Conference Paper

de Jong, G.; Daly, A.J.; Pieters M.; van der Hoorn A.I.J.M. (2005) *The logsum as an evaluation measure: review of the literature and new results* In: 45th Congress of the European Regional Science Association, 23rd -27th August 2005, Vrije Universiteit Amsterdam

The logsum as an evaluation measure: review of the literature and new results[1]

Gerard de Jong, RAND Europe and ITS Leeds

Andrew Daly, RAND Europe and ITS Leeds

Marits Pieters, RAND Europe

Toon van der Hoorn, Transport Research Centre, Dutch Ministry of Transport, Public Works and Water Management

Address for correspondence:

Gerard de Jong

RAND Europe

Newtonweg 1

2333 CP Leiden

The Netherlands

jong@rand.org

Key words: logsum, consumer surplus, discrete choice models, project assessment

Abstract:

The logsum is a measure of consumer surplus in the context of logit choice models. In spite of the very frequent use of logit models in transport, project assessment is only rarely done using logsums. Instead in project evaluation or appraisal, changes in transport costs and time (borrowing values of time from some source) are commonly used to get the traveller benefits. The paper contains a review of the theoretical and applied literature on the use of logsums as a measure of consumer surplus change in project appraisal and evaluation. It then goes on to describe a case study with the Dutch National Model System for transport in which the logsum method and the commonly used value of time method are compared for a specific project (high speed trains that would connect the four main cities in the Randstad: Amsterdam, The Hague, Rotterdam and Utrecht).

1. An earlier version of this paper was presented at the 45th Congress of the European Regional Science Association, 23-27 August 2005, Vrije Universiteit Amsterdam

1. Introduction

Transport infrastructure projects in The Netherlands are normally appraised ex ante by using cost-benefit analysis (CBA) procedures following the so-called 'OEI-guidelines' (CPB and NEI, 2000). The project benefits for travellers are incorporated in the form of changes in demand and changes in the generalised travel costs (using values of time from Stated Preference studies to monetise travel time savings), and applying the rule-of-a-half. This is the standard approach in many countries nowadays. The changes in demand are calculated using transport models, such as the Dutch National Model System, LMS, or the regional models, NRM (for a description, see Gunn, 1998 or Daly, 2000).

While a number of short-term improvements to the current procedures have been proposed (see Ecorys and 4cast, 2004), it is also interesting to consider a substantially different approach using explicit measures of consumer surplus, obtained by integrating the demand models directly[1].This is the topic of this paper.

The direct effects of a particular policy on the travellers can be measured as the change in consumer surplus that results from that policy (there can also be indirect and external effects that may not be covered in the consumer surplus change).

The consumer surplus associated with a set of alternatives is, under the logit assumptions, relatively easy to calculate. By definition, a person's consumer surplus is the utility, after conversion to money terms, that a person receives in the choice situation. If the unobserved component of utility is independently and identically distributed extreme value and utility is linear in income, then the expected utility becomes the log of the denominator of a logit choice probability, divided by the marginal utility of income, plus arbitrary constants. This is often called the "*logsum*". Total consumer surplus in the population can be calculated as a weighted sum of logsums over a sample of decision-makers, with the weights reflecting the number of people in the population who face the same representative utilities as the sampled person. Assuming no change in the unobserved component of utility, the change in consumer

[1] In the longer run, the practicalities of a more radically different approach, using income-compensated welfare change measures (see section 2.3), should be further investigated.

surplus is calculated as the difference between the logsum under the conditions before the change and after the change (e.g. introduction of a policy). The arbitrary constants drop out.

The advantages that the logsums would give to the appraisal procedure would be that logsums can incorporate a degree of heterogeneity in the population, while also being theoretically more correct and in many cases easier to calculate. Also, the logsum incorporates various factors that influence the choices (e.g. those of mode and destination), such as different travel time and costs components, service quality, person and household attributes and integrates these into a common framework. However, to calculate this change in consumer surplus, the researcher must know the marginal utility of income. Usually a price or cost variable enters the representative (indirect) utility and, in case that happens in a consistent linear additive fashion, the negative of its coefficient is the marginal utility of income by definition. If the marginal utility of income is not constant with respect to income, as is the case in the Dutch National Model System LMS and the Dutch regional models NRM, a far more complex formula is needed, or an indirect approach has to be taken, while thought is also needed about the measure of surplus to be used.

In this context, the Transport Research Centre (AVV) of the Dutch Ministry of Transport, Public Works and Water Management commissioned RAND Europe to undertake a study comparing the conventional approach to the use of the logsum change as a measure of the change in consumer surplus that would result from a transport infrastructure project. The paper is based on the work conducted in this study. This paper has two objectives: first it gives a summary of the existing theoretical and empirical literature on the use of the logsum as an evaluation measure. Secondly, it provides new results for a specific project, "Rondje Randstad" (a proposed high speed –possibly MAGLEV- train project that would connect the four main cities in the Randstad: Amsterdam, The Hague, Rotterdam and Utrecht). For this project, two methods are compared:

a. the 'classical' standard practice approach of measuring the change in generalised cost, with external values of time, invoking the rule-of-a-half; and
b. the logsum approach (where we also report on some practical issues that need to be dealt with when applying logsums in evaluation).

After having introduced the basic concepts in section 2, this paper reviews the theoretical literature on the use of the logsum as an evaluation measure, including both the original papers on this from the seventies and the work on the income effect in the nineties (section 3). Also recent application studies that used the logsum for evaluation purposes are reviewed (section 4). It then goes on in section 5 to describe the "Rondje Randstad" case study applying the LMS. Different methods for monetising the logsum change are compared as well. A summary and recommendations are provided in section 6.

2. Logsums and other welfare measures: the basic concepts

In this section we provide an introduction to the concept of logsums and related welfare measures. A more extensive introduction can be found in the textbooks on discrete choice models (e.g. Train, 2003).

2.1 Logsums

The utility that decision maker *n* obtains from alternative *j* is decomposed into an observed and an unobserved (random) component:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \tag{1}$$

Where:

U_{nj} is the utility that decision maker *n* obtains from alternative *j* (n = 1,..N ; j = 1,…,J),

V_{nj} = "representative utility";

ε_{nj} captures the factors that affect utility, but are not observable by the researcher.

In a standard multinomial logit (MNL) model, with ε_{nj} i.i.d. extreme value with standard variance ($\pi^2/6$), the choice probabilities are given by:

$$P_{ni} = \frac{e^{V_{nj}}}{\sum_j e^{V_{nj}}}. \tag{2}$$

The "logsum" is the log of the denominator of this logit choice probability. It gives the expected utility from a choice (from a set of alternatives) and can be used to link different choices (as in nested logit models, e.g. of mode and destination choice). The logsum can also be used in project evaluation in an expression for the consumer benefits. This is explained below.

In the field of policy analysis, the researcher is mostly interested in measuring a change in consumer surplus that results from a particular policy. By definition, a person's consumer surplus is the utility (also taking account of the disutility of travel time and costs), in money terms, that a person receives in the choice situation. The decision-maker *n* chooses the alternative that provides the greatest utility, so that, provided that utility is linear in income, the consumer surplus (CS_n) can be calculated in money terms as:

$$CS_n = (1/\alpha_n)\, U_n = (1/\alpha_n) \max_j (U_{nj}\ \forall\ j) \qquad (3)$$

where

α_n is the marginal utility of income and equal to $dU_{nj}/\,dY_n$ if j is chosen,

Y_n is the income of person n, and

U_n the overall utility for the person n

Note that the division by α_n in the consumer surplus formula, translates utility into money units (e.g. dollars, euros) since $1/\alpha_n = dY_n / dU_{nj}$.

If the model is MNL and utility is linear in income (that is α_n is constant with respect to income), then expected consumer surplus becomes:

$$E(CS_n) = (1/\alpha_n) \ln \left(\sum_{j=1}^{J} e^{V_{nj}} \right) + C \qquad (4)$$

where C is an unknown constant that represents the fact that the absolute value of utility can never be measured. Aside from the division and addition of constants, expected consumer surplus in a standard logit model is simply the logsum.

Under the usual interpretation of distribution of errors, $E(CS_n)$ is the average consumer surplus in the subpopulation of people who have the same representative utilities as person n. Total consumer surplus in the population can be calculated as the weighted sum of $E(CS_n)$ over a sample of decision-makers, with the weights reflecting the number of people in the population who face the same representative utilities as the sampled person.

The change in consumer surplus is calculated as the difference between the calculation of $E(CS_n)$ under the conditions before the change and the calculation of $E(CS_n)$ after the change (e.g. introduction of policy):

$$\Delta E(CS_n) = (1/\alpha_n) \left[\ln \left(\sum_{j=1}^{J^1} e^{V^1_{nj}} \right) - \ln \left(\sum_{j=1}^{J^0} e^{V^0_{nj}} \right) \right] \qquad (5)$$

where superscript 0 and 1 refer to before and after the change.

Since the unknown constant C appears in the expected consumer surplus both before and after change, it drops out in calculating the changes in the consumer surplus. However, to calculate this change in consumer surplus, the researcher must know (or have estimated) the marginal utility in income α_n. Usually a price or cost variable enters the representative utility and, when that happens in a linear additive fashion, the negative of its coefficient is α_n by definition. The above equations for calculating the expected consumer surplus depend critically on the assumption that the marginal utility of income is constant with respect to income. If this is not the case, a far more complex formula is needed, in which α_n becomes a function of the change in attributes. However, for policy analysis absolute levels are not required, rather only changes in consumer surplus are relevant, and the formula for calculating the expected consumer surplus can be used if the marginal utility of income is constant over the range of implicit changes that are considered by the policy. So, for policy changes that change the consumer surplus by small amounts per person relative to their income, the formula can be used with a current average value for α even though in reality the marginal utility of income varies with income.

A different interpretation, namely the logsum as a measure of accessibility, is given in the textbook by Ben-Akiva and Lerman (1985). More complex discrete choice models (notably Generalised Extreme Value or GEV models) are discussed in the next chapter. The most general model of this type is the Network or Recursive Nested Extreme Value model, described by Daly and Bierlaire (2006), who give a recursive generalisation of the logsum formula as Corollary 8 of their main result.

2.2 Rule-of-a-half

Standard practice in project evaluation is to multiply the number of trips before and after a change in generalised cost by the change in generalised costs times one half (rule-of-a-half, RoH). The RoH also tries to measure consumer surplus change, but it only applies for small cost changes and if the demand curve is (approximately) a straight line. However, the implicit demand curves in transport models are generally not straight lines. Therefore, the rule-of-half may be considered as only a rough approximation of real welfare changes. These real changes can be estimated more precisely by deriving them from the transport models themselves, as in the logsum approach

2.3 Compensating and equivalent variation

The consumer surplus (CS) as defined above was first proposed by Marshall and is derived using the Marshallian or uncompensated demand curve. A price change generally leads to both a substitution effect (from one good to the other) and an income effect (more or less purchasing power for all goods because of an expansion or contraction of the budget). The Marshallian demand curve gives the substitution effect only; the Hicksian or compensated demand curve gives substitution and income effects. Hicks also proposed an alternative measure for the change in welfare, the **compensating variation** (CV).

The CV gives the maximum amount of money (just as the consumer surplus, it is a money measure of utility change) that can be taken from the consumer while leaving him or her just as well off as before the price reduction (willingness to pay for a price reduction). In case of an increase in the price, the CV is the minimum amount of money that must be given to the consumer to compensate him for the price increase (willingness to accept a price increase). A related money measure of utility change is the **equivalent variation** (EV). This is the minimum amount of money that must be given to a consumer to make him as well off as he could have been after the price reduction. For a price increase, EV is the maximum amount a consumer is willing to pay to prevent the price increase (Johansson, 1993). So, the CV uses the old utility level and the new prices, whereas the EV uses the new utility level and the old prices.

CV and EV are both areas under the compensated demand curve. These give different outcomes than the CS, which is based on the uncompensated demand curve. In general the CS will be between the CV and the EV. The measures based on the Hicksian demand curve are more attractive in theory, because these do not assume a constant marginal utility of income. In practice the CS is by far the most used measure, mostly through the RoH, sometimes as the logsum change (see below), because researchers felt CV or EV were very hard to compute given the available information (see Cherchi et al, 2004). Also, the use of CS is often justified by considering that income effects of transport project will usually be small given the small to moderate share of transport expenditures in the consumer budgets. In this paper, we shall restrict ourselves to a comparison of the logsum versus the RoH (except for the review of the theoretical literature in the next section), based on our view that a wider use of logsums is both desirable and feasible. But in the longer run, practical calculation rules for CV (or EV) measures should rank high on the research agenda, also given that some policy measures that are under consideration (e.g. widespread use of congestion charging) might have considerable income effects.

3. Review of the theoretical literature

This section provides an overview of the theoretical literature discussing the issue of calculating overall utility derived by diverse consumers facing a discrete choice and the role of the 'logsum' formula in that calculation. First a general description of the early literature (until the early nineties) is given. Then the issues addressed in the more recent literature (income effects and taste variation) are introduced.

It is supposed that consumers face a situation in which they must choose one of a finite number of mutually exclusive alternatives[2]. Each alternative has a utility and each consumer chooses the alternative that gives him or her the maximum utility. However, because the consumers are diverse, i.e. have different preferences, the alternative that gives maximum utility may be different for different consumers. Moreover, it is acknowledged that neither the analyst nor the consumer can measure the utilities with perfect precision; any predictions of choice can therefore be made only as probabilities.

This analysis gives the Random Utility Model (RUM) framework, in which consumers (i.e. travellers or freight shippers, in the transport context) are represented as maximising utility, but that this utility is considered random, either because the analyst cannot measure the utilities perfectly or that the consumer does not act consistently, e.g. by making mistakes. This framework has been questioned persistently by psychologists and other social scientists, but remains the only complete paradigm for modelling and evaluating choice behaviour. For the present work we shall remain within the RUM framework.

The theoretical literature reviewed falls into two phases. First, covering the period up from the early 1970's to the early 1990's, appraisal analysts working within the RUM paradigm constrained themselves to models which did not allow for any effect of income on choice,

[2] For the purposes of discussion, it is presumed that a market segmentation has been carried out such that consumers within one segment can be considered to choose from the same set of alternatives. This segmentation has the practical consequence that overall utilities must be calculated separately for each segment.

nor for any variation in tastes which was related to variables in the model. Later, from the mid-90's to the present day, attempts have been made to incorporate these two effects into appraisal models, following successful incorporation of such effects in choice models. It is fair to say that not all of the problems of extending the appraisal models have yet been solved. In any case, practical appraisal procedures up to the present day have almost exclusively been based on the simpler, earlier, models. A summary of the theoretical literature review is presented below, split into the two phases. The detailed reviews are in RAND Europe (2005). Two recent papers offer an overview of the field from different points of view: Bates (2003) and Daly (2004). Bates gives a complete overview of current practice[3] of transport policy appraisal, relating this to the relevant theory. In particular this paper gives an excellent discussion of the strengths and weaknesses of the 'rule-of-a-half' approximation to consumer surplus calculation. Daly reviews the early theory of RUM modelling, showing that all the important researchers were working on basically equivalent hypotheses, which include a constant marginal utility of money and the exclusion of taste variation in policy variables. He then goes on to discuss more recent work that abandons these restrictions and to discuss the consequences that the various approaches have for appraisal.

3.1 The early RUM literature

The key early papers in the RUM literature are McFadden's 1978 and 1981 publications, which form his most important contribution to the discrete choice literature and a major component of his Nobel work. In those papers he first set out the GEV theorem (1978) and then gave full mathematical detail of the links between RUM, choice models and welfare functions (1981), which form the basis for discussing this issue. Essentially, the GEV theorem gives the basis for deriving choice probabilities and overall utilities from a class of functions, which satisfy a list of conditions. The specific form of the expression giving the overall utility (the welfare function) is, in simple cases, the log of the sum of the exponentiated utilities of the alternatives, hence acquiring the name 'logsum'.

In many papers, the first publication advocating the use of the logsum as a measure of consumer surplus is stated to be Williams (1977). This was indeed the paper that made the

---

3 Perhaps, current British practice.

breakthrough in understanding the linkage between choice models and user benefit measures. However, Cochrane (1975) gives the logsum formula for total utility and refers to 1971 work by Neuberger and work parallel to his own by Koenig. Williams himself refers to Neuberger and to work by Wilson and Kirwan of 1969, in both cases as having used the logsum formula for evaluation. The logsum measure was also in practical use for appraisal before 1977 (by Daly and probably by others, as it is quite simple to derive as the integral of a logit demand function). Both Cochrane and Williams gave a complete theory of utility on which the logsum could be based, but Williams and Daly and Zachary (1978) took this further to establish that the logsum was the key 'composite cost' measure which could be used in further modelling to obtain tree (nested) logit models and derived extended logsum measures from tree logit models. McFadden's contribution in this context was to generalise further the models from which logsum-type measures could be derived and to extend and make more rigorous the theory on which their derivation was based.

McFadden's GEV theorem also gives the choice probabilities for the model. These are equal to the derivatives of the logsum with respect to the utilities of the alternatives. That is, the logsum is equal to the integral of any of the choice probabilities with respect to the utility of the corresponding alternative. Given that the choice probability is the expected demand for the alternative from each consumer, it can be seen that the logsum is thus – in some sense – the integral of the ordinary demand curve.

It would thus be convenient to identify the logsum with the Marshallian consumer surplus arising from the choice situation, which is conventionally presented as the integral of the demand curve. However, Marshallian surplus is defined in terms of the integral of demand with respect to the *price* of an alternative, while the logsum is defined as the integral with respect to the *utility* of an alternative. In a context where the marginal value of money is considered to be constant, this presents no problem. The literature up to the early 1990's, including McFadden, is based on this assumption, which is tantamount to ignoring any influence of income on choice. Most models simply do not deal with the impact of budget constraints on behaviour.

In McFadden's early theory, the key assumptions identifying the models are:

1. the AIRUM assumption (Additive Income RUM), which requires income to enter indirect utility in a specific linear additive form, precluding any income effect on choice behaviour; and
2. the invariant RUM assumption that the distribution of the random component of utility[4] is not affected by the values of the observable components – essentially, there is no unobserved taste variation.

A recent paper by Daly (2004) shows that all the key early researchers made these key assumptions (in so far as they discussed the role of income) and also made the same more technical assumptions that are necessary to make the models operational. The most general form of these technical assumptions allows for the possibility that the logsum should be scaled by a constant positive factor to express it on a monetary scale.

3.2 Income effect and taste variation

The early RUM literature was based on the two assumptions set out in the previous paragraph, that is additive income and an absence of unobserved taste variation. The more recent literature has sought to generalise the models to avoid these two assumptions. In some cases this requires a generalisation of the logsum approach, while in other cases a different approach is required.

Income effect

Income effect appears in a choice model when the predictions of the model can change when the choice-maker's income changes. Equivalently, the model contains income effect if its predictions change when the prices of all the alternatives are increased by the same amount. In terms of the utility functions of the alternatives, the model has no income effect if the marginal utility of income, α in the equations above, is constant across the alternatives and for all levels of price.

In a model without income effect, the ordinary and compensated demand functions coincide and there is no problem in defining consumer surplus unambiguously. However, in a model

[4] More strictly, the distribution of the *differences* of the random components.

with income effect, ordinary (Marshallian) and compensated (Hicksian) surplus diverge and we can find we are able to measure one more easily than the other. In any case, Karlström (2000) states that the logsum always gives the Marshallian surplus; however, we may not be able to convert this unambiguously into monetary terms when α is varying.

The impact of income on discrete choice has of course been considered in models of car ownership and other issues for many years, but it seems that McFadden (1996) was the first to propose acceptable procedures for calculating consumer surplus measures for models with income effects. This paper gives three methods for assessing consumer surplus with models that are nonlinear in income: a simulation procedure; an approximation based on a representative consumer approach, which he rejects as inaccurate; and some bounds on the true value of the surplus. Herriges and Kling (1999) test these approaches on real data, concluding that the calculation of bounds is inconvenient and may be inaccurate but are unable to choose decisively among the other McFadden approaches and more approximate methods.

Daly (2004) notes the findings of McFadden, supported by Cherchi *et al*. (2004), concerning the linearisation approach and of Herriges and Kling concerning the bounds approach, which leads Daly to suggest that the simulation procedure is the most promising approach. McFadden (1996) concluded that this approach was computationally burdensome, but Daly points out that if non-satiation in income is assumed, then the income term in the indirect utility functions can be inverted and the surplus obtained more explicitly. The computational effort in this case is quite limited and it seems that the simulation approach would then be the most attractive choice. This simulation approach can be used to derive surplus in any random utility model for which the distribution of utility can be sampled. The surplus calculation is based on calculating the income change necessary to derive the same utility in base and changed circumstances, i.e. the calculation is essentially Hicksian and cannot be used to derive Marshallian surplus in models in which choice is predicted as a function of income.

Karlström (2000) presents a function that generalises the logsum and gives Hicksian surplus for GEV models without sampling. In principle this method is quicker in application than

McFadden's (1996) sampling method. He also notes that the logsum-equivalent in GEV models gives Marshallian surplus in any case.

Daly (2004) also gives a function which generalises the logsum and which can be used to obtain surplus in any 'Invariant' random utility model – that is, a model for which the distribution of the utilities does not depend on the utility values, i.e. models without taste variation. However, neither Karlström nor Daly offer any procedure for converting a logsum (or generalised logsum) change to monetary value in models with income effect.

In summary, it is possible to develop logsum-like methods for models that are more general than GEV, whether or not they exhibit income effect; however, the monetisation of these measures can be difficult when income effect is present. McFadden's method can be used for a still wider range of models and yields monetary values as its primary output; however, these measures are Hicksian and there is no explicit measure corresponding to the logsum.

The second key requirement for the early models is that they fall within the Invariant RUM class. Models not satisfying this requirement are not of the GEV class (as they fail to satisfy the homogeneity requirement of McFadden's (1978) theorem) and theory for the justification of logsums is therefore not well developed.

Unobserved taste variation

A simple approach to deriving benefit measures from models with unobserved taste variation is to build on work by McFadden and Train (2000) which showed that any RUM can be approximated by a random mixture of multinomial logit models and calculations using the model can then be made by drawing from the mixing distribution. If we view each such draw from the mixing distribution as representing a small fraction of the population, then for that small fraction the logsum can be calculated and by integrating over the mixing distribution we obtain a population value. Thus if unobserved taste variation is represented in the model by using mixed logit models, then the 'mixed logsum' is a valid construct and is justified on the same basis as the simple logsum.

Provided the model does not exhibit income effect, the logsum integrals thus obtained can be converted to money equivalents. When the price coefficient itself is subject to random variation, the logsum must be converted to a monetary equivalent *before* averaging. However, with this reservation, there appears no further problem in this procedure for calculating mixed logsums. Moreover the process is quite simple to apply.

In summary, provided that a model with unobserved taste variation can be approximated conveniently by a mixed logit (or a mixed GEV) it is theoretically valid and reasonably easy to calculate the mixed logsum measures. Few applications of such procedures have been seen in the literature, the work of Von Haefen (2003) being an exception.

The McFadden (1996) simulation procedure can be used to calculate surplus for models with unobserved taste variation, whether or not there is also income effect. This approach can be useful in cases where income effect is present or when the model cannot conveniently be formulated as a mixed GEV model, but it is less straightforward than the calculation of a logsum measure. However, once again we note that McFadden's procedure yields a Hicksian measure which does not bear any direct relation to the logsum.

4. Review of the applied literature

Although the theory on the use of the logsum change as a measure of the change in the consumer surplus was published in the late seventies and early eighties, the application of this theory in practical appraisals of transport projects has been very limited. Most applications in transport evaluation that the authors are aware of have been undertaken only recently (after 2000). The applications that were reviewed are summarised in Table 1. All applications use models that include mode choice. Some logsum applications in project evaluation also use models for destination choice and/or departure time choice. Logsums are first calculated for each individual decision-maker in the sample, and then aggregated over groups of decision-makers. Various segmentations are used, also depending on the segmentation used for the time or cost coefficients used later on to convert from the utility scale (measured in, say, 'utils') to money or time. A common segmentation for the logsum calculations and outputs is by travel purpose. The applications of the logsum concept in transport project appraisal all use the relatively simple formulation with constant marginal utility of money. It could be dangerous to assume that the marginal utility of income would be constant over a wide income range (it is more likely that it will decline with increasing income). Theory has moved beyond that in the nineties, but the later formulations are not in practical use. Also note that the Sacramento application uses the assumption of a constant marginal utility of money only within a number of distinct income/worker categories. This provides a solution, which we also used for the LMS applications reported later in this paper, where cost coefficients also differ between five income groups. Two applications (Castiglione et al, 2003 and Koopmans and Kroes, 2004) do not convert the (dimensionless) logsum change directly into money units, but convert to time in minutes. The other applications use one or more cost coefficients to get outcomes in money units.

Table 1. Summary of applications of the logsum in transport project appraisal

Model application	Choices included	Segmentation	Marginal utility of income	Conversion method of utility into money
San Francisco (Castiglione et al, 2003)	Mode choice[5]	By zone pair and 9 segments based on household size and car availability	constant	Using a common in-vehicle time coefficient to get outcomes in minutes
Europe (EXPEDITE, 2002)	Mode-destination choice	By almost 1,000 person segments and 5 travel purposes	constant	Using an implied cost coefficient per purpose to get outcomes in euros
Austin (Gupta et al., 2004; Kalmanje and Kockelman, 2004)	Mode, destination and departure time choice	4 trip purposes, calculated for an individual resident	constant	Using a cost coefficient per purpose to get outcomes in dollars
The Netherlands LMS (Koopmans and Kroes, 2004; De Raad, 2004)	Mode, destination and departure time choice	8 travel purposes	not constant	Using time coefficients per purpose to get minutes, then using value of time to get euros
Oslo (Odeck et al, 2003)	Mode-departure time choice	By trip purpose	constant	Using a cost coefficient per purpose to get outcomes in Kroner
The Netherlands TIGRIS (RAND Europe, 2004)	Mode, destination and departure time choice	8 travel purposes	not constant	No conversion to money used
Sacramento (USDoT, 2004)	Mode choice	Household segments base on income/worker categories	constant	Using a cost coefficient per segment to get outcomes in dollars

[5] The San Francisco County Transportation Authority (SFCTA) model could also provide destination and time-of-day choice, but the rules for project evaluation of the US Federal Transit Administration (FTA) only allow the mode choice logsum to be used for user benefit calculation and also prescribe scaling to in-vehicle time units.

5. Case study for The Netherlands

In this case study, we are using results from two different runs that were carried out with the Dutch National Model System LMS:

- the reference situation 2020; and
- the project situation 2020 (the same as the reference situation, except for the implementation of 'Rondje Randstad', with particular speed and frequency increases for a number of train links between the big cities in the Randstad, and reductions on some of the minor train links), reflecting a high speed rail or MAGLEV network between Amsterdam, Roterdam The Hague and Utrecht.

Below are the results for the number of tours by train travellers from the two LMS runs. Note that for home-based business and 'other' travel, the number of train tours is predicted to decrease. This is due to the fact that the project variant does not provide train times that are better than the reference situation 2020 in all cases; for some origin-destination relations, the train times in the reference situation are shorter. The travel times by train in the project variant always at least as good as in the base year, but the reference situation also includes some improvements in train travel times compared to the base year, some of which (especially stop trains) are not in the 'Rondje Randstad' variant.

Table 2. Number of tours by train on an average working day in 2020

	Reference 2020	Project 2020	Project 2020 (Reference=100)
Commuting	505571	513088	101.5
home-based business	16659	16509	99.1
non-home-based business	18959	18978	100.1
education	170121	171955	101.1
shopping	37499	37584	100.2
other	112678	112455	99.8
total	861487	870569	101.1

5.1 The 'classic' method

In the 'classic' method, the benefits from the project are calculated as follows. For instance for commuters the number of travellers that stay in the train is taken to be 505571. The number of 'new' travellers is taken to be 513088 – 505571 = 7571 (in fact this is substitution from other modes). The average travel time (train in-vehicle-time) in the reference situation is 62.26 minutes for commuting. With the project this is 61.62 minutes. This time gain is used as the benefit for all stayers: 505571* (62.26-61.62) * (value of time for train commuters). This value of time comes from the 1997/1998 stated preference (SP) surveys that Hague Consulting Group carried out for AVV. For the new travellers the gain is (rule-of-a-half): 7571* (62.26-61.62) * (value of time for train commuters) * 0.5. Repeating this for all purposes gives the traveller benefits as in Table 3 (first row).

As can be seen from the above example and the numbers of train travellers in Table 2, the benefits to travellers are completely dominated by the benefits of those that stayed in the train. The new train passengers make up only about 1% of the total number of train passengers in the project situation. Furthermore the benefit for a new train passenger is calculated as only half that of a remaining train passenger.

In the second row of Table 3 are outcomes when including not only the train in-vehicle-time benefits, but also the gains in terms of in-vehicle-time (bus) during the access to the train station and during the egress from the train station (this is exogenous project input). This more than doubles the time benefits of the project. The increase is caused by the fact that the (large) railway stations that will be used more in the 'Rondje Randstad' variant have better bus/tram/metro connections, so the access and egress times will be shortened. The benefits in Table 3 are between 0.10 and 0.23 Euro per train tour.

Table 3. Traveller benefits (project minus reference) for the full year 2020 in millions of 2003 Euros[6]

Method:	Traveller benefits
Classic method, train in-vehicle time only	24
Classic method, including access/egress time	58
Logsum using SP values of time	44-51
Logsum using average costs	56

We did not calculate additional benefits for the car users due to the reduction of congestion on the roads (that would be caused by substitution from road to rail), because the substitution from car driver to rail was so small (about a third of the 'new' train travel) that the average travel time by road did not change (even taking account of the fact that the volume-delay functions in the model are non-linear such that small changes in traffic volumes can result in substantial changes in travel time).

5.2 The logsum method

Logsums (dimensionless) were first calculated for the reference situation 2020 and for the project situation with shorter train times because of the 'Rondje Randstad' project. Then, logsum differences for the difference between the reference and the project situation were produced. The logsums and logsum differences were originally calculated per tour. These outcomes were aggregated/expanded to logsums and logsum differences for combinations of travel purpose and income class (with five income classes, as used in the LMS).

For each of these two logsum types of differences (NL, MNL), we applied two different methods for the conversion to money:

- Method 1: Translate the logsum differences to minutes, using the LMS travel time coefficients (by purpose) and then translate from minutes to 1995 money values by using the recommended values of time (from the 1997/1998 stated preference (SP) surveys that Hague Consulting Group carried out for AVV) by purpose and income

[6] The calculations were originally made in 1995 prices (as used in the current LMS version). For business travel a factor 1.28 was used to go from 1995 to 2003 (average contractual wage rate increase) and 1.30 for the increase in the values of time between 2003 and the forecast year 2010 (Ecorys and 4Cast, 2004). For commuting these values were 1.23 (consumer prices) and 1.15, and for other travel the values were 1.23 and 1.11. For the conversion from an average working day to a full year we used the factor 285 (Ecorys and 4Cast, 2004).

group. Because the project studied (Rondje Randstad) is a rail project, and rail users are affected most, we used the time coefficients (for in vehicle time and other time components) of rail here. The values of time used are those by income class and travel purpose (not by mode, but over all modes). This method has a consistency problem: it uses one set of implied values of time from the LMS to get the transport demand impacts in minutes and another set of information on values of time from SP surveys to get the transport demand impacts in money.

- Method 2: Divide the logsums by the product of the LMS cost coefficients and the expected value of (1/cost) per tour (separate for each income class and purpose, but averaged over all relevant origin-destination-mode combinations). In a linear cost model, division by the cost coefficients would have been sufficient for conversion to money units. But here the costs enter the calculation because of the use of logarithmic costs in the LMS mode-destination choice models. Moreover, the use of the expectation of (1/cost) is only approximately correct. On the other hand, this method does not use the information on values of time from the SP survey and therefore does not have the inconsistency problem that Method 1 has.

The cost coefficients in the LMS are the same for all modes (but differ between purposes and income groups), but a problem is the treatment of modes and population groups with zero costs (slow modes, car passengers, students). The LMS itself uses zero if there are no cost and ln(cost) for positive cost. For our conversion to money in method 2, we need to divide by costs, and have to avoid division by zero. To calculate this, we used the lowest observed cost (we found that this is just below 1 guilder, approximately € 0.45, and used 1 guilder here) per tour for modes and groups with zero costs, so that these have will have a small impact on the final results.[7]

Below the monetisation methods are described formally (for a given purpose and person type):

We have utility functions of the form:

[7] A cost formulation of the form ln(cost+1) in the LMS would have been more convenient. This also gives zero when cost is zero and a proper derivative for zero costs (1 guilder).

$$U = \beta \ln[C] + \chi T + \ldots \quad (6)$$

in which:

C is cost in 1995 guilders

T is time in minutes

β: 1 guilder is β utils, or 1 util is $1/\beta$ guilders

χ: 1 minute is χ utils, or 1 util is $1/\chi$ minutes

We also define:

LS = logsum value in utils

Now method 1 and 2 work as follows:

<u>Method 1 (Value of time)</u>:

LS/χ = logsum in minutes

$$\text{Logsum in guilders}^{8} = (LS/\chi).VoT \quad (7)$$

VoT comes from an external model, estimated on stated preference data.

<u>Method 2 (expectation of 1/ cost):</u>

The starting point is the consumer surplus equation (3). For a population which chooses j with probability p_j the average marginal utility then is:

$$A = \Sigma p_j \, dU_{n\,j}/ \, dY_n \quad (8)$$

in which:

A: average marginal utility

Now:

$$dU_{n\,j}/ \, dY_n = - dU_{n\,j}/ \, dC_n \quad (9)$$

in which:

C_n: the price of travel by alternative j

Therefore we obtain:

$$E(\partial U/\partial C) = [- A = - \Sigma p_j \, dU_{n\,j}/ \, dY_n =] \, \Sigma p_j \, dU_{n\,j}/ \, dC_n = \Sigma p_j \, \beta / C_j \quad (10)$$

8 When the model is a two-level nested logit (such as the LMS mode-destination choice models for most purposes), the time coefficient needs to be multiplied by the logsum coefficient (the differential of the logsum with respect to time is $\alpha_t.\gamma$, where γ is the logsum or tree coefficient that needs to be between 0 and 1 for global consistency with utility maximisation).

This calculation (10) needs to be made over all alternatives but the problems arise when the cost is zero (intrazonal) or non-existent (slow modes and car passenger).

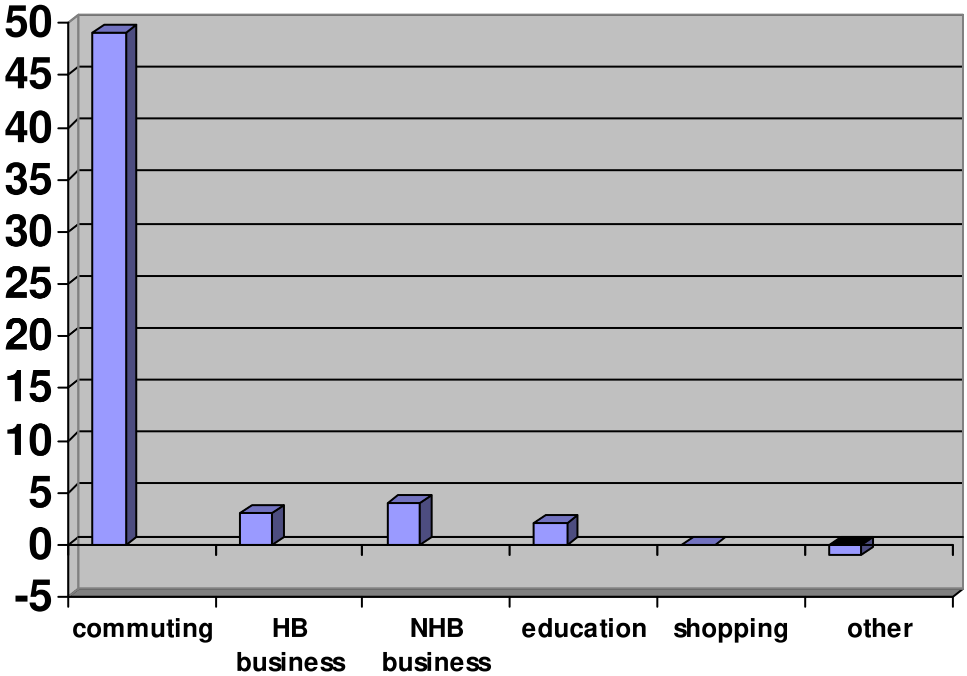
For the monetised logsum we now get:

$$\text{Logsum in guilders} = LS/(\Sigma\, p_j\, \beta /\, C_j) \tag{11}$$

The outcomes for the logsum approach are shown in Table 3 above. The outcomes are in the same range for both ways of monetising the logsum. Method 1 with the current LMS we have different results depending on whether we do the calculations by income group (first value given) or for all incomes at the same time (second value given). Differences between the first and second method are due to the fact that the first method uses external values of time (not from the LMS), whereas the second only uses information from the LMS. If values of time as implied by the LMS coefficients would have been used, both methods would have produced approximately the same total monetary change. Generally speaking the SP values of time are larger than the implied average LMS values of time for commuting for the lowest income classes. Also the SP values of time exceed the LMS values for business travel, shopping and other purposes. For commuting for the higher income classes (these are important categories for train travel), the SP values of time are lower than those implied by the LMS.

The project benefits for travellers (according to method 2) by travel purpose are illustrated in Figure 1.

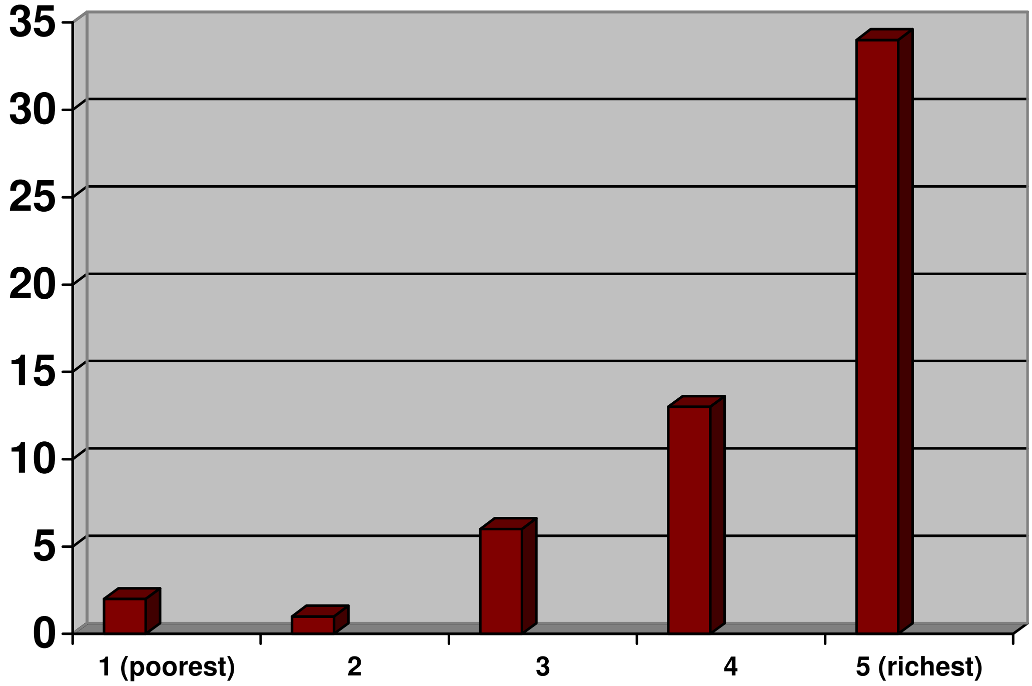
Figure 1. Distribution of logsum change for Rondje Randstad over purposes (in millions of 2003 Euros for the full year 2020)



It is clear that most of the benefits of Rondje Randstad are accruing to commuters, with business travellers second and travellers to school/university third. The division over purposes is similar when using Method 1. One should keep in mind that about 58% of the train tours and 55% of the train kilometres (on an average working day) in the LMS are made by commuters.

Furthermore, the higher incomes groups are enjoying most of the benefits, as can be seen from Figure 2. Again the figure is for the second, but for both method 1 the distribution over income groups is very similar. Allowing for projected income growth, the share of the highest income group in total train tours in these LMS forecasts is 55% and it is also 55% in total train kilometres.

Figure 2. Distribution of logsum change for Rondje Randstad over income groups[9] (in thousands of 2003 Euros for the full year 2020)



The logsum outcomes for the two monetisation methods are similar (see Table 3) to those using the in train in-vehicle time and access/egress time gains and SP values of time ('classic' method). This is just a coincidence. The logsums take into account the changes in all components of the utility functions: in-vehicle time, access/egress times, but also wait time. In the logsum approach, the LMS time coefficients are used to go from gains in utils to gains in minutes. These coefficients are not consistent with the SP results used in the 'classic' method. Most of the values of time from the SP are higher than those of the LMS (at the average costs). Substantial differences can be found especially for shopping, other purposes and for business (the LMS value of time only includes the traveller's value of time, whereas the values used in the official cost-benefit analyses include for business travel both the employee's and the employer's value of time).

[9] The income bands for net annual household income are: 0-€11300; €11300-18200; €18200-29500; €29500-38600; >€38600.

6. Summary and recommendations

6.1 Summary

At present, the method used in The Netherlands for quantifying the benefit for travellers of a transport project consists of calculating the change in consumer surplus (in terms of a reduction of generalised travel costs) for both the current users of the directly affected alternative and for new users. For the latter group the rule-of-a-half is used. This procedure has a basis in welfare analysis. For projects at a national scale, the LMS is often used to produce demand changes and the resulting benefits in travel time and costs. For regional projects, the NRM (new regional models) are regularly used, which use essentially identical demand functions as the LMS. In this study, an alternative approach is taken: instead of consumer surplus in terms of generalised costs the "logsum" is used to calculate user benefits.

The theory on the use of the logsum change as a measure of the change in the consumer surplus, to be used in project appraisal, was published in the late seventies and early eighties. Nevertheless, the application of this theory in practical appraisals of transport projects has been very limited, and most applications in transport evaluation that the authors are aware of have been undertaken only recently (after 2000). It is not easy to find the reasons for the inertia to use the theory in applied work. To some extent it can be related to the complexity of some of the theoretical literature, but the basic logsum concept (with constant marginal utility of money) is fairly straightforward to apply. It may also have to do with the fact that in many countries there is no (national) model system based on disaggregate random utility models. For the computation of logsum changes, disaggregate Generalised Extreme Value (GEV) models, such as the multinomial logit and the nested logit, are required, although in the EXPEDITE project it proved possible to go back from a more aggregate model to the implied underlying utility models. National disaggregate transport models are in use in Scandinavia, the Netherlands and Italy and regional and urban models using these concepts can be found in the same countries, France, the United Kingdom, Australia, Israel and especially the United States. It is therefore not surprising that the logsums applications in evaluation took place in the USA, Scandinavia and The Netherlands. It is unlikely that the computer run times for the calculation have been a major obstacle for the use of logsums in evaluation, since all the required inputs are already computed in the standard procedures for application of disaggregate models (calculation of individual probabilities in sample enumeration).

All applications reviewed use models that include mode choice. Some logsum applications in project evaluation also use models for destination choice and/or departure time choice. The applications of the logsum concept in transport project appraisal all use the relatively simple formulation with constant marginal utility of money. It could be dangerous to assume that the marginal utility of income would be constant over a wide income range (it is more likely that it will decline with increasing income). Theory has moved beyond that in the nineties, but the later formulations are not in practical use. Similarly, recent work on taste variation in policy variables has not become practical for application studies.

A case study in the logsums as an evaluation measure was carried out with the Dutch National Model System (LMS), for a rail project in the Randstad area ('Rondje Randstad'). We applied two different ways of monetisation of the logsum change in utils: method 1 that uses SP values of time and method 2 that uses the expectation of (1/cost).

For the rail project studied, we found that the application of the conventional rule-of-a-half approach leads to very different results depending on whether only the train in-vehicle time changes or also the access/egress time changes are taken into account. The logsum results for this project also vary between the two different monetisation methods which were tested, but the differences in outcomes are rather small. Most of the project benefits accrue to commuters and the highest income group, who each make more than half of the train tours and kilometres.

6.2 Recommendations

We think that replacing this approach by the logsum approach would provide a number of advantages:

- When using logit models as in the LMS, the logsum change also gives the change in the consumer surplus, and in a more exact way than the rule-of-a-half does, since it is based on a linearisation.
- At present there is an inconsistency in the evaluation procedure: for calculating the changes in travel demand. The LMS is used, which has its own set of implied values

of time. Then the resulting time changes are monetised using a different set of values of time (from Stated Preference surveys, SP). When using logsums we can avoid the use of external values of time (except in a method, which we called 'Method 1', of monetisation that expresses the logsum change in minutes first, and then through SP values of time in money). On the other hand, the SP studies might contain information that the LMS is lacking and it would be even better to estimate the transport demand models on a combination of the available Revealed Preference (RP) and SP data.

- The logsum method might seem to be much more complicated than the rule-of-a-half, but in fact a major advantage of logsums is the ease of calculation. Particularly when several alternatives are changing, e.g. in a destination and time period choice when traffic is reassigned in response to a project, the rule-of-a-half calculations can get very complicated while the logsum ones are easy and need to be done anyway to get demand. The logsum method can also easily give results per population group (the conventional approach can do this as well, but this is often a lot of extra work).
- The rule-of-a-half cannot deal with situations in which the number of choice alternatives changes (e.g. when a new mode is introduced), whereas the logsum approach can.

An advantage of the conventional approach is that it is more transparent (but only in simple situations) and more intuitive and therefore easier to explain to non-experts. On the other hand, the transport models that produce the logsums are already common practice.

We believe the advantages of the logsum approach can outweigh the disadvantages, but we also think that further testing of the logsum method is required. We recommend testing the logsum approach for other schemes, especially for highway schemes (or a combined road and public transport scheme) where the purpose and income mix is likely to be more representative of the travelling population as a whole and where the substitution effects might be more important. The monetisation of logsum changes through external values of time (as in Method 1) does not seem attractive, because that would reintroduce the consistency problem. Therefore at this stage we would prefer to use the monetisation of the logsum with the expectation of (1/cost) (Method 2).

References

Bates, J. (2003) Economic evaluation and transport modelling: theory and practice, resource paper for 10th International Conference on Travel Behaviour Research, Lucerne.

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

Castiglione, J., Freedman, J. and Davidson W. (2003) Application of a tour-based microsimulation model to a major transit investment, San Francisco County Transportation Authority and PBConsult, San Francisco.

Cherchi, E., Polak, J. and Hyman, G. (2004) The impact of income, tastes and substitution effects on the assessment of user benefits using discrete choice models, presented to *European Transport Conference*, Strasbourg.

Cochrane, R. A. (1975) A possible economic basis for the gravity model, *J. Transport Economics and Policy*.

CPB and NEI (2000) *Evaluatie van infrastructuurprojecten: leidraad voor kosten-batenanalyse* (Evaluation of infrastructural projects: guide for cost-benefit analysis), CPB/NEI, The Hague/Rotterdam.

Daly, A.J. (2000) National models, in D.A. Hensher and K.J. Button (eds.), *Handbook of Transport Modelling*, Elsevier Science, Oxford.

Daly, A. (2004) Properties of random utility models of consumer choice, presented to TraLog Conference, Molde.

Daly, A. and Bierlaire, M. (2006) A general and operational representation of GEV models, *Transportation Research B* 40, pp. 285-305, May.

Daly, A.J. and S. Zachary (1978) Improved Multiple Choice Models, in D.A. Hensher and M.Q. Dalvi (eds.), *Determinants of Travel Choice*, Saxon House.

Ecorys Transport and 4cast (2004) Verbeteren van bepaling directe effecten uit LMS/NRM output, naar een efficiëntere en betere werking van modeloutput (Improving the determination of direct effects from LMS/NRM output, towards a more efficient and better use of model output), eindrapport voor AVV, Ecorys, Rotterdam.

EXPEDITE Consortium (2002) EXPEDITE Final Publishable Report, Report for European Commission, DGTREN, RAND Europe, Leiden.

Gunn, H.F. (1998) An overview of European national models, presented at the Seminar on National Transport Models, Stockholm.

Gupta, S., Kalmanje S., and Kockelman, K.M. (2004) Road Pricing Simulations: Traffic, Land Use and Welfare Impacts for Austin, Texas. Paper submitted for presentation at the 84th Annual Meeting of the Transportation Board and for Publication *Transportation Research Record*.

Herriges, J. A. and Kling, C. L. (1999) Nonlinear income effects in random utility models, *Review of Economics and Statistics*, 81 (1), pp. 62–72.

Johansson, P.-O. (1993) *Cost-benefit analysis of environmental change*, Cambridge University Press, Cambridge.

Kalmanje, S. and Kockelman, K.M. (2004), Credit-Based Congestion Pricing: Travel, Land Value and Welfare Impacts, Paper to be presented at the 2004 Annual Meeting of the Transportation Board and recommended for Publication in *Transportation Research Record.*

Karlström, A. (2000) Non-linear value functions in random utility econometrics, presented to IATBR Conference, Gold Coast, Australia.

Koopmans, C.C. and Kroes, E.P. (2004) Werkelijke kosten van files tweemaal zo hoog (The real cost of queues twice as high), article in ESB April 2004, pp. 154-155

McFadden, D. (1978) Modelling the choice of residential location. In Karlqvist, A., Lundqvist, L., Snickars, F. and Weibull, J. (eds) *Spatial Interaction Theory and Residential Location*. North-Holland, Amsterdam.

McFadden, D. (1981) Econometric models of probabilistic choice. In Manski, C. and McFadden, D. (eds) *Structural Analysis of Discrete Data: With Econometric Applications*. The MIT Press, Cambridge, Massachusetts.

McFadden, D. (1996) On computation of willingness-to-pay in travel demand models, Dept. of Economics, University of California, Berkeley (unpublished).

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

Odeck, J., Rekdal, J. and Hamre, T.N. (2003) The Socio-economic benefits of moving from cordon toll to congestion pricing: The case of Oslo. Paper presented at *TRB Annual meeting.*

Raad, P.M. de (2004) Congestion costs closer examined, Monetarising adaptive behaviour using the Logsum-method, MSc Thesis, TU Delft.

RAND Europe (M. Pieters and B. Zondag) (2004) Voorstudie koppeling LMS en TIGRIS XL (Pre-study interaction between LMS and TIGRIS XL), commissioned by AVV, RAND Europe, Leiden.

RAND Europe (G. de Jong, M. Pieters, A. Daly, I. Graafland, E. Kroes and C. Koopmans) (2005) Using the logsum as an evaluation measure: literature and case study, Report for AVV, RAND Europe, Leiden.

Train, K. (2003) Discrete choice methods with simulation; Cambridge University Press, Cambridge

US Department of Transportation, Federal Highway Administration (2004?) Case Study: Sacramento, California Methodology Calculation of User Benefits, USDoT, http://www.fhwa.dot.gov/planning/toolbox/sacramento_methodology_user.htm.

von Haefen, R. H. (2003) Incorporating observed choice into the construction of welfare measures from random utility models, *J. of Environmental Economics and Management*, 45, pp. 145-165.

Williams, H.C.W.L. (1977) On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, 9 (3), pp. 285-344.