration demand equations by traveler sex for different modes (car drive alone; mass transit), trip purposes (work; shopping), regions (Greater Paris; large American Standard Metropolitan Areas; Metropolitan France) and US socio-cultural backgrounds (White; Black or Hispanic). For each such market, two broad sets of calculated derived results of interest have been focused on.

Relative Moment Valuations

In the first set of results, we extracted from explained endogenous durations T the marginal rates of substitution among their concatenated moments of various orders (no greater than the fourth, and all determinate in deference to Christian Berg's indeterminacy finding), understood without formal proof to match Maurice Allais' famous rates of substitution among all (determinate) moments of a random prospect with respect to the first. That is because we assume that T constructs reflect preferences for actual continuous distributions of trip time characterized solely by their empirical moments, without any resort to their exact analytical shape or even concern for their possible lack of unicity, a so-called Multiple Moment Dependent Utility (MMDU) modeling posture.

The internal structure of such traveler-built multi-moment duration concatenations varies notably by market but they all reveal risk aversion and generally indicate that: (a) women tend to value the second moment more, and the third less, than men; (b) for work trips, the second moment is more valued in large US cities and in Metropolitan France than in the Greater Paris Region, a trade-off shown to be socio-culturally sensitive and to imply locally different "fast fill-up" effects for new infrastructure; (c) trips to shopping centers reveal lower reliability moment valuations than work trips.

Absolute Moment Demands and the Preference for Traffic Super-Peaks

In the second set of results, the sensitivity of the constructed duration moment concatenations has been measured by derivatives (and elasticities) of the first three moments of durations T with respect to service S, first and foremost of the explanatory variables. As only its first moment is available here, duration moment responses to changes in mean service close the traffic model, originally built under the assumption of a certain distribution of home departure times, by modifying that continuous distribution profile. We hope to analyze duration reactions to higher moments of service when the data are available and expect results similar to those found for the mean.

Currently, we find that improved mean origin-destination service (speed) reduces mean trip durations, but less than proportionately, through safety margin time adjustments and that these utility gains from mean travel time are always augmented by concomitant changes in the three higher time reliability

moments. They were here measured by elasticities, analytical (2nd and 3rd moments) or qualitative (4th moment). With shorter durations, users show up in bunched manner (the so-called fast fill-up effect) and push up the acuity, or "peakedness", of the super-peak (a desirable metronomic effect) but slightly decrease prudence (a counter effect). If speed falls, the opposite occurs. Super-peaks are then not part of the problem but of the multi-moment demand solution and their suppression will fail, misconceived due to *first-moment-of-travel-time-blinkers*. But exactly how is a mean analysis certainly erroneous?

New Student Error

Again, mean service betterment shortens average trip duration and induces further reliability reactions. These jointly amount to a "reliability echo" boost, a new "error of a mean", that differs across markets. Its role is ignored in project assessment studies based on Q&MTT models limited to first moment time effects analyzed only on aggregate flows, without due concern for the multi-moment composition and mix of trips that are quite distinct by purpose and traveler market.

ACKNOWLEDGEMENTS

ENDNOTES

1. Abraham formulated the problem of the binary choice probability between alternatives described by utility functions with errors and then derived a Probit under the assumption of their Normal distributions and the Linear Probability model under that of their Rectangular distributions. McFadden repeated these steps 11 years later and added derivations of the Logit from independent Weibull distributions and of the Arctan from Cauchy distributions of the same utility function errors.

2. Koenig (1974, 1975) first showed that the logarithm of its denominator, or *Logsum*, denotes the average utility of the p_i.

3. If y is normally distributed and $\alpha > 0$, the distribution of $|y|^\alpha$ is determinate for $0 < \alpha \leq 4$ and indeterminate for $\alpha > 4$. That demonstration by the author must also hold for $|y|^{(\lambda)}$, at least with $0 < \lambda \leq 4$ and perhaps even with BCT $0 \leq \lambda \leq 4$, the BCT being applied only to strictly positive y samples, as done below in (6-A).

4. Declared by Jean-Martin Charcot in his lectures between October 13, 1885 and February 28, 1886, dates of Sigmund Freud's stay in Paris and attendance of Charcot's lectures. Freud quoted this statement by

Charcot, made in response to a student claiming that a particular theory ruled out the possibility of one of Charcot's claims, numerous times and notably in his eulogy of Charcot in 1893 (Freud, 1984, p. 23).

5. If the function *f* in (4-A) is additive, the duration concatenation $\left[r_1\hat{e}(\hat{T}_n) + r_2\hat{\sigma}(\hat{T}_n) + r_3\hat{\gamma}(\hat{T}_n) \right]$ may be viewed as a joint choice of ($\hat{d}_{e(n)} = \hat{t}_{e(n)} + \overline{t}_{e(n)}$), ($\hat{d}_{\sigma(n)} = \hat{t}_{\sigma(n)} + \overline{t}_{\sigma(n)}$) and ($\hat{d}_{\gamma(n)} = \hat{t}_{\gamma(n)} + \overline{t}_{\gamma(n)}$), with Allais' *r(y)* valuation weights applied.

6. Ekelund & Hébert (1999) suspect that Marshall's use of the sample notion in his *Principles of Economics* of 1890 was a case of plagiarism by an author very familiar with French ENPC authors and even quoting Dupuit. Should his then faked story of an inspired *ηὕρηκα!* wake-up in the middle of the night in a Palermo hotel count towards the "*Entente cordiale*"?

7. As documented in the program manual of the method we purport to use (Tran *et al.* 2008, Section 2.5), this surprise can be partly understood by noticing that any effect on a given moment mechanically implies structured effects on higher moments because, once $\partial e(y)/\partial X_k$ is known, it is easy to deduce $\partial \sigma(y)/\partial X_k$ and $\partial \gamma(y)/\partial X_k$ with Jacobians of transformations from one moment to the next.

8. In Box-Cox regression, the residual cannot be strictly normally distributed because *y* must be strictly positive. Should one then specify a weighted Likelihood function to take the probability to be below the lower limit (and symmetrically for the upper limit) into account? No: macroeconomic models explaining national GNP components C, I or G should not be rejected on the grounds that their errors cannot be strictly normally distributed. Consequently, our algorithm is constructed in the following way: (i) one first assumes normality and the *a priori* absence of limit observations, in which case the Likelihood function of the two-limit Tobit specified by Rosett & Nelson (1975) reduces to that of Box & Cox (1964) themselves; (ii) one then verifies *ex post* the reasonableness of the latter assumption by calculating an index of the probability (see Appendix 1, Part III, line 3 of source paper IV) of each fitted value to be at the limits   and   (defined by the user) as recommended by Olsen (1978) on the lines of remarks by Draper & Cox (1969) on this point.

9. It is possible, in certain circumstances, to apply a BCT to an X_k containing zeroes, as we do in the source papers.

10. Tobin's original formulation (Tobin, 1958) had different limits across observations. It does not require that the sample actually contain limit observations: our Likelihood function formulation assumes none are present.

11. The application of OLS in Case 1 involves no transformation of *y*. If it is transformed, only the intercept is rescaled.

12. A Normal law has 2 parameters (mean and variance) but the Poisson law has only one due to the equality constraint (mean of *y*) = (variance of *y*). The general heteroskedasticity formulation that allows to impose the Poisson restriction on our model (6-A) requires an additional equation for the heteroskedasticity of w_t (*cf.* the discussion in Gaudry & Dagenais, 1979, or the program manual by Tran *et al.*, 2008), not shown here, as demonstrated by Fridstrøm (1999, 2000).

13. Censuring short trips, as in Figure 3, to limit the effect of rounded stated departure and arrival times, makes women's trip durations closer to men's than they are when all trips are shown (*e.g.* Picard *et al.*, 2013, Table 1, for both modes).

14. This includes Small *et al.* (2005) who note women's preference for tolled highway options, as opposed to free ones, a result consistent with their stronger dislike of driving than men.

15. It is calculated in source paper III for two reference path times (36,6 and 12,2 minutes) and, to compute *U*, with the coefficient of in-vehicle time from the RATP 6-path model (-0,22), with in-vehicle time at 24,4 minutes and out-of-vehicle time at 27,5 minutes (both obtained from transit assignments).

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.

Aigner, D. J. (1971). *Basic Econometrics.* Prentice-Hall, Englewood Cliffs, N. J.

Allais, M. (1953). Le Comportement de l'Homme Rationnel Devant le Risque, Critique des Postulats et Axiômes de l'École Américaine. *Econometrica*, 21(4), 503–546.

Allais, M. (1987). Allais Paradox. In Atwell, J., Milgate, M. & Newman, P. (eds), *The New Palgrave Dictionary of Economics*, 1, 80-82. The Macmillan Press, London.

Allen, R. G. D. (1968). *Macro-Economic Theory: A Mathematical Treatment*. Macmillan, p420.

Ashenfelter, O., & Greenstone, M. (2004). Using Mandated Speed Limits to Measure the Value of a Statistical Life. *Journal of Political Economy*, 112(1), 226-267.

Beesley, M. E. (1965). The Value of Time Spent in Travelling: Some New Evidence. *Economica*, 32, 174-185.

Beesley, M. E. (1970). The Value of Time Spent in Travelling: Some New Evidence. Ch. 11 in Quandt, R.E. (ed). *The Demand for Travel: Theory and Measurement*, pages 221-234. Heath Lexington Books, Lexington, Massachusetts.

Belsley, D., Kuh, E., & Welsh, R. E. (1980). *Regression Diagnostics*. John Wiley, New York.

Berg, C. (1988). The cube of a normal distribution is indeterminate. *The Annals of Probability*, 16(2), 910-913.

Berndt, E. K., Hall, B. H., Hall, R. E., & Hausman, J. A. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, 3, 653–665.

Blondel, M. (1904). *Histoire et Dogme*. La Quinzaine, Montligeon.

Blum, U., & Gaudry, M. (2000). The SNUS-2.5 Model for Germany. Ch. 3 in Gaudry, M. & Lassarre, S. (eds), *Structural Road Accident Models: The International DRAG Family*. Pergamon, Elsevier Science, Oxford, 67-96.

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

Brockett, P. L., & Garven, J. R. (1998). A Reexamination of the Relationship between Preferences and Moment Orderings by Rational Risk Averse Investors. *Geneva Papers on Risk and Insurance Theory*, 23(2), 127-137.

Cirillo, C. (2005). *Optimal Design for Mixed Logit Models*. Groupe de recherche sur les transports, Facultés Universitaires Notre-Dame de la Paix, Namur, p11. Association for European Transport and contributors.

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.

Dagenais, M. G., & Dufour, J.-M. (1994, August). Pitfalls of Rescaling Regression Models with Box-Cox Transformations. *The Review of Economics and Statistics*, 76(3), 571-75.

Dagenais, M. G., Gaudry, M. J. I., & Liem, T. C. (1987). Urban Travel Demand: The Impact of Box-Cox Transformations with Nonspherical Residual Errors. *Transportation Research B*, 21(6), 443-477.

Davidson, R., & MacKinnon, J. G. (1985, August). Testing linear and loglinear regressions against Box-Cox alternatives. *The Canadian Journal of Economics / Revue canadienne d'économique*, 18(3), 499-517.

de Lapparent, M., Frei, A., & Axhausen, K. W. (2009, October). *Choice of mode for long distance travel: current SP-based models from three European countries*. 38 pages, July. Paper presented at the European Transport Conference, Leeuwenhorst.

Demsetz, H. (1970, October). The private production of public goods. *Journal of Law and Economics,* 13(2), 293-306.

Draper, N. R., & Cox, D. R. (1969). On distributions and their transformations to normality. *Journal of the Royal Statistical Society, Series B*, 31, 472-476.

Dupuit, J. (1844). De la mesure de l'utilité des travaux publics. *Annales des Ponts et Chaussées,* 2° série, Mémoires et Documents, n°116, t. VIII, pp 332-375.

Dutzig, T., Bradford, A., Weissman, G., Frontier Group, Aragon, L., & U.S. PIRG Education Fund. (2017, April). *Highway Goondoggles 3. Big Projects, Bigger Price Tags, Limited Benefits*. U.S. PIRG Education Fund and Frontier Group, p47.

Ekelund, R. B., & Hébert, R. F. (1999). Secret Origins of Modern Microeconomics: Dupuit and the Engineers. *The University of Chicago Press*, p468.

Erkel-Rousse, H. (1995). Détection de la multicolinéarité dans un modèle linéaire ordinaire: Quelques éléments pour un usage averti des indicateurs de Belsley, Kuh et Welsch. *Revue de Statistique Appliquée*, 43(4), 19–42.

FCAR. (1991). Rapport d'activité sur les 10 mois d'opération du Programme multidimensionnel en sécurité routière, Projet 5.2. Estimations dans le cadre du modèle DRAG, Tâche 1. Action concertée sur la sécurité routière FCAR-67. Fonds pour la formation des chercheurs et l'aide à la recherche. Québec.

Fosgereau, M. (2015, September 15). The valuation of travel time variability. Draft Discussion Paper 2015-04, International Transport Forum, *OECD*, p22.

Foster, C. D., & Beesley, M. E. (1963). Estimating the Social Benefit of Constructing an Underground Railway in London. *Journal of the Royal Statistical Society, Series A*, 126, 46-79.

Freud, S. (1984). *Freud présenté par lui-même [1925-1935]*. Gallimard, Paris.

Fridstrøm, L. (1999). *Econometric Models of Road Use, Accidents, and Road Investment Decisions*. Report 457 292, Institute of Transport Economics, Norwegian Centre for Transport Research, TØI, Oslo.

Fridstrøm, L. (2000). The IRPOSKML Procedure of Estimation. Ch. 13 in Gaudry, M. & Lassarre, S. (eds). *Structural Road Accident Models: The International DRAG Family*. Pergamon, Elsevier Science, Oxford, 325-334.

Gallant, A. R. (1981). On the Bias in Flexible Forms and an Essentially Unbiased Form: the Fourier Flexible Form. *Journal of Econometrics*, 15(2), 211-245.

Gaudremeau, J., & Gaudry, M. (2017a). *La demande de durée et de fiabilité des trajets domicile-travail à l'heure d'affluence et la mesure de l'utilité des hyperpointes de trafic américaines et françaises*. Publication AJD-164, Agora Jules Dupuit, Université de Montréal, p50, le 3 août.

Gaudremeau, J., & Gaudry, M. (2017b). *New urban road fast fill-up and traffic super-peak echoes in metropolitan areas of the United States and France*. Publication AJD-165, Agora Jules Dupuit, Université de Montréal, p50, le 4 août.

Gaudry, M. (1997, 2000). Some Perspectives on the DRAG Approach and Family of National Road Safety Models. Ch. 12 in Holst, H.v., Nygren, Å. & Andersson, Å.E. (eds). *Transportation, Traffic Safety and Health – Prevention and Health*. Third International Conference, Washington, D.C., U.S.A, 1997, pp. 123-168. Springer Verlag, Heidelberg, p198, 2000.

Gaudry, M. (2006). Life, Limb and Bumper Trade-Offs Calculable from Road Accident Models: An Empirical Multi-Moment Portfolio Analysis and Life Asset Pricing Model (LAPM). *Transport Reviews*, 26(4), 501-520.

Gaudry, M. (2010). *Quebec-Windsor Corridor High Speed Rail Market Forecast Profiles in Context: Level-of-Service Response Curvature Sensitivity and Attitude to Risk or to Distance in Forty Logit Core Model Applications of the Law of Demand*. 100 pages, Publication AJD-127, Agora Jules Dupuit, Université de Montréal and Working Paper DEST N° 12, Institut National de Recherche sur les Transports et leur Sécurité, Université Paris-Est.

Gaudry, M. (2015). *La demande de durée et de fiabilité des trajets de transport: tests d'une méthode à trois moments sur les durées domicile-travail constatées en Île-de-France*. Rapport d'expertise déposé en version finale par la Mission d'appui scientifique à la Société du Grand Paris pour l'utilisation des modèles de transport (MAPA 2013-32), Étape 2, Oikometra (Montréal) et Stratec

(Bruxelles), 134 p., le 28 décembre 2015. Publication AJD-153, Agora Jules Dupuit, Université de Montréal.

Gaudry, M. (2016a). *The demand for journey duration and reliability: Paris Region work trips by mode and sex, 2010-2011*. Publication AJD-154, Agora Jules Dupuit, Université de Montréal, p 35, Version 2 of 11th January.

Gaudry, M. (2016b, January). Appendix 1 of: Gaudry, M. (2016). *The demand for journey duration and reliability: Paris Region work trips by mode and sex, 2010-2011*. Publication AJD 154, Agora Jules Dupuit, Université de Montréal. p19.

Gaudry, M. (2016c). *Préférences MDU pour des distributions caractérisables par leurs seuls moments: applications aux transports et réinterprétation des fonctions de pondération des probabilités et de valeur en théorie cumulative des perspectives*. Publication AJD-156, Agora Jules Dupuit, Université de Montréal, p25, le 8 mars.

Gaudry, M. (2018a, November). The utility of journeys, from Dupuit's constant-time bridge crossing hops to commutes of chosen duration and reliability in the Paris region. *Transport Policy*, 70(15), 53-68.

Gaudry, M. (2018b). *Manifeste sur l'utilité des pointes et hyperpointes de trafic (ou sur les bons moments du temps de transport)*. Publication AJD-176, Agora Jules Dupuit, Université de Montréal, p15, 1er juin.

Gaudry, M. (2018c, September). The problem of congestion unreliability cost: The certain error of a mean and the measure of travel time utility gains from public works. Publication AJD-167, Agora Jules Dupuit, Université de Montréal, 4p3. Forthcoming in *Journal of Strategic Innovation and Sustainability*.

Gaudry, M., & Dagenais, M. G. (1979). Heteroscedasticity and the Use of Box-Cox Transformations. *Economics Letters*, 2(3), 225-229.

Gaudry, M., & Laferrière, R. (1989). The Box-Cox Transformation: Power Invariance and a New Interpretation. *Economics Letters*, 30(1), 27-29.

Gaudry, M., & de Lapparent, M. (2015, January). *Attitudes to Distance, Time and Cost in Logit Transport Choice Models*. Publication AJD-146, Agora Jules Dupuit, Université de Montréal and Report TRANSP-OR 150113, Transport and Mobility Laboratory, École Polytechnique Fédérale de Lausanne, p10.

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

Gaver D. P. (1968). Headstart strategies for combating congestion. *Transportation Science*, 2(3), 172-181.

Goldberger, A. S. (1968). The interpretation and estimation of Cobb-Douglas functions. *Econometrica*, 36(3), 464-472.

Griffiths, T., & Tenenbaum, J. (2006, September). Optimal predictions in everyday cognition. *Psychological Science*, 17(9), 767-773.

Gut, A. (2002). On the moment problem. *Bernouilli*, 8(3), 407-421.

Hendrickson, C., & Plank, E. (1984). The flexibility of departure times for work trips. *Transportation Research A,* 18(1), 25-36.

Hensher, D. A., & Johnson, L. W. (1979, December). A Random Coefficient Model of the Determinants of Frequency of Shopping Trips. *Australian Economic Papers*, 18(33), 322-336.

Jackson W. B., & Jucker J. V. (1982) An empirical study of travel time variability and travel choice behaviour. *Transportation Science*, 16(4), 460-475.

Johnson, L. W. (1977). Stochastic parameter estimation: an annotated bibliography. *International Statistical Review / Revue Internationale de Statistique*, 45(3), 257-272.

Johnson, L. W. (1978). Regression with Random Coefficients. *Omega*, 6(1), 71-81.

Johnson, L. W. (1979). An introduction to generalised functional form and random coefficients in transport modelling. *Environment and Planning A*, 11(9), 1029-1037.

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

Kane, E. J. (1968). *Economic Statistics and Econometrics, An Introduction to Quantitative Economics*. Harper & Row, p437.

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Houghton, Mifflin.

Knight, T. E. (1974). An approach to the evaluation of changes in travel unreliability: a "safety margin" hypothesis. *Transportation*, 3(4), 393-408.

Koenig, G. (1974). Théorie économique de l'accessibilité urbaine, *Revue Économique*, 25(2), 275-297.

Koenig, J-G. (1975). *A theory of urban accessibility: a new working tool for the urban planner*. Urban Traffic Models, Annual Summer Meeting, Planning and Transport Research and Computation Co, Warwick.

Koster, P., Kroes, E., & Verhoef, E. (2011). Travel time variability and airport accessibility. *Transportation Research B*, 45(10), 1545-1559.

Kouider, E., & Chen, H. (1995). Concavity of Box-Cox log-likelihood function. *Statistics and Probability Letters,* 25(2), 171-175.

Lam, T. C., & Small, K. A. (2001). The value of time and reliability: measurement from a value pricing experiment, *Transportation Research Part E*, 37(2-3), 231-251.

Lave, C. A. (1985, December). Speeding, Coordination and the 55 MPH limit. *The American Economic Review*, 75(5), 1159-1164.

Leblond, V., & Langlois, C. (2013). *Modelling detailed passengers' behaviours in a public transportation network*. DIT-IDD-PRI, Régie Autonome des Transports Parisiens (RATP), p. 12, presented at the European Transport Conference, Frankfurt, Sept. 30-Oct. 2, 2013.

Mandel, B. (2014, April). *Contemporary Airport Demand Forecasting Choice Models and Air Transport Forecasting*. OECD/ITF Roundtable on Forecasting Airport Demand, Paris, France. Discussion Paper No. 2014-07, p35.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5), 447-470.

Ng, W-S., & Aker, A. (2018, February). *Understanding Urban Travel Behaviour by Gender for efficient and Equitable Transport Policies*. ITF Discussion Paper No. 2018-01, OECD/International Transport Forum, Paris, p19.

Olsen, R. J. (1978). *The method of Box and Cox, a pitfall*. Institution for Social and Policy Studies, Yale University, Working Paper No 804.

Orro, A., Novales, M., & Benitez, F. G. (2005). *Nonlinearity and Taste Heterogeneity Influence on Discrete Choice Model Forecasts*. Association for European Transport and contributors, p18.

Orro, A., Novales, M., & Benitez, F. G. (2010, September). *Box-Cox mixed Logit model for travel behaviour analysis*. Paper presented at the 8th International Conference of Numerical Analysis and Applied Mathematics, Rhodes, 4 pages.

Picard, N., de Palma, A., & Dantan, S. (2013). Intra-household discrete choice models of mode choice and residential location. *International Journal of Transport Economics*, 40(3), 419-445.

Picard, N., Dantan, S., & de Palma, A. (2018, March). Mobility decisions within couples. *Theory and Decision*, 84(2), 149-180.

Polak, J. (1987). A more general model of individual departure time choice. *PTRC Summer Annual Meeting, Proceedings of Seminar C*.

Prat, C., & Leblond, V. (2014, September 29-October 1). *Passengers' density impact on route choice in urban public transport*. DIT-IDD-PRI, Régie Autonome des Transports Parisiens (RATP), p21, presented at the European Transport Conference, Frankfurt.

Prato, C., Rasmussen, T. K., & Nielsen, O. A. (2014). Estimating Value of Congestion and of Reliability from Observation of Route Choice Behavior of Car Drivers. *Transportation Research Record* No. 2412, Transportation Research Board of the National Academies, Washington, D.C., pp. 20–27.

Rassam, P. R., Ellis, R. H., & Bennett. J. C. (1970, 1971). The N-Dimensional Logit Model: Development and Application. Peat Marwick and Mitchell & Co, Washington, D.C., p42, 1970. Also in *Highway Research Record*, 369, 135-147, 1971.

Rich, J. (2018, February). A spline function class suitable for demand models. *Econometrics and Statistics*, p29.

Rich, J., & Mabit, S. L. (2016). Cost damping and functional form in transport. *Transportation*, 43(5), 889-912.

Rosett, R. N., & Nelson, F. D. (1975). Estimation of the two-limit Probit regression model. *Econometrica*, 43(1), 141-146.

Samuelson, P. A. (1954, October). The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*, 36(4), 387-389.

Samuelson, P. A. (1955, November). Diagrammatic Exposition of a Theory of Public Expenditure. *The Review of Economics and Statistics*, 37(4), 350-356.

Samuelson, P. A. (1969, February). Contrast Between Welfare conditions for Joint Supply and for Public Goods. *The Review of Economics and Statistics*, 51(1), 26-30.

Scott, R. C., & Horvath. P. A. (1980). On the Direction of Preference for Moments of Higher Order than the Variance. *Journal of Finance*, 35(4), 915-919.

Senna, L. A. D. S. (1994). The influence of travel time variability on the value of time. *Transportation*, 21(2), 203-228.

Seshadri, R., & Srinivasan, K. K. (2017): Robust traffic assignment model: Formulation, solution algorithms and empirical application. *Journal of Intelligent Transportation Systems*, 21(5), 1-18.

Setec. (1959). *Tunnel sous la Manche – Étude de Trafic et de Rentabilité : Partie B – Annexe*. The Economist Intelligence Unit Ltd (Londres), De Leuw, Cather and C^o (Chicago) et Société d'Études Techniques et Économiques (Paris), pages iii-ix (pages 196-202, Volume 3), novembre.

Sévène, L. (1877). *Notes prises par les élèves au cours de chemin de fer, session 1876-1877*. École Nationale des Ponts et Chaussées, Paris.

Small, K. A., Winston, C., & Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4), 1367–1382.

Spitzer, J. J. (1984). Variance Estimates in Models with the Box-Cox Transformation: Implications for Estimation and Hypothesis Testing. *The Review of Economics and Statistics*, 66(4), 645-652.

Starkie, D. N. M. (1971). Modal Split and the Value of Time - a Note on Idle Time. *Journal of Transport Economics and Policy*, 5(2), 216-220.

Swamy, P. A. V. B. (1970). Efficient Inference in a Random Coefficient Regression Model. *Econometrica*, 38(3), 311-323.

Tavernier, R. (1889). Note sur les principes de tarification et d'exploitation du trafic voyageurs. *Annales des Ponts et Chaussées, Mémoires et Documents, 6ème série*, 17(2), 559-564.

Tobin, J. (1957/1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25(2), 65-86.

Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica*, 26(1), 24-36.

Tobin, J. (1965). The Theory of Portfolio Selection. Ch. 1 in Hahn, F. & Brechling, F. (eds), *The Theory of Interest Rates*, 3-51. Macmillan, London.

Tran, C-L., Gaudry, M., & Dagenais, M. (2008, April). *LEVEL: The L-1.6 estimation procedures for BC-GAUHESEQ (Box-Cox Generalized AUtoregressive HEteroskedastic Single EQuation) regression and multi-moment analysis*. Publication AJD-105, Agora Jules Dupuit, Université de Montréal, p59.

Vickrey, W. S. (1969). Congestion theory and transport investment. *American Economic Review*, 59(2), 251-261.

von Nördling, W., (1886). Le prix de revient des transports par chemin de fer et la question des voies navigables en France, en Prusse et en Autriche. *Annales des Ponts et Chaussées, Mémoires et Documents, 6ème Série*, 11(1), 292-303.

Warner, S. L. (1962). *Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice*. Northwestern University Press, Evanston, III.

White, M. J. (1977). A model of residential location choice and commuting by men and women workers. *Journal of Regional Science*, 17(1), 41-52.

White, M. J. (1986, May). Sex differences in urban commuting patterns. *The American Economic Review*, 76(2), 368-372.

APPENDIX 1. ANALYTICALLY DEFINED MOMENTS OF A RANDOM VARIABLE

TABLE 12
DEFINITONS OF THE FIRST AND OF THE NEXT THREE
CENTRAL MOMENTS OF A RANDOM VARIABLE y

1st	Mean	$e = E(y_g)$	2nd	Standard error	$\sigma = \left\{ E\left[(y_g - E(y_g))^2 \right] \right\}^{1/2}$
3rd	Skewness	$\gamma = \dfrac{E\left[(y_g - E(y_g))^3 \right]}{\sigma^3}$	4th	Kurtosis (Fisher)*	$\kappa = \left\{ \dfrac{E\left[(y_g - E(y_g))^4 \right]}{\left\{ E\left[(y_g - E(y_g))^2 \right] \right\}^2} \right\} - 3 = \dfrac{e_4}{\sigma^4} - 3$

*Fisher's definition, with the shift of -3, implies a kurtosis of zero for the Normal distribution, in contrast with Pearson's for which that distribution has a kurtosis of 3: it therefore denotes acuity in excess of that the Normal distribution.

APPENDIX 2. CONCATENATION OF TRIP TIME MOMENT DEMAND SCHEDULES

FIGURE 6
DURATION DEMAND, A VERTICAL CONCATENATION OF
MOMENT OF TIME DEMAND SCHEDULES

The demand for intended duration $D(T_c^*)$ is a vertical sum of demands (with slopes alternating in signs) for standard error (d_σ), mean (d_ε) and skewness [D-($d_\sigma + d_\varepsilon$)].
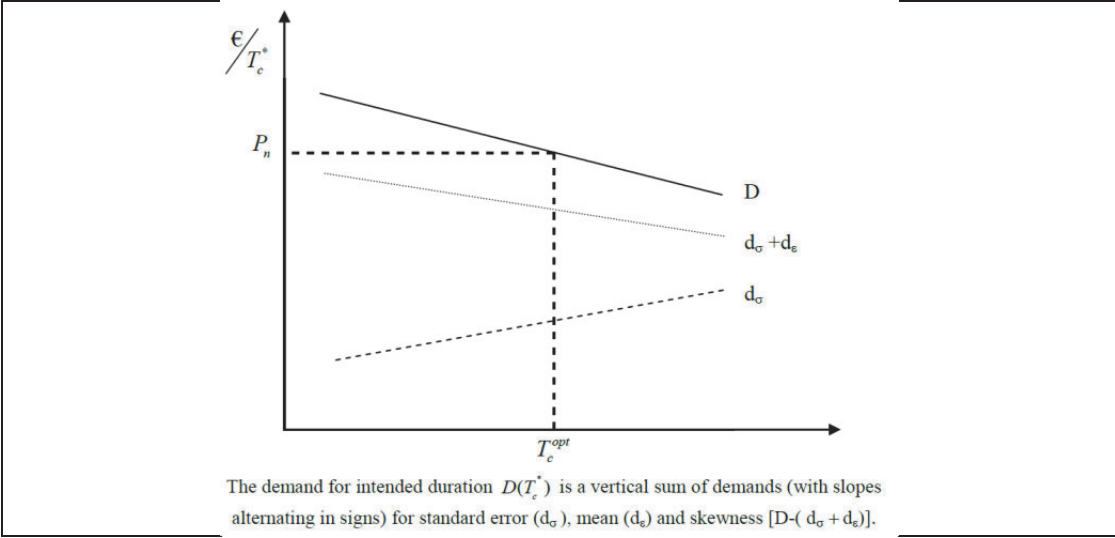
APPENDIX 3. ELASTICITY NOTIONS FOR CONTINUOUS AND DUMMY VARIABLES

TABLE 13
FOUR NOTIONS OF THE ELASTICITY η OF y WITH RESPECT TO A CONTINUOUS OR A DUMMY VARIABLE X_k

Name of elasticity concept		A. Point elasticity for ordinary continuous variable X_k	B. Arc elasticity approximation* for dummy variable X_k
1.	Sample elasticity of y with respect to X_k	$\eta^s(y, X_k) = \dfrac{\partial y}{\partial X_k} \cdot \dfrac{X_k}{y}$	$\eta^{s+}(y, X_k^+) = \eta^s(y, X_k) \cdot \dfrac{\bar{X}_k^+}{\bar{X}_k}$
2.	Elasticity of $e(y)$ with respect to X_k	$\eta^e(y, X_k) = \dfrac{\partial e(y)}{\partial X_k} \dfrac{X_k}{e(y)}$	$\eta^{e+}(y, X_k^+) = \eta^e(y, X_k) \cdot \dfrac{\bar{X}_k^+}{\bar{X}_k}$
3.	Elasticity of $\sigma(y)$ with respect to X_k	$\eta^\sigma(y, X_k) = \dfrac{\partial \sigma(y)}{\partial X_k} \dfrac{X_k}{\sigma(y)}$	$\eta^{\sigma+}(y, X_k^+) = \eta^\sigma(y, X_k) \cdot \dfrac{\bar{X}_k^+}{\bar{X}_k}$
4.	Elasticity of $\gamma(y)$ with respect to X_k	$\eta^\gamma(y, X_k) = \dfrac{\partial \gamma(y)}{\partial X_k} \dfrac{X_k}{\gamma(y)}$	$\eta^{\gamma+}(y, X_k^+) = \eta^\gamma(y, X_k) \cdot \dfrac{\bar{X}_k^+}{\bar{X}_k}$

* In Column B, $\bar{X}_k$ denotes the sample mean of the dummy variable and $\bar{X}_k^+$ its sample mean recalculated only on its positive values (usually all equal to 1, but any arbitrary constant will do). These expressions for dummy variables are studied in Dagenais *et al.* (1987), their source.

TABLE 14
BOX-COX MODEL DERIVATIVES OF MOMENTS OF y_t WITH RESPECT TO X_{kt}, WITH SPHERICAL* w_t ERRORS

$$D^e_{X_{kt}} = \frac{\partial E(y_t)}{\partial X_{kt}}^{cf\ note} = \beta_k X_{kt}^{\lambda_k-1}\int_{R_w} y_t(w)\varphi(w)dw, \text{ where } R_w = \text{domain of integration}$$

$-\infty < y_t < \infty$	$\lambda_y = 1$	$\lim\limits_{w_t \to -\infty}\beta_k X_{kt}^{\lambda_k-1}\int_{w_t}^\infty \varphi(w)dw = \beta_k X_{kt}^{\lambda_k-1}$
$y_t > 0$		$\beta_k X_{kt}^{\lambda_k-1}\int_{w_t}^\infty \varphi(w)dw$
	$\lambda_y > 0$	$\beta_k X_{kt}^{\lambda_k-1}\int_{w_t}^\infty y_t\,\varphi(w)dw$
	$\lambda_y = 0$	$\beta_k X_{kt}^{\lambda_k-1}E(y_t)$
	$\lambda_y < 0$	$\beta_k X_{kt}^{\lambda_k-1}\left[\int_{-\infty}^{w_t(v)} y_t\,\varphi(w)dw + v\int_{w_t(v)}^\infty \varphi(w)dw\right]$

$$D^\sigma_{X_{kt}} = \frac{\partial \sigma(y_t)}{\partial X_{kt}} = \frac{1}{2\sigma(y_t)}\left[\frac{\partial E(y_t^2)}{\partial X_{kt}} - 2E(y_t)\frac{\partial E(y_t)}{\partial X_{kt}}\right]$$

$-\infty < y_t < \infty$		0
	$\lambda_y = 1$	$\dfrac{\beta_k X_{kt}^{\lambda_k-1}}{\sigma(y_t)}\int_{w_t}^\infty [y_t(w)-E(y_t)]\varphi(w)dw$
$y_t > 0$	$\lambda_y > 0$	$\dfrac{\beta_k X_{kt}^{\lambda_k-1}}{\sigma(y_t)}\int_{w_t}^\infty [y_t(w)]^{1-\lambda_y}[y_t(w)-E(y_t)]\varphi(w)dw$
	$\lambda_y = 0$	$\beta_k X_{kt}^{\lambda_k-1}\sigma(y_t)$
	$\lambda_y < 0$	$\dfrac{\beta_k X_{kt}^{\lambda_k-1}}{\sigma(y_t)}\left[\int_{-\infty}^{w_t(v)}[y_t(w)]^{1-\lambda_y}[y_t(w)-E(y_t)]\varphi(w)dw + v\int_{w_t}^\infty \varphi(w)dw\right]$

$$D^\gamma_{X_{kt}} = \frac{\partial \gamma(y_t)}{\partial X_{kt}} = \left\{\frac{1}{\sigma^3(y_t)}\left[\frac{\partial e_3(y_t)}{\partial X_{kt}} - a(y_t)\frac{\partial \sigma^3(y_t)}{\partial X_{kt}}\right]\right\}$$

$$= \left\{\frac{1}{\sigma^3(y_t)}\left\{\left[\frac{\partial E(y_t^3)}{\partial X_{kt}}\right] - \left\{\frac{\partial E(y_t)}{\partial X_{kt}}\right\}[3E(y_t^2)-2(E(y_t))^2] + E(y_t)\left[3\frac{\partial E(y_t^2)}{\partial X_{kt}} - 4E(y_t)\frac{\partial E(y_t)}{\partial X_{kt}}\right]\right\} - a(y_t)\left[3\sigma^2(y_t)\frac{\partial \sigma(y_t)}{\partial X_{kt}}\right]\right\}^\circ$$

$-\infty < y_t < \infty$		0
	$\lambda_y = 1$	0
$y_t > 0$	$\lambda_y > 0$	$\{\ \}^\circ$, where $E(y_t^r) = \int_{w_t}^\infty y_t^r \varphi(w)dw$
	$\lambda_y = 0$	$\{\ \}^\circ$, where $E(y_t^r) = \int_{-\infty}^\infty y_t^r \varphi(w)dw$
	$\lambda_y < 0$	$\{\ \}^\circ$, where $E(y_t^r) = \int_{-\infty}^{w_t(v)} y_t^r \varphi(w)dw + v^r\int_{w_t(v)}^\infty \varphi(w)dw$

Note: in the three sections of the table, derivatives of the first three non central moments of (6-F) are calculated by:

$$\partial(y_t^r)/\partial X_{kt} = r\beta_k X_{kt}^{\lambda_k-1}\int_{R_w}[y_t(w)]^{r-\lambda_y}\varphi(w)dw, \quad r=1,2,3;\text{ where } R_w \text{ denotes the domain of integration.}$$

* See the manual (Tran *et al.*, *op.cit.*, Section 2.5) for derivatives in the presence of heteroskedasticity and autocorrelation of *w*ₜ.

APPENDIX 5. US STANDARD METROPOLITAN AREAS (SMA) WITH AT LEAST 3
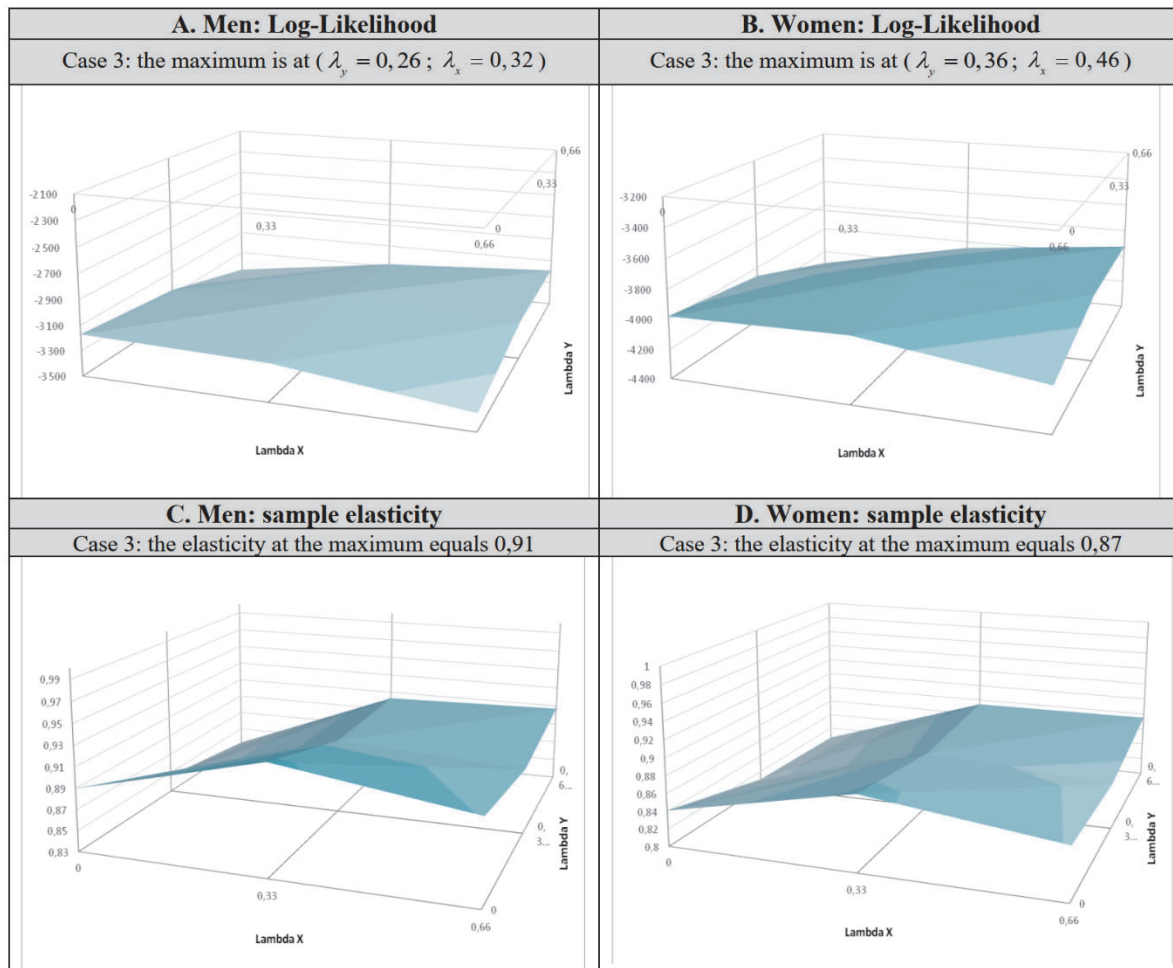
Million Inhabitants in 2009

We selected trips from Standard Metropolitan Areas (SMA) of at least 3 million inhabitants, as indicated by membership of SMA size class 5 in the NPTS 2009. These SMA correspond to the following Consolidated Metropolitan Statistical Areas (CMSA):

Atlanta, GA	Miami--Fort Lauderdale, FL
Boston--Worcester--Lawrence, MA--NH--ME--CT	New York--Northern New Jersey--Long Island, NY--NJ--CT--PA
Chicago--Gary--Kenosha, IL--IN--WI	Philadelphia--Wilmington--Atlantic City, PA--NJ--DE--MD
Dallas--Fort Worth, TX	Phoenix--Mesa, AZ
Detroit--Ann Arbor--Flint, MI	San Francisco--Oakland--San Jose, CA
Houston--Galveston--Brazoria, TX	Seattle--Tacoma--Bremerton, WA
Los Angeles--Riverside--Orange County, CA	Washington--Baltimore, DC--MD--VA--WV

APPENDIX 6. NEIGHBORHOOD OF THE LOG-LIKELIHOOD MAXIMUM (USA)

To verify that the high elasticity values (0,87 and 0,91) for the USA in Table 8 are really smaller than 1, examine in Figure 7 the behavior of the maximized Log-Likelihood (LL) value and of η^s $(\approx \eta^e)$ in the domain defined by $1/3 \le \lambda_x \le 2/3$ and $0 \le \lambda_y \le 2/3$ in the neighborhood of the highest LL value, a domain enclosing the maximum for men ($\lambda_y = 0,26$; $\lambda_x = 0,32$) and women ($\lambda_y = 0,36$; $\lambda_x = 0,46$).

FIGURE 7
WORK TRIP LOG-LIKELIHOOD AND SAMPLE SERVICE ELASTICITIES
IN THE NEIGHBORHOOD OF THE MAXIMUM (USA)

A. Men: Log-Likelihood	B. Women: Log-Likelihood
Case 3: the maximum is at ($\lambda_y = 0,26$; $\lambda_x = 0,32$)	Case 3: the maximum is at ($\lambda_y = 0,36$; $\lambda_x = 0,46$)

C. Men: sample elasticity	D. Women: sample elasticity
Case 3: the elasticity at the maximum equals 0,91	Case 3: the elasticity at the maximum equals 0,87

Remember also that a wider domain including negative values of λ_y would be irrelevant because, in (6-F), $E(y_t^r)$ does not then exist when, with $\varepsilon \to 0$ *and* $v \to \infty$, one approaches limit values for observations. Figure 7 confirms that the two maxima of the LL for the USA are indeed global and that there exists in their neighborhoods no tendency for elasticities to approach unity even if, at $\lambda_y = 0$ (where $\lambda_x = 0,33$), the elasticity values rise somewhat to 0,93 for men and 0,89 for women. This is not surprising, in view of asymptotic Student *t*-values equal to 100 for their underlying β_k coefficients, calculated here by the method of first derivatives of Berndt *et al.* (1974).

The source papers present information to complement that found in Table 9 on MRS among moments. Firstly, elasticities of substitution among them and secondly measures of fit for the moments defined in equations (6-C), (6-D) and (6-E). Both sets are found in this enriched version of Table 9:

TABLE 15
WORK TRIP DURATION MOMENT RATES AND ELASTICITIES OF SUBSTITUTION BY REGION, CASE 3 OF (8-D)

		A. Greater Paris Region					
		M (variant 22)			W (variant 35)		
Moment i\j		e	σ	γ	e	σ	γ
MRS^{m_e,m_σ}, MRS^{m_e,m_γ} MRS^{m_σ,m_γ}	e	1	+ 3.5	− 400	1	+ 4.5	− 220
	σ		1	− 110		1	− 49
	γ			1			1
η^{m_e,m_σ}, η^{m_e,m_γ} η^{m_σ,m_γ}	e	1	+ 1.1	− 8.2	1	+ 1.3	− 4.1
	σ		1	− 7.2		1	− 3.1
	γ			1			1
Sample moment value		42	24	1.50	35	19	1.40
Fitted value at the means		42	14	0.88	36	10	0.67
Mean of fitted values		42	19	0.98	35	16	1.30

		B. US Standard Metropolitan Areas						C. Metropolitan France					
		M (variant H-0)			W (variant F-0)			M (variant H-0)			W (variant F-0)		
Moment i\j		e	σ	γ	e	σ	γ	e	σ	γ	e	σ	γ
MRS^{m_e,m_σ}, MRS^{m_e,m_γ} MRS^{m_σ,m_γ}	e	1	+ 7.9	− 380	1	+ 8.6	− 250	1	+ 9.5	− 190	1	+ 9.0	− 160
	σ		1	− 47		1	− 29		1	− 20		1	− 18
	γ			1			1			1			1
η^{m_e,m_σ} η^{m_e,m_γ} η^{m_σ,m_γ}	e	1	+ 1.4	− 3.8	1	+ 1.6	− 2.8	1	+ 1.9	− 2.16	1	+ 1.8	− 2.2
	σ		1	− 2.8		1	− 1.8		1	− 1.1		1	− 1.2
	γ			1			1			1			1
Sample moment value		37	23	1.4	31	19	1.6	29	18	1.3	25	16	2.9
Fitted value at the means		38	6.4	0.38	6.4	9.38	0.35	30	6.0	0.32	25	4.9	0.33
Mean of fitted values		37	22	1.5	31	19	1.9	29	17	1.4	25	15	3.1

Elasticities of Substitution

To avoid dealing with units of MRS measures in (5-D.1)-(5-D.3), one may instead use elasticities of substitution η^{m_τ,m_s}, $for(\tau \neq s = 1,3)$, namely (5-D.1*)-(5-D.3*), all pure numbers equal to the MRS multiplied by matching inverses of moments, evaluated below at sample means of variables. The corresponding expressions are as follows:

(D.1) $MRS^{m_e,m_\sigma} = \dfrac{\partial e(y_t)}{\partial \sigma(y_t)} = \dfrac{\partial e(y_t)/\partial X_{kt}}{\partial \sigma(y_t)/\partial X_{kt}} = \dfrac{\partial e(y_t)/\partial X_{\ell t}}{\partial \sigma(y_t)/\partial X_{\ell t}}$;	(D.1*) $\eta^{m_e,m_\sigma} = MRS^{m_e,m_\sigma} \dfrac{m_\sigma}{m_e} = \dfrac{\partial m_e}{\partial m_\sigma} \dfrac{m_\sigma}{m_e}$,
(D.2) $MRS^{m_e,m_\gamma} = \dfrac{\partial e(y_t)}{\partial \gamma(y_t)} = \dfrac{\partial e(y_t)/\partial X_{kt}}{\partial \gamma(y_t)/\partial X_{kt}} = \dfrac{\partial e(y_t)/\partial X_{\ell t}}{\partial \gamma(y_t)/\partial X_{\ell t}}$;	(D.2*) $\eta^{m_e,m_\gamma} = MRS^{m_e,m_\gamma} \dfrac{m_\gamma}{m_e} = \dfrac{\partial m_e}{\partial m_\gamma} \dfrac{m_\gamma}{m_e}$,
(D.3) $MRS^{m_\sigma,m_\gamma} = \dfrac{\partial \sigma(y_t)}{\partial \gamma(y_t)} = \dfrac{\partial \sigma(y_t)/\partial X_{kt}}{\partial \gamma(y_t)/\partial X_{kt}} = \dfrac{\partial \sigma(y_t)/\partial X_{\ell t}}{\partial \gamma(y_t)/\partial X_{\ell t}}$;	(D.3*) $\eta^{m_\sigma,m_\gamma} = MRS^{m_\sigma,m_\gamma} \dfrac{m_\gamma}{m_\sigma} = \dfrac{\partial m_\sigma}{\partial m_\gamma} \dfrac{m_\gamma}{m_\sigma}$,

Higher Moments than the First

Concerning fit statistics at the bottom of each Part of Table 15, note that the first sample moment is better adjusted than the second, itself being better adjusted than the third. This no doubt occurs because model construction is guided by intuition of first moment response: indeed, how many ever add an explanatory variable to specifically improve 3rd moment fit? Few indeed – perhaps because current measures of model fit tend to be based only on calculated first moment measure analogs of R^2 and neglect higher moment fit statistics such as those calculated here.